The Far-Infrared Radio Correlation at low radio frequency with LOFAR/*H*-ATLAS

S. C. Read [1], D. J. B. Smith [1], G. Gürkan [1,2] M. J. Hardcastle [1] W.
L. Williams [1], P. N. Best,[3] E. Brinks [1], G. Calistro-Rivera [4], K. T. Chyży,[5]
K. Duncan [4], L. Dunne,[3,6] M. J. Jarvis,[7,8] L. K. Morabito [7], I. Prandoni [9], H. J.
A. Röttgering,[4] J. Sabater [3] and S. Viaene[1,10]

[1]*Centre for Astrophysics Research, School of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield, Herts, AL10 9AB, UK*
[2]*CSIRO Astronomy and Space Science, 26 Dick Perry Avenue, Kensington, Perth, 6151, WA, Australia*
[3]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK*
[4]*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands*
[5]*Astronomical Observatory, Jagiellonian University, ul. Orla 171, 30-244 Kraków, Poland*
[6]*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, UK*
[7]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, England*
[8]*Physics and Astronomy Department, University of the Western Cape, Bellville 7535, South Africa*
[9]*INAF-IRA, Via P. Gobetti, 101 40129 Bologna, Italy*
[10]*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281, B-9000 Gent, Belgium*

ABSTRACT

The radio and far-infrared luminosities of star-forming galaxies are tightly correlated over several orders of magnitude; this is known as the far-infrared radio correlation (FIRC). Previous studies have shown that a host of factors conspire to maintain a tight and linear FIRC, despite many models predicting deviation. This discrepancy between expectations and observations is concerning since a linear FIRC underpins the use of radio luminosity as a star-formation rate indicator. Using LOFAR 150 MHz , FIRST 1.4 GHz , and *Herschel* infrared luminosities derived from the new LOFAR/*H*-ATLAS catalogue, we investigate possible variation in the monochromatic (250 μm) FIRC at low and high radio frequencies. We use statistical techniques to probe the FIRC for an optically selected sample of 4082 emission-line classified star-forming galaxies as a function of redshift, effective dust temperature, stellar mass, specific star formation rate, and mid-infrared colour (an empirical proxy for specific star formation rate). Although the average FIRC at high radio frequency is consistent with expectations based on a standard power-law radio spectrum, the average correlation at 150 MHz is not. We see evidence for redshift evolution of the FIRC at 150 MHz, and find that the FIRC varies with stellar mass, dust temperature, and specific star formation rate, whether the latter is probed using MAGPHYS fitting, or using mid-infrared colour as a proxy. We can explain the variation, to within 1σ, seen in the FIRC over mid-infrared colour by a combination of dust temperature, redshift, and stellar mass using a Bayesian partial correlation technique.

Key words: galaxies: star formation – infrared: galaxies – radio continuum: galaxies.

1 INTRODUCTION

The far-infrared luminosities of star-forming galaxies have long been known to correlate tightly and consistently with synchrotron radio luminosity across many orders of magnitude in infrared and radio luminosities, independent of galaxy type and redshift (van der Kruit 1971; de Jong et al. 1985; Condon, Anderson & Helou 1991; Yun, Reddy & Condon 2001; Bell 2003; Bourne et al. 2011).

The existence of some relation should not be surprising since the basic physics relating the emission in each waveband to the presence of young stars is well understood. Young stars heat the dust within their surrounding birth clouds, which radiate in the infrared

(Kennicutt 1998; Charlot & Fall 2000). The supernovae resulting from the same short-lived massive stars accelerate cosmic rays into the galaxy's magnetic field thereby contributing non-thermal radio continuum emission over $\approx 10^8$ yr (Blumenthal & Gould 1970; Condon 1992; Longair 2011). However, the fact that the Far-Infrared Radio Correlation (FIRC) has consistently been found to have low scatter (de Jong et al. 1985; Helou, Soifer & Rowan-Robinson 1985; Condon 1992; Lisenfeld, Voelk & Xu 1996b; Wong et al. 2016) is surprising. Such tight linearity is consistent with a simple calorimetry model (Voelk 1989), whereby cosmic ray electrons lose all of their energy before escaping the host galaxy and where all UV photons are absorbed by dust and re-radiated in the infrared. This results in synchrotron radiation being an indirect measure of the energy of the electron population and infrared luminosity being proportional to young stellar luminosity. Therefore, assuming calorimetry, the ratio of these two measures will remain constant as they are both dependent on the same star formation rate. The FIRC can therefore be used to bootstrap a calibration between a galaxy's star formation rate and its radio luminosity (e.g. Condon 1992; Murphy et al. 2011), but only if there is no additional contribution from AGN.

The physics required to model the FIRC is complex. For example, the timescale of the electron synchrotron cooling that produces the radio emission is thought to be longer than the timescale for the escape of those electrons (Lisenfeld, Voelk & Xu 1996a; Lacki, Thompson & Quataert 2010) for normal spirals, and starlight is only partially attenuated in the UV (Bell 2003). Therefore, it is reasonable to suppose that the calorimetry interpretation must be at least partially inaccurate and that there should be some observable variation in the FIRC over the diverse population of star-forming galaxies. In particular, due to their strong magnetic fields, we expected starburst galaxies to be good calorimeters and therefore have a correlation with a slope that is much closer to one than other star-forming galaxies (Lacki et al. 2010).

However, since synchrotron emission depends strongly on magnetic field strength, the assumption about how this changes with galaxy luminosity is crucial to explain the correlation. Alternatives to the calorimetry model have also been proposed, e.g. (i) the model of Niklas & Beck (1997), where the FIRC arises as the by-product of the mutual dependence of magnetic field strength and star-formation rate upon the volume density of cool gas, and (ii) Schleicher & Beck (2016), where the FIRC is based on a small-scale dynamo effect that amplifies turbulent fields from the kinetic turbulence related to star formation. There are a number of reasons to expect the FIRC to vary with the parameters that control synchrotron and dust emission, but it seems that both infrared and radio synchrotron must fail as star formation rate indicators in such a way as to maintain a tight and linear relationship over changing gas density. The model detailed by Lacki & Thompson (2010) and Lacki et al. (2010) suggests that although normal galaxies are indeed electron and UV calorimeters, conspiracies at high and low surface density, Σ_g, contrive to maintain a linear FIRC. At low surface density, many more UV photons escape (and therefore lower observed infrared emission) due to decreased dust mass but at the same time, because of the lower gravitational potential, more electrons escape without radiating all their energy, decreasing the radio emission. Meanwhile, at high surface densities, secondary charges resulting from cosmic ray proton collisions with ISM protons become important (Torres 2004; Domingo-Santamaria & Torres 2005). Synchrotron emission from those electrons and positrons may dominate the emission from primary cosmic ray electrons. However, the FIRC is maintained due to the increased non-synchrotron losses from bremsstrahlung and inverse Compton scattering at higher densities.

These conspiracies rely on fine-tuning of many, sometimes poorly known, parameters in order to balance the mechanisms that control the linearity of the FIRC. If we expect variation over star-forming galaxies due to differences in gas density, stellar mass, and redshift (to name a few), then we should probe the FIRC over known star-forming sequences such as those found in colour-magnitude (Bell et al. 2004) and mid-infrared colour–colour diagrams (e.g. Jarrett et al. 2011; Coziol, Torres-Papaqui & Andernach 2015), and the star formation rate – stellar mass relation (Brinchmann et al. 2004; Noeske et al. 2007; Peng et al. 2010; Rodighiero et al. 2011).

Naively, we might also expect some variation of the FIRC with redshift. At the very least, radio luminosity should decrease with respect to infrared luminosity due to inverse Compton losses from cosmic microwave background (CMB) photons (Murphy 2009). The CMB energy density increases proportional to $(1 + z)^4$ (Longair 2011), so the ratio of infrared to radio luminosity should noticeably increase with redshift even at relatively local distances, assuming a calorimetry model and that CMB losses are significant.

However, this is one of the key areas of dispute between different observational studies. While the many works find no evidence for evolution (e.g. Garrett 2002; Appleton et al. 2004; Seymour et al. 2009; Sargent et al. 2010), there are exceptions (e.g. Seymour et al. 2009; Ivison et al. 2010; Michałowski, Watson & Hjorth 2010b; Michałowski, Hjorth & Watson 2010a; Basu et al. 2015; Calistro-Rivera et al. 2017; Delhaize et al. 2017). Particular among those studies, Calistro-Rivera et al. (2017) find a significant redshift trend at both 150 MHz and 1.4 GHz when using the *Low Frequency Array* (LOFAR ; van Haarlem et al. 2013) data taken over the Boötes field. The FIRC has been studied extensively at 1.4 GHz (de Jong et al. 1985; Condon et al. 1991; Bell 2003; Jarvis et al. 2010; Bourne et al. 2011; Smith et al. 2014) but rarely at lower frequencies. These low frequencies are particularly important, since new radio observatories such as LOFAR are sensitive in the $15 - 200$ MHz domain, where at some point the frequency dependence of optical depth results in the suppression of synchrotron radiation by free–free absorption (Schober, Schleicher & Klessen 2017), causing the radio SED to turn over. As a result, there will be some critical rest-frame frequency below which we can expect a substantially weaker correlation between a galaxy's radio luminosity and its star formation rate.[1] Moreover, at the higher frequencies probed by *Faint Images of the Radio Sky at Twenty centimetres* (FIRST; Becker, White & Helfand 1995) (1.4 GHz), there may be a thermal component present in the radio emission (Condon 1992), which tends to make the correlation between infrared and higher radio frequencies more linear. However, due to the poor sensitivity of FIRST to star-forming galaxies with low brightness temperatures (galaxies with $T_{bright} < 10$ K will not be detected by FIRST), we cannot expect the thermal components of detected sources to help linearize the FIRC at 1.4 GHz. At low frequencies, these effects become less important and so the perspective they provide is useful in disentangling the effect of thermal contributions and lower luminosity galaxies on the FIRC. Given the potential ramifications for using low-frequency radio observations as a star formation indicator, this possibility must be investigated. Indeed, Gürkan et al. (2018) have found that a broken power-law is a better calibrator for radio continuum luminosity to star-formation rate, implying the existence of some other

---

[1]This frequency at which a galaxy's radio SED turns over will depend heavily upon gas density and ionization, and so we expect it to vary from galaxy to galaxy.

additional mechanism for the generation of radio-emitting cosmic rays.

Furthermore, lower radio frequencies probe lower-energy electrons, which take longer to radiate away their energy than the more energetic electrons observed at 1.4 GHz, and this results in a relationship between the age of a galaxy's electron population and the radio spectral index (Scheuer & Williams 1968; Blundell & Rawlings 2001; Schober et al. 2017). Therefore, even if the FIRC is linear at high frequencies due to some conspiracy, this will not necessarily be the case at low frequencies. An investigation of the FIRC at low frequency will test models of the FIRC that rely on spectral ageing to maintain linearity (e.g. Lacki et al. 2010).

Combined with the fact that radio observations are impervious to the effects of dust obscuration, this makes low-frequency radio observations a very appealing means of studying star formation in distant galaxies, providing that the uneasy reliance of SFR estimates on the FIRC can be put on a more solid footing. The nature of the FIRC conspiracies varies over the type of galaxy and its star formation rate (Lacki & Thompson 2010). The detection of variation in the FIRC over those galaxy types, or lack thereof, will provide important information about the models that have been constructed (e.g. Lacki & Thompson 2010; Schober et al. 2017). Several methods are used to distinguish galaxy types for the purposes of studying the FIRC, particularly to classify these into star-forming galaxies and AGN such as BPT diagrams (Baldwin, Phillips & Terlevich 1981), panchromatic SED-fitting with AGN components (Berta et al. 2013; Calistro-Rivera et al. 2016; Ciesla et al. 2016), and classification based on galaxy colours. Among these, galaxy colours provide a readily accessible method to distinguishing galaxy types or act as proxies for properties such as star formation rate. Diagnostic colour–colour diagrams are commonplace in galaxy classification; infrared colours in particular have been widely used to distinguish between star-forming galaxies and AGN (Lacy et al. 2004; Stern et al. 2005; Jarrett et al. 2011; Mateos et al. 2012; Coziol et al. 2015). In order to investigate the potential difference in the FIRC over normal galaxies as well as in starbursts, we use the Mid-Infrared Diagnostic Diagram (MIRDD; Jarrett et al. 2011). Constructed from the *Wide-field Infrared Survey Explorer* (WISE; Wright et al. 2010) [4.6] − [12] and [3.4] − [4.6] colours, SWIRE templates (Polletta et al. 2006, 2007) and GRASIL models (Silva et al. 1998) can be used to populate the MIRDD with a range of galaxy types spanning a redshift range of $0 < z < 2$. This MIRDD not only distinguishes AGN and SFGs but also describes a sequence of normal star-forming galaxies whose star formation rate increases to redder colours.

Past 1.4 GHz surveys such as FIRST and the *NRAO VLA Sky Survey* (NVSS; Condon et al. 1998) have been extremely useful in studying star formation, though there are inherent problems in using them to do this. NVSS is sensitive to extended radio emission on the scale of arcminutes. However, its sensitivity of ∼0.5 mJy beam^{-1} and resolution of 45 arcsec means that it has trouble identifying radio counterparts to optical sources and its flux limit means that it will peferentially detect bright or nearby sources. FIRST has both a higher resolution and a higher sensitivity than NVSS (5 arcsec with ∼0.15 mJy beam^{-1}). However, due to a lack of short baselines, FIRST resolves out the extended emission frequently present in radio-loud AGN and in local star-forming galaxies (Jarvis et al. 2010). This makes it difficult to remove galaxies dominated by AGN and to directly compare star-forming galaxies over different wavelengths. Meanwhile, LOFAR offers the best of both worlds: a large field of view coupled with high sensitivity on both small and large scales and high resolution (van Haarlem et al. 2013) at

frequencies between 30 and 230 MHz. Operating at 150 MHz, LOFAR contributes a complementary view to the wealth of data gathered at higher frequencies. The sparsely examined low-frequency regime offered by LOFAR combined with its increased sensitivity and depth relative to other low-frequency instruments allows us to probe the FIRC in detail, and to test predictions of its behaviour relative to relations at higher frequencies that we measure with FIRST.

This study will analyse the nature of the FIRC at low and high frequencies and over varying galaxy properties. How does the FIRC evolve with redshift? Does it vary as a function of WISE mid-infrared colour? Do the specific star-formation rate (as fit by MAGPHYS) and stellar mass impact these questions? We answer these questions for our data set and compare these metrics with those found at higher frequencies and with literature results using different selection criteria.

This work uses the same base data set as Gürkan et al. (2018). The same aperture-corrected fluxes extracted from *Herschel*, LOFAR, and FIRST images are used here. Our investigation differs from theirs in that we concentrate on the observed variation of the FIRC over dust properties whereas Gürkan et al. (2018) focus on the direct characterization of radio star-formation rates. In Section 2, we describe our data sources and the method of sample selection. In Section 3 we outline our methods for calculating *K*-corrections, luminosities, and the methods used to characterize the variation of the FIRC. We present and discuss the results of these procedures in Section 4 and summarize our conclusions in Section 5.

We assume a standard ΛCDM cosmology with $H_0 = 71 \,\mathrm{km \, s^{-1} \, Mpc^{-1}}$, $\Omega_M = 0.27$, and $\Omega_\Lambda = 0.73$ throughout, and for consistency with Jarrett et al. (2011), all magnitudes are in the Vega system.

2 DATA SOURCES

The data set we use here is the same as Gürkan et al. (2018) in that the infrared and radio aperture-corrected fluxes are drawn from the same catalogue. However, due to two effects listed below, our star-forming sample is selected using a different method. First, a potential contamination of AGN will have a large effect on the detected variation of infrared-to-radio luminosity ratio over mid-infrared colours. We therefore require stronger signal-to-noise criteria (5σ detections in the BPT optical emission lines) than the one in use in Gürkan et al. (2018) (3σ). Secondly, using the Gürkan et al. (2018) star-forming selection criterion but with a 5σ requirement results in too few star-forming galaxies with reliable 5σ detections in the first three WISE bands. In order to increase our sample size but maintain robust classification, we employ methods detailed below.

2.1 Sample selection

To avoid introducing a possible bias by selecting our sample from far-infrared and/or radio catalogues, our sample is drawn from the MPA-JHU catalogue (Brinchmann et al. 2004) over the region of the *North Galactic Pole* (NGP) field covered by the LOFAR/*H*-ATLAS survey, which is described in Sections 2.2 and 2.3. The MPA-JHU catalogue uses an optimized pipeline to re-analyse all SDSS (York 2000) spectra, resulting in a sample with reliable spectroscopic redshifts, improved estimates of stellar mass, and star formation rate, as well as emission line flux measurements for each galaxy. We use their latest analysis performed on the SDSS DR7 release (Abazajian et al. 2009) to obtain optical emission line fluxes and spectroscopic redshifts for *K*-corrections.
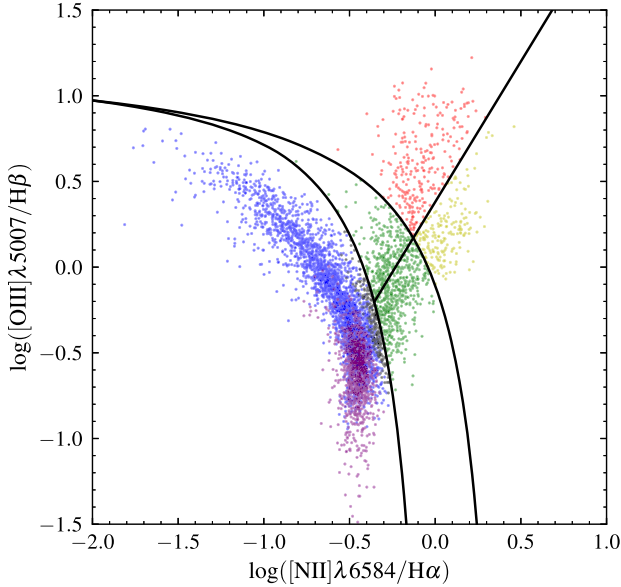
Figure 1. Emission line ratio diagnostic diagram. The coloured points represent Seyfert 2s (in red), star-forming galaxies (blue), transition objects (green), and LINERs (yellow). The black points are those galaxies whose 5σ upper limit on $[OIII]\lambda5007$ flux would not classify them as purely star-forming (not included in our sample). The purple points show those additional galaxies whose upper limits in $[OIII]\lambda5007$ still classify them as star-forming. The upper and lower solid black lines used to distinguish between populations are from Kewley et al. (2001) and Kauffmann et al. (2003), respectively.

To select our star-forming sample, we first obtain all optically selected 15003 sources in the MPA-JHU catalogue with reliable (ZWARNING = 0) spectroscopic redshifts $z < 0.7$ in the region covered by our LOFAR/H-ATLAS data. Since we are interested in studying the FIRC, we wish to focus only on star-forming galaxies, and remove those sources with evidence for contamination by emission from an active galactic nucleus (AGN). Our priority is to seek an unbiased sample at the cost of such a sample not necessarily being complete. We do this using the BPT (Baldwin et al. 1981) emission line classification method, requiring fluxes detected at $\geq5\sigma$ in $H\alpha$, $H\beta$, $[OIII]\lambda5007$, and $[NII]\lambda6584$, together with the star-forming/composite line defined by Kewley et al. (2001). 3082 galaxies, with redshifts $z < 0.4$, are identified as star-forming in this manner.

To give us the largest possible sample of star-forming galaxies, we include those galaxies with 5σ detections in $[NII]\lambda6584$, $H\alpha$, and $H\beta$, provided that the upper limit on the $[OIII]\lambda5007$ flux in the MPA-JHU catalogue enables us to unambiguously classify them as star-forming. By using this method, we can be sure that they lie below the star-forming/composite line from Kewley et al. (2001) in Fig. 1. We identify an additional 1012 star-forming galaxies using this criterion, and they are shown in purple in Fig. 1. In addition, we remove the 12 sources which lie within the QSO box defined in Jarrett et al. (2011). This provides us with our main sample of 4082 star-forming galaxies with $z < 0.4$ for use in comparing the FIRC at high and low frequencies.

We constructed the MIRDD (Jarrett et al. 2011) based on *WISE All Sky Survey* (WISE; Cutri 2012) fluxes (with no *K*-correction applied) to identify the location of our galaxies compared to a range of sources of different types. Since we are binning across WISE colour spaces, we construct a second sample for the
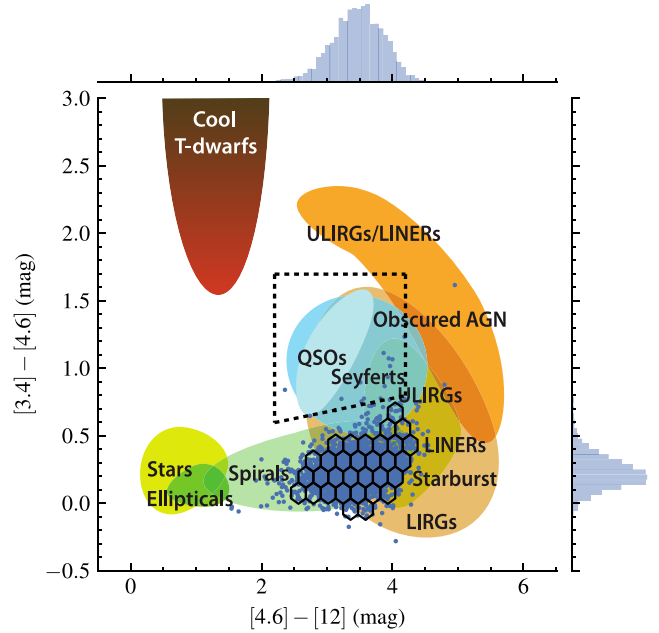

Figure 2. The mid-infrared sub-sample (5σ WISE detections, shown as blue points) overlaid on the Jarrett et al. (2011) MIRDD that uses the magnitudes at three WISE wavelengths $W1$ at 3.4 μm, $W2$ at 4.6 μm, and $W3$ at 12 μm. The coloured regions are as published in Wright et al. (2010), and intended to show the approximate locations of galaxies of a range of different types. The hexagonal bins over the region centred on $[4.6] - [12] \approx3.5$, and $[3.4] - [4.6] \approx0.25$ are used to trace q_{250} in later sections of this paper, and are shown here to provide context. The QSO box defined by Jarrett et al. (2011) is depicted as a dashed box. Number counts over both colours are shown as blue histograms.

Table 1. Number of star-forming galaxies within each sub-sample detected with *Herschel* at 250 μm, LOFAR at 150 MHz, and FIRST at 1.4 GHz.

		1. All SFGs	2. WISE detected SFGs
$>3\sigma$	*Herschel*	3351	2673
	LOFAR	2436	2016
	FIRST	1438	1098
	Both radio bands	1008	863
$>5\sigma$	*Herschel*	2616	2209
	LOFAR	1876	1627
	FIRST	835	640
	Both radio bands	533	455
Total		4082	2901

mid-infrared analysis only, requiring 5σ detections in the first three WISE bands (centred on 3.4, 4.6, and 12 μm). This results in a sub-sample of 2901 sources for use in tracing the FIRC over the mid-infrared colour space depicted in Fig. 2. Our sample sizes are shown in Table 1.

We do not use the catalogue of detected sources summarized by Table 1 for our analysis here. Such a catalogue will inevitably become contaminated with noise spikes. Instead, we employ averaging techniques described below in order to treat non-detections and detections in the same manner. We don't make any signal-to-noise cuts beyond those imposed on the BPT emission lines used in the star-forming classification. In addition, our samples are drawn from the MPA-JHU catalogue and so this imposes a strong optical prior on the location of a given source. This allows us to conduct

forced aperture photometry, in order to estimate radio fluxes (see Section 2.4), for our entire sample with a high degree of confidence that the aperture is correctly placed.

2.2 Infrared data

The far-infrared data used in this study come from the *H*-ATLAS survey (Eales et al. 2010; Valiante et al. 2016; Smith et al. 2017b; Furlanetto et al. 2018; Maddox et al. 2018). *H*-ATLAS is the largest extragalactic *Herschel* survey, covering a total of 510 deg^2 in five infrared bands with the *Photoconductor Array Camera and Spectrometer* (PACS; Ibar et al. 2010; Poglitsch et al. 2010) and *Spectral and Photometric Imaging Receiver* (SPIRE ; Griffin et al. 2010; Pascale et al. 2011; Valiante et al. 2016) instruments (sampling wavelengths of 100, 160, 250, 350, and 500 μm). The *H*-ATLAS catalogues have a 5σ noise level of 33.5 mJy at 250 μm, which is the most sensitive band (Ibar et al. 2010; Rigby et al. 2011; Smith et al. 2011, 2012b; Smith et al. 2017a). In this study, we focus on the *H*-ATLAS observations covering 142 deg^2 of the NGP field.

2.3 LOFAR data from LOFAR/*H*-ATLAS

LOFAR has observed the *H*-ATLAS NGP field at the sensitivity and resolution of the LOFAR Two-Metre Sky Survey (LoTSS; Shimwell et al. 2017; Duncan et al. in preparation; Williams et al. in preparation). Whilst the first implementation of the LOFAR/*H*-ATLAS survey Hardcastle et al. (2016) used a facet-calibration technique, this paper uses data calibrated by a significantly improved method. The new direction-dependent calibration technique uses the methods of Tasse (2014a,b). The calibrations are implemented in the software package KILLMS and imaged with DDFACET (Tasse et al. 2018), which is built to apply these direction-dependent calibrations. The LOFAR/*H*-ATLAS data were processed using the 2016 December version of the pipeline, DDF-PIPELINE2^2 (Shimwell et al. 2017, in preparation). This reprocessing yields a higher image fidelity and a lower noise level than the process detailed by Hardcastle et al. (2016). It not only increases the point-source sensitivity and removes artefacts from the data, but also allows us to image at (slightly) higher resolution. The images used here (as in Gürkan et al. 2018) have a restoring beam of 6 arcsec FWHM, and 50 per cent of the newly calibrated LOFAR/*H*-ATLAS field has an RMS below ∼0.25 mJy beam^{-1} and 90 per cent is below ∼0.85 mJy beam^{-1}.

2.4 Photometry

Since we used optical data to select our sample, flux limited catalogues from the LOFAR, FIRST, or *H*-ATLAS surveys do not contain photometry for every source in our sample, since some of our sources are not formally detected (e.g. to ≥5σ). Moreover, some sources are larger than the *Herschel* beam and so matched filter images are not preferred. Instead, the data set used here (from Gürkan et al. 2018) follows Jarvis et al. (2010), Smith et al. (2014), and Hardcastle et al. (2016), by measuring LOFAR, FIRST, and *Herschel* flux densities using forced aperture photometry.

In order to have consistent flux densities across radio and infrared bands, we use 10 arcsec radius circular apertures, centred on each source's optical position, finding that this size of aperture is optimal since it is small enough to limit the influence of confusion noise,

---

^{2}See http://github.com/mhardcastle/ddf-pipeline for the code.

and large enough to mean that aperture corrections are small. The uncertainties on both LOFAR and FIRST flux densities were estimated using their respective r.m.s. maps: scaling the noise value in the image at the pixel coordinate of each source by the square root of the number of beams in the aperture. We do not correct for thermal contributions, whereby the thermal SED also contributes at radio frequencies, in FIRST or LOFAR. In the *Herschel* bands, we add the recommended calibration uncertainties in quadrature (5 per cent for PACS and 5.5 per cent for SPIRE) (Valiante et al. 2016; Smith et al. 2017b).

3 METHODS

3.1 Low-frequency luminosities

We calculate *K*-corrected 150 MHz luminosity densities for every source in our sample assuming that $S_\nu \propto \nu^\alpha$, with a spectral index of −0.71 (Condon 1992; Mauch et al. 2013):

$$L_\nu = 4\pi d_L^2(z) S_{\nu,\mathrm{obs}}(1+z)^{-\alpha-1}, \tag{1}$$

where the additional factor of $(1 + z)^{-1}$ accounts for the bandwidth correction, and $d_{\mathrm{L}}(z)$ is the luminosity distance in our adopted cosmology.

There is an additional uncertainty on the *K*-corrected luminosity densities due to assuming a constant spectral index; we attempt to account for this by bootstrapping based on the Mauch et al. (2013) distribution of star-forming spectral indices. For each galaxy we draw 1000 spectral indices from the prior distribution centred on −0.71 with an RMS of 0.38. The luminosity densities are calculated using equation (1) with uncertainties estimated based on the standard deviation of the bootstrapped distribution, however we note that the *K*-corrections and their uncertainties derived for our sample are small since all sources are below $z = 0.4$.

3.2 Far-infrared luminosities

To estimate the intrinsic far-infrared luminosity densities, we assume an optically thin greybody for the dust emission:

$$S_\nu \propto \frac{\nu^{3+\beta}}{\exp{(h\nu/kT)} - 1}, \tag{2}$$

where T is the dust temperature, k is the Boltzmann constant, h is the Planck constant, and β is the emissivity index. The dust emissivity varies as a power law over frequency and its inclusion as the constant β attempts to summarize the varying dust compositions into a single galaxy-wide isothermal component. Taking $\beta = 1.82$ has been found to provide an acceptable fit to the infrared SEDs of galaxies in the *H*-ATLAS survey (Smith et al. 2013) and so we assume the same value for β here. We fit equation (2) to the *Herschel* PACS /SPIRE fluxes at 100, 160, 250, 350, and 500 μm. We include the PACS wavelengths despite their reduced sensitivity since they have been found to be important in deriving accurate temperatures (Smith et al. 2013).

We use the Python package EMCEE (Foreman-Mackey et al. 2013), which is an implementation of the Goodman & Weare (2010) Affine Invariant MCMC Ensemble Sampler (AIMCMC). AIMCMC is known to sample from degenerate and highly correlated posterior distributions with an efficiency superior to traditional Metropolis techniques (Goodman & Weare 2010). For each galaxy, 10 walkers are placed at initial temperatures drawn from a prior normal distribution centred at 30 K with a standard deviation of 100 K. We find

that altering the width of the temperature prior does not affect our results.

The walkers sample the probability distribution set by the least squares likelihood function. At each temperature that the walkers sample, the resultant grey-body is redshifted to the observed frame and propagated through the *Herschel* response curves. We ran the sampler for 500 steps with the 10 walkers and a burn-in phase of 200 steps. Each galaxy therefore has 3000 informative samples to contribute to the probability distributions. In addition to the temperature for each MCMC step, we recorded the modelled intrinsic luminosity densities, modelled observed fluxes, and *K*-corrections for each infrared wavelength. This allowed us to find the probability distributions for these parameters and hence their uncertainties in a Bayesian manner.

3.3 Calculating the FIRC

The FIRC is traditionally parametrized by the log of the ratio of infrared to radio luminosity, q (Helou et al. 1985; Bell 2003; Ivison et al. 2010). However, the lack of PACS $60\,\mu m$ coverage and a small number of sources (<5 per cent) with WISE $22\,\mu m$ fluxes in the *H*-ATLAS NGP field prohibits an accurate estimation of q based on total dust luminosity for a statistically significant sample. Therefore, we calculate a *K*-corrected monochromatic q_{250} in the SPIRE $250\,\mu m$ band following Jarvis et al. (2010) and Smith et al. (2014).

$$q_{250} = \log_{10}\left(\frac{L_{250}}{L_{\rm rad}}\right). \qquad (3)$$

The uncertainties on our monochromatic q_{250} estimates are found by propagating uncertainties from the *K*-corrected luminosity densities in the radio and $250\,\mu m$. We note that in all of the following sections, we calculate q_{250} using $L_{\rm rad}$ calculated at $150\,MHz$ in the rest-frame.

In addition to the individual q_{250} found for each galaxy, we use a stacking method to evaluate trends across colour spaces, redshift, and temperature. Averaging q_{250} is fraught with problems such as underestimation caused by AGN contamination, undesirable influence by outlier sources, and amplification of those effects by using the average of the ratio of luminosity densities rather than the ratio of the average luminosity densities (luminosity stacking). To make matters worse, selection in either the radio or infrared band used to evaluate the FIRC introduces an inherent SED-related bias (Sargent et al. 2010). Here we have mitigated the effects of such biases by selecting in an independent optical band. To mitigate the effect of outliers and AGN, the ratio of the median luminosity densities has previously been used (e.g. Bourne et al. 2011; Smith et al. 2014). Median averaging is sometimes preferred since it is more resistant to outliers (e.g. residual low-luminosity AGN that may not have been identified by the emission line classifications), and since the median often remains well-defined even in the case of few individual detections (e.g. Gott et al. 2001). However, the distributions of luminosity density even in finite-width bins of redshift are skewed. We find that a median-stacked q_{250} calculated for the whole star-forming sample does not agree with the likewise-stacked q_{250} in bins of redshift (in that the median of the medians is not close the global median – this is not the case with the mean). If we use the mean-stacked q_{250}, we arrive at an agreement between the global and binned q_{250} across redshift. Due to this counter-intuitive disagreement between measures of q_{250} and the importance of being able to quantify a change in the FIRC over redshift, we use the ratio of the mean luminosity densities (mean-stacked) to evaluate q_{250}. Although we may sidestep issues regarding skewed distributions by

using the mean, we are now potentially more affected by outliers and AGN. We will discuss the possible influence of AGN on our results in more detail in the coming sections.

To calculate our mean q_{250} values, we use a method similar to Smith et al. (2014) and take the quotient of the mean radio and $250\,\mu m$ luminosity density for each bin. Uncertainties are estimated on each stacked q_{250} using the standard deviation of the distribution resulting from re-sampling this mean 10000 times with replacement (bootstrapping). This bootstrapped uncertainty of q_{250} is representative of the distribution of the luminosity densities being stacked. To complement the parametrization of the FIRC with q_{250}, we also fit the FIRC as a power law with finite intrinsic width[3] to the data using equation (4)

$$L_{\rm radio} = kL_{250}^{\gamma}, \qquad (4)$$

where k is the normalization and γ is the slope of the FIRC. We take into account non-detections by re-sampling from each data point's uncertainty and discarding the negative-value realizations. We use EMCEE to fit the power law with 6000 steps and 32 walkers. Fitting a power law allows us to probe the physical mechanisms of radio continuum emission generation. A value of the slope close to one indicates that the conditions required for calorimetry are satisfied and the FIRC is linear. A super-linear slope might result from an escape-dominated scenario whereby cosmic rays escape before emitting in the radio. At sub-linear slopes, losses from cooling processes such as inverse-Compton dominate (Li et al. 2016).

We have discussed two methods of quantifying the FIRC (mean-stacked q_{250} and power law fit). In addition, there are three types of uncertainty in the FIRC that we discuss in this analysis:

(i) The uncertainty in q_{250}, calculated as the width of the boot-strapped distribution of stacked q_{250}.

(ii) The uncertainty in the slope of the FIRC, γ, quantified by MCMC fit.

(iii) The change in stacked q_{250}, γ, and other statistical results due to the presence of misclassified AGN.

We estimate the change in our results due to misclassified AGN in Section 4.5, where we run our analysis again, this time including the BPT-AGN. This test will be of limited use since BPT-AGN galaxies may not be similar in luminosity nor in temperature to those galaxies that host a low-luminosity AGN. We resort to this method since we are investigating the FIRC itself and so we cannot use the FIRC to distinguish low-luminosity AGN from star-forming galaxies.

4 RESULTS AND DISCUSSION

4.1 Isothermal fits

Before proceeding to investigate the variation of the FIRC with redshift and other parameters, we undertake several checks to ensure that our temperature estimates and *K*-corrections are reliable. As a means of testing goodness-of-fit, we calculate the Gelman-Rubin $\mathcal{R}$ statistic for the sampled temperature and reduced χ^2 for each object. Fig. 3 shows the distributions of $\mathcal{R}$ and reduced χ^2 for our full sample of star-forming galaxies.

---

[3]The intrinsic width of the power-law fit, $\log[\sigma]$, is defined as the logged fractional width of the Gaussian over the power-law line, which we define as $L_{150\rm MHz} \sim (kL_{250\mu m}^{\gamma})(1 + \epsilon)$, where $\epsilon \sim \mathcal{N}(0, \sigma)$. We fit the parameter σ along with γ and k in our MCMC run.
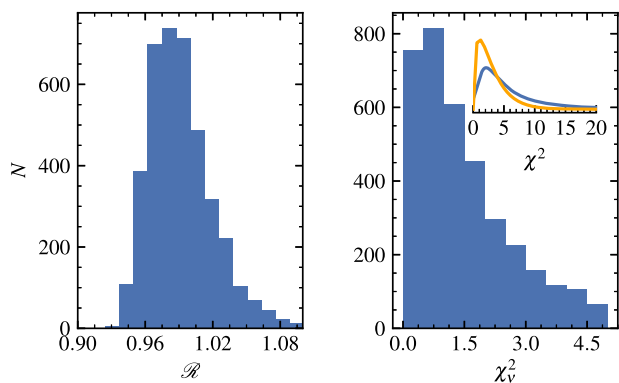
Figure 3. Fit diagnostics for our full star-forming sample. The Gelman–Rubin convergence statistic histogram is shown on the left indicating that all of our fits have converged. The reduced χ^2 distribution of the sample is shown on the right side as the blue histogram. The distribution of χ^2 is also shown in the inset in blue, along with the χ^2 distribution expected for 3 degrees of freedom for comparison in orange.

An $\mathcal{R} \approx 1$ signifies that all chains are sampling from the same distribution and have therefore converged (see Gelman & Rubin 1992, for a full description); all sources in our sample have $0.9 < \mathcal{R} < 1.1$, indicating that the fits have converged.

The χ^2 distribution of our sample, which we fit by least squares regression, has 3 degrees of freedom consistent with our 1-parameter model (normalisation is not fit and is instead optimized with χ^2 minimization) when fitting with five bands of far-infrared observations. In addition, 83 per cent of our total sample have a reduced $\chi^2 < 2$. Conducting this experiment with only those sources with reduced $\chi^2 < 2$ does not affect the conclusions presented here.

Smith et al. (2013) found that median likelihood estimators in greybody fitting are less susceptible to bias with *H*-ATLAS data than the best fit. Therefore, in what follows we adopt the median likelihood value from the MCMC fits as a galaxy's effective temperature for use in equation (2), along with uncertainties estimated according to the 16th and 84th percentiles of the derived distribution. Fig. 4 shows an example fit.

Fig. 5 shows that our sample of emission-line classified star-forming galaxies exhibits a dust temperature distribution centred around ∼23 K with a standard deviation of ∼10 K. The total aggregated temperature probability distribution for all galaxies, also shown in Fig. 5, is slightly wider than the median likelihood temperature histogram. This is due to the fact that the aggregated distribution includes the uncertainty from each galaxy rather than just reporting the average median likelihood temperature.

4.2 The global FIRC at different radio frequencies

To compare the values of q_{250} obtained at 150 MHz and 1.4 GHz, we extrapolate the FIRST luminosity densities to 150 MHz assuming a power law with a spectral index of -0.71. For clarity, we label this transformed q_{250} as q_{150MHz}^{FIRST} to distinguish it from the related quantity at its measured frequency, $q_{1.4GHz}^{FIRST}$. Though it isn't especially instructive due to the large range of redshifts included in our study, we find an average value of $q_{1.4GHz}^{FIRST} = 2.30 \pm 0.04$ (which is equivalent to $q_{150MHz}^{FIRST} = 1.61 \pm 0.04$), which is consistent with previous studies (Ivison et al. 2010; Smith et al. 2014) to within 1σ. We find that the average FIRC is not consistent between low and high radio frequencies, with $q_{150MHz}^{LOFAR} = 1.42 \pm 0.03$ and $q_{150MHz}^{FIRST} = 1.61 \pm 0.04$.
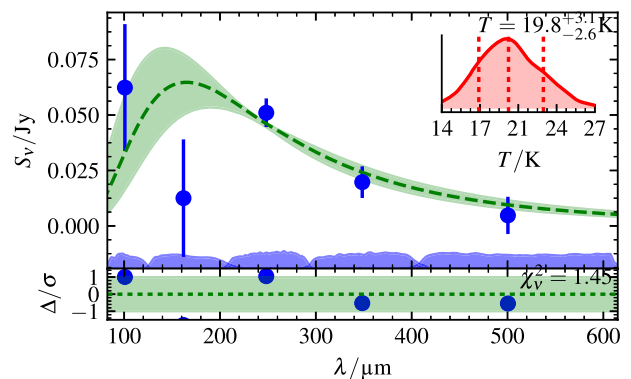

Figure 4. A sample isothermal fit to an SDSS star-forming galaxy at $\alpha_{J2000} = 12^h49^m46.1^s$, $\delta_{J2000} = 31°35'30''$. The probability distribution for temperature is shown in the top right with the 1σ equal-tailed credible interval as dashed lines around the median temperature of 19.8 K with a reduced χ^2 of 1.45. The Herschel flux measurements and their uncertainties are shown as blue errorbars. The fit observed-frame isothermal greybody with its own 1σ credible interval is shown as the green curve. The differences between the estimated flux and the measurement are shown along the bottom axis. The filter transmission profiles are also shown in blue along the bottom for each wavelength.
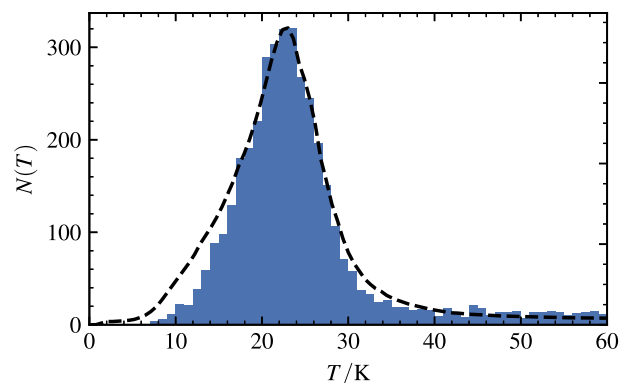

Figure 5. The distribution of median temperatures for our sample of emission-line classified star-forming galaxies (blue histogram), overlaid with the sum of temperature distributions for every galaxy obtained by MCMC (dashed line). No radio or infrared detection threshold is applied to arrive at this sample of galaxies.

These values of aggregate q_{250} are inclusive of all our star-forming-classified sources. A spectral index calculated from detected sources will be unreliable and a bias towards flatter spectral indices would be introduced due to the differing sensitivities and depths of LOFAR and FIRST. Free–free absorption is also an issue at low frequency, where it flattens the radio SED, and so may have an effect on q_{250}, but we do not correct for its influence here. To check whether the difference in q_{250} between low and high frequency is due to spectral index, we find the value of α, which allows $q_{150MHz}^{LOFAR} - q_{150MHz}^{FIRST} = 0$ for sources detected at 3σ in both bands. The value for the spectral index that we find from the mean-stacked q_{250} of these sources is -0.58 ± 0.04 (Gaussian distributed), which is in agreement with Gürkan et al. (2018). We note that we do not use this value for the spectral index in our analysis because it will be biased by only considering the brighter sources that are 3σ detected. Instead we continue to use the value of -0.71 from Mauch et al. (2013) as originally stated.

We fit the slope of the FIRC to our star-forming sample for LO-FAR and FIRST using equation (4). We find that the FIRC measured with LOFAR is described by $L_{150}^{\text{LOFAR}} = 10^{-0.77\pm0.19}L_{250}^{0.97\pm0.01}$ with an intrinsic width of 0.89 ± 0.02 dex. This is slightly below the value of unity quoted for pure calorimetry. The FIRC measured with FIRST is described by $L_{150}^{\text{FIRST}} = 10^{2.94\pm0.25}L_{250}^{0.83\pm0.01}$ with an intrinsic width of 1.04 ± 0.03 dex. We show these fits graphically in Fig. 6 and include supplementary fits to the FIRC over different ranges of mid-infrared colour and specific star-forming rates in Appendix B.

4.3 The evolution of the FIRC

As discussed in Section 1, there have been numerous studies of the redshift evolution of the FIRC.

Fig. 7 shows the evolution of 250 μm and radio luminosity densities over our redshift range for context. To quantify the evolution of temperature and q_{250} with redshift, we fit a straight line using the Bayesian method detailed in Hogg, Bovy & Lang (2010) and implemented with PyMC3 (Salvatier, Wiecki & Fonnesbeck 2016). We show these redshift relationships in Fig. 8.

To calculate the effective temperature in each bin, the *Herschel* fluxes are mean-stacked and their uncertainties are derived from bootstrapping. Uncertainties on the mean redshift and mean fluxes are propagated through the MCMC fit to gain an effective temperature for each bin and its uncertainty. The uncertainty on the mean flux is small in bins with large numbers of sources, resulting in temperature uncertainties of order 2 K. Due to significance cuts made with BPT line ratios, Fig. 8 lacks the higher redshift galaxies present in the work of Smith et al. (2014), hence there is a large uncertainty above $z = 0.25$ (not shown). However, in the bins where the uncertainty on the dust temperature is small (<2 K), there is no statistically significant trend with redshift, consistent with Smith et al. (2014). With an MCMC trace of 50000 samples for each fit, we find strong evidence of a decrease in q_{250} over our low redshift range for LOFAR (gradient $= -1.0^{+0.2}_{-0.3}$) but no such strong evidence of such a decline with FIRST (gradient $= -0.5^{+0.5}_{-0.3}$), despite being consistent with LOFAR to within 1σ. It is worth noting that using the median stacking results in gradients that are consistent with the gradients calculated using the mean to within 1σ. We discuss the difference between the mean and median results (and lack of impact on our results) further in Section 4.5. A lack of evolution seen with FIRST is in line with the 250 μm result from Smith et al. (2014), the 70 μm result from Seymour et al. (2009), and the 70 μm and 24 μm result from Sargent et al. (2010). Calistro-Rivera et al. (2017) detect an evolution at both frequencies in the Boötes field and our result is consistent with theirs at redshifts below 0.25 at both frequencies. However, it is important to note that Calistro-Rivera et al. (2017) find curved radio SEDs, suggesting that a constant slope between 150 MHz and 1.4 GHz is not realistic.

At 3 GHz, Molnár et al. (2018) find no evidence for evolution in the total infrared-radio correlation in disc-dominated galaxies up until $z \sim 1.5$ (though Delhaize et al. 2017 find such an evolution in q using total infrared luminosity densities at redshifts ≥ 6). Together with Fig. 8, we therefore find tentative evidence for a frequency dependence of the evolution of q_{250} over redshift. However, Molnár et al. (2018) also find that an evolution in q_{250} over redshift is present in spheroids and is consistent with other studies of star-forming galaxies in general. They suggest that AGN activity not identified with traditional diagnostics is the cause. Extending their conclusion to our star-forming sample may imply that the cause of
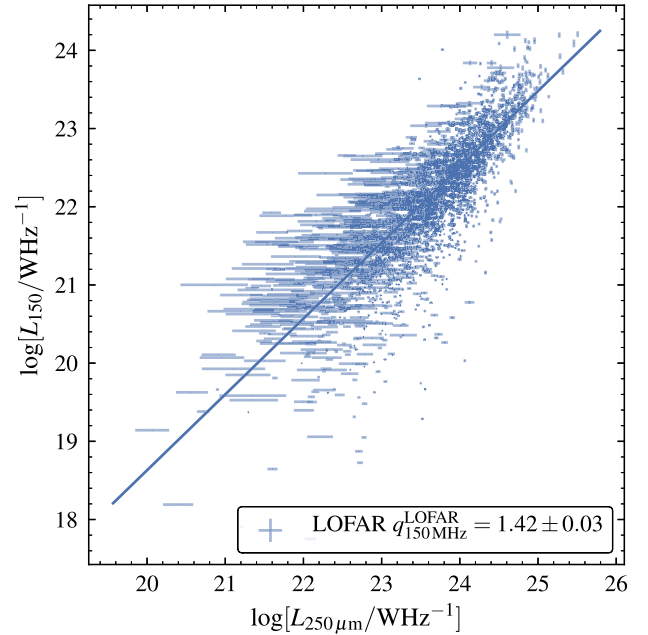
(a) The FIRC as measured with LOFAR at 150 MHz
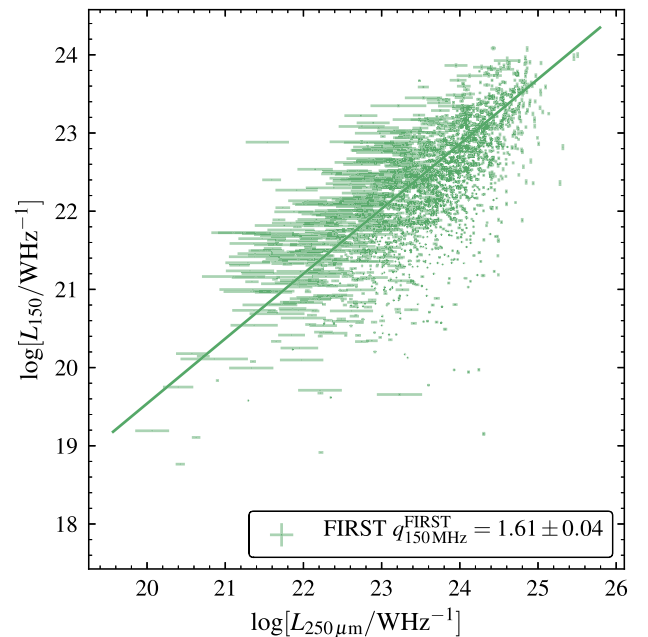
(b) The FIRC as measured with FIRST at 1.4 GHz

Figure 6. The FIRC for LOFAR (blue) and FIRST (green). The points shown are $>3\sigma$ detected in radio and infrared fluxes, showing two clear but distinctly different correlations at 1.4 GHz and 150 MHz. The fit lines are power-law fits to all sources in our star-forming sample including non-detections. For the purpose of comparison the FIRST 1.4 GHz luminosity densities have been transformed to 150 MHz assuming a power law with spectral index from Mauch et al. (2013).

the evolution found here is also low-level AGN activity, with AGN prevalence increasing with redshift.

Fig. 9 shows the evolution of q_{250} versus temperature. For comparison, $q_{150\text{MHz}}^{\text{LOFAR}}$, $q_{150\text{MHz}}^{\text{FIRST}}$, and the results of Smith et al. (2014) transformed to 150 MHz ($q_{150\text{MHz}}^{\text{Smith+14}}$) are shown together. Assuming
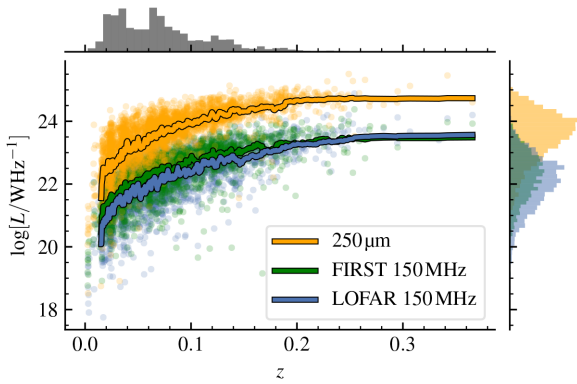
Figure 7. Distributions of Herschel SPIRE 250 μm (yellow), FIRST (green), and LOFAR (blue) luminosity densities over redshift for our main star-forming sample. A rolling mean (inclusive of all non-detections) with a window size of 200 points is plotted to guide the eye.
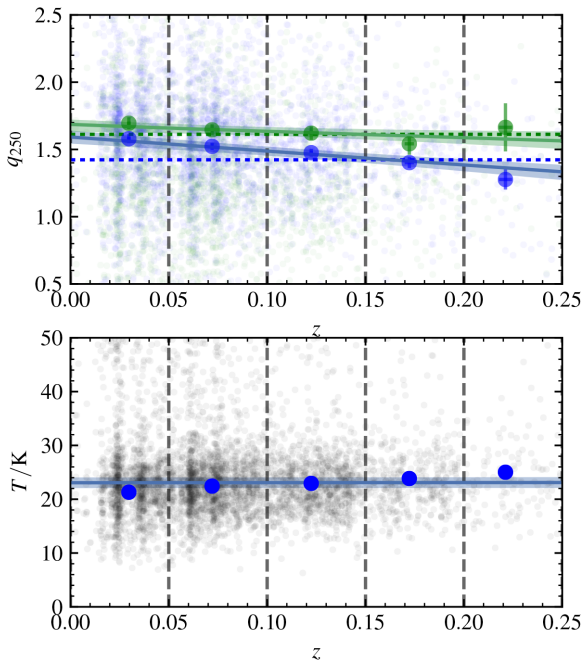

Figure 8. **Top:** Evolution of q_{250} over redshift measured with LOFAR at 150 MHz (blue) and FIRST transformed to 150 MHz (green). The dashed horizontal line in the upper plot is the mean-stacked q_{250} for all star-forming galaxies taken from Fig. 6 for FIRST and LOFAR at 150 MHz. The coloured lines indicate the straight line fit to all galaxies in our sample binned in redshift for LOFAR and FIRST. **Bottom:** The temperature in each bin, calculated by constructing an infrared SED from the average K-corrected flux of each source in every band and fitting equation (2) to the result. The temperature and uncertainties are overlaid with a straight line fit to the data. The vertical dashed lines represent bin edges.

a spectral index of -0.71, the trend of decreasing q_{250} with increasing temperature is found with both LOFAR and FIRST, agreeing within uncertainties when transformed to the same frequency at higher temperatures. Cold cirrus emission is not associated with recent star formation and so the ratio of infrared to radio luminosity (and hence q) will be larger for galaxies with colder integrated dust temperatures Smith et al. (2014). We discuss the deviation at lower temperatures in Section 4.5.
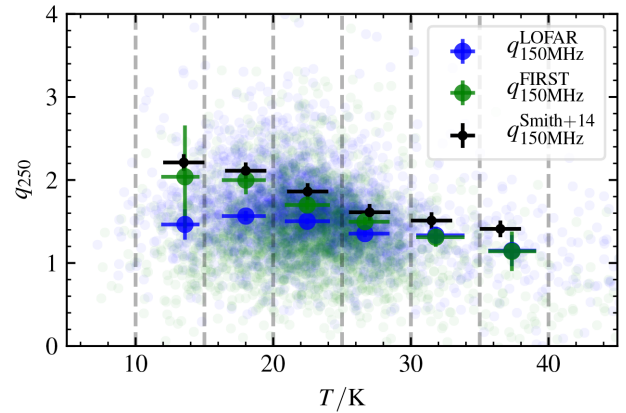

Figure 9. The temperature dependence of q_{250} compared between high and low frequency. The background dots are the individual q_{250} calculated from the LOFAR 150 MHz (blue) and FIRST (green) luminosity densities. The q_{250} calculated from stacked LOFAR and SPIRE luminosity densities described earlier is plotted in bold points with errorbars derived from bootstrapping the luminosity densities within the depicted dashed bins 10000 times. The temperature uncertainties in each bin are calculated from the 16th and 84th percentiles. The same calculation from Smith et al. (2014) is shown as the black errorbars for comparison.

The origin of the evolution of q_{250}^{LOFAR} with redshift is uncertain but we show here that the dependence of luminosity density upon redshift cannot account for all of the evolution measured in q_{250}^{LOFAR}. The bottom panel of Fig. 8 shows that the average dust temperature does not depend on redshift, when averaging across the whole sample. Therefore, if stacked 250 μm luminosity density is correlated with dust temperature (and Smith et al. 2014 show that same dependency at 250 μm) in our sample, then the dependency of stacked q_{250}^{LOFAR} upon redshift cannot only be due to a luminosity dependence on redshift.

4.4 Variation over the mid-infrared colour–colour diagram

In this section we focus solely on the sample of 2901 star-forming galaxies with 5σ WISE detections in order to construct the MIRDD of Jarrett et al. (2011). This sample covers part of the star-forming region defined by Wright et al. (2010) as shown in Fig. 2. When showing q_{250} variation of this sub-sample, we zoom in on this region.

We calculate the mean values of temperature and q_{250} as described in Section 3 over hexagonal bins in the WISE colour space. We show only those bins that contain more than 50 galaxies and have a stacked q_{250} with $SNR > 3$. When these conditions are applied, 33 and 29 contiguous bins remain for LOFAR and FIRST, respectively, all with a high SNR in binned $q_{150\,MHz}^{LOFAR}$, $q_{150\,MHz}^{FIRST}$ of at least 7 and 3, respectively. Fig. 10 shows the mean isothermal temperature in each bin. There is a clear and smooth increase in temperature towards redder $[4.6] - [12]$ and $[3.4] - [4.6]$ colours. The isothermal temperature of our sample increases towards the area populated mainly by starburst and Ultra-Luminous Infrared (ULIRG) galaxies. Our sample is positioned away from the Jarrett et al. (2011) AGN area, shown as a dashed box in Figs 10 and 11, although we note that radiatively inefficient radio-loud AGN may populate other regions of this plot (Gürkan et al. 2018).

The trend in temperature over mid-infrared colour is reflected in Figs 10 and 11, where the q_{250} measured using both FIRST and LOFAR decreases with redder WISE colours in a similar fashion to temperature. The higher sensitivity of LOFAR in comparison to
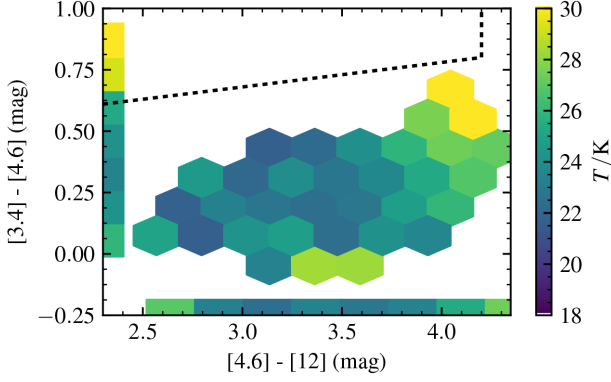
Figure 10. Mean isothermal temperature across the Jarrett et al. (2011) MIRDD. Bins are hexagonal and are coloured linearly between 18 K and 30 K described by the colour bar. All bins have an SNR in $q^{\rm LOFAR}_{\rm 150\,MHz} > 7$ and contain more than 50 galaxies each. Also plotted are the marginal bins summarizing horizontal and vertical slices of the entire plane. These slices also obey the two conditions set on the hexagonal bins. For reference, the box described by Jarrett et al. (2011) to contain mostly QSOs is marked by dotted lines.
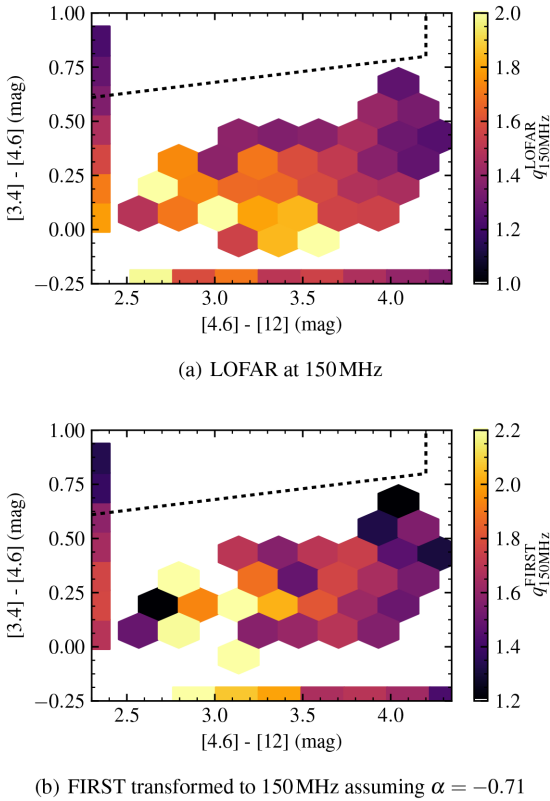

(a) LOFAR at 150 MHz

(b) FIRST transformed to 150 MHz assuming $\alpha = -0.71$

Figure 11. Mean-stacked q_{250} across the Jarrett et al. (2011) MIRDD. Bins are hexagonal and are coloured linearly according to the scale shown on the right. All bins have an SNR in $q_{250}{>}3$ and contain more than 50 galaxies each. Also plotted are the marginal bins summarizing the horizontal and vertical slices of the entire plane. These slices also obey the two conditions set on the hexagonal bins. For reference, the box described by Jarrett et al. (2011) to contain mostly QSOs is marked by dotted lines.

FIRST is reflected in the much smoother relation between binned q_{250} and mid-infrared colours.

Both the q_{250} parameter (for both frequencies) and the temperature change smoothly across mid-infrared colour. We interpret this smooth variation of the temperature over $[\,4.6\,] - [\,12\,]$ colour towards more heavily star-forming galaxies as tracing the specific star formation rate of a population of normal star-forming galaxies.

To quantify the observed trend with mid-infrared colour, we use a Bayesian method to find the correlation coefficients of stacked q_{250} against both WISE colours. From Fig. 11, q_{250} clearly correlates with both $[\,3.4\,] - [\,4.6\,]$ and $[\,4.6\,] - [\,12\,]$. However, since redshift is also highly correlated with $[\,3.4\,] - [\,4.6\,]$ and q_{250} is independently correlated with redshift, it is necessary to control for the effects of redshift using partial correlation (Baba, Shibata & Sibuya 2004) in order to quantify the effect of mid-infrared colour on q_{250}. We also control for isothermal temperature and stellar mass to see if all of the variation in q_{250} over mid-infrared colour can be accounted for by covariances with those parameters.

Our method consists of fitting a trivariate normal distribution to $[\,4.6\,] - [\,12\,]\,(x)$, $[\,3.4\,] - [\,4.6\,]\,(y)$, and q_{250} to obtain correlation-coefficient estimates (ρ_x and ρ_y). We estimate the correlation coefficients for q_{250} without controlling for any other parameters ($\rho_{x\cdot\emptyset}$ and $\rho_{y\cdot\emptyset}$) and for the residuals in q_{250} obtained from fitting a linear relationship to q_{250} against z, T_{eff}, and M_*. We fit the correlation coefficients with an LKJ prior (Lewandowski, Kurowicka & Joe 2009) using the PyMC3 (Salvatier et al. 2016) model specification along with EMCEE Ensemble sampler used above. LKJ distributions represent uninformative priors on correlation matrices and their inclusion allows us to randomly sample correlation coefficients.

To represent the correlation of q_{250} over the two dimensions of WISE colour space, Fig. 12 shows the the marginalized probability distributions for each correlation coefficient. The top panel of Fig. 12 shows the effect of controlling for redshift, temperature, and stellar mass independently as well as a naive fit that accounts for no other influential variables. The bottom panel of Fig. 12 shows the probability distribution of the correlation coefficients when controlling for redshift, temperature, and stellar mass at the same time. Initially, the distribution of q_{250} is highly correlated with both MIR colours (-0.5 ± 0.1 and -0.7 ± 0.1 for $[\,4.6\,] - [\,12\,]$ and $[\,3.4\,] - [\,4.6\,]$ colours, respectively). Fig. 12 as a whole shows that the variation of q_{250} with either WISE colour cannot be satisfactorily explained by a dependence on temperature, redshift, or stellar mass individually, but by all three at once. This results in correlation coefficients of 0.1 ± 0.2 and 0.2 ± 0.2 for $[\,4.6\,] - [\,12\,]$ and $[\,3.4\,] - [\,4.6\,]$, respectively.

Using the model described above, we find that the effects of stellar mass, dust temperature, and redshift upon q_{250} explain 16, 36, and 48 per cent of the total explainable correlation of q_{250} over the $[\,3.4\,] - [\,4.6\,]$ and 8, 71, and 21 per cent over $[\,4.6\,] - [\,12\,]$, respectively. However, the effects of these parameters on the variation of q_{250} are not independent of each other. Indeed, there are non-zero covariances between these parameters, e.g., the effect of stellar mass and dust temperature upon q_{250} at once is not equivalent to the sum of their independent effects.

Luminosity in 250 μm and both radio bands increases towards redder WISE colours and hotter temperatures, consistent with evidence of a luminosity-temperature relation found by Chapman et al. (2003) and Hwang et al. (2010), and in the radio by Smith et al. (2014). Given that the temperature evolution over redshift in our sample is consistent with being flat to within 1σ, we can conclude that such a luminosity-temperature relation is not simply due to
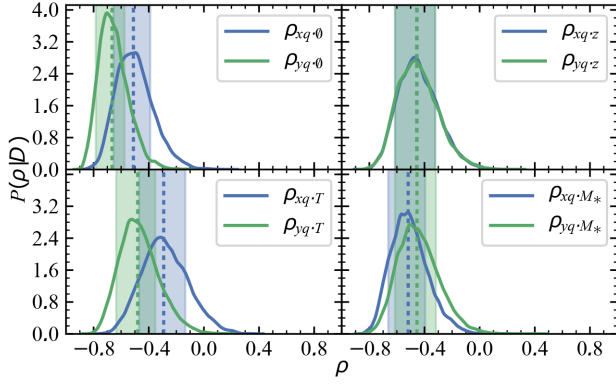
(a) Independent partial correlation coefficient PDFs
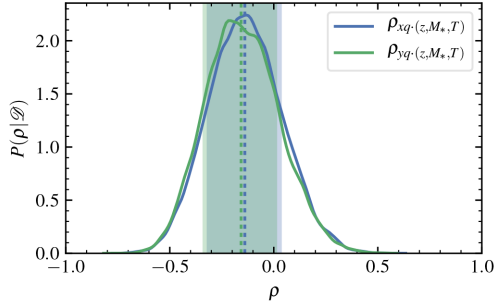
(b) Partial correlation coefficient PDF controlling for all variables at once.

Figure 12. The marginalized probability density, $P(\rho|\mathcal{D})$, distributions for the correlation coefficients (ρ) of $[4.6] - [12]$ (blue) and $[3.4] - [4.6]$ (green) against stacked q_{250}^{LOFAR}. $\rho = (-)1$ corresponds to maximal (anti-)correlation, whilst $\rho = 0$ corresponds to no correlation. **Top left (a):** The correlation coefficient PDFs calculated assuming that q_{250}^{LOFAR} does not depend on other variables. **Top right (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon redshift. **Bottom left (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon effective temperature. **Bottom right (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon stellar mass. **Bottom panel (b):** The correlation distribution when controlling for all three parameters at once. The vertical lines mark the median value for the correlation coefficient with the shaded areas marking the $16-84$th percentile range. A Gaussian kernel was used to smooth the probability distributions.

redshift effects. This is more evidence of the trend in q_{250} tracing the specific star formation rate.

To test our assumption that the $[4.6]-[12]$ colour traces specific star formation rate, we use the specific star formation rates obtained from MAGPHYS fits (Smith et al. 2012a). Fig. 13 shows a highly significant trend (both gradients are non-zero with a significance above 3σ) between MAGPHYS specific star formation rate and q_{250} for both FIRST and LOFAR (low sSFR is discussed below). The gradients of the trend at high and low frequency are consistent within 1σ.

Gürkan et al. (2018) have found that above a stellar mass of $10^{10.5}$ M$_\odot$, a strong mass dependence of radio emission, inferred to be non-AGN in origin, emerges. We show here that for the variation of q_{250} over the MIRDD to be explained, the effects of stellar mass and specific star-formation rate (for which isothermal temperature is an effective proxy) must be taken into account since they independently explain 25 and 38 per cent of the total correlation, respectively.
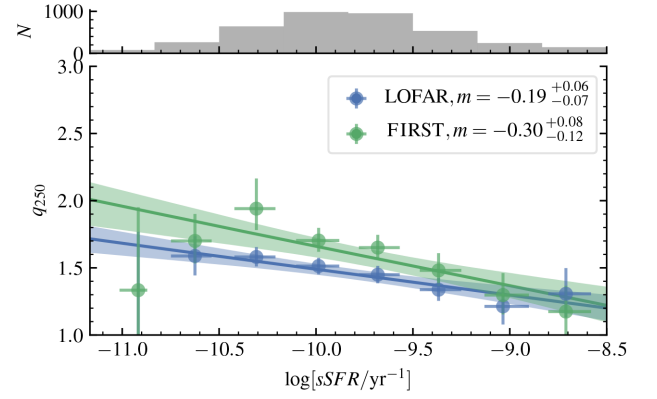

Figure 13. q_{250} for LOFAR (blue) and FIRST (green) at 150MHz against the specific star formation rate in 8 bins of width 0.3 dex. The uncertainties on q_{250} are calculated via bootstrapping within each bin. The uncertainties on sSFR are calculated from the 16th and 84th percentiles in each bin. Straight line fits are shown as coloured lines with 1σ credible intervals shown as shaded regions. The top histogram shows the number of galaxies in each bin.

4.5 Potential AGN contamination

BPT classification identifies AGN based on emission line ratios. However, star formation and AGN activity are not mutually exclusive (Jahnke et al. 2004; Trump et al. 2013; Rosario et al. 2013) and one ionization process can mask the other. Indeed, the BPT diagram shows a population of Seyfert 2 objects seamlessly joined to the star-forming branch (Baldwin et al. 1981; Kewley et al. 2006). Obscured AGN SEDs are bright in the mid-infrared due to the re-radiated emission from their obscuring structure (Antonucci 1993; Stern et al. 2005). In particular, radiatively efficient QSOs and obscured AGN are expected to be detected by WISE and to be located in the reddest space on the WISE MIRDD (Jarrett et al. 2011).

4.5.1 Searching for hidden AGN

Whilst it may be difficult to exclude composite galaxies based purely on line ratios, spectra can be searched for AGN features and radio images inspected for signs of jets or compact cores. The angular resolutions of FIRST and LOFAR are too low to distinguish AGN cores from compact starbursts, but we can rule out obvious radio loud contamination. To look for signs of physical differences between the low and high q_{250} areas and to check for the impact of radio-mode AGN, we take two samples of galaxies. The first sub-sample, named 'WISE-blue', we take from the region of highest q_{250} and bluer WISE colours. This region is described by the conditions $2.5 < [4.6] - [12] < 3.25$ and $0.0 < [3.4] - [4.6] < 0.4$ and so should correspond to lower luminosity star-forming galaxies.

The second sub-sample, named 'WISE-red', we take is described by the conditions $3.75 < [4.6] - [12] < 4.5$ and $0.2 < [3.4] - [4.6] < 0.6$, and is characterized by the lowest values of q_{250}. This is the area most likely to be contaminated by AGN, given its proximity to the QSO box defined in Jarrett et al. (2011) and low value of q_{250}. We note that we have removed the 12 sources that lie within the QSO box before conducting the analysis here.

We visually inspected the FIRST and LOFAR images of 100 randomly chosen galaxies from the WISE-blue and WISE-red sub-samples for signs of cores and jets. However, although the sub-samples are selected based on their position in the MIRDD, they also correspond to different redshift ranges. The higher redshift
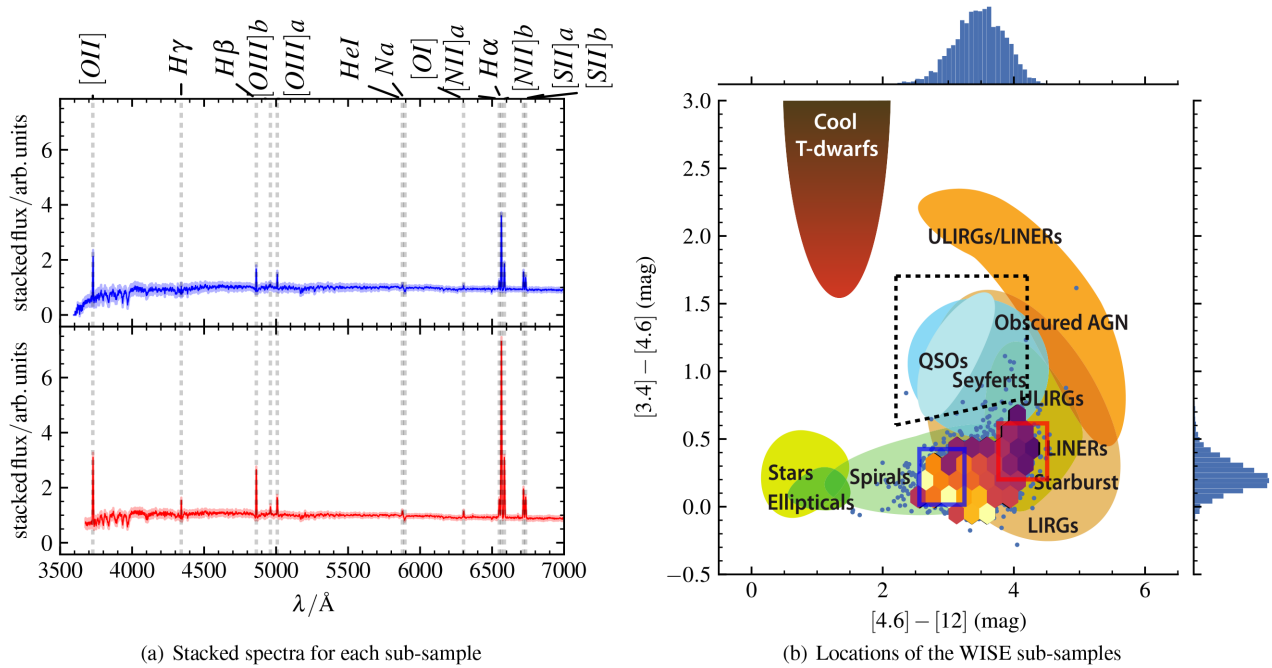
(a) Stacked spectra for each sub-sample

(b) Locations of the WISE sub-samples

Figure 14. The median-stacked rest-frame spectra of the WISE-blue (blue) and WISE-red (red) sub-samples (left) and their locations on the MIRDD (right). The spectra were blueshifted to their rest frame and interpolated to a common wavelength grid between 3500 Å and 8000 Å. Each spectrum was normalized based on the median interpolated flux between 5450 and 5550 Å. The 1σ uncertainties in the median spectra are shown as light-shaded regions around the median. The MIRDD is the same as in Fig. 2 with the binned $q_{150\,\mathrm{MHz}}^{\mathrm{LOFAR}}$ taken from Fig. 11 overlaid. Blue and red boxes indicate the boundaries for the WISE-blue and WISE-red sub-samples, respectively.

sources are selected at redder WISE colours and therefore the most luminous radio sources are selected in the WISE-red sub-sample and, conversely, the WISE-blue sub-sample consists of some of the least luminous radio sources. As a result, the WISE-red sub-sample tends to have extended and brighter radio emission at 150 MHz than our WISE-blue sub-sample. Therefore, if there is any significant AGN presence, they are more likely to populate the WISE-red sample than in the WISE-blue sample. We find little evidence of AGN activity due to compact cores or jet structures in either sample. However, the extended emission due to the luminosity bias mentioned above makes it difficult to compare the two sub-samples. Fig. 14 shows the rest-frame spectra, median-stacked using the method found in Rowlands et al. 2012, for the WISE-red sub-sample in red and the WISE-blue sub-sample in blue. Taking the difference between the spectra of the two sub-samples indicates the potential AGN (WISE-red sub-sample) have brighter emission lines relative to their continuum and have strong $H\gamma$ and [*OIII*] lines relative to the WISE-blue sub-sample. The $H\gamma$ line is found to have an equivalent width of 1 Å which is below what would be expected for broad-line AGN (Peterson 1997). Given the increased infrared luminosity, this seems to indicate that the ionization required to excite $H\gamma$ is generated by star formation. Moreover, if we position each median spectrum on the BPT diagram, they are both firmly within the star-forming region.

However, a significant fraction of the radio AGN population lack characteristic emission lines (Jackson & Rawlings 1997; Sadler et al. 1999; Best et al. 2005; Evans et al. 2006) and hence cannot be identified using a BPT classification. Such Low Excitation Radio AGN (LERAGN) could explain the decrease in q_{250} with their additional contribution to radio luminosity. LERAGN have traditionally not been reconciled with the AGN unification model proposed by Antonucci (1993). However, there are significant differences be-

tween LERAGN and their standard high excitation counterparts (HERAGN) such as black hole masses (e.g. McLure & Jarvis 2004; Smolcic 2009), depending on sample selection (Fernandes et al. 2015). Hardcastle, Evans & Croston (2006) have suggested that LERAGN are the consequence of a different accretion mechanisms whereby LERAGN accrete in a radiatively inefficient mode. LERAGN will have excess radio luminosity for their $H\alpha$ star formation, and so lie beneath the star-forming FIRC. The radio-loud fraction of galaxies that are LERAGN or HERAGN has been found to correlate with stellar mass, colour, and star-formation rate (Janssen et al. 2012). However, 98 per cent of our star-forming sample have stellar masses below $10^{11}\,\mathrm{M_\odot}$, where Janssen et al. (2012) report a radio-loud fraction of LERGS below 0.001.

In addition, Gürkan et al. (2015) found that radio-loud AGN tend to have lower star formation rates than star-forming galaxies, and that radio-quiet AGN also exhibit SFRs that are also offset from the star-forming galaxies. We use star formation rates and stellar masses fit by MAGPHYS (da Cunha, Charlot & Elbaz 2008; Smith et al. 2012a) to test whether our sample exhibits this offset. Fig. 15 shows our entire sample including BPT-classified AGN. Our BPT-classified AGN, in yellow, clearly lie below the star-forming SFR-mass relation found by Gürkan et al. (2015) whereas our star-forming sample (in green) lie on it. Furthermore, our sub-sample drawn from the WISE-red region (red) lie above the rest of the galaxies suggesting again that q_{250} decreases with increased specific star formation rate. Fig. 15 shows that the galaxies within the WISE-red sub-sample (red) are representative of starbursts given their increased star formation rate. The fact that our BPT-classified AGN are offset in star formation and stellar mass gives some reassurance that the effect of radio-quiet AGN is minimized in our WISE sample. Since using the mean instead of the median may well increase the effect of outliers/AGN, we test the effect of using the median on
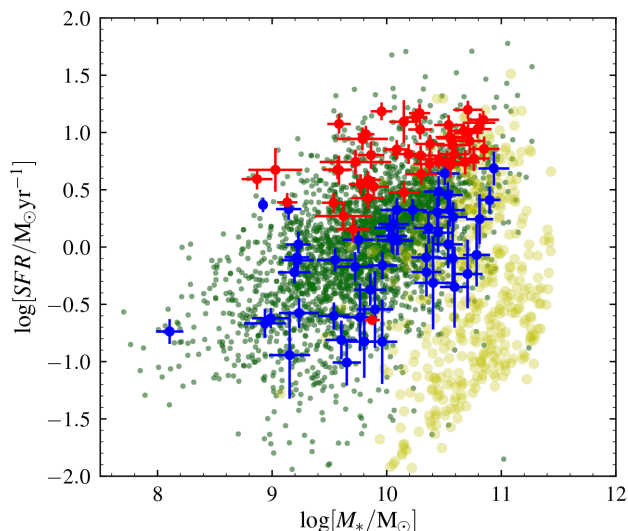
Figure 15. Stellar mass versus star formation rate for our sample. Our full sample is plotted in green. The 100 galaxies selected from both the WISE-red sub-sample (red) and the WISE-blue sub-sample (blue) are shown as the larger points with errorbars. For comparison, AGN classified by BPT are plotted in yellow.

our results. We find that the discrepancy between q_{250} measured with LOFAR and FIRST, the evolution of q_{250} over redshift, and the variation of q_{250} over the MIRDD are not affected. In addition, we check for obvious outliers in the luminosity distributions for LOFAR and FIRST, finding no such source. This allows us to gain some degree of confidence that we are not plagued by outlying AGN. However, we note that the global value of q_{250} is increased by 0.2 dex for both LOFAR and FIRST when using the median.

4.5.2 Testing with the inclusion of known AGN

If we include the 447 5σ BPT-classified AGN and rerun our analysis, we find a large effect on both global q_{250} values and distributions over redshift, temperature, and specific star-formation rate (which we show in Appendix A). The AGN decrease the whole-sample q_{250} to 1.23 ± 0.01 and 1.25 ± 0.10 for LOFAR and FIRST, respectively. The relative increase in radio luminosity is unsurprising and demonstrates that AGN can have a large impact on q_{250}.

We find that the addition of BPT-AGN affect q_{250} calculated in low temperature bins most severely but that the difference in q_{250} at high temperature is marginal and consistent with q_{250} calculated with our star-forming sample. The greatest effect is seen in the lowest temperature bin where there is a drop of 1.2 dex in q_{250}. Indeed, even when we include BPT-selected AGN we find that $q_{250} - T$ trend in Fig. 9 is still consistent with Smith et al. (2014) at high temperatures.

The largest effect that the BPT-selected AGN have on q_{250} is found at very low specific star-formation rates. These are likely to be LERAGN since low specific star-formation rates imply that they are quenched. In fact, at $\log[sSFR/\mathrm{yr}^{-1}] > -10.5$ there are very few BPT-AGN; we find that a q_{250} evaluated in this regime with AGN included is in agreement with a q_{250} calculated with our star-forming sample.

The evolution of q_{250} with redshift at high and low frequency remains qualitatively the same (decreasing with redshift) with an average offset of ≈ 0.2. Moreover, when the problematic low temperature and low specific star-formation rate bins are removed from

the calculation, the evolution of q_{250} over redshift at 150 MHz and 1.4 GHz is unaffected.

When BPT-AGN are included in the analysis, the variation in q_{250} over the WISE colour $[\,4.6] - [12]$ can be no longer be completely explained by the combined dependence on redshift, temperature, and stellar mass. The direction of this new correlation is not towards the location of QSO box at redder $[\,3.4] - [4.6]$ colour and actually inverts the correlation direction from negative to positive (see Fig. A3, for the equivalent Fig. 12 with the addition of known AGN). This is likely due to the fact that we require 5σ detections in the MIRDD WISE bands and so only bright AGN are identified and positioned at bluer $[\,4.6] - [12]$ colours (since the $12\mu m$ WISE band is less sensitive than $2.4\ \mu m$ and $3.4\ \mu m$). As a result, no BPT-AGN are found towards the reddest colours on the MIRDD (see Fig. A5). If such a small number of included BPT-AGN can alter the correlation in the positive direction, it is unlikely that hidden AGN are the root cause of the negative correlation found in our star-forming sample. Since the correlation of q_{250} with WISE colours in our star-forming sample is only just explained by these factors to 1σ, this could signal either a different misclassified population of objects in our supposed star-forming sample or a feature of star-formation itself. However, AGN that are not detected but have a BPT classification may be positioned differently on the WISE MIRDD and have different emission properties (see discussion above). We therefore cannot rule out low-level AGN contamination.

Based on the above analysis, q_{250} should only be marginally affected and therefore considered relatively trustworthy in galaxies with medium to high specific star-formation rates.

4.6 Reconciling with star-forming models

Through the model developed by Lacki & Thompson (2010), energy loss in starbursts is mainly due to bremsstrahlung and ionization. This would increase q if it were not for the effect of secondary charge radiation. Though the exact contribution of secondary charges – resulting from cosmic ray proton collisions with ISM protons – that is needed for a consistent FIRC is model-dependent, their addition allows the high-Σ_g conspiracy to be maintained (i.e. a linear FIRC more or less unchanging over surface density, Σ_g). Our results show that q_{250} decreases with specific star formation rate and hence high gas density. Such a decrease in q_{250} is at odds with the expected behaviour that q_{250} should remain constant (especially at high specific star formation rates), derived from the standard model described by Lacki & Thompson (2010). However, there are numerous reasons why the high-Σ_g conspiracy could break down detailed by Lacki & Thompson (2010).

If the magnetic field is assumed to be dependent on volume density rather than surface density, synchrotron cooling becomes dominant and q will decrease with increasing density. The high-Σ_g conspiracy also depends on the assumption that the escape time for cosmic rays is the same in all starbursts and normal star-forming galaxies. If the vertical (with respect to the disc) cosmic ray diffusion scale height is constant instead, then the escape time would be two orders of magnitude smaller for starbursts than for normal star-forming galaxies. However, we expect this effect to be much stronger than the variation in q_{250} we see in our result and advective transport by galactic winds may dominate in spiral galaxies (Heesen et al. 2018). Indeed, the fact that the variation is smooth and can be adequately explained by a combination of three parameters, for normal and highly star-forming galaxies, suggests that the same mass/temperature-dependent mechanism is responsible.

This work is based on the monochromatic q_{250} and is therefore not sampling all of the reprocessed light from recent star formation. It is therefore possible that a FIRC based on integrated infrared luminosity does not vary as the FIRC at 250 µmm does. In a two-component model of the dust SED (Charlot & Fall 2000), a warm but low-mass stellar birth cloud will outshine a cold but more massive ISM at 250 µm. If the warm stellar birth clouds are more dominant at higher isothermal temperature, then the FIRC will be a less accurate calibrator of star formation rate at lower temperatures, assuming a direct relation between synchrotron and recent star formation. Some of the variation of q_{250} with star formation rate could then be attributed to the effect of not using integrated dust luminosities.

There are many effects to consider when modelling the FIRC and the conspiracies listed in Lacki & Thompson (2010) depend on a subset of models. We cannot say precisely which effect will reconcile their standard model with this result.

5 CONCLUSIONS

We have used a catalogue of optically selected, BPT-classified star-forming galaxies from Gürkan et al. (2018) to study variation in the FIRC over redshift and other parameters. We calculate the monochromatic FIRC, parametrized as q_{250}, for 150 MHz and compare it to that found for 1.4 GHz, using forced aperture photometry. We obtained the photometry (fluxes were measured using 10 arcsec radius circular apertures centred on the optical positions) for all of these sources, including those that are not formal detections at LOFAR, FIRST, and *Herschel* wavelengths. To avoid introducing bias to our findings, we make no significance cuts on infrared or radio fluxes.

Knowing about possible variation in the FIRC is of great importance, since a constant FIRC underpins the use of radio luminosity estimates as a star formation rate indicator. Our main results are summarized as follows:

(i) q_{250} at 1.4 GHz for our sample is found to be consistent with previous studies (Ivison et al. 2010; Jarvis et al. 2010; Smith et al. 2014).

(ii) The FIRC for 150 MHz is found not to be consistent with that for 1.4 GHz assuming a standard power law with spectral index of -0.71 (0.1 dex lower).

(iii) We find evidence for a decreasing q_{250} with redshift at 150 MHz (gradient of $-1.0^{+0.2}_{-0.3}$). By comparing to the results of Molnár et al. (2018), we also find tentative evidence that the slope of this evolution becomes shallower with increasing frequency. An increase in radio luminosity of star-forming galaxies with redshift will be useful for high-redshift SFG detection, assuming that this evolution is maintained above $z = 0.5$, as has also been suggested by FIRC studies conducted with LOFAR at higher redshifts (e.g. Calistro-Rivera et al. 2017).

(iv) We corroborate the q_{250}-temperature variation discovered by Smith et al. (2014) at high frequency. We find that this relation also applies at low frequency to within 1σ, but only at temperatures above 20 K.

(v) We find that q_{250} varies across a two-dimensional mid-infrared colour–colour space, at both radio frequencies, and within the star-forming region defined by Jarrett et al. (2011). By using a hierarchical correlation model, we find that all of the correlation between q_{250} with [4.6] − [12] and [3.4] − [4.6] colours can be attributed to the combined effects of the correlations that we measure between q_{250} and stellar mass, redshift, and isothermal temperature,

to within 1σ. We note that the variation is not explained by redshift, temperature, or stellar mass alone but by all three in conjunction.

(vi) Using the indicative locations of different galaxy types within the WISE colour–colour plot from Jarrett et al. (2011) – e.g. spirals etc. – we see that the trend to lower q_{250} appears to reflect the transition from spirals to LIRGs to starbursts. q_{250} decreases with redder [4.6] − [12] colour and with increasing specific star formation rate. Indeed, the lowest values of q_{250} are seen in the region of the MIRDD occupied by the Polletta et al. (2007) starburst templates. Moreover, the region where LIRGs overlap with normal spirals ($3 < [4.6] − [12] < 4$) is the region where the largest gradient in q_{250} (relative to WISE colour) is seen.

(vii) To test the possible influence of AGN contamination on our results, we re-ran our analysis but this time included the BPT-classified AGN. The only significant change in our results was at the lowest dust temperatures, and lowest specific star formation rates; the other regions of parameter space, and therefore our conclusions, are unchanged. We can be confident, therefore, that our results are robust to the inclusion of detectable AGN, and it is tempting to attribute this variation to hitherto unknown physics of the FIRC. However, we cannot totally rule out the possibility that widespread low-level AGN have some influence (though we see no evidence of high ionization and/or broad emission lines indicative of their presence in stacked rest-frame optical spectroscopy for subsets of our BPT-classified SFG sample). We also test for residual AGN contamination by analysing the radio images for WISE-red and WISE-blue sub-samples, finding no clear evidence for obvious AGN jet structure in either group. We also test that our choice of aggregate statistic (the mean) of the parameter q_{250} is not affected by outliers by performing the same analysis with the median. We find that the trends that we report of q_{250} over redshift, sSFR, temperature, and mid-infrared colours remain unaffected by the choice of aggregate statistic with only the global value of q_{250} changing.

Taken together, these results indicate that the monochromatic FIRC varies strongly across the full range of BPT-classified star-forming galaxies in a manner dependent upon their mid-infrared colours (which are widely used as an empirical probe of galaxies' star formation properties), even at fixed redshift.

We do not draw conclusions from our results alone about the efficacy of using the FIRC to calibrate radio star-formation rates, however Gürkan et al. (2018) used the same sample of galaxies, along with a full analysis of energy-balance-derived stellar mass and star formation rate estimates, to investigate the low-frequency radio luminosity star-formation rate relation directly. The broken power-law relation between SFR and 150 MHz luminosity found in that work – which they suggest may indicate the presence of an additional mechanism for the generation of radio-emitting cosmic rays – is consistent with the possibility of residual low-level AGN contamination, and the FIRC behaviour we observe at low specific star formation rates. Indeed, this suggests that calibrations such as those proposed in Brown et al. (2017) may need to be more nuanced than they currently are.

Although our results underline the exquisite combined power of *Herschel* and LOFAR for studying star-forming galaxies (and in particular the high quality of the maps produced by the LoTSS pipeline), it will be of great interest to investigate the star-formation and AGN content of galaxies in more detail with even more sensitive, high-resolution data in the coming years, as we enter the era of the Square Kilometre Array.

ACKNOWLEDGEMENTS

This work makes use of data products from the *Herschel*Astrophysical Terahertz Large Area Survey *(H*-ATLAS), the Low Frequency Array (LOFAR), the Sloan Digital Sky Survey (SDSS) data release 7; and the Wide-field Infrared Survey Explorer (WISE).

H-ATLAS is a project with *Herschel*, which is an European Space Agency (ESA) space observatory with science instruments provided by European-led principal investigator consortia and with important participation from the National Aeronautics and Space Administration (NASA). The *H*-ATLAS website is http://www.h-atlas.org/.

LOFAR was designed and constructed by the Netherlands Institute for Radio Astronomy (ASTRON) and has facilities in several countries that are owned by various parties (each with their own funding sources) and that are collectively operated by the International LOFAR Telescope (ILT) foundation under a joint scientific policy. The LOFAR website is http://www.lofar.org/.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS website is http://www.sdss.org/.

WISE is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. The WISE website is http://wise.ssl.berkeley.edu/.

This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration 2013). The Astropy website is http://www.astropy.org/.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Antonucci R., 1993, ARA&A, 31, 473
Appleton P. N. et al., 2004, ApJS, 154, 147
Astropy Collaboration, 2013, A&A, 558, A33
Baba K., Shibata R., Sibuya M., 2004, Aust. N. Z. J. Stat., 46, 657
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Basu A., Wadadekar Y., Beelen A., Singh V., Archana K. N., Sirothia S., Ishwara-Chandra C. H., 2015, ApJ, 803, 51
Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559
Bell E. F., 2003, ApJ, 586, 794
Bell E. F. et al., 2004, ApJ, 608, 752
Berta S. et al., 2013, A& A, 551, A100
Best P. N., Kauffmann G., Heckman T. M., Ivezic Z., 2005, MNRAS, 362, 9
Blumenthal G. R., Gould R. J., 1970, Rev. Mod. Phys., 42, 237
Blundell K. M., Rawlings S., 2001, in ASP Conf. Ser., Vol. 250. Astron. Soc. Pac., San Francisco, p. 363
Bourne N., Dunne L., Ivison R. J., Maddox S. J., Dickinson M., Frayer D. T., 2011, MNRAS, 410, 1155
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Brown M. J. I. et al., 2017, ApJ, 847, 136
Calistro-Rivera G., Lusso E., Hennawi J. F., Hogg D. W., 2016, ApJ, 833, 98
Calistro-Rivera G. et al., 2017, MNRAS, 469, 3468
Chapman S. C., Blain A. W., Ivison R. J., Smail I., 2003, Nature, 422, 695
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Ciesla L. et al., 2016, A& A, 585, A43
Condon J. J., 1992, ARA&A, 30, 575
Condon J. J., Anderson M. L., Helou G., 1991, ApJ, 376, 95
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Coziol R., Torres-Papaqui J. P., Andernach H., 2015, AJ, 149, 192
Cutri R. M., 2012, VizieR Online Data Catalog, 2311
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
de Jong T., Klein U., Wielebinski R., Wunderlich E., 1985, A&A, 147, L6
Delhaize J. et al., 2017, A& A, 602, A4
Domingo-Santamaria E., Torres D. F., 2005, A&A, 444, 403
Eales S. et al., 2010, PASP, 122, 499
Evans D. A., Worrall D. M., Hardcastle M. J., Kraft R. P., Birkinshaw M., 2006, ApJ, 642, 96
Fernandes C. A C. et al., 2015, MNRAS, 447, 1184
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Furlanetto C. et al., 2018, MNRAS, 476, 961
Garrett M. A., 2002, A&A, 384, 4
Gelman A., Rubin D. B., 1992, Stat. Sci., 7, 457
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Gott J. Richard I., Vogeley M. S., Podariu S., Ratra B., 2001, ApJ, 549, 1
Griffin M. J. et al., 2010, A&A, 518, L3
Gürkan G. et al., 2015, MNRAS, 452, 3776
Gürkan G. et al., 2018, MNRAS, 475, 3010
Hardcastle M. J., Evans D. A., Croston J. H., 2006, MNRAS, 370, 1893
Hardcastle M. J. et al., 2016, MNRAS, 462, 1910
Heesen V. et al., 2018, MNRAS, 476, 158
Helou G., Soifer B. T., Rowan-Robinson M., 1985, ApJ, 298, L7
Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)
Hwang H. S. et al., 2010, MNRAS, 409, 75
Ibar E. et al., 2010, MNRAS, 409, 38
Ivison R. J. et al., 2010, A&A, 518, L31
Jackson N., Rawlings S., 1997, MNRAS, 286
Jahnke K. et al., 2004, ApJ, 614, 568
Janssen R. M. J., Rottgering H. J. A., Best P. N., Brinchmann J., 2012, A&A, 541, A62
Jarrett T. H. et al., 2011, ApJ, 735, 112
Jarvis M. J. et al., 2010, MNRAS, 409, 92
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kennicutt J., 1998, ARA&A, 36, 189
Kewley L. J., Heisler C. A., Dopita M. A., Lumsden S., 2001, ApJS, 132, 37
Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961
Lacki B. C., Thompson T. A., 2010, ApJ, 717, 196
Lacki B. C., Thompson T. A., Quataert E., 2010, ApJ, 717, 1
Lacy M. et al., 2004, ApJS, 154, 166
Lewandowski D., Kurowicka D., Joe H., 2009, J. Multivariate Anal., 100, 1989
Li J.-T. et al., 2016, MNRAS, 456, 1723
Lisenfeld U., Voelk H. J., Xu C., 1996a, A&A, 306, 677
Lisenfeld U., Voelk H. J., Xu C., 1996b, A&A, 314, 745
Longair M. S., 2011, High Energy Astrophysics. 3rd edn., Cambridge Univ. Press, Cambridge; New York
Maddox S. J. et al., 2018, ApJS, 236, 30
Mateos S. et al., 2012, MNRAS, 426, 3271

Mauch T., Klöckner H.-R., Rawlings S., Jarvis M. J., Hardcastle M. J., Obreschkow D., Saikia D. J., Thompson M. A., 2013, MNRAS, 435, 650

McLure R. J., Jarvis M. J., 2004, MNRAS, 353, L45

Michałowski M., Hjorth J., Watson D., 2010a, A&A, 514, A67

Michałowski M. J., Watson D., Hjorth J., 2010b, ApJ, 712, 942

Molnár D. C. et al., 2018, MNRAS, 475, 827

Murphy E. J., 2009, ApJ, 706, 482

Murphy E. J. et al., 2011, ApJ, 737, 67

Niklas S., Beck R., 1997, A&A, 320, 54

Noeske K. G. et al., 2007, ApJ, 660, L43

Pascale E. et al., 2011, MNRAS, 415, 911

Peng Y. et al., 2010, ApJ, 721, 193

Peterson B. M., 1997, An Introduction to Active Galactic Nuclei. Cambridge Univ. Press, Cambridge

Poglitsch A. et al., 2010, A&A, 518, L2

Polletta M. et al., 2006, ApJ, 642, 673

Polletta M. et al., 2007, ApJ, 663, 81

Rigby E. E. et al., 2011, MNRAS, 415, 2336

Rodighiero G. et al., 2011, ApJ, 739, L40

Rosario D. J. et al., 2013, A&A, 560, A72

Rowlands K. et al., 2012, MNRAS, 419, 2545

Sadler E. M., McIntyre V. J., Jackson C. A., Cannon R. D., 1999, Publ. Astron. Soc. Aust., 16, 247

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, Peer J Comput. Sci., 2, e55

Sargent M. T. et al., 2010, ApJS, 186, 341

Scheuer P. A. G., Williams P. J. S., 1968, ARA&A, 6, 321

Schleicher D. R. G., Beck R., 2016, A&A, 593, A77

Schober J., Schleicher D. R. G., Klessen R. S., 2017, MNRAS, 468, 946

Seymour N., Huynh M., Dwelly T., Symeonidis M., Hopkins A., McHardy I. M., Page M., Rieke G., 2009, MNRAS, 398, 1573

Shimwell T. W. et al., 2017, A&A, 598, A104

Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103

Smith D. J. B. et al., 2011, MNRAS, 416, 857

Smith D. J. B. et al., 2012a, MNRAS, 427, 703

Smith M. W. L. et al., 2012b, ApJ, 748, 123

Smith D. J. B. et al., 2013, MNRAS, 436, 2435

Smith D. J. B. et al., 2014, MNRAS, 445, 2232

Smith M. W. L. et al., 2017a, ApJS, 233, 26

Smith M. W. L. et al., 2017b, ApJS, 233, 26

Smolcic V., 2009, ApJ, 699, L43

Stern D. et al., 2005, ApJ, 631, 163

Tasse C., 2014a, preprint (arXiv:1410.8706) [astro-ph]

Tasse C., 2014b, A&A, 566, A127

Tasse C. et al., 2018, A&A, 611, A87

Torres D. F., 2004, ApJ, 617, 966

Trump J. R., Hsu A. D., Fang J. J., Faber S. M., Koo D. C., Kocevski D. D., 2013, ApJ, 763, 133

Valiante E. et al., 2016, MNRAS, 462, 3146

van der Kruit P. C., 1971, A&A, 15, 110

van Haarlem M. P. et al., 2013, A&A, 556, A2

Voelk H. J., 1989, A&A, 218, 67

Wong O. I. et al., 2016, MNRAS, 460, 1588

Wright E. L. et al., 2010, AJ, 140, 1868

York D. G., 2000, AJ, 120, 1579

Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

The appendices A and B are to be supplementary data for online viewing only. They are submitted here as a separate supplementary PDF.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: FIRC RELATIONS WITH BPT-AGN

Further to the discussion in Section 4.5, we present versions of Figs 8, 9, and 12, which are generated from a sample of 4541 sources formed by merging our star-forming sample with 447 BPT-classified AGN detected at 5σ in BPT emission lines. This material can be found online.
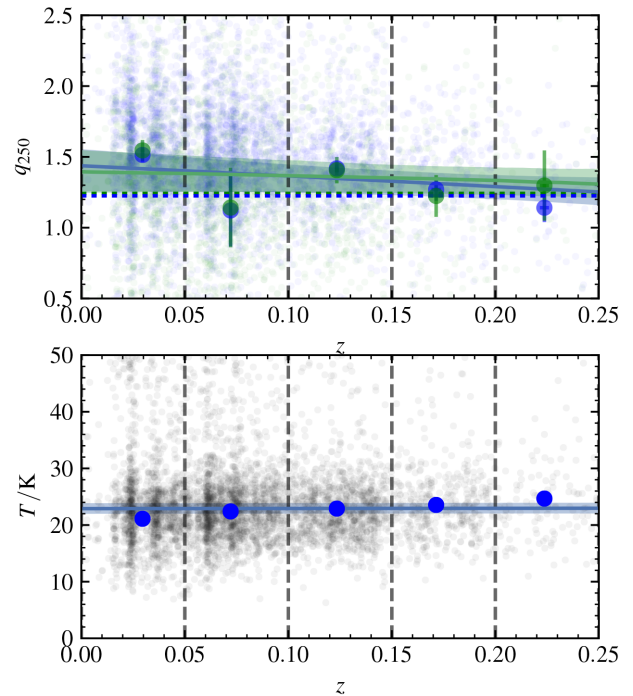

Figure A1. Top: Evolution of q_{250} over redshift measured with LOFAR at 150 MHz (blue) and FIRST transformed to 150 MHz (green). This plot uses our star-forming sample merged with BPT-AGN whose emission lines are detected at 5σ. The dashed horizontal line in the upper plot is the mean-stacked q_{250} taken from Fig. 6 for FIRST and LOFAR at 150 MHz. The coloured lines indicate the straight line fit to all galaxies in our sample binned in redshift for LOFAR and FIRST. **Bottom:** The temperature in each bin is calculated by constructing an infrared SED from the average K-corrected flux of each source in every band and fitting equation (2) to the result. The temperature and uncertainties are overlaid with a straight line fit to the data. The vertical dashed lines represent bin edges.
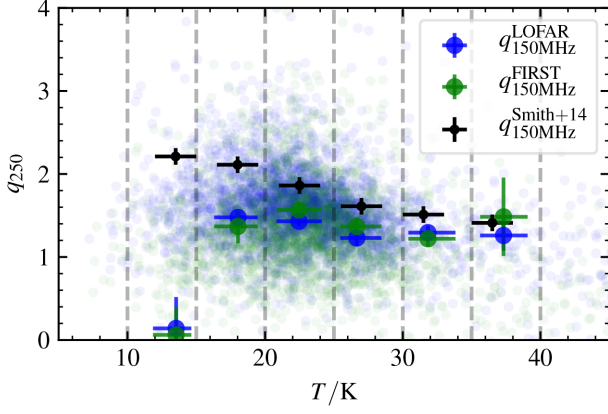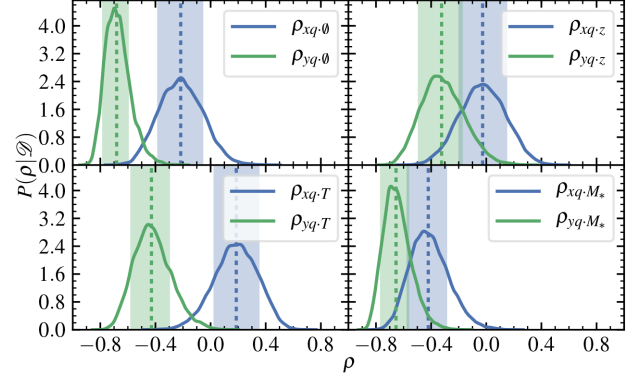
Figure A2. The temperature dependence of q_{250} compared between high and low frequency using our star-forming sample merged with BPT-AGN whose emission lines are detected at 5σ. The background dots are the individual q_{250} calculated from the LOFAR 150 MHz (blue) and FIRST (green) luminosity densities. The q_{250} calculated from stacked LOFAR and SPIRE luminosity densities described earlier is plotted in bold points with errorbars derived from bootstrapping the luminosity densities within the depicted dashed bins 10000 times. The temperature uncertainties in each bin are calculated from the 16th and 84th percentiles. The same calculation from Smith et al. (2014) is shown as the black errorbars for comparison.

(a) Independent partial correlation coefficient PDFs
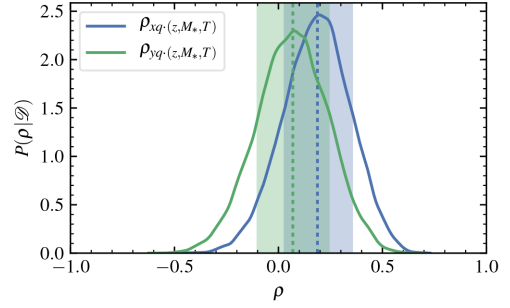
(b) Partial correlation coefficient PDF controlling for all variables at once.

Figure A3. The marginalized probability density, $P(\rho|\mathcal{D})$, distributions for the correlation coefficients of $[4.6] - [12]$ (blue) and $[3.4] - [4.6]$ (green) against stacked q_{250}^{LOFAR} (using our star-forming sample merged with BPT-AGN whose emission lines are detected at 5σ). **Top left (a):** The correlation coefficient PDFs calculated assuming that q_{250}^{LOFAR} does not depend on other variables. **Top right (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon redshift. **Bottom left (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon effective temperature. **Bottom right (a):** The correlation coefficient PDFs after controlling for a linear dependence of q_{250}^{LOFAR} upon stellar mass. **Bottom panel (b):** The correlation distribution when controlling for all three parameters at once. The vertical lines mark the median value for the correlation coefficient with the shaded areas marking the $16-84$th percentile range. A Gaussian kernel was used to smooth the probability distributions.
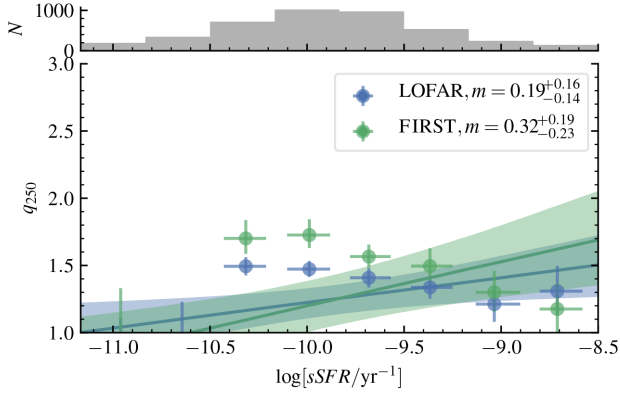
Figure A4. q_{250} for LOFAR (blue) and FIRST (green) at 150 MHz against the specific star formation rate in 8 bins of width 0.3 dex. This plot uses our star-forming sample merged with BPT-AGN whose emission lines are detected at 5σ. The uncertainties on q_{250} are calculated via bootstrapping within each bin. The uncertainties on sSFR are calculated from the 16th and 84th percentiles in each bin. Straight-line fits are shown as coloured lines with 1σ credible intervals shown as shaded regions. The top histogram shows the number of galaxies in each bin.
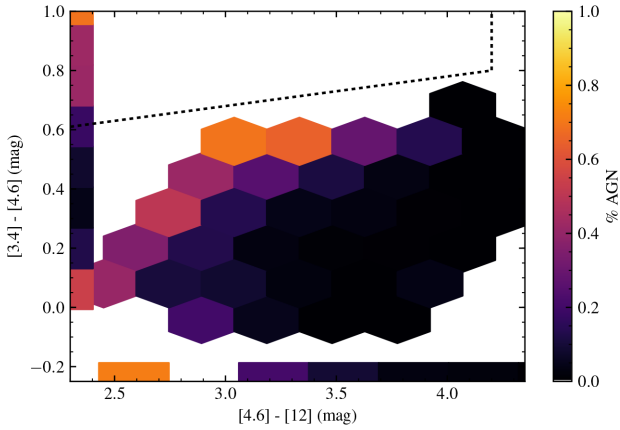
Figure A5. Mean fraction of BPT-AGN across the Jarrett et al. (2011) MIRDD using our star-forming sample merged with BPT-AGN whose emission lines are detected at 5σ. Bins are hexagonal and are coloured linearly according to the scale shown on the right. All bins have an SNR in $q_{250\mu m}^{\mathrm{LOFAR}} > 3$ and contain more than 50 galaxies each. Also plotted are the marginal bins summarizing the horizontal and vertical slices of the entire plane. These slices also obey the two conditions set on the hexagonal bins. For reference, the box described by Jarrett et al. (2011) to contain mostly QSOs is marked by dotted lines.
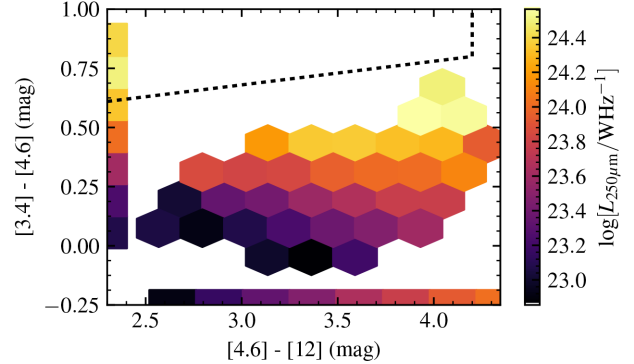
APPENDIX B: SUPPLEMENTARY FIGURES

This material can be found online.
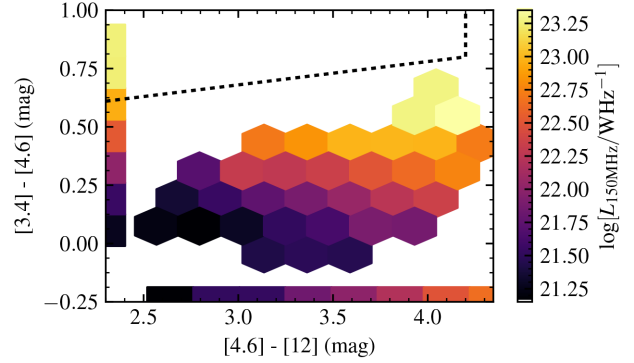
(a) LOFAR at 150 MHz
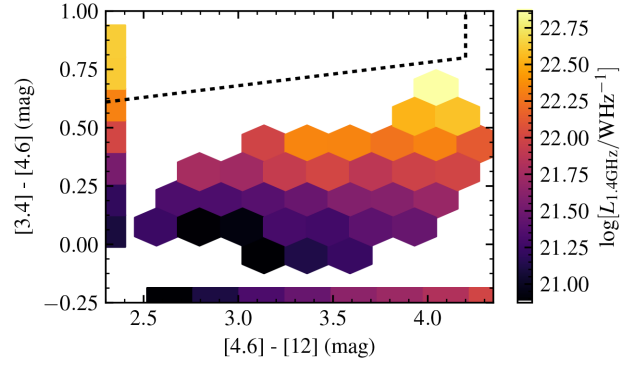
(b) FIRST at 1.4 GHz

(c) 250μm

Figure B1. Mean luminosity density across the Jarrett et al. (2011) MIRDD. Bins are hexagonal and are coloured linearly according to the scale shown on the right. All bins have an SNR in $q_{250\mu m}^{\mathrm{LOFAR}} > 3$ and contain more than 50 galaxies each. Also plotted are the marginal bins summarizing the horizontal and vertical slices of the entire plane. These slices also obey the two conditions set on the hexagonal bins. For reference, the box described by Jarrett et al. (2011) to contain mostly QSOs is marked by dotted lines.
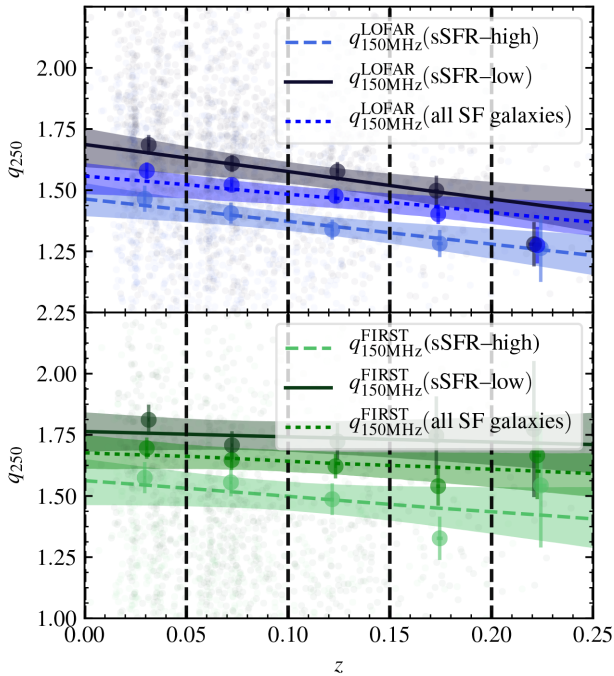
Figure B2. Evolution of q_{250} over redshift for the high-sSFR (light points, dashed lines), low-sSFR (darker points, solid lines), and all star-forming galaxies (dotted line) measured with LOFAR at 150 MHz (blue) and FIRST transformed to 150 MHz (green). The coloured lines indicate the straight-line fit to all galaxies in our sample binned in redshift for LOFAR and FIRST.
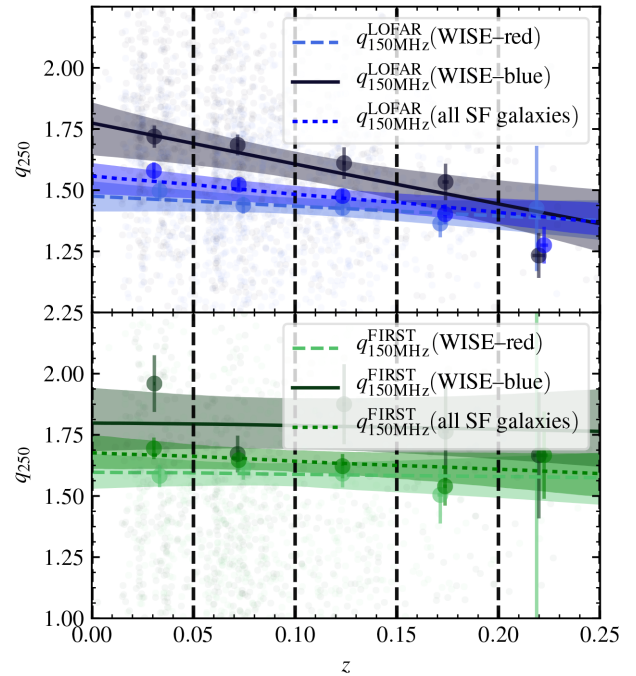
Figure B3. Evolution of q_{250} over redshift for the WISE-red (light points, dashed lines), WISE-blue (darker points, solid lines), and all star-forming galaxies (dotted line) measured with LOFAR at 150 MHz (blue) and FIRST transformed to 150 MHz (green). The coloured lines indicate the straight-line fit to all galaxies in our sample binned in redshift for LOFAR and FIRST.
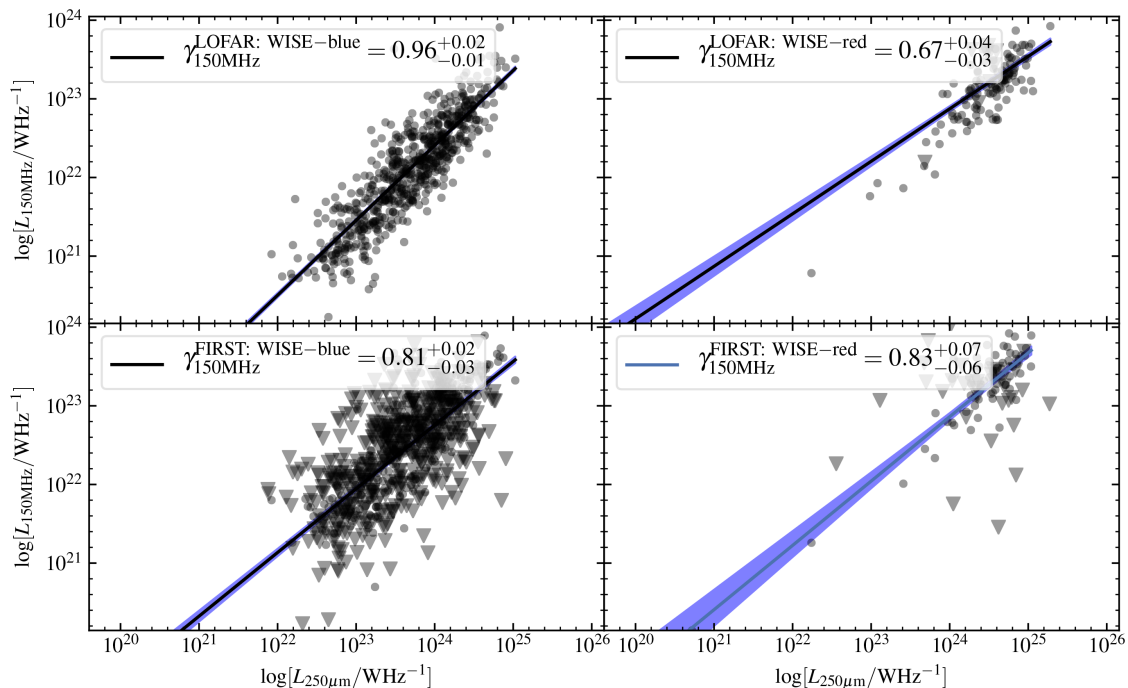

Figure B4. Fits to the FIRC measured by LOFAR and FIRST for the WISE-red and WISE-blue sub-samples. 3σ detections are shown as black points. 3σ upper limits for sources for which there is not formal 3σ detection are shown as black triangles The fit lines are power-law fits to all sources in our star-forming sample including non-detections. For the purpose of comparison the FIRST 1.4 GHz luminosity densities have been transformed to 150 MHz assuming a power law with spectral index from Mauch et al. (2013).
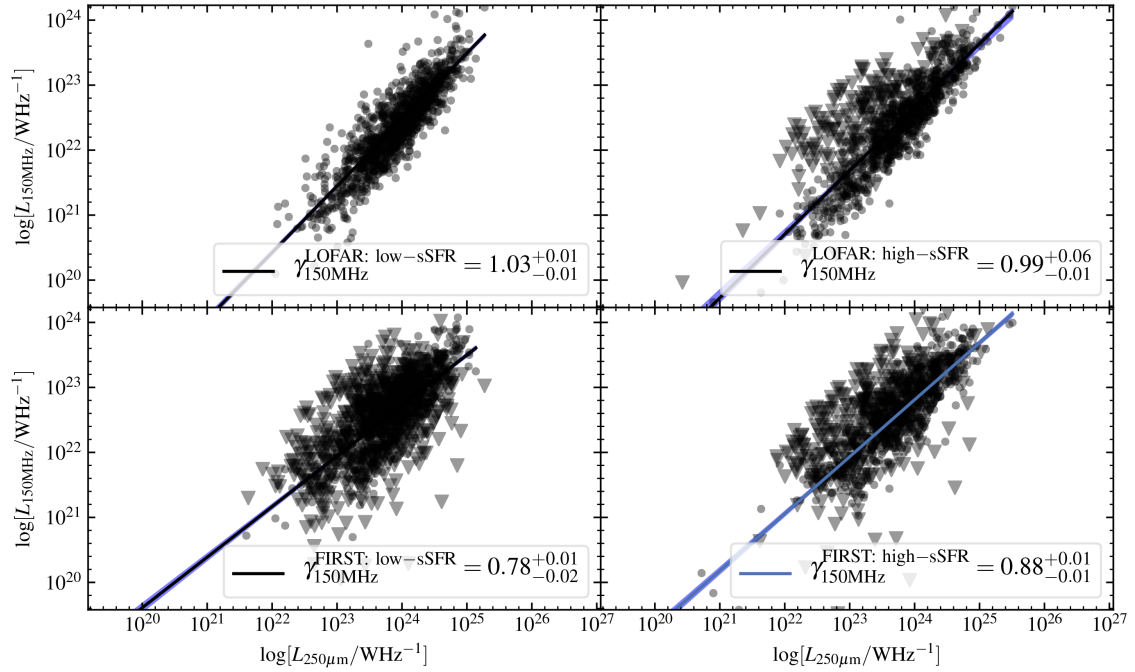
Figure B5. Fits to the FIRC measured by LOFAR and FIRST for galaxies in our star-forming sample with MAGPHYS fit specific star-formation rates above $-9.75\,\mathrm{yr}^{-1}$ (high-sSFR) and below $-9.75\,\mathrm{yr}^{-1}$ (low-sSFR). 3σ detections are shown as black points. 3σ upper limits for sources for which there is not formal 3σ detection are shown as black triangles The fit lines are power-law fits to all sources in each sample including non-detections. For the purpose of comparison the FIRST 1.4 GHz luminosity densities have been transformed to 150 MHz assuming a power law with spectral index from Mauch et al. (2013).

This paper has been typeset from a TEX/LATEX file prepared by the author.