



HAL
open science

Piercing the dusty veil of hyper-luminous infrared galaxies: Sub-arcsecond 144 MHz ILT observations of HLIRGs in the Lockman Hole

F. Sweijen, Y. Lyu, L. Wang, F. Gao, H. Röttgering, R. van Weeren, L. Morabito, P. Best, K. Malek, W. Williams, et al.

► **To cite this version:**

F. Sweijen, Y. Lyu, L. Wang, F. Gao, H. Röttgering, et al.. Piercing the dusty veil of hyper-luminous infrared galaxies: Sub-arcsecond 144 MHz ILT observations of HLIRGs in the Lockman Hole. *Astronomy and Astrophysics - A&A*, 2023, 671, pp.A85. 10.1051/0004-6361/202244378 . hal-04505841

HAL Id: hal-04505841

<https://hal.science/hal-04505841>

Submitted on 16 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Piercing the dusty veil of hyper-luminous infrared galaxies: Sub-arcsecond 144 MHz ILT observations of HLIRGs in the Lockman Hole

F. Sweijen¹, Y. Lyu¹, L. Wang^{2,3}, F. Gao^{2,3}, H. J. A. Röttgering¹, R. J. van Weeren¹, L. K. Morabito^{4,5},
P. N. Best⁶, K. Małek^{7,8}, W. Williams¹, I. Prandoni⁹, M. Bonato^{9,10}, and M. Bondi⁹

¹ Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
e-mail: sweijen@strw.leidenuniv.nl

² Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands

³ SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD Groningen, The Netherlands

⁴ Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham, South Road, Durham DH1 3LE, UK

⁵ Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

⁶ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

⁷ National Centre for Nuclear Research, ul. Pasteura 7, 02-093 Warszawa, Poland

⁸ Aix-Marseille Univ., CNRS, CNES, LAM, rue Frédéric Joliot-Curie, 13388 Marseille cedex 13, France

⁹ INAF – Istituto di Radioastronomia, Via P. Gobetti 101, 40129 Bologna, Italy

¹⁰ Italian ALMA Regional Centre, Via Gobetti 101, 40129 Bologna, Italy

Received 29 June 2022 / Accepted 21 October 2022

ABSTRACT

Context. Hyper-luminous infrared galaxies (HLIRGs) are among the most extreme systems in the Universe. With infrared (IR) luminosities of $L_{\text{IR}} > 10^{13} L_{\odot}$ they can have IR-derived star formation rates (SFRs) exceeding $10^3 M_{\odot} \text{ yr}^{-1}$. Theoretical models have a hard time reproducing the observed number densities of such extreme star-forming systems. It is known that at least part of the population harbours active galactic nuclei (AGNs), but their prevalence and relative contribution to the IR output is still debated. Assessing this is further complicated by the heavy dust obscuration.

Aims. We aim to investigate the HLIRG population in the Lockman Hole field to assess whether they are truly highly star-bursting systems or whether notable AGN activity is present. A substantial AGN population could help resolve the tension between the HLIRG number densities obtained from observations and predicted by galaxy formation models by lowering the amount of truly extremely star-forming galaxies.

Methods. Starting from a highly complete *Herschel*-selected sample, we made use of recent wide-field sub-arcsecond 144 MHz International LOFAR Telescope (ILT) observations of the Lockman Hole field to probe AGN activity in HLIRGs in a dust-unobscured way. AGN presence was deduced through means of the brightness temperature (T_{b}). Brightness temperature measurements were made to determine the fraction of HLIRGs harbouring a radio AGN. This identification was then compared to the classification based on spectral energy distribution (SED) fitting based identification from the LOFAR Deep Fields project, the radio-excess q and IRAC infrared colours.

Results. We detect 33% of previously identified HLIRGs at sub-arcsecond resolution. All but one of the detected sources is found to exceed the T_{b} threshold for pure star formation, showing 98% of detections to contain a radio AGN, even though lower-resolution observations had classified them as star-forming galaxies (SFGs). The remaining source is concluded to be consistent with having no AGN activity. All of the sources in our sample that were previously classified as radio AGNs, radio-quiet AGNs or those that were unclassified were detected as high- T_{b} objects (16% of the total sample or 47% of the detected sub-arcsecond detections). In addition, we identify AGNs through T_{b} measurements in 20% of sources that were classified as SFGs through SED fitting, raising the overall fraction of AGNs in the total sample from 16% to 32%. AGNs identified through brightness temperature measurements are also found to be more likely to be mid-IR AGNs.

Conclusions. 98% of our detected sources are found to likely host radio-AGNs, raising the number of identified AGNs among the HLIRG population in this field from 16% to 32%. This increased number of AGNs is not sufficient to bring observations and predictions of HLIRG number densities in agreement, however. Even at cosmic noon around $z \sim 2$, where the tension is lowest, it remains at a factor of just a few. The identification of radio AGNs in supposed SFGs highlights the value of high-resolution radio observations in studying dusty objects such as HLIRGs. Broad consistency is seen between T_{b} and the other AGN indicators, and the observed relation between SFR and T_{b} is seen as indicative of co-evolution between stellar mass build-up and black hole growth.

Key words. radio continuum: galaxies – galaxies: starburst – galaxies: high-redshift – galaxies: active – galaxies: evolution

1. Introduction

Observations at radio wavelengths offer a unique perspective on astronomical objects. They probe particle-acceleration pro-

cesses such as free-free and synchrotron emission, tracing star formation (SF) activity or activity from super-massive black holes (SMBHs), while their long wavelengths make them free them from the effects of dust attenuation that hinder shorter

(non-radio) wavelengths. This allows observations at radio wavelengths to provide information about the source of emission in areas where optical and infrared (IR) observations can not penetrate. The non-thermal behaviour of synchrotron emission makes it such that radio sources are generally brighter at lower frequencies. This makes low-frequency radio telescopes a prime choice for studying the radio-source population.

For that purpose, the Low-Frequency Array (LOFAR; van Haarlem et al. 2013) has been conducting the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017, 2019): a sensitive low-frequency survey of the northern sky between 120 and 168 MHz. It reaches a median sensitivity of $71 \mu\text{Jy beam}^{-1}$ with 8 h' observing time per pointing. In addition to covering the full northern sky using 8-h pointings, select areas of the sky are being observed for long time periods, aiming for several hundreds of hours. These fields, colloquially called LOFAR's "Deep Fields", are the Lockman Hole (Lockman et al. 1986), Boötes (part of the NOAO Deep Wide-Field Survey; Jannuzi & Dey 1999), ELAIS-N1 (European Large Area Infrared Space Observatory Survey-North 1; Oliver et al. 2000), and NEP (North Equatorial Pole) fields. They were chosen for their wealth of available ancillary data from X-ray to infrared. The first three of these fields have had ~ 100 h of data reduced, currently reaching sensitivities $\leq 30 \mu\text{Jy beam}^{-1}$ (Sabater et al. 2021; Tasse et al. 2021).

The bulk of luminous radio-emitting objects consist of objects whose activity is driven by a central SMBH, the so-called active galactic nuclei (AGNs). The dominant population here are the radio-loud (RL) AGNs (Padovani 2016). These typically launch powerful jets, readily giving away the presence of an AGN by their morphology in spatially resolved observations. AGNs can still be identified within spatially unresolved observations; however, as they are also characterised by an excess of non-thermal radio emission compared to what would be expected from their star formation rates (SFRs; if available), or a radio luminosity that is high compared to their luminosity in other bands such as the optical or infrared (IR; Hardcastle et al. 2019). At faint flux densities around $S_{1.4\text{GHz}} \sim 1 \text{ mJy}$ (corresponding to $S_{144\text{MHz}} \sim 6 \text{ mJy}$ at $\alpha = -0.8$), the population composition starts to shift from being dominated by RL AGNs to being dominated by star-forming galaxies (SFGs), where instead of being driven by AGN activity, the emission is driven by supernova remnants (Padovani 2016). Among this faint radio population, there are still AGNs that do not show large-scale jets or a notable excess of radio emission, making them easily mistakable for SFGs in radio observations that do not spatially resolve the galaxy. As such, it is important for our understanding of this population to be able to confidently separate AGN activity from star formation.

Inferring the presence of an AGN can be tricky due to the multitude of ways their signatures can manifest. It is therefore best done by leveraging broad, multi-wavelength spectral coverage. Extensive work has been put into classifying the sources detected in the Deep Fields, using such available ancillary data. This is described further in Sect. 2. One class of objects for which the distinction between star formation and AGN activity is particularly important is that of the luminous infrared galaxies. Discovered by the Infrared Astronomical Satellite (IRAS, Neugebauer et al. 1984), luminous infrared galaxies emit the majority of their energy at IR wavelengths. They come under four luminosity-based categories: luminous IR galaxies ($L_{\text{IR}} > 10^{11} L_{\odot}$), ultra-luminous IR galaxies (ULIRGs, $L_{\text{IR}} > 10^{12} L_{\odot}$, Cutri et al. 1994), hyper-luminous IR galaxies (HLIRGs, $L_{\text{IR}} > 10^{13} L_{\odot}$, Rowan-Robinson 2000), and extremely luminous IR galaxies (ELIRGs, $L_{\text{IR}} > 10^{14} L_{\odot}$, Tsai et al. 2015), and they are

believed to be powered by a combination of (merger-induced) star formation and black hole accretion (see also the introduction of Wang et al. 2021 and references therein Armus et al. 1987; Smith et al. 1998; Rowan-Robinson et al. 2018). Here, L_{IR} refers to the integrated 8–1000 μm luminosity. Based on their IR and sub-millimetre luminosities, LIRGs appear to reach SFRs in excess of $10^3 M_{\odot} \text{ yr}^{-1}$ (e.g., Rowan-Robinson et al. 2018).

The extreme nature of these objects poses a problem for our current understanding of galaxy formation. One problem is that reproducing their extreme SFRs is difficult (e.g., Davé et al. 2010; Narayanan et al. 2015). Furthermore, the predicted number densities from simulations are difficult to reconcile with observations (e.g., Hayward et al. 2013; Lovell et al. 2021; Wang et al. 2021). Their association with massive galaxies ($M_{*}^{\text{median}} \sim 10^{12} M_{\odot}$ between redshifts 1 and 6; Gao et al. 2021) further complicates the situation. How do such massive systems form so early on in the Universe? Although a shared contribution between star formation activity and AGN activity seems likely, the relative contributions of intense star formation and AGN activity to the emission of these objects is still under debate (Farrah et al. 2002; Ruiz et al. 2013; Gao et al. 2021). Finally, their often heavy obscuration makes it difficult to detect them at shorter wavelengths such as optical light. To determine the origin of the observed emission, one thus has to rely on a variety of selection techniques such as colour-colour cuts or luminosity ratios, depending on the data available, resulting in notable discourse about the origin mechanism in the literature, as succinctly summarised in the introduction of Farrah et al. (2017).

In this paper, we hope to address part of the tension with galaxy formation models found by Wang et al. (2021). To do so, we investigate AGN activity in the HLIRG population in the Lockman Hole field from a high-resolution, low-frequency radio perspective. Unaffected by dust, this should offer a clear view into the central engine of these sources. This allows radio observations to be used to infer the presence of an AGN in two ways. If SFR estimates are available, one can look for excess radio emission compared to what is expected from pure star formation. Best et al. (in prep.) used this criterion to identify radio AGN. The second method uses the measured brightness temperature T_{b} . There is an (frequency-dependent) upper limit that pure star formation can reach (Condon et al. 1991). Emission from a source exceeding this limit can therefore not originate purely from star formation, and this criterion was used to successfully identify AGNs using very long baseline interferometry (VLBI; e.g., Radcliffe et al. 2021).

The almost 2000 km long baselines of the International LOFAR Telescope (ILT) offer an angular resolution of $\theta \sim 0.3''$ at the central frequency of 144 MHz. Recent developments have made it possible to calibrate the ILT at its native resolution (Jackson et al. 2022; Morabito et al. 2022a) and to exploit its full field of view (Sweijen et al. 2022). This now allows us to investigate a sample of HLIRGs from a low-frequency, high-resolution perspective, adding this brightness temperature criterion as an additional AGN classifier. Using the classification by Best et al. (in prep.), we can distinguish already known AGNs. From the remaining SFGs and unclassified sources, their brightness temperature will be used to look for AGN-related emission, which can then tell us how many are truly SFGs and how many harbour an AGN after all, refining their earlier classification. An example of AGN identification in a wider population of radio sources detected in the LOFAR Deep Fields can be found in Morabito et al. (2022b). In the context of this work, we use the term AGN to refer to sources that display signs of AGN activity based on one or more (broadband) criteria. This means that

they are either classified as RQ AGN, HERG, or LERG by Best, et al. (in prep.; explained in the following section), or exceed the brightness temperature thresholds used in this work.

Section 2 outlines the data used. In Sect. 3, we outline the methods used to analyse this data. Section 4 presents our findings, which are discussed in Sect. 5. Finally, Sect. 6 covers the conclusions.

The assumed cosmology is that of the 2016 *Planck* results (Ade et al. 2016), with $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.308$, and $\Omega_\Lambda = 0.692$. Additionally, the spectral index α is defined as $S_\nu \propto \nu^\alpha$, where S_ν is the flux density at a given frequency ν .

2. Data

2.1. Radio data and source classification

We used 144 MHz radio data from Lockman Hole observations taken as part of the LOFAR Two-metre Sky Survey Deep Fields programme. The pointing centre of this field is J2000 $10^{\text{h}}47^{\text{m}}00^{\text{s}} +58^{\text{d}}05^{\text{m}}00^{\text{s}}$ (LT10_012, L659948; PI: Best). Standard-resolution $6''$ data is provided through a catalogue created from ~ 100 h of data combined (Tasse et al. 2021; Kondapally et al. 2021).

Fields observed as part of the LOFAR Deep Fields have valuable, deep, ancillary multi-wavelength data available from other surveys. The multi-wavelength cross-identification and catalogues are described by Kondapally et al. (2021). Photometric redshift estimates were derived for the vast majority of $6''$ radio detections by Duncan et al. (2021), with a normalised median absolute deviation (NMAD) with respect to spectroscopic redshifts of $\sigma_{\text{NMAD}} = 1.6\text{--}2\%$ for galaxies and $\sigma_{\text{NMAD}} = 6.4\text{--}7\%$ for AGNs. Best et al. (in prep.) employ four different spectral-energy distribution (SED) fitting codes to arrive at ‘consensus’ values for stellar masses and SFRs. The codes were CIGALE (Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019), MagPhys (da Cunha et al. 2008), BAGPIPES (Carnall et al. 2018, 2019), and AGNFitter (Calistro Rivera et al. 2017). Additionally, galaxies were classified as SFGs, radio-quiet (RQ) AGNs, low-excitation radio galaxies (LERGs), or high-excitation radio galaxies (HERGs). LERGs and HERGs are both RL AGNs, with the latter hosting a so-called radiative-mode AGN. These are identifiable using, for example, emission lines, mid-IR colours, IR excess, and SED fitting. Ultimately, the HERG classification was determined by a combination of the estimated AGN fractions from each SED-fitting code and a comparison of the goodness of fit from codes that include AGN templates (CIGALE and AGNFitter) versus those that do not (Magphys and BAGPIPES). Sources that did not show a radio excess, but were found to host a radiative-mode AGN, were classified as RQ AGN. When they neither contained a radiative-mode AGN nor displayed a radio excess, based on a cut in the $L_{150\text{MHz}} - \text{SFR}$ plane, sources were classified as SFGs. Finally, sources for which the classification was not clear were designated unclassified. This classification provides a rich new dataset to further our understanding of the properties of radio galaxies and the various relations between them.

For the high-resolution data, we used the catalogue obtained from an eight-hour observation at $0.3'' \times 0.4''$ angular resolution by Sweijen et al. (2022) (see the acknowledgements about data availability). It covers the central 6.6 deg^2 with a point-source sensitivity at the pointing centre of $25 \mu\text{Jy beam}^{-1}$, which is comparable to the $6''$ image at $23 \mu\text{Jy beam}^{-1}$. PyBDSF¹

(Mohan & Rafferty 2015) was used for source-finding and to derive source properties such as flux densities and source sizes by means of multi-Gaussian fits to the source’s brightness distribution. The resulting catalogue was then cross-matched against the $6''$ Deep Fields catalogue. This yielded a total of 2430 cross-matched sources² in the high-resolution image with a peak intensity exceeding $5\sigma_{\text{rms,local}}$, where $\sigma_{\text{rms,local}}$ is the local noise estimate in the high-resolution image by PyBDSF.

As the sub-arcsecond image was smaller, only the central 6.6 deg^2 covered by the high-resolution image was considered for the low-resolution data, using the multi-order coverage (MOC) map of the LoTSS-Deep-HD observation (LDHD hereafter). Finally, both the LoTSS-Deep and the LDHD images used the same inner uv -cut of 80λ , corresponding to a largest angular scale (LAS) of $\sim 46'$ at 144 MHz. In other words: the same short baselines are present in both datasets, and thus the recovery of diffuse emission is limited by surface-brightness sensitivity rather than being resolved out.

2.2. HLIRG sample

The HLIRG sample was first selected in Wang et al. (2021) and then further studied in Gao et al. (2021). In Wang et al. (2021), a parent sample of IR luminous galaxies was identified based on *Herschel* blind catalogues. They cross-matched *Herschel* sources with the LoTSS Deep Fields first data release (Duncan et al. 2021; Kondapally et al. 2021; Sabater et al. 2021; Tasse et al. 2021) in order to obtain multi-wavelength counterparts for them. Thanks to the tight correlation between FIR and radio (e.g., Calistro Rivera et al. 2017; Wang et al. 2021) and the superb quality of LOFAR imaging, they successfully matched LOFAR sources for $>90\%$ of *Herschel* sources with $250 \mu\text{m}$ flux density above 40 mJy in the Lockman Hole field. After adopting the SED fitting code CIGALE to derive galaxy properties, they selected 164 HLIRGs in the Lockman Hole field that have IR luminosity due to star formation activity above $10^{13} L_\odot$. The number of HLIRGs increased to 259 in Gao et al. (2021) when they required the total IR luminosity from both star formation activity and AGN activity to be above $10^{13} L_\odot$. Gao et al. (2021) used two different SED fitting codes in order to study the properties of HLIRGs, CYGNUS, and CIGALE, and three different AGN models from Efstathiou & Rowan-Robinson (1995; hereafter E95), Fritz et al. (2006; hereafter F06), and Stalevski et al. (2006; hereafter S12). Most relevant to this work are the SFR, IR luminosity, and AGN fraction derived using these SED fits. The IR luminosity L_{IR} is the integrated rest-frame $8\text{--}1000 \mu\text{m}$ luminosity. The AGN fraction f_{AGN} denotes the fraction of this IR luminosity attributed to AGN activity. We refer the reader to Wang et al. (2021) and Gao et al. (2021) for a detailed description of this HLIRG sample. For conciseness, we only show quantities obtained from CYGNUS with the Efstathiou & Rowan-Robinson (1995) AGN model (hereafter CyE95) and CIGALE with the Stalevski et al. (2012) model (hereafter CiS12) in the main text as these showed the largest differences in fitting results and, therefore, provide an indication of the uncertainty. The results from the Fritz et al. (2006) model and other combinations are shown in Appendix A.

Something to keep in mind is that all of these SED models are ‘torus models’. In such models, the AGN is surrounded by a

² This is slightly fewer than reported in Sweijen et al. (2022) due to accidental reference to an earlier version of the catalogue before publication. It has not impacted the results and the online catalogue contains the here-quoted number of sources.

¹ Python Blob Detector and Source Finder.

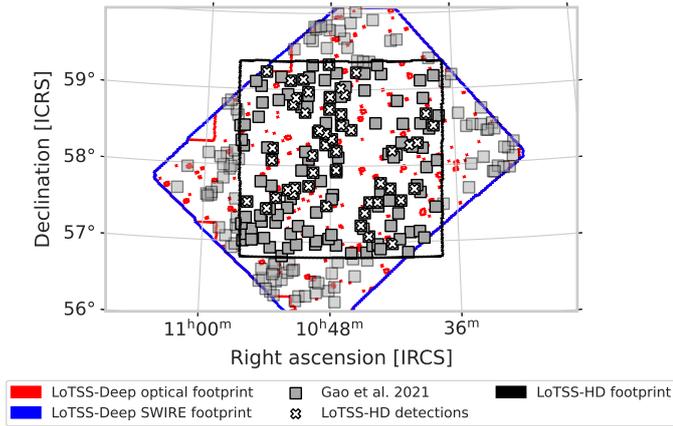


Fig. 1. Map showing the distribution of HLIRGs and, respectively, all sources in the sample (squares), the LoTSS-Deep-HD footprint (black line), sources that fall inside this footprint (opaque squares) and sources with a $\sigma_{\text{rms}}^{\text{local}} > 5$ detection in the LoTSS-Deep-HD image (white crosses). The red and blue outlines show the footprints of the optical and SWIRE coverage of this field.

dusty parsec-scale, torus-like entity (see e.g., Hönl 2019). This obscures the central engine, depending on the observer’s viewing angle, and affects how radiation from near the black hole affects the rest of the galaxy. The torus is confined to the nucleus of the galaxy, resulting in a relatively small volume of dust that is directly heated by the AGN. However, Symeonidis et al. (2016) and Symeonidis (2017), and early on Sanders et al. (1989), suggested that luminous AGNs could heat dust much further out, to kiloparsec scales. Not accounting for such large-scale heating could lead to an underestimation of the AGN contribution to the IR emission and subsequently an overestimation of the SFR.

The HLIRG catalogue was trimmed to the common sky area using the MOC, leaving 154 sources. These were cross-matched with the LDHD catalogue by matching their respective parent sources, which come from the LoTSS-Deep catalogue. Figure 1 illustrates the sky distribution of the full HLIRG sample, those that fall inside the LDHD footprint and those that are detected at the $5\sigma_{\text{rms}}$ confidence level. The redshift distribution of the sample is shown in Fig. 2. A Kolmogorov-Smirnov (KS) test was performed to test the null hypothesis of the redshift distributions of detected and undetected sources being equal, using SciPy’s `scipy.stats.ks_2samp`. This resulted in a KS-statistic of $D = 0.28$ and a p -value of $p = 0.007$. This p value corresponds approximately to a 2.7σ interval, where σ is the standard deviation of a normal distribution. We therefore cannot confidently reject the null hypothesis and conclude no statistically significant difference between the two.

In Fig. 3, the estimated IR luminosity is shown as a function of redshift. A KS test on the distributions of the estimated IR luminosities yield values of D between 0.18 and 0.25 with p values between 0.02 and 0.2. We therefore conclude no significant difference between the distributions of IR luminosities of LDHD-detected and LDHD-undetected sources either. Although both the redshift and IR-luminosity distributions of detected and undetected sources appear similar, we point out that the sample of detected sources is nonetheless inherently biased towards AGNs, because of the high angular resolution and lower surface brightness sensitivity compared to the lower angular resolution images.

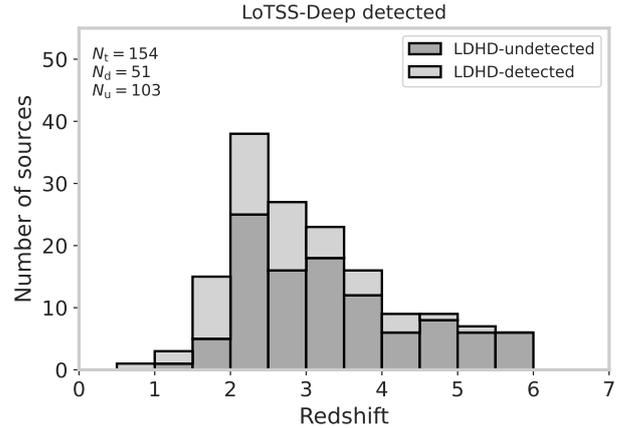


Fig. 2. Stacked redshift distributions for the full sample (N_t , full-height of the stacked bars), LDHD-detected (N_d , light-grey bars) and LDHD-undetected (N_u , dark-grey bars) sources. The majority are photometric estimates by Duncan et al. (2021), but some have spectroscopically confirmed redshifts. No significant difference is seen between detected and undetected sources.

3. Methods

Uncertainties on S_{ν}^{LoTSS} measured in the LoTSS-Deep image (where applicable) are taken conservatively at 20% (the uncertainty on the LoTSS flux density scale; see Shimwell et al. 2019).

3.1. Brightness temperature

To determine whether a source is a radio AGN, we made use of the brightness temperature T_b . Following Condon et al. (1991), a starburst galaxy will satisfy

$$T_b = T_e (1 - e^{-\tau_{\text{ff}}}) \left(1 + 10 \left(\frac{\nu}{1 \text{ GHz}} \right)^{0.1+\alpha} \right), \quad (1)$$

where $T_e \sim 10^4$ K is the electron temperature, τ_{ff} is the free-free optical depth, ν is the observing frequency, and α is the synchrotron spectral index (note the change in sign with respect to Condon et al. (1991) due to our definition of α). For $\nu = 144$ MHz, $T_e = 10^4$ K, $\alpha = -0.8$, and $\tau_{\text{ff}} \rightarrow \infty$ this gives a maximum brightness temperature of $T_b^{\text{max}} = 4 \times 10^5$ K, or $\log_{10} T_b^{\text{max}} = 5.6$. An HLIRG exceeding this threshold must be characterised by something else besides star formation powering its emission. Equation 1’s dependence on α makes it sensitive to changes in the source’s spectral index. Varying this changes the threshold to $\log_{10} T_b^{\text{max}} = 5.5$ for $\alpha = -0.7$ and $\log_{10} T_b^{\text{max}} = 5.7$ for $\alpha = -0.9$.

The sample was cross-matched with the deep 1.4 GHz observation of Prandoni et al. (2018). This yields a sub-sample of 39 sources for which a spectral index can be determined with respect to LoTSS-Deep. A Gaussian fit to data binned in $\Delta\alpha = 0.05$ wide bins yields a mean spectral index of $\alpha_{1.4 \text{ GHz}}^{1.4 \text{ GHz}} = -0.68 \pm 0.12$, where the uncertainty is the standard deviation of the fitted Gaussian. Only five sources (13%) are steeper than -0.8 . Without spectral information for all sources, we therefore consider the T_b threshold for $\alpha = -0.8$ as a reasonably conservative assumption for this sample.

We calculate the brightness temperatures of our sources by assuming their intensity profiles to be two-dimensional Gaussians. Furthermore, the intensity I is calculated using its definition S/Ω rather than the measured peak intensity I_{peak} ,

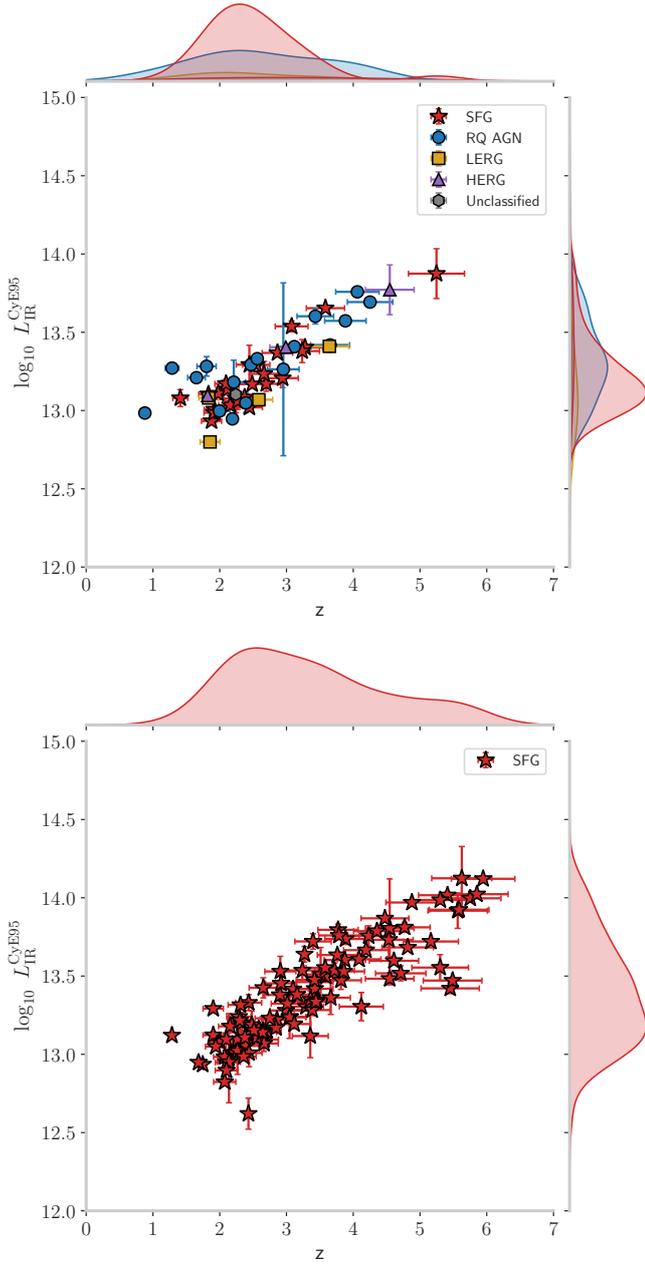


Fig. 3. Estimated IR luminosity from SED fitting using the CyE95 model versus measured photometric and spectroscopic redshifts. *Top:* LDHD-detected sources. *Bottom:* LDHD-undetected sources.

because time and bandwidth smearing will have artificially reduced I . Here, S is the source flux density and Ω is the source solid angle. The rest-frame brightness temperature can then be expressed as

$$T_b = 1.22 \times 10^6 \times (1+z) \times \left(\frac{\nu}{1 \text{ GHz}} \right)^{-2} \times \left(\frac{\theta_{\text{maj}} \theta_{\text{min}}}{1 \text{ arcsec}^2} \right)^{-1} \times \frac{S_{\nu}^{\text{ILT}}}{1 \text{ Jy}} \text{ K}, \quad (2)$$

where ν is the observing frequency, z is the source redshift, θ_{maj} and θ_{min} are the full width at half maximum (FWHM) major and minor axes of the Gaussian fitted to the source in the image plane, and S_{ν}^{ILT} is the measured high-resolution flux density. We point out that time and bandwidth smearing will also artificially increase θ_{maj} and θ_{min} , leading to an overestimate of the true source size. The uncertainty on T_b is propagated from the 30% uncertainty on S_{ν}^{ILT} and the uncertainty in θ_{maj} and θ_{min} , plus

the most conservative uncertainty of 7% in z for photometric redshifts, from Sect. 2.1. Gaussian uncertainty propagation then yields the uncertainty on T_b given by Eq. (3). Uncertainties are dominated by the flux density term, followed by the size terms, and lastly the redshift term. The median unsquared contributions are, respectively, 30%, 14%, 12%, and 5%:

$$\sigma_T = T_b \sqrt{\left(\frac{\sigma_S}{S_{\nu}^{\text{ILT}}} \right)^2 + \left(\frac{\sigma_{\text{maj}}}{\theta_{\text{maj}}} \right)^2 + \left(\frac{\sigma_{\text{min}}}{\theta_{\text{min}}} \right)^2 + \left(\frac{\sigma_z}{1+z} \right)^2}. \quad (3)$$

3.2. Radio excess

We probe the radio excess compared to a pure star-formation-driven source via the so-called q value. This value is calculated using the total rest-frame IR luminosity L_{IR} determined from SED fitting by Gao et al. (2021) and the measured 144 MHz rest-frame luminosity,

$$L_{144 \text{ MHz}} = 4\pi D_L^2 S_{144 \text{ MHz}} (1+z)^{-(1+\alpha)}, \quad (4)$$

as

$$q_{150} = \log_{10} \frac{L_{\text{IR}}}{3.75 \times 10^{12} \text{ Hz} \times L_{144 \text{ MHz}}}, \quad (5)$$

where L_{IR} is the total rest-frame dust luminosity obtained via Bayesian analysis between 8 and 1000 μm and $L_{144 \text{ MHz}}$ the radio luminosity at 144 MHz. Using Gaussian uncertainty propagation, the uncertainties $\sigma_{144 \text{ MHz}}$ and σ_q can be expressed as

$$\sigma_{144 \text{ MHz}} = L_{144 \text{ MHz}} \sqrt{\left(\frac{\sigma_S}{S_{144 \text{ MHz}}} \right)^2 + \left(\frac{1+\alpha}{1+z} \sigma_z \right)^2}, \quad (6)$$

$$\sigma_q = \frac{1}{\ln 10} \times \sqrt{\left(\frac{\sigma_{\text{IR}}}{L_{\text{IR}}} \right)^2 + \left(\frac{\sigma_{144 \text{ MHz}}}{L_{144 \text{ MHz}}} \right)^2}, \quad (7)$$

where $\sigma_{144 \text{ MHz}}$ and σ_{IR} indicate the uncertainties in their respective luminosities. Due to the lack of spectral indices for the sources, a fiducial value of $\alpha = -0.8$ is chosen to calculate the luminosity and no uncertainty is assumed. For the redshift, again a 7% uncertainty for photometric redshifts is chosen. To ensure the full radio luminosity was captured, the flux density of the 6'' LoTSS-Deep catalogue, $S_{144 \text{ MHz}}^{\text{LoTSS}}$, was used to calculate $L_{144 \text{ MHz}}$.

4. Results

4.1. Fraction of radio detections and brightness temperature

Of the full sample of 154 HLIRGs within the LDHD coverage, 51 sources are detected with a peak intensity of $I_{\text{peak}} > 5\sigma_{\text{rms}}^{\text{local}}$, where $\sigma_{\text{rms}}^{\text{local}}$ is the local root-mean-square noise level around the source, as was estimated by PyBDSF when the catalogue was built. Figure 4 compares the low-resolution flux density and peak intensity distributions between high-resolution detected and undetected sources. The mean and median values of LDHD-undetected HLIRGs lie close to or below the estimated limit (see Sect. 5). A minor fraction of nine sources (6% of the total sample) has a possible detection in the $3-5\sigma_{\text{rms}}$ range, but these were left out in favour of reliability. Despite these potential detections, since the estimated limiting values are still at an overall $>5\sigma$ level, we interpret the non-detections as probably being predominantly extended sources with no significant compact emission. An intermediate $\sim 1''$ resolution image would be able to confirm this.

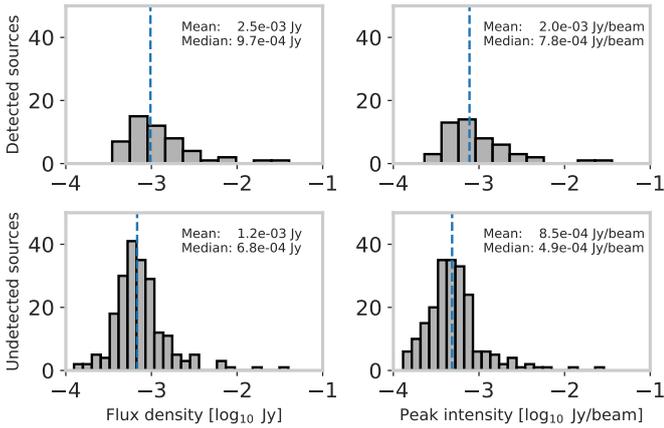


Fig. 4. LoTSS-Deep $6''$ flux density (*left*) and peak intensity (*right*) distributions for LoTSS-Deep-HD detected (*top*) and undetected (*bottom*) sources. The blue dashed line indicates the median of each distribution. LoTSS-Deep-HD-undetected sources have systematically lower values that are close to or below the 5σ surface brightness limit of the high-resolution image.

The brightness temperature distribution for the sample, as calculated by Eq. (2), is shown in Fig. 5. Theoretical upper limits derived by Condon et al. (1991) for starburst activity are indicated as well, for several synchrotron spectral indices. Conservatively, we find that 40 out of 51 sources (78%) have $\log_{10} T_b > 5.6$. These sources are likely hosting an AGN. With a less conservative cut of $\log_{10} T_b > 5.5$, this fraction rises to 98%. Out of these 51 sources, 24 are known AGNs from Best et al. (in prep.). From the remaining 27 sources, 11 exceed $\log_{10} T_b = 5.7$, 18 exceed $\log_{10} T_b = 5.6$ and 26 exceed $\log_{10} T_b = 5.5$, leaving 16, 9 and 1 below the AGN threshold, respectively. Of the full HLIRG sample, 16% are known AGNs following the criteria of Best et al. (in prep.). This can be increased to 32% by additionally including a brightness temperature cut. These results suggest that the majority of LDHD-detected HLIRGs in our sample cannot be explained by pure star formation.

In Table 1 we summarise the detection fraction of HLIRGs in redshift bins of $\Delta z = 0.5$. The tension with galaxy formation models found by Wang et al. (2021) starts around $z = 2$. We see a decline in the ratio of detected to undetected HLIRGs in each bin as redshift increases. Table 2 summarises the classification of the detected HLIRGs in the first two columns. The number of sources within that classification exceeding the two brightness temperature limits for $\alpha_{\text{syn}} = -0.8$ and -0.7 are reported in the third and fourth columns, respectively. Ignoring the known AGNs, 27 sources remain that were classified as either SFGs or unclassified. All of the 103 sources below the detection threshold are classified as SFGs by Best et al. (in prep.).

4.2. Brightness temperature compared to SED fitting

4.2.1. AGN fraction

The AGN fraction f_{AGN} gives the estimated IR luminosity due to an AGN component and it is defined through $L_{\text{AGN}} = f_{\text{AGN}} L_{\text{IR}}$. Figure 6 shows the estimated AGN fraction versus brightness temperature. We see both sources with a high AGN fraction and a high brightness temperature, but also sources with a low AGN fraction and a high brightness temperature. A general trend where the brightness temperature rises with increasing AGN fraction can be seen, but the estimated AGN fraction varies

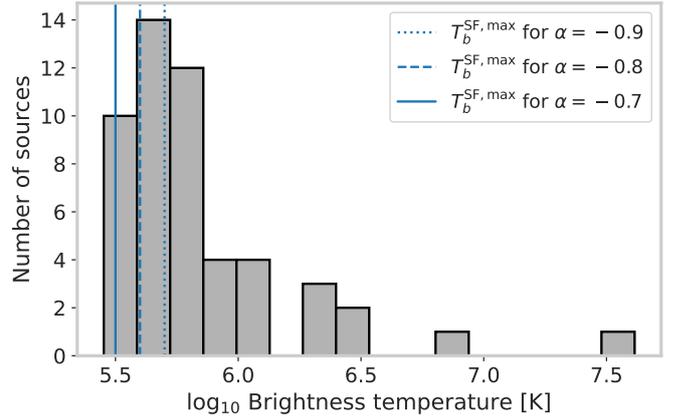


Fig. 5. Distribution of brightness temperatures in the LoTSS-Deep-HD detected sources. The solid, dashed, and dotted blue lines indicate the maximum brightness temperature from starbursts for a synchrotron spectral index α of -0.7 , -0.8 , and -0.9 , respectively, as per Condon et al. (1991). Values to the right of this line cannot be driven by pure star formation.

notably between the different models and fitting codes. The reliability of estimated $f_{\text{AGN}} < 0.01$ was deemed questionable, and hence those values were offset by 0.01 to aid visualisation.

4.2.2. Star formation rate

Figure 7 shows, for detected sources, the star formation rate derived from SED fitting by Gao et al. (2021) versus the measured brightness temperature. The colour and shape of the points indicate their classification by Best et al. (in prep.). Finally, the theoretical brightness temperature limits for star formation from Condon et al. (1991) are indicated for values of the synchrotron spectral index of -0.7 , -0.8 , and -0.9 , respectively. One difference between the CyE95 and CiS12 models is that the former allows for lower SFRs. The other model combinations in Fig. A.1 show that this is particular to the CyE95 model. Other combinations all show distributions similar to the CiS12 result. We also note a ‘zone of avoidance’. The HLIRGs with high T_b values appear to avoid the low-SFR region, or, vice versa, sources with low SFRs appear to avoid the high- T_b region.

4.2.3. Distance to the main sequence

The idea of a star-forming main sequence results from the observed correlation between the stellar mass and SFR of galaxies. We compared the difference between the estimated SFRs to their respective ‘main-sequence’ SFRs, that is, the distance to the main sequence $\Delta\text{MS} = \text{SFR} - \text{SFR}_{\text{MS}}$, for our LDHD-detected sample, using the relation of Speagle et al. (2014). An anti-correlation between the AGN fraction derived from their SED fitting and the distance to the main sequence was found by Gao et al. (2021). Figure 8 plots the distance to the main sequence versus the measured brightness temperature. Negative values of ΔMS indicate an SFR below that of the main sequence, while positive values indicate SFRs higher than expected from the main sequence. Visually, the distribution hints at a potential anti-correlation among SFGs and RQ AGNs, for which sources with estimated SFRs below that of the main sequence show higher brightness temperatures and less scatter than those near or below the main sequence. Closer inspection, however, reveals that only SFGs for the CyE95 and CiS12 models show a

Table 1. HLIRG detections grouped in redshift bins of $\Delta z = 0.5$.

Redshift	$N_{\text{det.}} / N_{\text{tot.}}$	$T_{\text{b}} > 10^{5.7} \text{ K}$	$T_{\text{b}} > 10^{5.6} \text{ K}$	$T_{\text{b}} > 10^{5.5} \text{ K}$
$0.0 \leq z < 0.5$	–	–	–	–
$0.5 \leq z < 1.0$	1/1	1	1	1
$1.0 \leq z < 1.5$	2/3	1	2	2
$1.5 \leq z < 2.0$	10/15	5	6	10
$2.0 \leq z < 2.5$	13/38	3	8	12
$2.5 \leq z < 3.0$	11/27	8	11	11
$3.0 \leq z < 3.5$	5/23	3	4	5
$3.5 \leq z < 4.0$	4/16	3	3	4
$4.0 \leq z < 4.5$	3/9	3	3	3
$4.5 \leq z$	2/22	2	2	2
All	51/154	29	40	50

Notes. Column 1 indicates the detection fraction, Cols. 2–4 indicate the number of sources exceeding the respective T_{b} thresholds.

Table 2. Classification of detected HLIRGs.

Class	N_{obj}	$T_{\text{b}} > 10^{5.7}$	$T_{\text{b}} > 10^{5.6}$	$T_{\text{b}} > 10^{5.5}$
SFG	25 (103)	10	17	24
RQ AGN	17	11	15	17
LERG	4	4	4	4
HERG	3	3	3	3
Unclassified	2	1	1	2
Total	51 (103)	29	40	50

Notes. Column 1 states the number of objects in the given class, Cols. 2 and 3 indicate the number of sources exceeding the respective T_{b} thresholds. Numbers in parentheses indicate LDHD-undetected sources.

potential anti-correlation. The Pearson r and Spearman ρ correlations³ for them are $r = -0.13$, $\rho = -0.32$ and $r = -0.30$, $\rho = -0.41$, respectively. The p values are insufficiently constraining, however, at only $p = 0.52$, $p = 0.12$, $p = 0.14$, and $p = 0.044$, respectively. Therefore, we conclude that there is no statistically significant correlation present in the current sample.

4.3. Brightness temperature compared to radio excess

Figure 9 shows the 144 MHz radio excess q_{144} , as calculated by Eq. (5), versus the brightness temperature of each detected HLIRG. A clear anti-correlation between T_{b} and q_{144} is seen. Such a correlation is not unexpected, as a lower value of q_{144} indicates a higher radio-excess, which for these high-SFR objects translates to a higher brightness temperature. The classification of Best et al. (in prep.) reveals this correlation to be primarily driven by HERGs and LERGs, and to a lesser extent by RQ AGN. The SFGs and RQ AGNs predominantly cluster together around the typical value of $q_{144\text{MHz}} \approx 1.7$ as found by, for example, Calistro Rivera et al. (2017). This is also the approximate value below which the correlation becomes apparent.

4.4. Brightness temperature compared to IRAC detections

Figure 10 displays a colour-colour diagram of the HLIRG sample using the four IRAC channels: 3.5 μm , 4.5 μm , 5.7 μm , and 8.0 μm , respectively. The black solid lines constrain a region of

³ Calculated using `scipy.stats.pearsonr` and `scipy.stats.spearmanr`.

this mid-IR colour space defined by Donley et al. (2012; hereafter the Donley region) to yield highly complete samples of AGNs with little contamination from SFGs. We refer to sources within this region as IRAC AGNs and sources outside this region as IRAC non-AGNs. We find 58 out of 154 HLIRGs to be IRAC AGNs. Among the LDHD detections, 22 out of 51 can be classified as IRAC AGNs, while of the non-detections, 36 out of 103 sources are IRAC AGNs.

Figure 11 shows the fraction of IRAC AGNs in sources that exceed a given limiting brightness temperature. The fraction of IRAC AGNs is seen to increase with limiting brightness temperature, as shown in black. In red and blue, the fraction of, respectively, IRAC AGNs and IRAC non-AGNs exceeding this limit is shown. We see two trends. Firstly, the fraction of sources exceeding the threshold decreases as the threshold increases. This is expected as there are fewer sources with higher brightness temperatures. Secondly, the IRAC AGNs show a systematically higher fraction of sources exceeding the limit compared to IRAC non-AGNs.

It is clear that the Donley cut does well in terms of selecting AGNs, especially those identifiable by the criterion of Best et al. (in prep.). This is not surprising, given that the mid-IR region was a crucial component in the SED fitting by Best et al. (in prep.). Most of the sources in the Donley region are previously identified AGNs, while the remaining sources exceed our brightness temperature thresholds. However, LDHD-detected SFGs outside this region are all radio-AGN candidates according to the brightness temperature criterion, except for one.

5. Discussion

One of the caveats of high-resolution observations is the decreased surface brightness sensitivity. This will affect the detection of extended source structure and thus completeness. Here, we discuss this effect. We also put our high-resolution radio observations into context with other AGN indicators and the general idea of co-evolution of black holes and star formation.

5.1. Sensitivity to extended source structure, completeness, and detection bias

At the given sensitivities per beam, the LoTSS-Deep-HD image will be ~ 326 times less sensitive to diffuse emission compared to the LoTSS-Deep image. A back-of-the-envelope estimate of

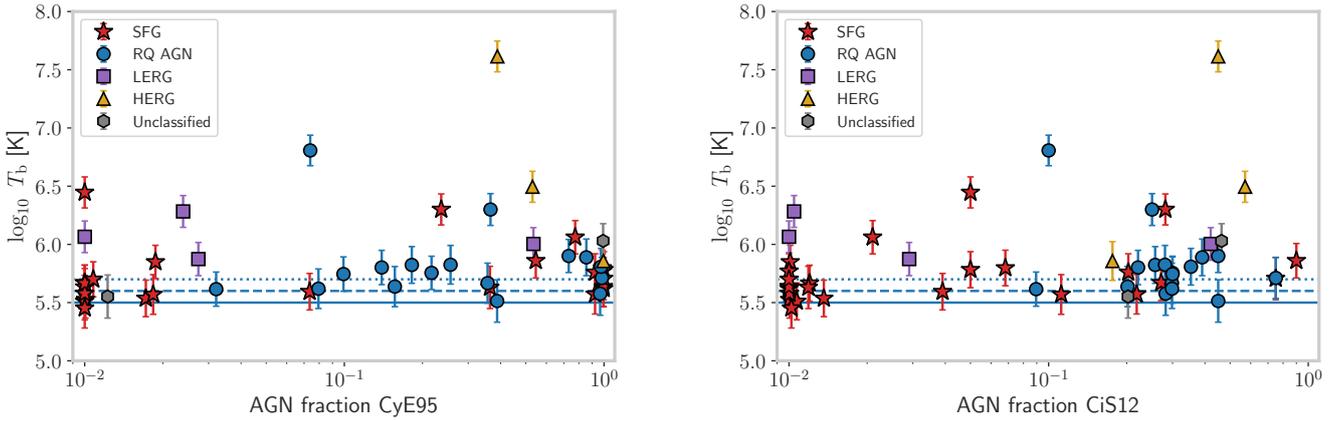


Fig. 6. Brightness temperature versus AGN fraction. *Left:* AGN fraction derived from the CyE95 fit. *Right:* AGN fraction derived from the CiS12 fit. The solid, dashed, and dotted blue lines indicate the maximum brightness temperature from starbursts for a synchrotron spectral index α of -0.7 , -0.8 , and -0.9 , respectively, as per Condon et al. (1991). Points above these lines cannot be driven by pure star formation. Estimates with $f_{\text{AGN}} < 0.01$ were offset by 0.01 for visual clarity. Points are coloured by the classification of Best et al. (in prep.).

this limiting surface brightness can be made as follows. For a patch of extended emission to be detected at the 5σ level, in the $0.3'' \times 0.4''$ resolution image, it would need an intensity of $\geq 125 \mu\text{Jy beam}^{-1}$ at the phase-centre of the field, or higher towards the edges. A one arcsec² area would consist of 7.1 beam elements, implying a flux density limit of roughly 0.9 mJy for a uniform patch of emission of that size, and thus an identical peak intensity for a $6''$ -unresolved source. We therefore expect sources with flux densities of the order of a milliJansky and that are extended to be easily missed in the LDHD image, or at best be partially recovered.

Only one third of the HLIRGs are recovered above the detection limit, making it clear that the full LDHD sample is incomplete with respect to the LoTSS detections. In Fig. 12, we show the fraction of detected sources as a function of their LoTSS-Deep $6''$ flux density. From this it can be seen that 81% of sources brighter than 2 mJy are detected. The completeness then starts to drop, before a steep drop in completeness occurs between 0.7 and 0.8 mJy. In the top left panel of Fig. 9 of Best et al. (in prep.), it is seen that SFGs become the dominant population below 1.6 mJy. This is of the same order as the earlier estimate. Combined with the generally larger scale and more diffuse nature of the emission of SFGs, we therefore interpret the non-detections as being closer to SFGs than AGNs.

Finally, the reduced sensitivity to extended structure introduces an obvious but significant bias towards compact sources. This biases sub-arcsecond imaging to inherently prefer the detection of compact nuclear components such as AGNs and distant unresolved or barely resolved starbursting galaxies. Imaging at an intermediate angular resolution around $1\text{--}2''$ would aid in determining the compactness of the LDHD non-detections.

5.2. Addressing the tension between HLIRGs and galaxy formation models

To relieve the tension between the GALFORM galaxy formation model (which assumes all the IR emission is reprocessed stellar emission; Gutcke et al. 2015) and the number densities of extreme star-forming systems found by Wang et al. (2021), those densities would have to come down by orders of magnitude. Here, we discuss our findings in this context.

Our results show that even conservatively and after removing AGNs identified by indicators other than brightness tem-

perature, the majority of the remaining HLIRGs classified as SFGs are likely to contain AGNs. This is higher compared to Morabito et al. (2022b), where only 12% of the more general SFG population was found to have a high brightness temperature, but expected due to the extreme nature of HLIRGs. The high prevalence of AGNs is consistent with other literature studies as well, finding that a substantial fraction of HLIRGs display signs of AGN activity (e.g., Farrah et al. 2002; Ruiz et al. 2013; Gao et al. 2021). However, Rowan-Robinson et al. (2018), for example, concluded that the majority of IR emission from their sample was dominated by star formation, while Symeonidis & Page (2018) proposed that AGNs can explain the full IR luminosity of sources with $L_{\text{IR}} > 10^{13.5} L_{\odot}$. We see a similar ‘threshold’ in Fig. 13, where all LDHD-detected HLIRGs with $L_{\text{IR}} > 10^{13.5} L_{\odot}$ exceed our most conservative threshold of $\log_{10} T_{\text{b}} > 5.7$ for every model, except in the case of CyE95 where one falls below it, and have substantial estimated AGN fractions as well. A difference is that in our sample these sources lie at $z > 3$, while Symeonidis & Page (2018) studied sources in the range of $z = 1\text{--}2$. A substantial AGN fraction in combination with a brightness temperature in excess of our most conservative threshold could support the idea that AGNs contribute substantially to the IR output in high- z high- L_{IR} HLIRGs. However, among the LDHD non-detections the SED fits yield between 18 and 35 HLIRGs also having $L_{\text{IR}} > 10^{13.5} L_{\odot}$, spanning the same range of estimated redshifts of $z > 2.9$. Thus, not all of these extreme high- L_{IR} sources necessarily harbour a radio-loud, high- T_{b} AGN.

To see if the tension between the observed number density of extreme star formers and theoretical models can be resolved, we can, as a logical extreme, consider sources that exceed the brightness temperature threshold to be 100% powered by an AGN. We consider the sole source falling below all thresholds to be a ‘pure SFG’ for this argument, and it has an estimated redshift of $z \approx 2.1$. This falls in the $z = 2.0\text{--}2.2$ bin of Fig. 10 of Wang et al. (2021). The tension here is ‘smallest’ at only an order of magnitude. Four LDHD-detected sources and nine LDHD-undetected sources fall in this redshift range. Only the aforementioned SFG falls below all brightness temperature thresholds. It also has an estimated AGN fraction consistent with zero for both the CyE95 and CiS12 SED-fitting models. Of the undetected HLIRGs, two have notable AGN fractions of 59% and 20% (CiS12) or 28% and 44% (CyE95). If for argument’s sake we assume all three

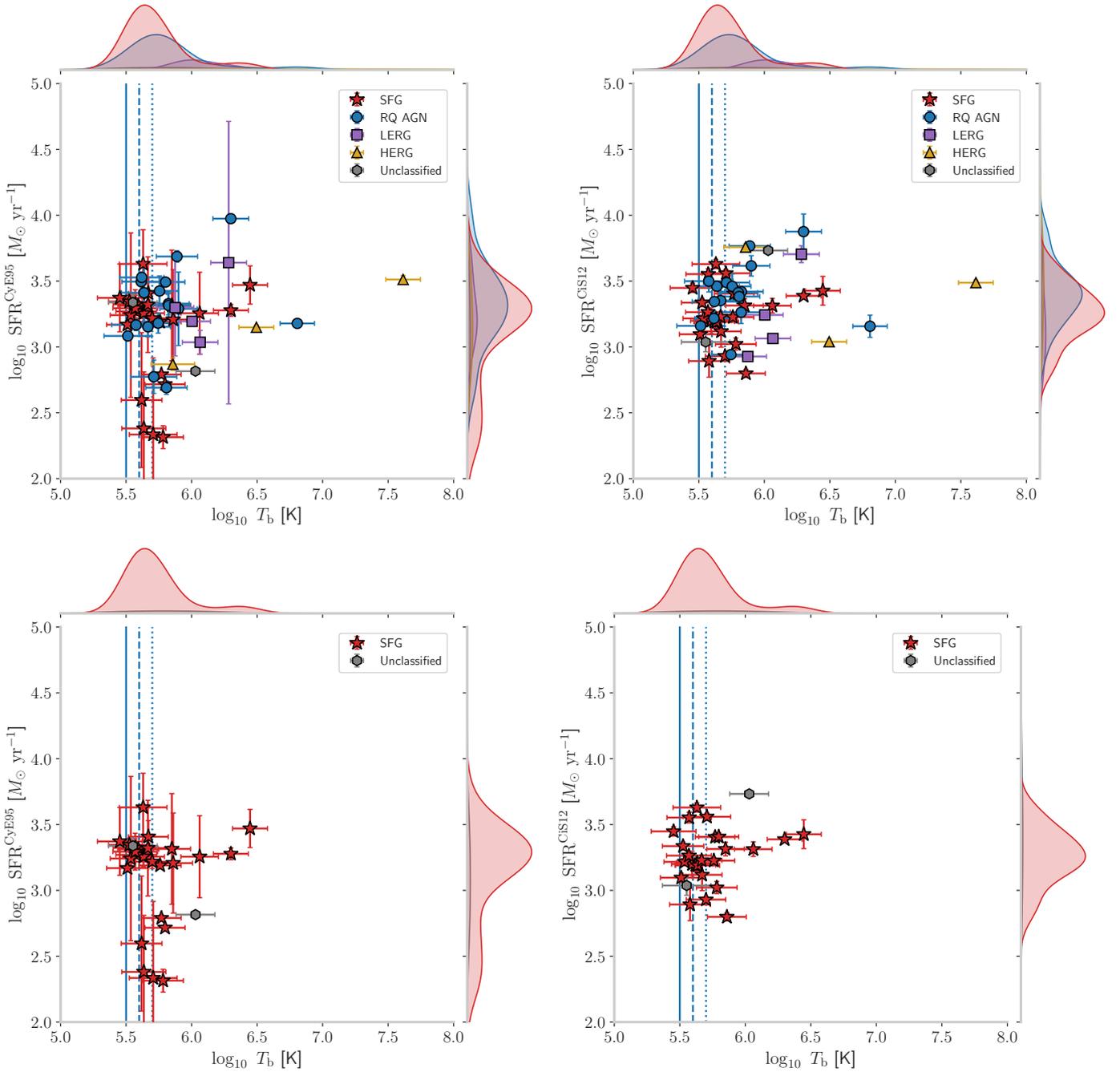


Fig. 7. Star formation rate versus brightness temperature of the 51 HLRIGs detected in LDHD image at the $5\sigma_{\text{rms}}$ level. *Top:* full sample, coloured according to the classification of Best et al. (in prep.). *Bottom:* showing only SFGs and unclassified sources; 27 in total. *Left:* SFR as estimated from the CyE95 model. *Right:* SFR estimated from the CiS12 model. Marginalised distributions over the SFR and brightness temperature are shown on top and to the right of each panel. The colours correspond to the colours of the points.

of the high- T_b sources to be ‘pure AGNs’, the number of ‘pure SFGs’ in this bin would drop from 13 to 10, which is a factor of 1.3 or 23%. In Fig. 10 of Wang et al. (2021), it can be seen that this is insufficient to reconcile the number counts with the predictions from the GALFORM model. In higher redshift bins this reduction can be of the order of a factor of 2, but by $z = 3$ the discrepancy is already two orders of magnitude and increasing. Furthermore, realistically only the brightest HLRIGs may be fully powered by AGNs, lessening this ‘correcting factor’ again. The tension between models and observations cannot be resolved, even if we assume all our detections to be completely AGN-driven.

If we include the non-detections of the same $z = 2.0$ – 2.2 bin and also assume they are fully powered by an AGN, the required order of magnitude to reconcile theory and observations can be reached. Such a scenario may be possible under different model assumptions, as found by Symeonidis et al. (2016), where the AGN heats dust to kiloparsec scales and the actual SFR is notably lower than the IR luminosity. This might be implied under the model assumptions in this work (Symeonidis 2022; Symeonidis et al. 2022). Incidentally, the E95 model does incorporate heating of dust beyond the torus and in unobscured directions (dubbed ‘conical dust’), and in general it yields some of the lowest SFRs of the models explored for our sample, as

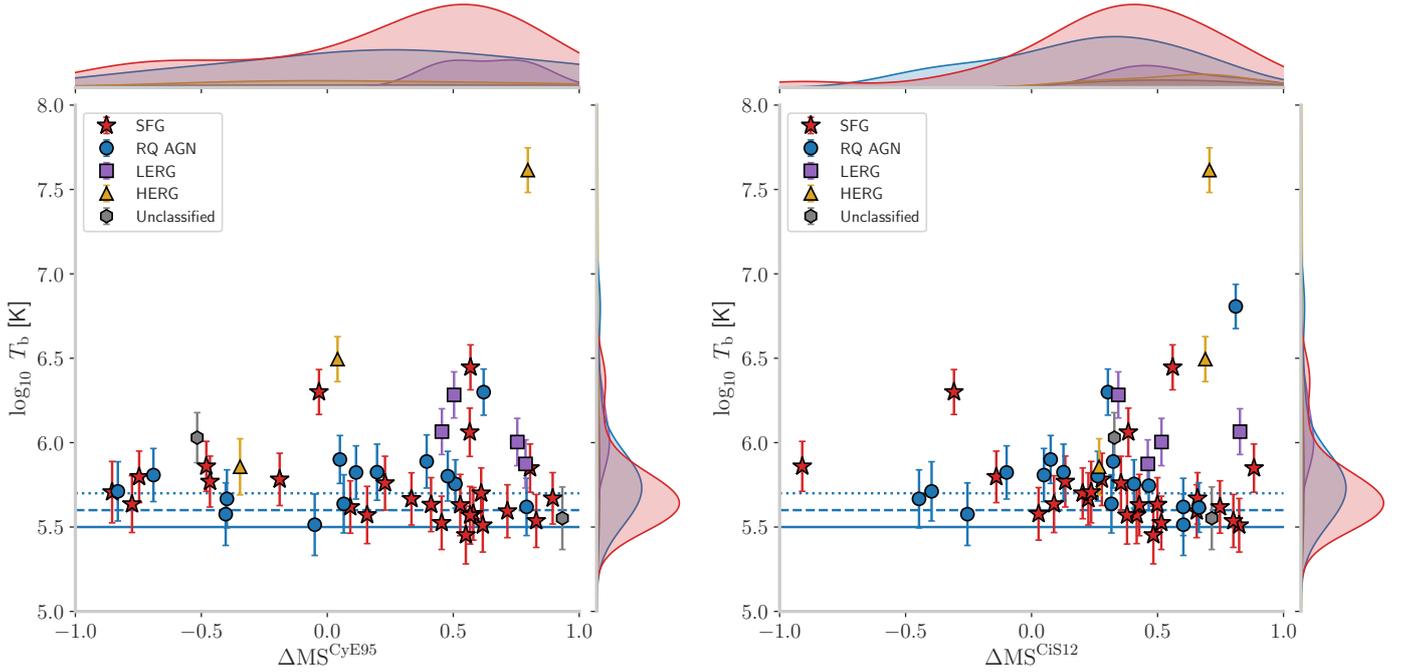


Fig. 8. Brightness temperature T_b versus distance to the main sequence ΔMS for HLIRGs above the LDHD detection threshold. ΔMS is shown for the CyE95 (left) and CiS12 (right) models. Points are coloured by their classification according to Best et al. (in prep.). The solid, dashed, and dotted blue lines indicate the maximum brightness temperature from starbursts for a synchrotron spectral index α of -0.7 , -0.8 , and -0.9 , respectively, as per Condon et al. (1991). Points above these lines cannot be driven by pure star formation.

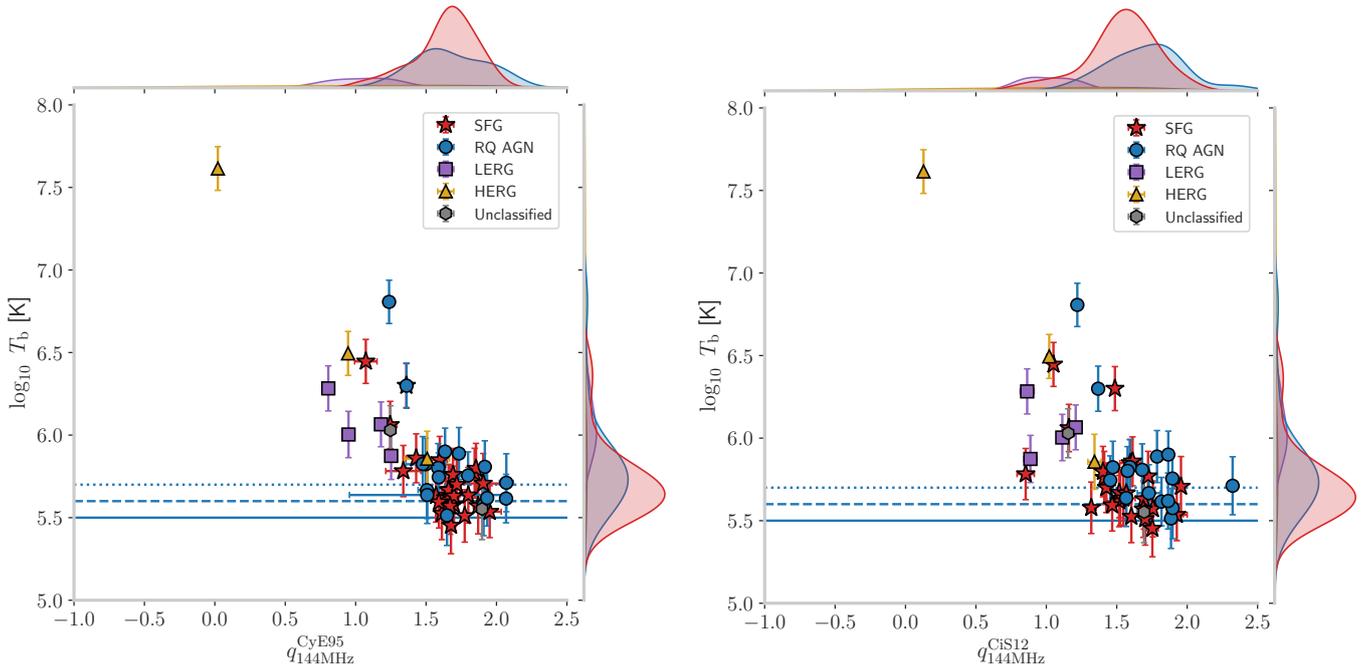


Fig. 9. Brightness temperature versus radio excess. q was determined using IR luminosities from the CyE95 (left) and CiS12 (right) models. The solid, dashed, and dotted blue lines indicate the maximum brightness temperature from starbursts for a synchrotron spectral index α of -0.7 , -0.8 , and -0.9 , respectively, as per Condon et al. (1991). Points above these lines cannot be driven by pure star formation.

seen in Fig. 7. Exploring different models such as those by Symeonidis et al. (2016), where the AGN heats dust on larger scales, would be an interesting follow-up to the work presented here.

Finally, we provide a note about the photometric redshift estimates. Photometric redshifts can be an appreciable source of uncertainty. Incorrect redshift estimates will affect the red-

shift distribution of sources and thus ultimately the tension. This issue is not easily resolved until most sources in the sample have had spectroscopic follow-up. For a few sources, spectroscopic follow-up has been conducted with Subaru and the Northern Extended Millimeter Array (NOEMA; priv. comm. L. Wang and Gao et al. 2022). The photometric estimates agree well with these spectroscopic values lying within 5% of each other. This is

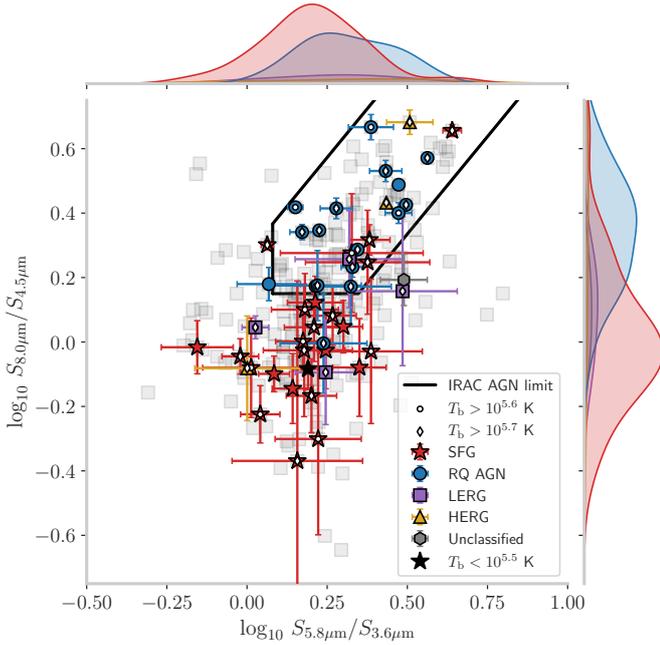


Fig. 10. IRAC colour-colour diagram for LDHD-detected sources. The black line denotes the region in which IRAC AGN are expected, as per the Donley et al. (2012) criteria. Faded grey squares indicate the LDHD-undetected sample, while the coloured symbols indicate LDHD-detected sources. White circles indicate sources with $\log_{10} T_b > 5.6$ and white diamonds indicate sources with $\log_{10} T_b > 5.7$. Sources with no dot or diamond have $5.5 \leq \log_{10} T_b \leq 5.6$, with the exception of the black star, which indicates the single source with $\log_{10} T_b < 5.5$.

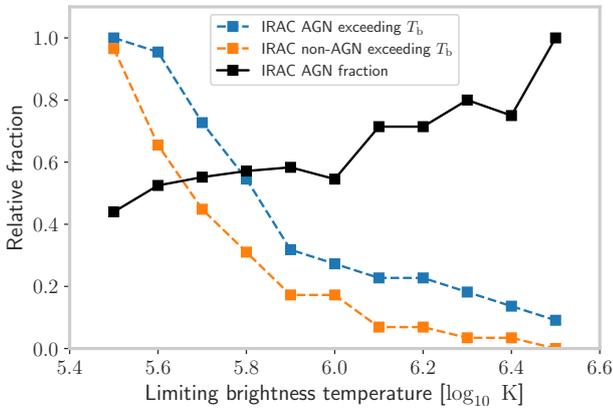


Fig. 11. Fraction of IRAC sources as function of certain brightness temperature limit. The black solid line indicates the fraction of IRAC AGN in sources that exceed the limiting brightness temperature. The red and blue dashed lines indicate the fraction of IRAC AGN and IRAC non-AGN, respectively, that exceed this limit.

consistent with the 7% uncertainty set on the photometric estimates by Duncan et al. (2021). For now, we therefore conclude that the photometric redshifts appear to be of good quality, but the potential for uncertainty here needs to be kept in mind, and more work is needed to understand the whole population.

5.3. Brightness temperature as a dust-unobscured tracer of AGN activity

To broadly assess the use of brightness temperatures as an AGN indicator in these HLIRGs, the brightness temperature measure-

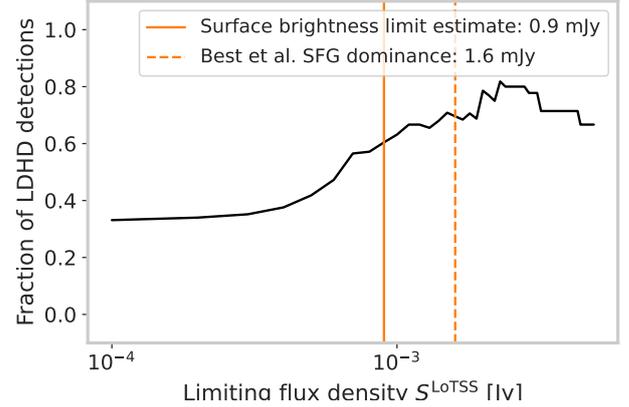


Fig. 12. Fraction of LDHD detections versus LoTSS flux density S_{LoTSS} (black). At $S_{\text{LoTSS}} > 2$ mJy, a maximum of 81% is reached. The orange lines indicate the estimated surface-brightness sensitivity limit towards extended emission (solid) and the point where SFGs become dominant (dashed, Best et al. in prep.).

ments were compared to three other indicators. We remind the reader that the convolved source sizes (i.e. the intrinsic source size convolved with the point spread function) were used to estimate brightness temperatures and that these overestimate the true source size both due to smearing and instrumental effects (convolution with the point spread function). This effectively turns the measured brightness temperatures into lower limits.

Best et al. (in prep.) use four well-known SED-fitting codes (CIGALE, MagPhys, BAGPIPES and AGNFitter) to classify sources into either SFGs, RQ AGNs, LERGs, and HERGs. The SED-fitting from Gao et al. (2021), based on broad-band photometry SED-fitting using the CIGALE code and CYGNUS radiative transfer models, provides estimates of various parameters such as L_{IR} and f_{AGN} for various combinations of SED-fitting codes and models. Here, we compare these classifications and quantities with our brightness temperature measurements.

All HLIRGs classified by Best et al. (in prep.) as radio-loud AGNs (i.e. HERGs and LERGs) are recovered as high- T_b objects well exceeding the threshold value for AGN activity (see Table 2). Additionally, of the two unclassified sources, one clearly exceeds the threshold and the other marginally exceeds the least conservative threshold. For SFGs and RQ AGNs, the story differs, as both types of sources have objects exceeding the thresholds to various degrees. All RQ AGNs are detected and exceed the least conservative threshold. In the case of SFGs, it demonstrates how high-resolution radio observations can complement existing classification to identify additional AGNs that cannot be picked up through the SED fitting. These sources are likely RQ AGNs, where the AGN is weak and does not dominate the SED, but may still be a significant component of the overall IR light.

Comparing T_b with L_{IR} and f_{AGN} , the estimated IR luminosities and AGN fractions estimated by CYGNUS and CIGALE both show general correlation with brightness temperature. We do see some differences between the different SED fitting methods, notably with respect to the difference in estimated IR luminosities. This is discussed in Appendix B.

The existence of a correlation between radio emission and IR emission from star-forming galaxies has long been known (de Jong et al. 1985; Helou et al. 1985). This has since been used as a way to trace star formation through the radio-to-IR ratio q (Bell 2003; Calistro Rivera et al. 2017; Delvecchio et al. 2021). At the same time, this ratio can be used to search for

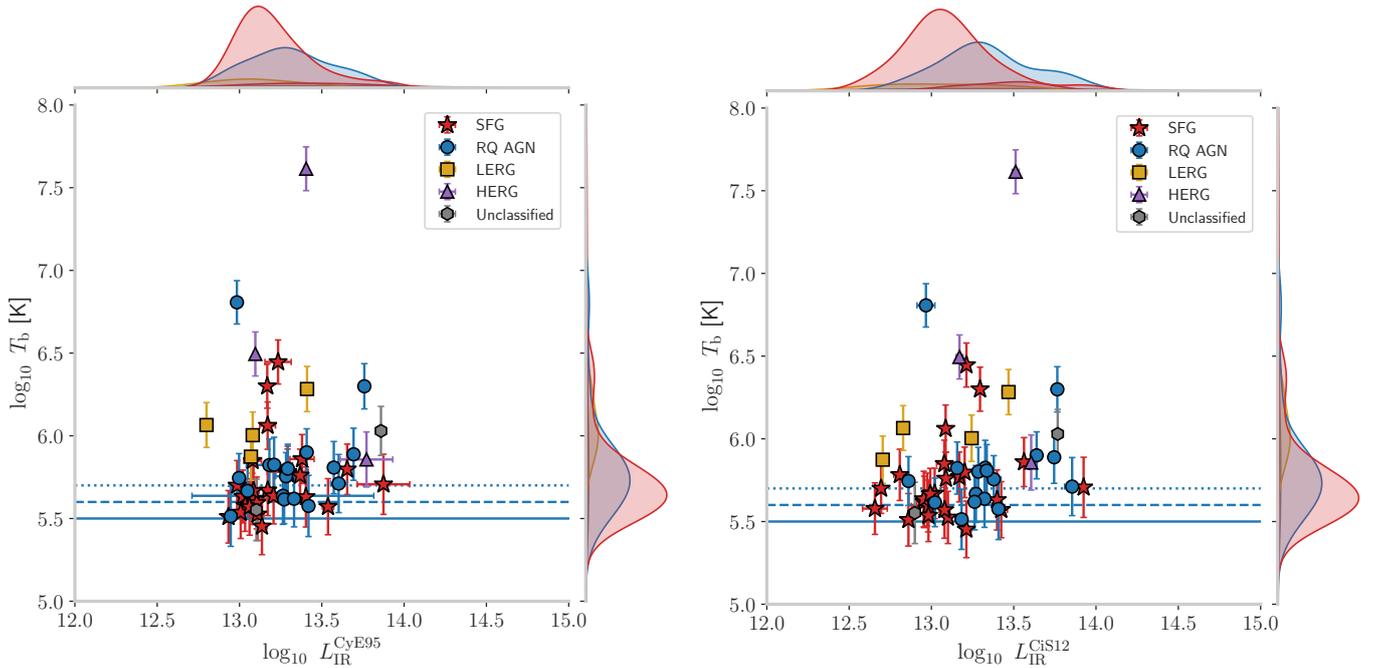


Fig. 13. Brightness temperature versus infrared luminosity determined from the CyE95 (*left*) and CiS12 (*right*) models. The solid, dashed, and dotted blue lines indicate the maximum brightness temperature from starbursts for a synchrotron spectral index α of -0.7 , -0.8 , and -0.9 , respectively, as per [Condon et al. \(1991\)](#). Points above these lines cannot be driven by pure star formation.

AGNs, by looking for an excessive q -value (e.g., [Bonzini et al. 2015](#); [Delvecchio et al. 2017](#)). A lower value of q_{144} indicates a higher 144 MHz luminosity compared to a given IR luminosity. A higher brightness temperature is the result of a brighter and/or more compact radio source. One would thus expect a correlation between q_{144} and T_b if both trace AGN activity well. Figure 9 shows the measured brightness temperature versus the calculated q -value for the CyE95 and CiS12 SED fits. A clear correlation of a higher radio excess being associated with a higher brightness temperature is seen, as expected. The correlation is most obvious for high values of T_b and the HERGs and LERGs. For lower values of T_b and the SFGs and RQ AGNs it is less obvious. This correlation supports the idea that brightness temperatures is a sensible tracer of AGN activity and reaffirms either one’s ability to be used for AGN identification.

In Fig. 10, the LDHD-detected sources were compared to their mid-IR colours. Almost all IRAC AGNs are classified as AGNs by Best et al. (in prep.) as well. They are predominantly RQ AGNs. The SFGs in this region all exceed the most conservative threshold of $\log_{10} T_b > 5.7$, strongly suggesting AGN activity in those sources as well. All but one of the LDHD-detected SFGs outside this region also exceed one of the brightness temperature thresholds. In total, we find one third of the HLIRGs within the LDHD footprint to be consistent with hosting an AGN. We find this consistent with the literature in that a substantial amount of these sources are thought to harbour AGNs (e.g., [Farrah et al. 2002](#); [Rowan-Robinson et al. 2018](#); [Gao et al. 2021](#)) and that they could be a substantial driver behind the IR emission (e.g., [Symeonidis & Page 2018, 2021](#)). All of the undetected sources are classified as SFGs by Best et al. (in prep.). This is consistent with the limitations and biases in the LDHD image, as its high resolution leads to a lower surface brightness sensitivity, biasing it against the detection of more diffuse sources such as SFGs. We also note a number of these sources

do enter the Donley region. Best et al. (in prep.) elaborate on this by noting that this cut is incomplete for composite systems (which a substantial number of HLIRGs likely are). For $z > 2.5$, they add, it becomes difficult to distinguish AGN templates from galaxy templates, and 70% of non-detections within the LDHD footprint exceed that redshift.

On their own, HLIRGs above the LDHD detection threshold show no direct likelihood of also being IRAC AGNs. Compared to the undetected HLIRGs, however, they do appear more likely to be IRAC AGNs. To test the statistical significance of this trend, we bootstrapped the measured mid-IR flux densities 10,000 times and determined the fraction of IRAC AGNs each time. Bootstrapping was done by assuming each measurement to be drawn from a Gaussian distribution with a mean equal to the measured value and a standard deviation given by the uncertainty in the Deep Fields catalogue. This yielded fractions of IRAC AGNs of $38\% \pm 3\%$ for LDHD-detected sources and of $31\% \pm 3\%$ for LDHD-undetected sources. Compared to HLIRGs below the LDHD detection threshold, however, they do appear more likely to be IRAC AGNs at a 2.6σ difference in the respective fractions of IRAC AGNs. Bootstrapping indicates a mild statistical significance of 2.6σ difference between the detected and undetected sources. The mild tendency for high- T_b sources to also be IR AGNs may indicate a preference for radio jets to trigger during an obscured evolutionary phase, such as seen in, for example, [Patil et al. \(2020\)](#), who found young jets in their sample of luminous heavily obscured quasars.

Figure 11 shows an increasing fraction of IRAC AGNs with brightness temperature. A tendency for IRAC AGNs to exceed the brightness temperature threshold more often than IRAC non-AGNs is seen. This correlation between the IR and radio activity is also seen in Figs. 13 and A.3, where L_{IR} is seen to increase with T_b . Notably, the trend is most apparent from the RQ AGNs and SFGs, which may not always be identified as radio AGNs by low-resolution radio observations. This makes brightness

temperatures especially useful to pinpoint relatively weak AGNs that would escape cuts based on radio excess, for example, for these obscured sources.

5.4. Co-evolution of black hole growth and star formation

The growth of black holes and the build-up of stellar populations over cosmic time are remarkably alike (see e.g., the review of Heckman & Best 2014). This is not completely unexpected. After all, they have a common fuel source: the gas reserves of the galaxy.

In the top panel of Fig. 7 we see that, broadly speaking, the higher brightness temperature HLIRGs seem to be associated with higher SFRs and display a larger spread in T_b . Sources with lower SFRs display both lower values of T_b and a smaller spread. Furthermore, the detected HLIRGs seemingly avoid the region of parameter space corresponding to a high brightness temperature, but a low SFR. Figure A.6 shows the SFR versus brightness temperature for the full sample. From this it becomes clear that this region is predominantly inhabited by LERGs and HERGs. Of interest here is that this is not a general avoidance zone, but a radio-loud, AGN-dominated region. This raises the question of why the HLIRGs in our sample seem to avoid this particular combination of parameters. To sustain the incredibly high SFR in HLIRGs, these galaxies should have a large reservoir of gas to draw from. This could naturally supply the central engine with fuel from that same supply, leading to a simultaneous increase in both star formation activity and AGN activity. We therefore speculate that the tentative trend of brightness temperature with star formation rate supports the idea of co-evolution in these massive galaxies, where large gas reservoirs feed both the star formation process and AGN activity. If IR emission as a tracer of star formation breaks down at high L_{IR} , however, as raised by Symeonidis (2022) and Symeonidis et al. (2022), the interpretation of this correlation would be different, linking two indicators of AGN activity with each other. Further studies on the impact of the chosen and alternative SEDs on the IR-derived quantities and independent measures of AGN activity (e.g., from their radio luminosity Morabito et al. 2022a) would be needed to interpret this further.

6. Conclusion

Hyper-luminous infrared galaxies are difficult to reproduce in simulations, both in terms of their extreme SFR and observed number density. In this work, we studied the HLIRG population in the Lockman Hole field from a sub-arcsecond perspective at 144 MHz using the ILT. These high-resolution radio data were used to assess AGN activity in HLIRGs by means of brightness temperature measurements in order to see if the tension between observations and current galaxy formation models could be reduced. We conclude our findings as follows:

- The low-frequency brightness temperature T_b is an effective tool for revealing AGN activity in dusty objects such as HLIRGs. All known radio-loud and radio-quiet AGNs among the sample are recovered, clearly exceeding the upper limit for pure star formation determined by Condon et al. (1991); additional AGNs not readily picked up by the other methods discussed in this paper were identified.
- The total number of HLIRGs in this field harbouring bona fide AGNs could be increased from 16% to 26–32% (depending on how conservatively one sets the threshold). All of these additional high- T_b AGNs were initially classified as SFGs. One of the LDHD-detected sources remains consis-

tent with pure star formation. A notable amount of HLIRGs may thus harbour a radio AGN that is not picked up on when using UV-to-IR or low-resolution radio observations.

- The additional AGNs identified here cannot relieve the tension between the observed number density of HLIRGs and predictions from the GALFORM theoretical model. Conservatively, the number density of extreme star-forming systems can be only be reduced by a factor of $\sim 1-1.5$ at cosmic noon using this additional information and assuming AGN dominance for all detected sources. On the other hand, if we assume even the non-detections to be fully powered by an AGN (see the penultimate point), the required order of magnitude at cosmic noon can be reached. At higher redshifts the discrepancy is larger and remains so, however.
- Brightness temperature is seen to broadly correlate with AGN dominance probed by other indicators such as the radio excess, SED-fitted AGN fractions, and IR colours.
- Higher SFRs generally allow for more higher brightness temperatures. We see this as an indication of co-evolution between star formation and AGN activity. In extreme scenarios, the AGN contribution in sources with the highest SFRs could be significantly underestimated, such as that found by Symeonidis et al. (2016), which could change the interpretation of this trend.
- High- T_b AGNs show a mild preference towards being associated with IR AGNs.

In this work, we give a taste of what high-resolution radio observations can offer in terms of population studies of HLIRGs, but the sample size is still small, therefore limiting the robust identification of trends. Future sub-arcsecond surveys over large areas of the sky will provide detailed radio data for a large number of these extremely IR-luminous objects, penetrating the dusty curtain that enshrouds them, allowing us to peer deeper into the engine powering their high infrared luminosities.

Acknowledgements. This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT) under project code LT10_012. LOFAR (van Haarlem et al. 2013) is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefited from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d’Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland. IP acknowledges support from INAF under the SKA/CTA PRIN “FOReCaST” and the PRIN MAIN STREAM “SAuROS” projects. K.M. has been supported by the National Science Centre (UMO-2018/30/E/ST9/00082) PNB is grateful for support from the UK STFC via grant ST/V000594/1 RJvW acknowledges support from the ERC Starting Grant ClusterWeb 804208. WLW acknowledges support from the CAS-NWO programme for radio astronomy with project number 629.001.024, which is financed by the Netherlands Organisation for Scientific Research (NWO). This work has made use of the Dutch national e-infrastructure with the support of SURF Cooperative through grant e-infra 180169. Images and catalogues of both the deep 6”-resolution LoTSS Deep Fields data and the sub-arcsecond ILT data are publicly available at <https://lofar-surveys.org/deepfields.html> and <https://lofar-surveys.org/hdfields.html>, respectively. LKM is grateful for support from the UKRI Future Leaders Fellowship (grant MR/T042842/1).

References

- Ade, P. A. R., Aghanim, N., Arnaud, M., et al. 2016, *A&A*, 594, A13
 Armus, L., Heckman, T., & Miley, G. 1987, *AJ*, 94, 831
 Bell, E. F. 2003, *ApJ*, 586, 794
 Bonzini, M., Mainieri, V., Padovani, P., et al. 2015, *MNRAS*, 453, 1079
 Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, *A&A*, 622, A103

- Buat, V., Ciesla, L., Boquien, M., Malek, K., & Burgarella, D. 2019, *A&A*, **632**, A79
- Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, *MNRAS*, **360**, 1413
- Calistro Rivera, G., Williams, W. L., Hardcastle, M. J., et al. 2017, *MNRAS*, **469**, 3468
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, *MNRAS*, **480**, 4379
- Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, *ApJ*, **873**, 44
- Condon, J. J., Huang, Z. P., Yin, Q. F., & Thuan, T. X. 1991, *ApJ*, **378**, 65
- Cutri, R. M., Huchra, J. P., Low, F. J., Brown, R. L., & Vanden Bout, P. A. 1994, *ApJ*, **424**, L65
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, **388**, 1595
- Davé, R., Finlator, K., Oppenheimer, B. D., et al. 2010, *MNRAS*, **404**, 1355
- de Jong, T., Klein, U., Wielebinski, R., & Wunderlich, E. 1985, *A&A*, **147**, L6
- Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2017, *A&A*, **602**, A3
- Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2021, *A&A*, **647**, A123
- Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, *ApJ*, **748**, 142
- Duncan, K. J., Kondapally, R., Brown, M. J. I., et al. 2021, *A&A*, **648**, A4
- Efstathiou, A., & Rowan-Robinson, M. 1995, *MNRAS*, **273**, 649
- Farrah, D., Serjeant, S., Efstathiou, A., Rowan-Robinson, M., & Verma, A. 2002, *MNRAS*, **335**, 1163
- Farrah, D., Petty, S., Connolly, B., et al. 2017, *ApJ*, **844**, 106
- Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, *MNRAS*, **366**, 767
- Gao, F., Wang, L., Efstathiou, A., et al. 2021, *A&A*, **654**, A117
- Gao, F., Wang, L., Ramos Padilla, A. F., et al. 2022, *A&A*, **668**, A54
- Gutcke, T. A., Fanidakis, N., Macciò, A. V., & Lacey, C. 2015, *MNRAS*, **451**, 3759
- Hardcastle, M. J., Williams, W. L., Best, P. N., et al. 2019, *A&A*, **622**, A12
- Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, *MNRAS*, **428**, 2529
- Heckman, T. M., & Best, P. N. 2014, *ARA&A*, **52**, 589
- Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, *ApJ*, **298**, L7
- Hönig, S. F. 2019, *ApJ*, **884**, 171
- Jackson, N., Badole, S., Morgan, J., et al. 2022, *A&A*, **658**, A2
- Jannuzi, B. T., & Dey, A. 1999, in *Photometric Redshifts and the Detection of High Redshift Galaxies*, eds. R. Weymann, L. Storrie-Lombardi, M. Sawicki, & R. Brunner, *ASP Conf. Ser.*, **191**, 111
- Kondapally, R., Best, P. N., Hardcastle, M. J., et al. 2021, *A&A*, **648**, A3
- Lockman, F. J., Jahoda, K., & McCammon, D. 1986, *ApJ*, **302**, 432
- Lovell, C. C., Geach, J. E., Davé, R., Narayanan, D., & Li, Q. 2021, *MNRAS*, **502**, 772
- Mohan, N., & Rafferty, D. 2015, *Astrophysics Source Code Library* [record ascl:1502.007]
- Morabito, L. K., Jackson, N. J., Mooney, S., et al. 2022a, *A&A*, **658**, A1
- Morabito, L. K., Sweijen, F., Radcliffe, J. F., et al. 2022b, *MNRAS*, **515**, 5758
- Narayanan, D., Turk, M., Feldmann, R., et al. 2015, *Nature*, **525**, 496
- Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, *ApJ*, **278**, L1
- Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, *A&A*, **507**, 1793
- Oliver, S., Serjeant, S., Efstathiou, A., et al. 2000, in *ISO Survey of a Dusty Universe*, eds. D. Lemke, M. Stickel, & K. Wilke, 548, 28
- Padovani, P. 2016, *A&A Rev.*, **24**, 1
- Patil, P., Nyland, K., Whittle, M., et al. 2020, *ApJ*, **896**, 18
- Prandoni, I., Guglielmino, G., Morganti, R., et al. 2018, *MNRAS*, **481**, 4548
- Radcliffe, J. F., Barthel, P. D., Garrett, M. A., et al. 2021, *A&A*, **649**, L9
- Rowan-Robinson, M. 2000, *MNRAS*, **316**, 885
- Rowan-Robinson, M., Wang, L., Farrah, D., et al. 2018, *A&A*, **619**, A169
- Ruiz, A., Risaliti, G., Nardini, E., Panessa, F., & Carrera, F. J. 2013, *A&A*, **549**, A125
- Sabater, J., Best, P. N., Tasse, C., et al. 2021, *A&A*, **648**, A2
- Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, *ApJ*, **347**, 29
- Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, *A&A*, **598**, A104
- Shimwell, T. W., Tasse, C., Hardcastle, M. J., et al. 2019, *A&A*, **622**, A1
- Smith, H. E., Lonsdale, C. J., & Lonsdale, C. J. 1998, *ApJ*, **492**, 137
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, **214**, 15
- Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popović, L. Č. 2012, *MNRAS*, **420**, 2756
- Sweijen, F., van Weeren, R. J., Röttgering, H. J. A., et al. 2022, *Nat. Astron.*, **6**, 350
- Symeonidis, M. 2017, *MNRAS*, **465**, 1401
- Symeonidis, M. 2022, *MNRAS*, **509**, 3209
- Symeonidis, M., & Page, M. J. 2018, *MNRAS*, **479**, L91
- Symeonidis, M., & Page, M. J. 2021, *MNRAS*, **503**, 3992
- Symeonidis, M., Giblin, B. M., Page, M. J., et al. 2016, *MNRAS*, **459**, 257
- Symeonidis, M., Maddox, N., Jarvis, M. J., et al. 2022, *MNRAS*, **514**, 4450
- Tasse, C., Shimwell, T., Hardcastle, M. J., et al. 2021, *A&A*, **648**, A1
- Tsai, C.-W., Eisenhardt, P. R. M., Wu, J., et al. 2015, *ApJ*, **805**, 90
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, **556**, A2
- Wang, L., Gao, F., Best, P. N., et al. 2021, *A&A*, **648**, A8

Appendix A: Other code-model combinations

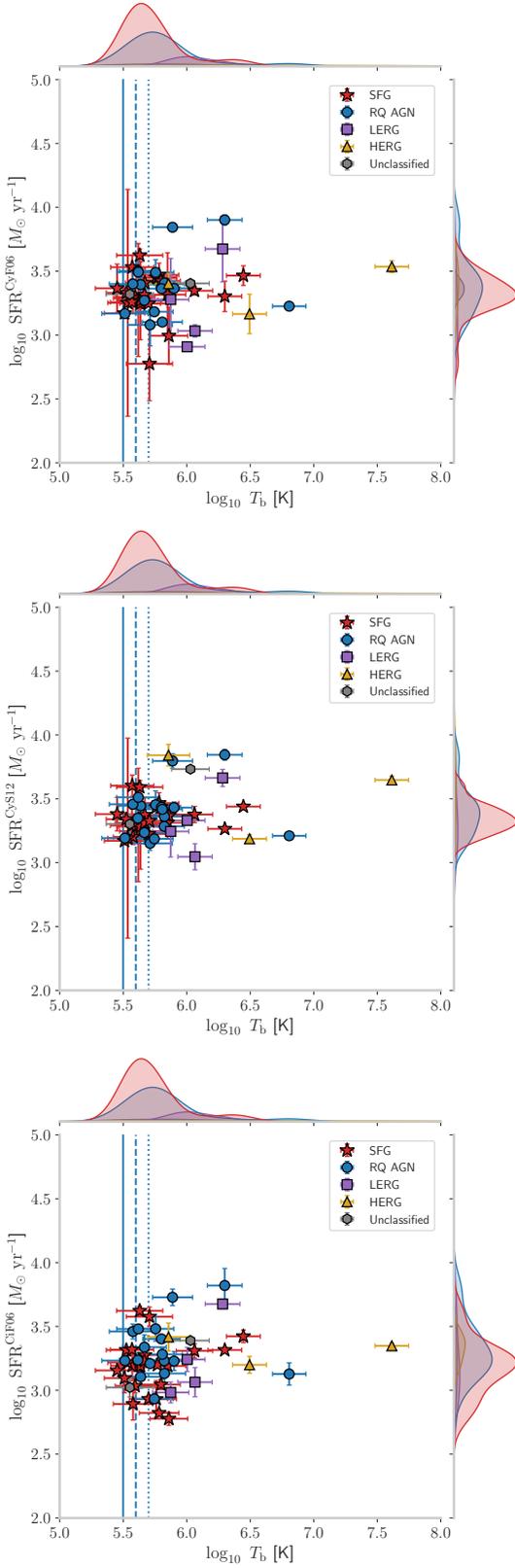


Fig. A.1. Star formation rate versus brightness temperature for CyF06, CyS12, and CiF06 models.

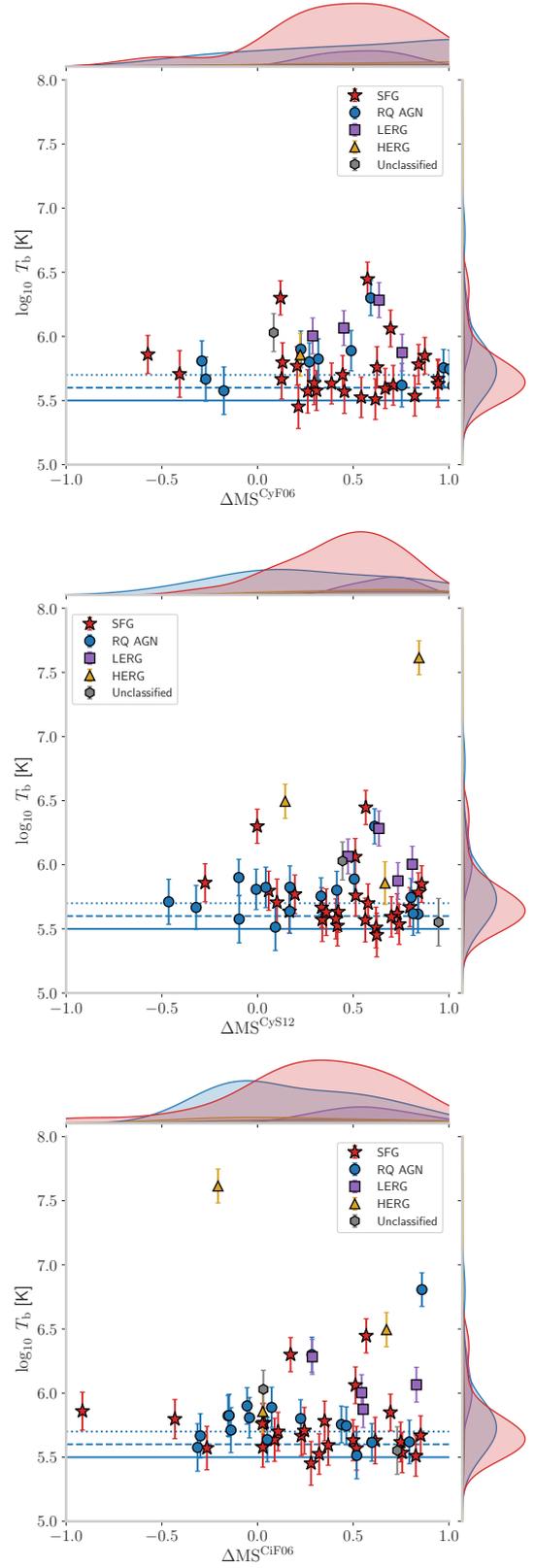


Fig. A.2. Brightness temperature versus distance to star-forming main sequence for the CyF06, CyS12, and CiF06 models.

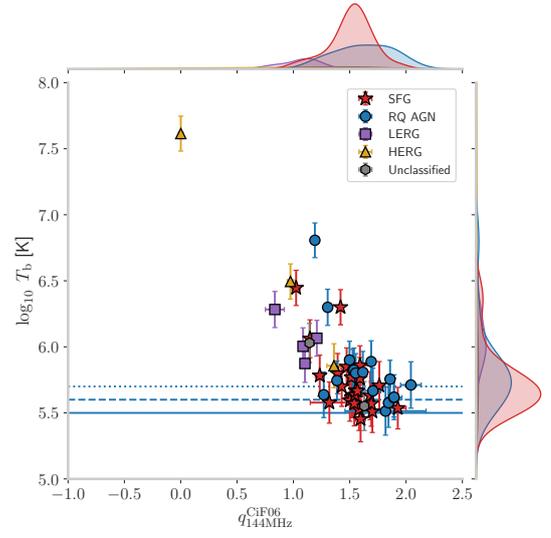
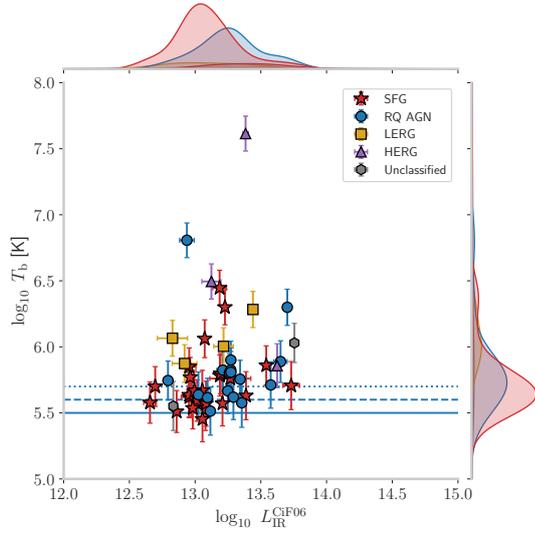
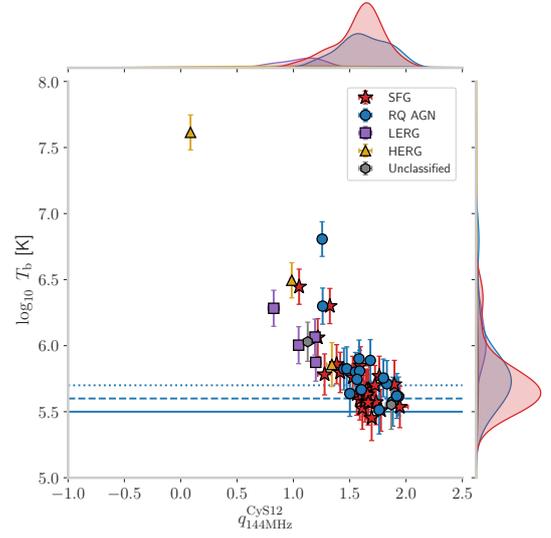
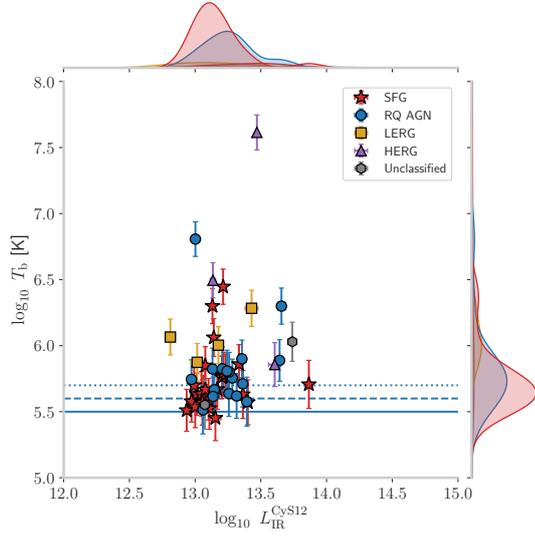
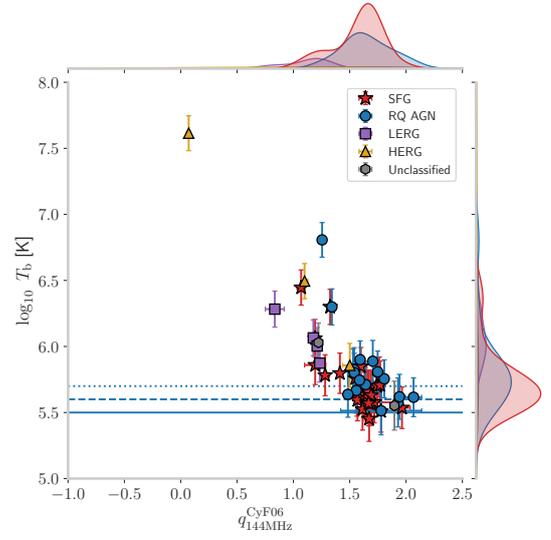
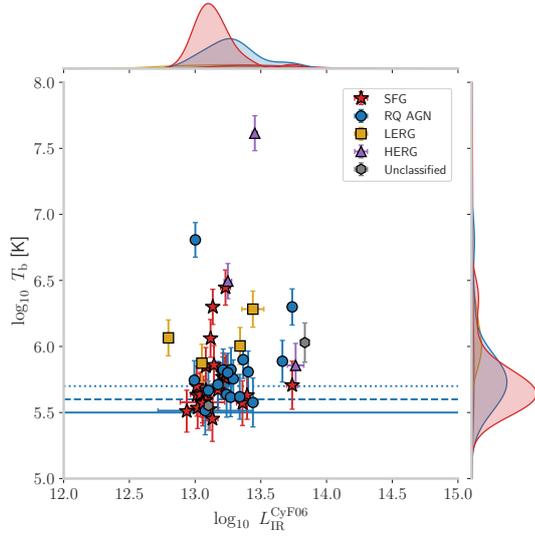


Fig. A.3. Brightness temperature versus infrared luminosity for the CyF06, CyS12, and CiF06 models.

Fig. A.4. Brightness temperature versus radio-excess for the CyF06, CyS12, and CiF06 models.

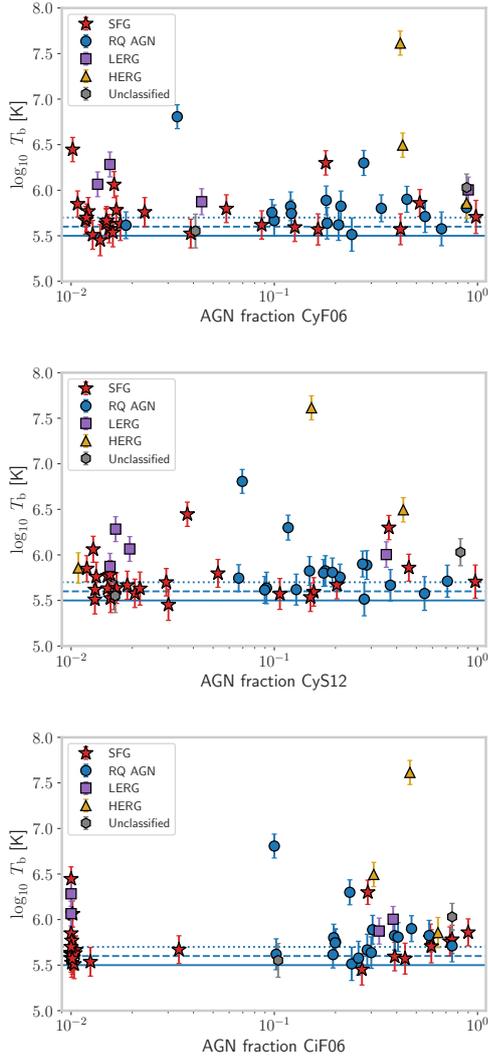


Fig. A.5. Brightness temperature versus AGN fraction for CyF06, CyS12, and CiF06 models.

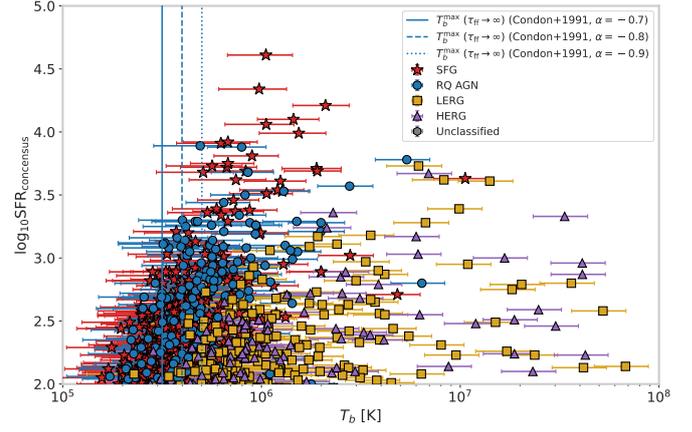


Fig. A.6. Star formation rate versus brightness temperature for the full sample of LDHD sources. Axis ranges are cropped to reflect the same range that our HLRG sample occupies.

Appendix B: IR luminosity difference

Figure B.1 plots brightness temperature and the estimated AGN fraction versus the difference in estimated IR luminosity $L_{\text{IR}}^{\text{CYGNUS}} - L_{\text{IR}}^{\text{CIGALE}}$. As this difference increases (that is, moves towards positive values), an increase in T_b for SFGs is seen (top panel). At the same time f_{AGN} strongly decreases as this difference increases (middle and bottom panel). This may be related to the assumptions made in the fitting codes and to the difficulty of disentangling low-level AGN activity from, for example, star formation activity. A detailed analysis of these trends and their significance is beyond the scope of this paper, but we note that Buat et al. (2019), for example, find that different dust attenuation laws should be used for their sample of star-bursting sources. This highlights the importance of model assumptions.

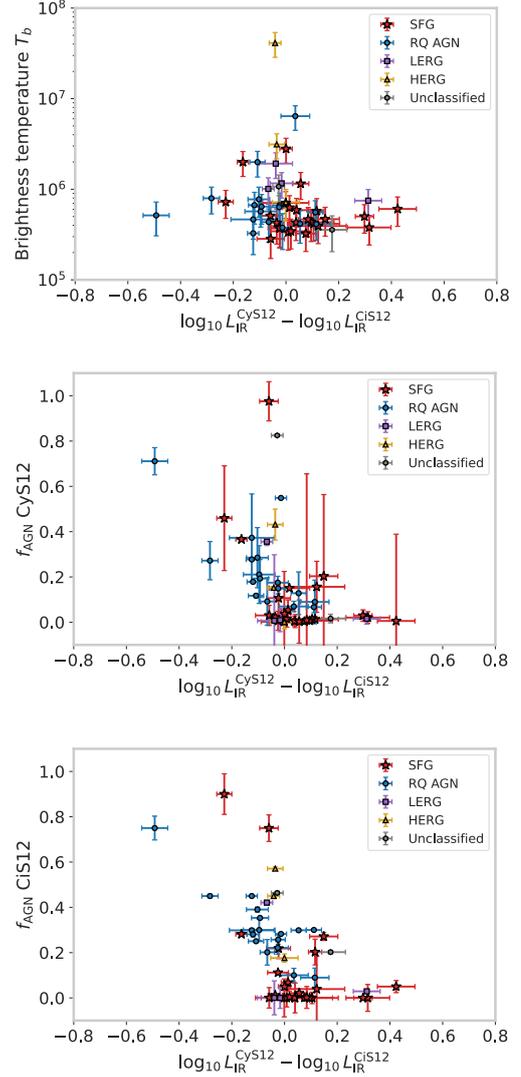


Fig. B.1. Brightness temperature and AGN fractions for the S12 model versus the difference in estimated IR luminosity.