THE UNIVERSITY *of* EDINBURGH

Role of AGN feedback in galaxy evolution

Rohit Kondapally

Doctor of Philosophy
The University of Edinburgh
July 2021

Abstract

Over the past two decades, detailed studies in the nearby Universe have shown that accreting supermassive black holes (SMBHs; or active galactic nuclei; AGN) can have a significant effect on their host galaxies, suppressing star formation and regulating their growth (known as AGN feedback). One of the most striking examples of AGN feedback in action comes from detailed studies of local AGN which exhibit powerful bi-polar jet outflows that can deposit significant energy into the galaxy halo, heating the surrounding gas and thereby regulating their own growth and suppressing star-formation activity. Therefore, studying the properties of AGN and the galaxies in which they reside is crucial in understanding and further developing our current models of galaxy formation and evolution.

AGN can be split into two distinct categories, based on the accretion efficiency of the SMBH: radiative-mode AGN, and jet-mode AGN. Radiative-mode AGN are associated with efficient accretion, typically from cold gas, resulting in the formation of a geometrically thin, optically thick accretion disk that is typically surrounded by a dusty obscuring structure; these AGN are known to drive powerful outflows. Jet-mode AGN are associated with inefficient accretion, typically from hot gas, and display powerful bi-polar synchrotron radio jets that emit the bulk of their energetic output into the surrounding medium in the form of mechanical energy; these AGN are identified as such based on radio observations, showing no signs of AGN activity (e.g. accretion disk or torus) at other wavelengths. Based on the nature of the excitation lines, the jet-mode and radio-loud radiative-mode populations are also known as low-excitation radio galaxies (LERGs) and high-excitation radio galaxies (HERGs), respectively. However, our understanding of these AGN and their feedback effect is built primarily from detailed local Universe observations. Determining the physical mechanisms underpinning triggering and fuelling of AGN and how this affects AGN feedback activity across cosmic time is crucial but lacking.

In this thesis, I address this shortcoming using deep observations carried out by the LOw

Frequency ARray (LOFAR) telescope: the LOFAR Deep Fields; this forms the deepest radio continuum survey to date at low frequencies. I generated key science-enhanced datasets using this survey and studied the cosmic evolution of AGN feedback from low-luminosity radio-AGN within the past 10 Gyrs and how this feedback affects the growth and evolution of galaxies.

In the first science chapter of the thesis, I detail the pipeline I developed to generate new, more robust multi-wavelength catalogues in the LOFAR Deep Fields. The existing catalogues in the literature either did not include the deepest available datasets in each survey field, or were created using different methodologies for detecting sources and measuring their fluxes; all this meant the catalogues were not sufficiently robust for the scientific aims of the thesis. To overcome these issues, I generated new catalogues in two of the Deep Fields by combining information from the ultraviolet to the mid-infrared wavelengths to detect sources and extract their properties in a clean and homogeneous manner. These are some of the best-studied regions of the sky and therefore these catalogues are also expected to provide a legacy value beyond the aims of the LOFAR surveys.

Then, in the next chapter, using the multi-wavelength catalogues generated, I identify the host-galaxy counterparts of the radio-detected sources in the LOFAR Deep Fields. Host-galaxy identification and characterisation is crucial, in particular for radio surveys, in determining the photometric redshifts and physical properties (e.g. stellar masses, luminosities, star-formation rates) of the radio-source host galaxies, greatly expanding the scientific scope of the survey. I identified the host-galaxy counterparts of the LOFAR sources using a combination of the statistical Likelihood Ratio method and a visual classification scheme, using a workflow to decide the most appropriate method of identification for each source. This process results in a value-added catalogue of over 80 000 radio sources with multi-wavelength counterparts identified for $> 97\%$ of them. In this chapter, I then also investigate the properties of host galaxies of the faint radio population in the LOFAR Deep Fields.

In the fourth chapter I focus on studying the evolution of the radio-AGN population and their properties in the LOFAR Deep Fields and how feedback from these AGN evolves across cosmic time. Of particular interest, and the focus of this chapter are the LERGs, which dominate at low radio luminosities and are thought to play a key role in the formation of massive galaxies in the local Universe; however, the evolution of this population beyond $z \sim 1$ is not well known. In this chapter, I present the first robust measurement of the LERG luminosity functions out to $z \sim 2.5$ and characterise the evolution of their host galaxy properties. This population shows relatively mild

evolution across the redshifts examined; this is explained by the different evolution of the LERGs hosted in star-forming galaxies and those hosted in quiescent galaxies. The evolution of the quiescent LERGs show a strong decline in their space densities with increasing redshift, in accordance with the available host galaxies, while there is also an increase in the characteristic luminosity. I also find that unlike in the local Universe, the bulk of the LERGs are hosted by star-forming galaxies at higher redshifts and that the AGN in these galaxies appear to be fuelled by a different mechanism, likely associated with the cold gas, as compared to the LERGs in quiescent galaxies.

In the final chapter I present the conclusions and look towards further exploration of this dataset. In particular, I discuss the characterisation of 3% of the radio-sources that were found to be completely invisible at optical and near-infrared wavelengths; an investigation of the far-infrared and radio properties of this subset found that the vast majority of these sources are likely high-redshift star-forming galaxies hosting a radio-AGN. To understand the nature of these extreme sources at early epochs, I present preliminary analysis from recent sub-millimetre follow-up via the sub-millimetre array (SMA) and the James Clerk Maxwell Telescope (JCMT). In addition, I also discuss the spectroscopic follow-up of the radio-AGN population found in the previous chapter with the upcoming multi-object WEAVE spectrograph on the William Herschel Telescope to study the prevalence of AGN activity as a function of different galaxy properties out to high redshifts. Finally, I also describe plans to compare the observational results found in the previous chapters with predictions from the latest cosmological simulations.

Lay Summary

Black holes are some of the most exotic objects in the known Universe. These objects are very compact and incredibly massive (weighing more than the mass of our Sun); such a high density of mass results in black holes having a gravitational pull that is so strong that not even light can escape from them, hence their name. An important finding within the past two decades has been the discovery of *supermassive* black holes, which are an even more extreme form of black holes that weigh up to a million to a billion times the mass of our Sun, and reside in the centre of all galaxies, including our own, the Milky Way. The pioneering work that lead to the most conclusive evidence of the existence of a supermassive black hole at the centre of the Milky Way, was also recently awarded the Nobel Prize in Physics in 2020.

Supermassive black holes go through phases in their lifetime where they consume surrounding material and are referred to as active galactic nuclei (AGN). During this phase, because the material being consumed is moving so fast and gets very hot as it falls into the black hole, AGN can release a vast amount of energy and are known to be some of the most luminous sources in the Universe. It is now widely believed that powerful emission from the AGN during this phase, in the form of winds or very high velocity jets, can have a significant effect on the galaxy in which they reside, such as halting the subsequent formation of stars within the galaxy (this is known as 'AGN feedback'). This is supported by AGN feedback being a crucial ingredient in the current state-of-the-art theoretical models, however, observational evidence about the nature of these AGN, and exactly how they impact the formation and evolution of galaxies, in particular in the early ages of the Universe, remains elusive.

Motivated by these shortcomings, in this thesis, I aim to obtain a better understanding of the role that AGN feedback plays in galaxy evolution over the past 10 billion years of the Universe. These results will also form critical refinements to our current theoretical models of galaxy formation.

Emission from AGN is seen across the electromagnetic spectrum, from X-rays to radio-waves. Astronomers therefore use a wide range of telescopes located both on top of high mountains (e.g. in Hawaii and Chile), and up in space (e.g. the Hubble Space Telescope), to characterise these objects using information from the entire electromagnetic spectrum. In this thesis, I make use of new, highly sensitive observations from the LOFAR telescope, which is a radio telescope located in Europe. The radio wavelength regime is particularly well suited to study AGN as radio waves can pierce through the vast amounts of dust that typically surrounds these objects, and can also efficiently detect the jets from the AGN which emit radiation at radio wavelengths.

In Chapter 2, I introduce the details of the radio and multi-wavelength observations, used throughout the analysis in this thesis. Catalogues based on the existing multi-wavelength data already exist in the literature, but were not sufficiently robust enough for the planned analysis. Therefore, in this Chapter, I have constructed new, higher quality multi-wavelength catalogues that contain information spanning the ultra-violet to the infrared regime, with measurements for over 5 million galaxies, all generated in a consistent manner. The LOFAR observations detect over 80 000 radio sources, but to fully characterise the physical nature of these sources, such as their rate of star formation, their brightness, the distance to them, and much more, it is vital to find their counterparts detected in other wavelengths, such as the visible spectrum. For this reason, the LOFAR observations were taken in the best-studied regions of the sky, where there are previous sensitive observations from other telescopes across the electromagnetic spectrum. In Chapter 3, I detail the process developed to identify the multi-wavelength counterparts (also known as host galaxies) of the radio sources detected by LOFAR. This was performed using the combination of an automated method based on using the statistical properties of the sources, and by visual inspection of the more complex structured sources. This process results in the identification of the host galaxies for over 97% of the LOFAR detected sources. This catalogue has since been used for a wide range of other scientific analyses. Finally, in Chapter 4, I use the datasets generated from the previous two Chapters to present the first study to date on the evolution of faint radio-detected AGN from the present day back to when the Universe was only about 3 billion years old. In Chapter 5, I present the conclusions of the thesis and outline some ongoing and future planned research projects to build upon the work present in the thesis.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Chapters 2 and 3 are based upon work published in Kondapally et al. (2021), and Chapter 4 is based on a paper currently under consortium review, that will be submitted to Monthly Notices of the Royal Astronomical Society in the near future. In all chapters, I use 'we' to acknowledge that this work was performed under the guidance of my supervisor and with the help of collaborators. Nevertheless, all work stems from first-author papers.

(*Rohit Kondapally, July 2021*)

Acknowledgements

This thesis has been possible due to the support and encouragement of many people over the past four years and beyond; you have made undertaking this PhD a thoroughly enjoyable experience.

Foremost, I would like to extend my sincere thanks to my supervisor, Prof. Philip Best, for his guidance, support, and encouragement throughout the PhD. Our weekly meetings over the past four years have been invaluable and I'm grateful for your endless patience and guidance throughout many aspects of the PhD, especially over the past year during the pandemic and in particular over the last few weeks of the PhD. I have been in awe of your seemingly boundless knowledge and passion about astronomy, which has been a constant source of motivation. I could not have asked nor hoped for a better supervisor. I look forward to continuing to work with you on more exciting projects in the future! Thank you so much Philip!

I would like to thank my collaborators here at the Institute for Astronomy and within LOFAR, for their support and discussions over the course of the PhD. I would also like to thank my mathematics tutor from high school, Mrs. Shalini Das: you were instrumental in getting me from a point of loathing and almost failing mathematics to a level of competency and enjoyment for the subject that has continued since.

I have been incredibly lucky to do my PhD at Edinburgh with its friendly and welcoming environment, thanks especially to my fellow PhD students who have made this PhD a pleasurable experience. I would particularly like to mention *Dr.* Rachel Cochrane, Dylan Robson, Nicole Thomas, Sarah Appleby, and Sophie Dubber – you have all been great, supportive friends and I have thoroughly enjoyed spending my time at Edinburgh with you, especially while finishing up my PhD over the past year under the extraordinary circumstances. To Hayley, I cannot thank you enough for your support throughout the PhD, and for indulging my interest in astronomy by listening to my barely comprehensible ramblings about any exciting plots or results. You have made this task of doing a PhD a whole lot easier and manageable.

To my parents Kavitha and Ramana, and my brother *Dr.* Harshit Kondapally, thank you for always supporting me and encouraging my interest in astronomy since childhood. From buying me telescopes to funding trips to educational outreach events and more; these have all helped me in growing my interest of all things space and I would not be here, finishing my PhD, without you.

I would also like to acknowledge support from the Science and Technology Facilities Council (STFC) through an STFC studentship via grant ST/R504737/1.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cosmological background

The current theory for structure formation in our Universe (the standard model) is the Λ Cold Dark Matter (ΛCDM) model. The predictions of this cosmological model are very tightly constrained by the observations of the temperature fluctuations in the cosmic microwave background (CMB) made by Wilkinson Microwave Anisotropy Probe (WMAP; Hinshaw et al. 2013) and Planck (Planck Collaboration et al., 2016, 2020) and complementary information from distant supernovae (Riess et al., 1998; Perlmutter et al., 1999). These observations tell us that ~69% of the energy content (density) of the Universe is composed of a cosmological constant (Λ or vacuum energy or dark energy) term that is responsible for the current accelerated expansion of the Universe (Riess et al., 1998). The remaining energy density is made up of matter that is usually split into baryonic matter (~3%) and dark matter (~28%). Baryonic matter makes up all of the matter we can see such as gas, planets, and stars. The nature of dark matter (DM) remains mysterious as it does not interact with the electromagnetic force, remaining 'invisible' to the electromagnetic spectrum, but plays a major role in the growth of structure as it does interact via the gravitational force.

In our ΛCDM universe, the growth of structure occurs via a hierarchical (bottom up) merging process where smaller DM haloes merge due to gravitational attraction to form larger, more massive DM haloes (Peebles & Yu, 1970; White & Frenk, 1991). The baryons follow the potential well of the DM haloes resulting in large scale filament-void structures, known as the 'cosmic web'. In addition to the CMB observations, the ΛCDM

model for structure formation has enjoyed other considerable observational success. For example, observations from galaxy surveys (eg. SDSS: York et al. 2000a; 2dFGRS: Colless et al. 2001) match well with numerical simulations of structure formation (Springel et al., 2005, 2006), along with the detection of the predicted gravitational lensing of the CMB (Smith et al., 2007; Das et al., 2011). In addition, simulations of the intergalactic medium (IGM) have also reproduced observations of the Ly-α forest formed from the absorption of quasar light along the line of sight (Cen et al., 1994).

Cosmological simulations studying structure formation are often run on scales of 10's - 100's of mega-parsec (Mpc). However, poorly understood processes such as star formation and feedback act on sub-parsec scales, below the resolution limit of most simulations, and are parametrised in hydrodynamical simulations as 'subgrid' physics (Schaye et al., 2015). Therefore, although recent simulations have benefited from the increase in computational power and higher resolution, the implementation of the subgrid physics remains uncertain. For example, Scannapieco et al. (2012) found that 13 different hydrodynamical codes that simulated the formation of a galaxy with the same initial conditions but a different prescription of treating subgrid physics produced a large variation in galaxy size, stellar mass, morphology, and gas content.

One such poorly understood feedback process that forms the focus of this research is the effect that accreting central super-massive black holes (SMBHs; also known as active galactic nuclei; AGN hereafter) have on their host galaxies. An important discovery over the past two decades has been the correlation between the mass of the central SMBH and the velocity dispersion of the stars in the central bulge of the galaxy (known as the M_{BH}-σ relation; Ferrarese & Merritt 2000; Gebhardt et al. 2000; see also Sect. 1.2.1), suggesting that the growth of the SMBH and the growth of the surrounding galaxy are fundamentally connected (see reviews by Fabian, 2012; Kormendy & Ho, 2013; Heckman & Best, 2014; Tadhunter, 2016; Hardcastle & Croston, 2020). One might naively not expect the SMBH that typically occupies < 1 parsec at the centre of the galaxy to affect the galaxy on scales of tens of kilo-parsecs; however, the effect that SMBHs could have on their host galaxy can be understood by considering the energetics of black hole accretion. Under theoretical arguments, assuming that the black hole is accreting matter with an accretion efficiency of 10%, the energy released as the black hole grows is given by $E_{\rm BH} \approx 0.1 M_{\rm BH} c^2$. The binding energy of the bulge component of a galaxy with velocity dispersion σ of stars and mass M_{Bulge} is given as $E_{\rm Bulge} \approx M_{\rm Bulge}\sigma^2$, where massive galaxies typically have $\sigma \sim 350\,{\rm km\,s^{-1}}$. The black hole mass correlates with the mass of the bulge as $M_{\rm BH} \approx 1.4 \times 10^{-3} M_{\rm Bulge}$ (Merritt & Ferrarese, 2001). Hence, the ratio of the energy released by the black hole as it grows,

to the binding energy of the bulge is of the order ~ 100. Therefore, if only a small fraction of the energy from the AGN is able to couple with the surrounding gas, this can have a significant impact on the evolution of the host galaxy; this has been a key driver for the increased momentum in AGN research in recent years.

Over the past two decades, it has been shown that AGN do indeed have a significant effect on their host galaxies (known as 'AGN feedback') by, for example, suppressing star formation due to their energy input, and by regulating galaxy growth. Perhaps the most striking example of the impact of the AGN on its environment comes from observations of bubbles or cavities in massive clusters, created as gas is swept away by the powerful radio jets from the AGN (see Sect. 1.4.2). The energy input from these jets can prevent gas from cooling and keep the galaxy from growing much further.

The growth of the central SMBH and the star-formation in the galaxy were both much more enhanced at early times than today. It is therefore useful to study how the properties of the AGN and its host galaxies evolve through cosmic time and the role of the AGN on the evolution of galaxies from near this peak of SMBH growth and star-formation activity; this can provide crucial tests of galaxy formation and evolution models.

1.2 Motivation and evidence for AGN feedback

The idea that the formation and evolution of the central SMBH and the host galaxy in which they reside are connected is relatively new. The existence of quasars or AGN (Schmidt, 1963) has been known for quite some time, and theories were developed to explain the powering mechanism behind these sources as accretion onto a SMBH (Hoyle & Fowler, 1963; Lynden-Bell, 1969; Lynden-Bell & Rees, 1971). However, until a few decades ago, even dynamical evidence that SMBHs with masses $\mathrm{M}_{BH} \sim 10^6 - 10^9 M_\odot$ exist was lacking, much less the link between the SMBH and the host galaxy. In the next four subsections, I will describe some of the main motivations for the connected evolution of AGN and their host galaxies.

1.2.1 The $\mathrm{M}_{BH} - \sigma$ Relationship

It is now a widely accepted notion that SMBHs affect the growth of their host galaxy. This notion came from the spatially resolved dynamical detections of BHs in galactic

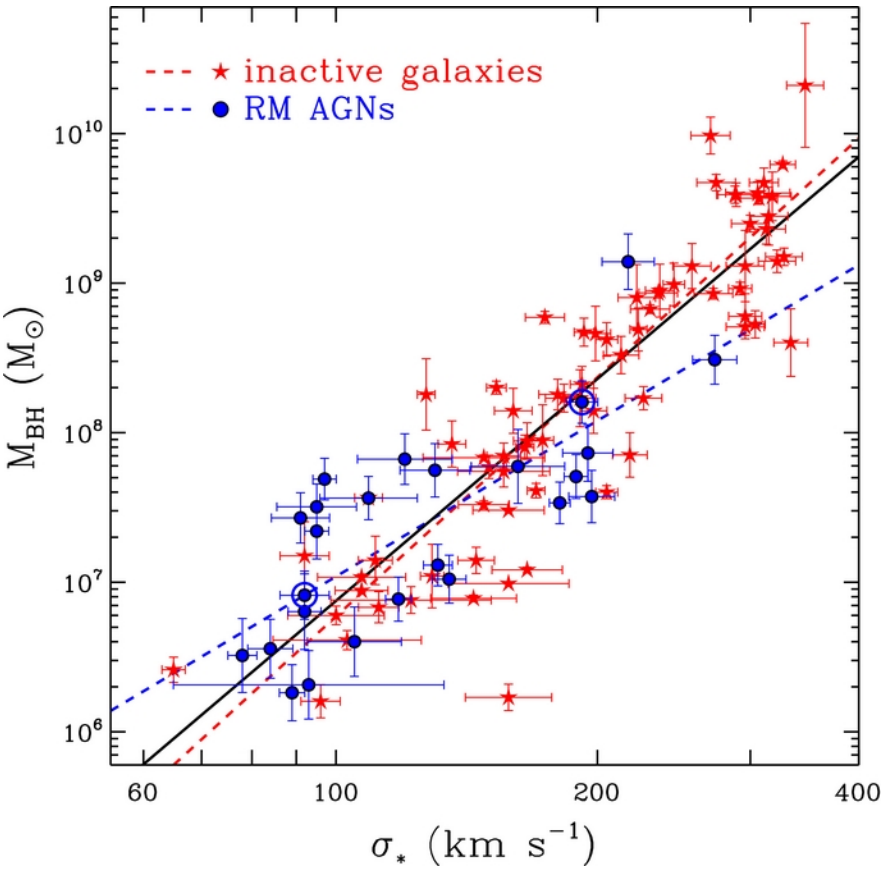

Figure 1.1 *The M_{BH}-σ relationship for galaxies in the local Universe with direct BH mass estimates. The red dashed line shows the relationship for the inactive (quiescent) galaxies (red stars) and the blue dashed line shows the same relationship for AGN (blue dots). When selection effects and the difference in M_{BH} measurements are accounted for, both the AGN and the quiescent galaxies are consistent with following the same relation (black line). Plot taken from Woo et al. (2013).*

centres, thanks to observations by the *Hubble Space Telescope* (HST). A tight correlation between the mass of the black hole and the velocity dispersion of the bulge σ was found by Ferrarese & Merritt (2000) and Gebhardt et al. (2000). A plot of this relationship from Woo et al. (2013) is shown in Figure 1.1 with updated measurements. The red stars and red dashed line show the observations and the best fit line for inactive (quiescent) galaxies. The same relationship is also shown in blue dots and dashed line, but for galaxies with an AGN. The relationship appears slightly steeper for the quiescent galaxies than for the AGN. However, when Woo et al. (2013) account for the uncertainties in the different methods of estimating M_{BH} and sample selection effects for the two classes of sources, they find that both AGN and quiescent galaxies follow the same relationship (black line). The existence of this correlation implies that the growth of the black hole and the build-up of mass in the bulge must be connected.

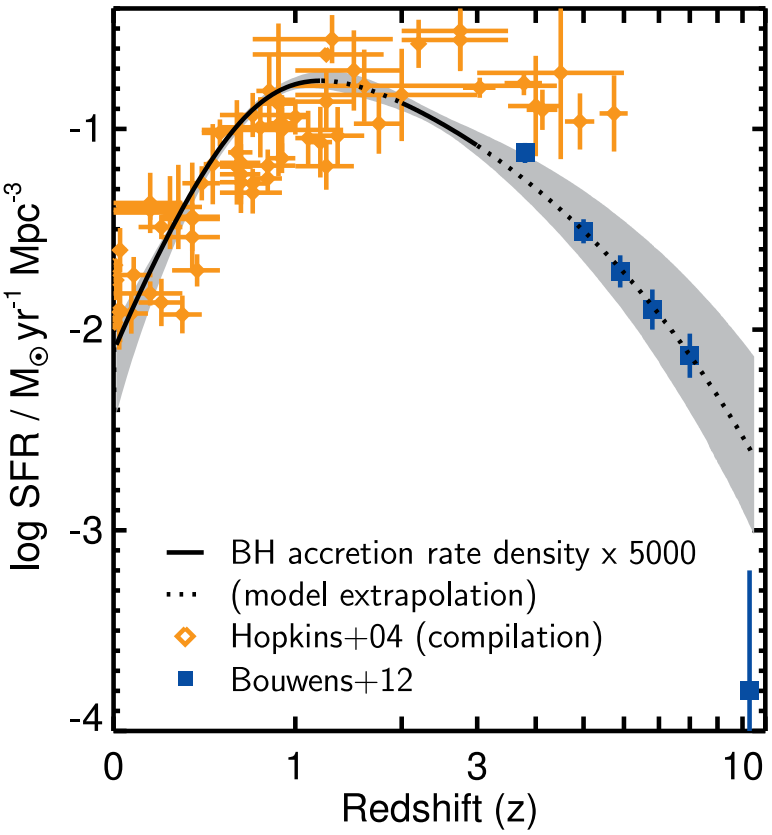

Figure 1.2 *Evolution of the volume-density of star-formation rate over cosmic time (based on data complied by Hopkins, 2004; Bouwens et al., 2012). Also shown as a black line is the volume density of the black hole (BH) accretion rate (based on data from Aird et al., 2010), scaled up by a factor of 5000, to aid the eye, over cosmic time which closely follows the SFR history. This figure was taken from Kormendy & Ho (2013), who adapted and updated the plot from Aird et al. (2010, their fig. 13).*

1.2.2 Cosmic star-formation rate and black hole growth

Another realisation of the co-evolution of SMBHs and their galaxies came from studies of how the black hole and the stars in the galaxy grow over cosmic time. Shown in Figure 1.2 is the volume-averaged star formation rate (yellow and blue points; Hopkins 2004; Bouwens et al. 2012) as a function of redshift. Also shown is the volume-density of the black hole accretion rate, scaled up by a factor of 5000 (black line; Aird et al., 2010). As evident from the plot, the star-formation rate (SFR) history closely tracks the black hole growth, increasing from today and peaking at $z \sim 2$ and then declining steadily at higher redshifts. This suggests that the growth of the black hole and the SFR have been connected across cosmic time.

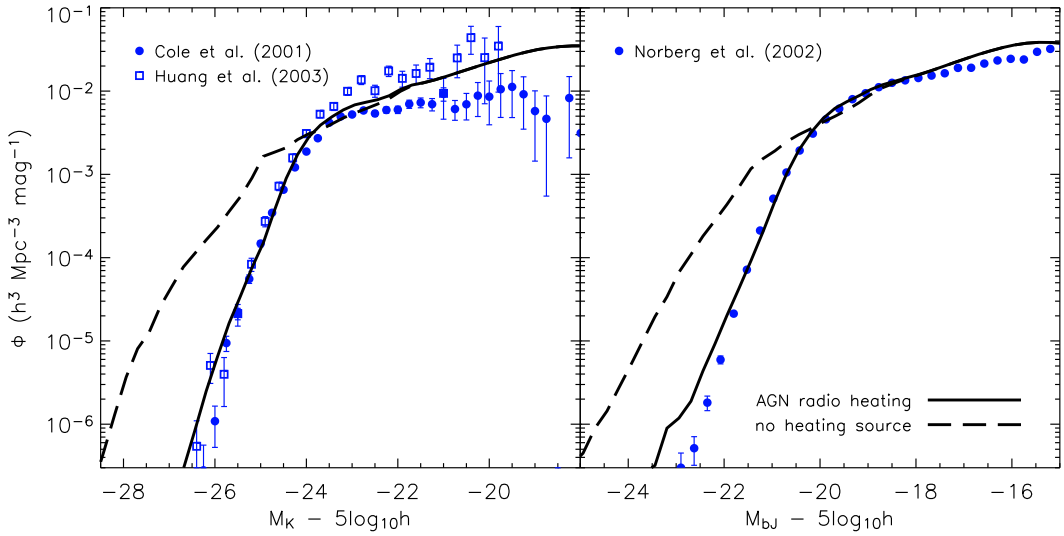

Figure 1.3 *Local galaxy luminosity function in the K-band (left) and J-band (right). The data points in each panel show observational estimates of the luminosity function, whereas the solid and dashed lines show the galaxy luminosity functions with and without heating from radio-mode AGN, respectively, as predicted by the semi-analytical model of Croton et al. (2006). This shows that feedback from radio-mode AGN is particularly important in reproducing the sharp decline at the high-luminosity end of the luminosity function. Figure taken from Croton et al. (2006, from their fig. 8).*

1.2.3 Galaxy Luminosity Function

A galaxy luminosity function describes the number of galaxies per unit comoving volume as a function of luminosity. Comparing the observed galaxy luminosity function with that from predictions based on ΛCDM can provide valuable tests of galaxy formation models.

Semi-analytical models have often been used to make detailed predictions of observed galaxy properties. Such models transform the dark matter halo mass function into galaxy luminosity functions based on an underlying cosmology (i.e. the ΛCDM paradigm), where the growth of dark matter halo structures (from initial instabilities hierarchically) and their abundances are combined with a prescription of the cooling of gas in the haloes, the formation of stars, and any feedback processes that affect the cooling of gas, and hence the star-formation efficiency (e.g. see reviews by Baugh, 2006; Benson, 2010). It is found that under a model where no feedback processes are employed, there is a mis-match between the observed and predicted number density of high luminosity (mass) galaxies: the observed galaxy luminosity function shows an exponential cut-off

at high luminosities, which is far steeper than that predicted based on a model with no heating from AGN. This can be visualised in Fig. 1.3, taken from Croton et al. (2006), with the various blue points showing observations of the local galaxy luminosity function in the K-band (*left panel*) and J-band (*right panel*). The dashed line shows the results predicted by a model in which there is no feedback from the AGN, whereas the solid line shows the expected luminosity function under a model where there is heating from the energetic output of AGN, and in particular from the so-called 'radio-mode AGN' (see Sect. 1.3). This extra source of heating prevents the galaxies from growing further in the most massive haloes, and is thus required to match the sharp drop seen in the luminosity function at high luminosities (or equivalently, high masses).

1.2.4 The cooling-flow problem

Galaxy clusters are one of the largest gravitationally bound structures in the Universe and as such form an important testbed for galaxy formation and evolution models.

There are large reservoirs of hot gas ($T \sim 10^7 - 10^8$ K) present within the intra-cluster medium of clusters, in particular surrounding the centres of galaxy clusters. This hot gas emits in the X-rays via Bremsstrahlung radiation, thereby resulting in gas cooling. Under the assumptions of an isothermal gas in hydrostatic equilibrium, we can determine the 'cooling time' of the gas $t_{\rm cool}$, which is an estimate of the time that the gas would take to radiate all of its thermal energy away, at its current radiation rate. This scales as

$$t_{\rm cool} \propto \frac{T^{1/2}}{n}, \tag{1.1}$$

where T is the temperature of the gas and n is the gas density. In the core of clusters where the gas density is the highest, the gas cooling time is expected to be relatively short compared to the Hubble time. As the gas in the centre of a cluster cools, the loss of thermal pressure in the centre causes matter to flow inwards to maintain hydrostatic equilibrium. This raises the central density, decreasing the cooling time further, and resulting in a so-called 'cooling flow' with a centrally peaked gas profile; simple calculations predict that, in the absence of other processes, this would result in star-formation rates of $> 1000\,{\rm M}_\odot\,{\rm yr}^{-1}$ (see review by Fabian, 1994). However, X-ray observations of galaxy clusters show a lack of such strongly peaked central surface brightness profiles, and that the amount of gas that cools is around an order of magnitude lower than predictions (e.g. Peterson et al., 2001; Peterson & Fabian, 2006; McNamara

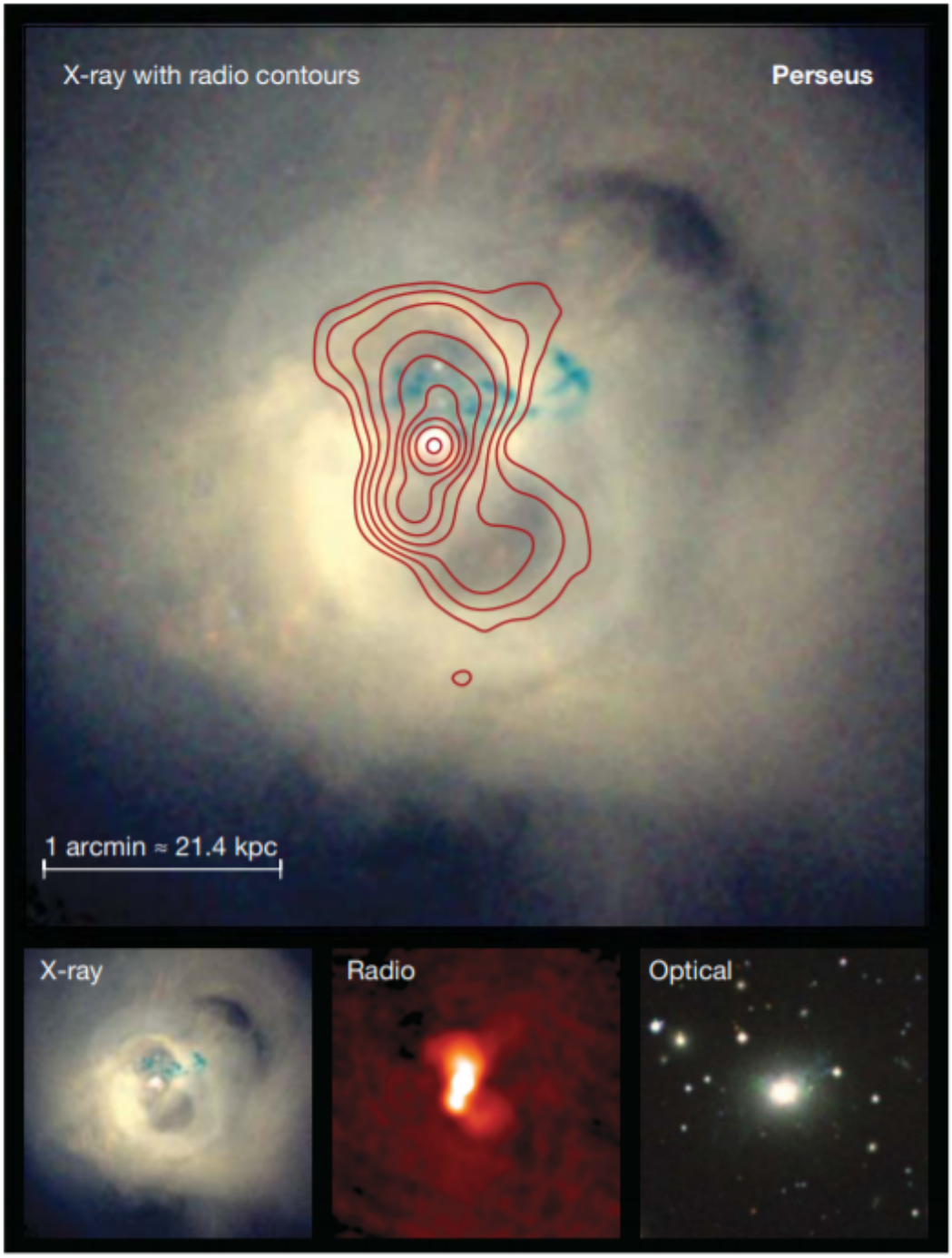

Figure 1.4 *Multi-wavelength cutouts of the Perseus cluster. The main panel shows the X-ray image of the Perseus cluster, with the radio emission overlaid as red contours. The radio emission traces the jets launched by the central AGN, and as these jets propagate through the surrounding gas, they evacuate 'cavities' that can be seen in the X-ray emission (as 'holes' or dimmer regions in the X-ray image). Figure taken from Cattaneo et al. (2009) and based on images and data from Boehringer et al. (1993); Jones et al. (2004); Fabian et al. (2006).*

& Nulsen, 2007). This discrepancy between the observations and predictions was termed the 'cooling flow problem'.

It is well known that the galaxies at the centres of 'cool-core' clusters very frequently host powerful AGN (e.g. Best et al., 2007) that are detected in the radio and display powerful collimated jets. The most striking examples of the feedback from these AGN, also known as 'radio-mode' AGN feedback (see Sect. 1.4.2 for more details) on the cluster gas medium comes from multi-wavelength observations of clusters, for example, Perseus, as shown in Fig. 1.4, where it can be seen that the radio jets have inflated large cavities in the hot X-ray gas. Sect. 1.4.2 discusses how these cavities can be used to estimate the heating rate of the AGN into the surrounding gas; this has been shown to be sufficient to balance the radiative cooling losses in the X-rays (see reviews by McNamara & Nulsen, 2007; Cattaneo et al., 2009; Heckman & Best, 2014), thereby resolving the cooling flow problem.

Another closely related issue pertains to the observation that the most massive galaxies in recent epochs tend to be passive with red colours, indicative of having finished their bulk of star-formation at early times compared to lower mass galaxies that are still currently forming stars. This is known as 'downsizing' (e.g. Cowie et al., 1996) and at first seemed to be in apparent contradiction to the hierarchical growth of structure.

One proposed solution to reconcile observations and models is that in the most massive galaxies, which are preferentially located in galaxy groups and clusters, 'radio-mode' AGN feedback keeps the gas in the halo hot, preventing the cooling of gas and the onset of star formation at late times. Moreover, the semi-analytical model of Croton et al. (2006), which was tuned such that the feedback from the 'radio-mode' AGN was scaled to balance the radiative cooling losses of the hot gas, also results in an agreement in the galaxy colour distributions, with the most massive galaxies being red and dead. A more realistic model of the heating and cooling cycle from the 'radio-mode' AGN was implemented for the more recent semi-analytical galaxy evolution (SAGE) model (Croton et al., 2016), which also reproduces many observed properties of the local galaxy population.

1.3 The two classes of AGN

Historically, AGN have been classified into many types based on the different observed properties, for example, blazars (AGN where jets are oriented close to our line-of-

sight), quasi-stellar objects (QSOs; luminous objects that appeared as point sources), Seyferts, etc. More recently, it has been noted that there are fundamentally two different types of AGN (Hardcastle et al., 2007; Best & Heckman, 2012; Heckman & Best, 2014; Hardcastle, 2018) defined based on the accretion efficiency of the black hole: radiative-mode AGN (radiatively efficient, also known as quasar-mode or cold-mode) and jet-mode AGN (radiatively inefficient, also known as radio-mode). The properties of these two classes of AGN are discussed in the sections below and a schematic drawing of these AGN is shown in Figure 1.5 taken from Heckman & Best (2014). The different modes of AGN are thought to be associated with the nature of accretion flow, fuelling mechanism and also the likely fuel sources; these are discussed in detail in Sect. 1.3.4.

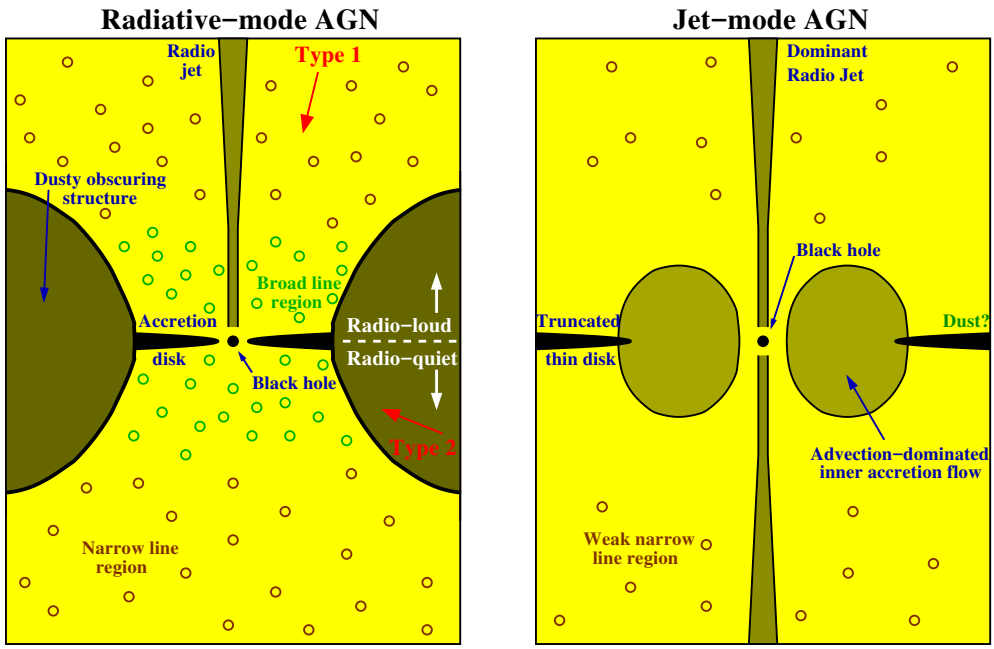

Figure 1.5 *Schematic diagrams of ~ parsecs around the central SMBH for typical radiative-mode (left) and jet-mode (right) AGN. Figure from Heckman & Best (2014).*

1.3.1 Radiative-Mode AGN

Radiative-mode AGN release most of their energy output in the form of radiation, resulting from the efficient accretion of matter onto the SMBH. As shown in Figure 1.5,

the central SMBH is surrounded by a geometrically thin, optically thick accretion disc from which matter accretes onto the SMBH. The formation of this accretion disc can be understood by considering the nature of the accretion flow. If the thermal energy, resulting from the conversion of potential energy from accretion of matter, can be radiated away efficiently, the orbiting gas becomes much cooler than the local virial temperature, and forms a thin-disc structure (Shakura & Sunyaev, 1973). The accretion disc radiates thermally, over a range of temperatures, with the inner regions of the disc being hotter than the outer regions and emitting predominantly in the ultraviolet (known as the 'Big Blue Bump' in the SED of quasars). Some photons are inverse-Compton scattered by the surrounding hot corona to high X-ray frequencies. This UV and X-ray radiation from the accretion disc ionises high density gas close to the black hole (see Figure 1.5, *left*) resulting in broad emission lines due to high velocity motions of the gas clouds (known as the broad-line region, hereafter BLR). At larger scales (few hundred to thousand parsecs), there exists a population of lower density gas clouds which can be ionised by radiation travelling along the polar axis of the SMBH; this also produces emission lines but with smaller widths (due to relatively lower velocity gas clouds); therefore, this is referred to as the narrow line region (NLR).

Surrounding the accretion disc is a dusty obscuring structure (assumed to take the shape of a torus) which absorbs the radiation from the accretion disc and hot corona and re-radiates them in the infrared part of the spectrum. This model has been used to explain that some of the 'different' types of AGN observed are simply due to orientation effects about the line of sight. For example, an observer viewing the AGN close along the polar axis of the AGN would probe the accretion disc, NLR, and the BLR; these sources are hence classed as Type 1 AGN. If, however, an observer viewed from a line of sight near the plane of the accretion disc, their observations would not directly observe the accretion disc, nor contain broad emission lines from the BLR as they are blocked by the dusty torus; these sources are hence classed as Type 2 AGN. This was proposed as the unified model of (radiative-mode) AGN (Antonucci, 1993).

As shown in Figure 1.5, there exist populations of both radio-loud and radio-quiet radiative-mode AGN; in the former, this traces the radio emission coming from jets (quasars that show radio jets). These are a small subset of the radiative-mode population, and can play an important role when considering the evolution of the radio-loud AGN population (see Sect. 1.5). The photoionisation of gas resulting from photons from the accretion disc that is present in these AGN results in high-excitation emission lines in their spectra. As such, historically, the subset of the radio-loud radiative-mode AGN have also been referred to as high-excitation radio galaxies (HERGs). The presence of

such emission lines in the spectra of these objects also allows for efficient identification of these sources using emission line diagnostics (e.g. Best & Heckman, 2012; Heckman & Best, 2014; Best et al., 2014; Pracy et al., 2016).

1.3.2 Jet-Mode AGN

The second category of AGN (Figure 1.5, *right*), undergoing radiatively inefficient accretion, do not show evidence of a geometrically-thin accretion disc or a dusty torus (Chiaberge et al., 2002) expected from a typical AGN, but do display powerful bi-polar jets that emit synchrotron radiation from the acceleration of charged particles to relativistic velocities than can be traced in the radio. Therefore, these sources are only identified as AGN in the radio regime. The lack of an optically-thick accretion disc means that these sources do not display high-excitation lines in their spectra, and as such, are commonly referred to as low-excitation radio galaxies (LERGs). These observational features can be explained by considering the nature of the accretion flow. In these AGN, most of the thermal energy liberated by the accretion process is carried in by the accreting gas (i.e. energy is advected), instead of being radiated away. This is known as an advection-dominated flow (Narayan & Yi, 1994, 1995) and leads to heating of the accreting gas, causing it to expand vertically about the plane of accretion and form a geometrically thick disc. Most of the energetic output of these AGN is in the form of kinetic (mechanical) energy from the two-sided jets, with the radio emission produced by the synchrotron radiation from charged particles accelerated to relativistic velocities. As such, these AGN are also sometimes referred to as 'radio-mode' AGN.

1.3.3 Host galaxies of AGN

It is now useful to consider how the SMBHs at the centres of galaxies hosting these two modes of AGN are fuelled, since the fundamental difference between the two modes appears to be the nature of the accretion flow (and therefore the accretion efficiency). It is also important to consider the properties of the host galaxies and how these influence the incidence of the different modes of AGN activity across the galaxy population.

Detailed characterisation of the host galaxy properties of these two AGN classes have revealed that the jet-mode AGN and radiative-mode AGN show a dichotomy in their host galaxy properties, that is thought to be driven by how the two classes of AGN are triggered (e.g. see reviews by Heckman & Best, 2014; Hardcastle & Croston, 2020).

One of the properties of the host galaxies of radiative-mode AGN, the AGN accreting matter efficiently, is the supply of dense cold gas. Heckman et al. (2004) studied host galaxy properties of 23,000 AGN selected using narrow emission lines. They found that the typical host galaxy stellar masses lie in the range of $M_* \sim 10^{10}$ - 10^{11} $M_\odot$, and that these AGN are typically hosted by moderate mass black holes ($M_{BH} \sim 10^{6.5}$ - 10^8 $M_\odot$). The concentration index C, which determines the fraction of light that originates from the central region of a galaxy, is low (peaking around $C \sim 2.2$) for these hosts, which is typical of disk galaxies (Conselice, 2014).

Best & Heckman (2012) studied the properties for a sample of 18 286 local radio-loud AGN by combining Sloan Digital Sky Survey (SDSS) data with radio observations over large areas from the NRAO VLA Sky Survey (NVSS) and the Faint Images of the Radio Sky at Twenty centimetres (FIRST) survey. They classified their sample into HERGs (i.e. the radio-loud subset of the radiative-mode AGN) and LERGs (i.e. the jet-mode AGN) using emission line diagnostics and find a dichotomy in the host galaxy properties; they found that LERGs tend to be typically hosted in massive red galaxies, with an old stellar population (as indicated by a higher 4000 Å break), and massive black holes whereas the HERGs tend to be hosted in less massive, bluer galaxies with more recent star-formation activity; these findings are largely consistent with other studies in the literature (e.g. Tasse et al., 2008; Smolčić, 2009; Janssen et al., 2012; Mingo et al., 2014; Ching et al., 2017; Whittam et al., 2018; Williams et al., 2018). Moreover, the LERGs are also more often found in rich group or cluster environments (e.g. Best, 2004; Tasse et al., 2008; Gendre et al., 2013; Sabater et al., 2013; Croston et al., 2019).

1.3.4 Fuelling mechanisms of AGN

The above results and the properties of the host galaxies discussed in Sect. 1.3.3 are consistent with the idea that the AGN activity in the two classes of AGN are triggered by different mechanisms related to the available fuel source; jet-mode AGN are thought to be primarily fuelled by the cooling hot gas from haloes present in their massive host galaxies, whereas the radiative-mode AGN tend to accrete efficiently, typically from cold gas, from either external (e.g. mergers or interactions) or internal (e.g. bars or instabilities) processes (see Sect. 1.3.4.1), with the more plentiful gas supply leading to the formation of an accretion disc (e.g. Allen et al., 2006; Hardcastle et al., 2007; Heckman & Best, 2014; Hardcastle, 2018; Hardcastle & Croston, 2020). In addition, it is also possible that the strength of the jet outflows emitted can be enhanced by the

spin of the black hole (McNamara et al., 2011) although debate on this still remains in literature.

1.3.4.1 Fuelling of radiative-mode AGN

One of the key ingredients in achieving the high accretion rates observed for radiative-mode AGN is the availability of a plentiful amount of cold gas, that can be brought into the galaxy via external or secular processes. However, there is still a lot of debate in the literature about the relative importance of these two processes; I discuss the evidence, both in support and against, of the importance of mergers and secular processes triggering radiative-mode AGN acticity.

Mergers of galaxies are commonly used in simulations to trigger powerful AGN activity that have significant implications for feedback (e.g., Croton et al. 2006; Hopkins et al. 2008). The tidal forces involved during a merging event can lead to cold gas inflows towards the centres of the galaxies (Barnes & Hernquist, 1996). The gas infall can lead to high densities of gas which can trigger black hole growth and activity and a burst of star formation (e.g. Di Matteo et al., 2005; Aird et al., 2019). Observationally, galaxies that have undergone a recent merger event show asymmetric (lopsided) structure and a burst of a recent star-formation episode. If mergers were solely responsible for the fuelling or the growth of the black hole, we may also expect a correlation between AGN luminosity and galaxy asymmetry. Reichard et al. (2009) found that while galaxies that host an AGN are usually more asymmetric than quiescent galaxies, only a small number of high luminosity AGN (accreting efficiently) have such asymmetric host galaxy structure. In addition, they also find that while the AGN luminosity correlates well with the age of the stellar population (older stars indicating a lack of recent star formation), with higher AGN activity occurring for younger stellar ages, the asymmetries can span a wide range of values.

However, recently, Sharma et al. (2021) studied how AGN activity correlates with mergers using a sample from the CANDELS survey across $0.5 < z < 3$. They found that AGN hosts do not show higher close-pair fractions than a mass-matched control sample of inactive galaxies. Furthermore, even the most luminous AGN were found to be no more likely than their inactive counterparts to be found in merging systems. They also generated synthetic *Hubble Space Telescope (HST)* observations similar to their CANDELS data, finding reasonable agreement with their observational results. These findings suggest that while mergers, in particular those that are gas-rich, may trigger AGN activity, other mechanisms likely play a significant role in fuelling these AGN.

These are similar to findings by other studies (Silva et al., 2021; Lambrides et al., 2021), indicating that all that is required for black hole growth is a supply of cold dense gas (traced by ongoing star formation) to the centre of the galaxy. The picture is less clear at higher luminosities; studies focusing on high-luminosity AGN at high redshift, selected based on IR diagnostics only, find that these AGN are more likely to be hosted in galaxies showing asymmetric or disturbed morphologies (by around 50–70%) and also to be hosted by galaxies that are more likely to be classified as interacting or merging systems compared to their lower-luminosity X-ray selected AGN (Donley et al., 2018).

Secular processes, for example, galaxies with bars, distortions or spiral arms, that form as a result of non axis-symmetric rotation, can drive gas inflow to the central regions of the galaxy and create a 'pseudo-bulge'. If secular processes are solely responsible for fuelling these AGN, then we would expect these AGN to be preferentially located in galaxies with bars or bulges. However, debate on this still remains in literature, where while some studies find no such link between the prevalence or power of AGN and the existence of bars or bulges (e.g., Lee et al. 2012; Cisternas et al. 2013; Goulding et al. 2017), others find that the presence of bars does induce AGN activity (e.g. Galloway et al., 2015; Kim & Choi, 2020).

1.3.4.2 Fuelling of jet-mode AGN

Best et al. (2005a) found that the prevalence of jet-mode AGN activity depends strongly on stellar mass ($\propto M_{\star}^{2.5}$), increasing from almost zero to more than 30% between stellar masses of 10^{10} $M_{\odot}$ to 5×10^{11} $M_{\odot}$, and also on black hole mass ($\propto M_{\bullet}^{1.6}$) in the local Universe (Best et al., 2005a; Janssen et al., 2012)). Recently, Sabater et al. (2019), using deeper, more sensitive radio data at 150 MHz, studied the dependence of radio-AGN activity on stellar mass and luminosity, and found that the most massive galaxies (i.e. $M_{\star} > 10^{11}$ $M_{\odot}$; dominated by the jet-mode AGN population) are always 'switched on', with radio luminosities of $L_{150\,\mathrm{MHz}} \geq 10^{21}$ W Hz^{-1}. These massive systems have large reservoirs of hot gas, and Best et al. (2005a) find that the black hole mass dependence on the rate at which gas cools in these haloes is very similar to the black hole mass dependence on the AGN activity; this suggests that the jet-mode AGN activity is related to accretion at low rates from the cooling of hot gas within haloes. This provides further evidence of the model discussed above, and suggests that the accretion mode (i.e.radiatively inefficient) and the fuel source (i.e. hot gas in haloes) may be closely linked. This also has interesting implications for feedback from these AGN (see Sect. 1.4.2).

Often, and in simulations in particular, the accretion of hot gas onto the SMBH is treated as being governed by Bondi accretion, with the accretion rate given as

$$\dot{M}_{\rm Bondi} = \epsilon_{\rm m} \frac{4\pi G^2 M_{\rm BH}^2 \rho}{c_s^3}, \tag{1.2}$$

where $M_{\rm BH}$ is the mass of the black hole, ρ is the density of the gas cloud, and c_s is the sound speed (Bondi, 1952). This relation is based on the assumption of spherically symmetry, lack of magnetic fields, and a uniform gas density, which are not necessarily held in reality. In recent years, improved models for the nature of the accretion flow from the hot phase suggest that under certain conditions, thermal instabilities in the hot gas medium can result in the formation of cold filaments of gas that condense out of the plasma and then 'rain' down onto the black hole (e.g. Sharma et al., 2012; Gaspari et al., 2013, 2015, 2017). This means that accretion onto the SMBHs in jet-mode AGN may also be from gas in a 'cold' state, and can lead to stochastic accretion rates that are around two orders of magnitude larger than the Bondi rate. Indeed there is now growing observational support for this scenario, with observations of a handful of jet-mode AGN being fuelled by cold gas (e.g. Tremblay et al., 2016; Maccagni et al., 2018; Ruffa et al., 2019b). This is supported by observations of large reservoirs of cold molecular gas ($\sim 10^8 - 10^{10}\,{\rm M}_\odot$) that have been found in the host galaxies of jet-mode AGN (e.g. Okuda et al., 2005; Ocaña Flaquer et al., 2010; Ruffa et al., 2019a,b). Interestingly, although interactions and merger events are expected to play a role in driving radiative-mode AGN activity by bringing cold gas into the galaxy (see discussion in Sect. 1.3.4.1), there is also evidence for enhanced jet-mode AGN activity in interacting galaxies, with higher merger fractions observed in the host galaxies of jet-mode AGN than in control samples (e.g. Sabater et al., 2013, 2015; Gordon et al., 2019; Gao et al., 2020).

1.3.4.3 Eddington-scaled accretion

Rather than solely the gas origin, it has been suggested that the accretion rate (in particular the Eddington-scaled accretion rate) onto the black hole may determine the formation of a radiatively efficient/inefficient AGN (e.g. Best & Heckman, 2012; Best et al., 2014; Hardcastle, 2018; Hardcastle & Croston, 2020). The Eddington limit defines the maximum luminosity that a black hole can attain while growing by accretion before the radiation pressure overcomes the gravitational force. This depends only on

the black hole mass $M_{\rm BH}$ and is given by

$$L_{\rm Edd} = \frac{4\pi G m_{\rm p} c}{\sigma_{\rm T}} M_{\rm BH} \tag{1.3}$$

where $m_{\rm p}$ is the mass of a proton, $\sigma_{\rm T}$ is the Thomson scattering cross-section and $L_{\rm Edd}$ is the maximum luminosity that a black hole of mass $M_{\rm BH}$ can achieve. The luminosity radiated by a black hole simply depends on the accretion efficiency ϵ and the accretion rate $\dot{M}$, such that $L = \epsilon \dot{M} c^2$.

In this picture, the different modes of AGN activity are governed by a switch in the Eddington-scaled accretion rates such that AGN that are accreting matter at high fractions of the Eddington rate $L/L_{Edd} \gtrsim 0.01$ are able to efficiently convert the potential energy to energy that can be radiated, and are classed as radiative-mode AGN. In contrast, the jet-mode AGN are accreting matter at low fractions of the Eddington rate $L/L_{Edd} \lesssim 0.01$, where the bulk of the energetic output is in the form of mechanical energy from collimated jets. Evidence in support of this comes from the analysis by Best & Heckman (2012) who find that in addition to a dichotomy in the host galaxy properties, there is also a dichotomy in the Eddington scaled accretion rates, with HERGs typically having accretion rates above $\sim 1\%$ of the Eddington rate, whereas LERGs have accretion rates below 1% of the Eddington rate. Further evidence comes from observations that (radio-quiet) quasars (which are radiative-mode AGN) are seen to show a range of Eddington-scaled accretion rates of $0.01 \lesssim L/L_{Edd} \gtrsim 1$ (e.g. Kollmeier et al., 2006; Trump et al., 2009), but are not found below the ~1% limit.

This model where a switch in the Eddington-scaled accretion rate determines the formation of a radiatively efficient and radiatively inefficient AGN can also explain the observed properties of these two modes. Fuelling by cooling of hot gas generally tends to occur at low accretion rates in massive galaxies which host massive black holes, leading to a jet-mode AGN, whereas low-mass galaxies with abundant cold gas supply (and hence ongoing star-formation) are likely to result in higher accretion rates, and hence in the formation of a radiative-mode AGN. However, due to the stochastic nature of accretion processes, accretion from cold gas can occur at low Eddington rates leading to a jet-mode AGN.

Whittam et al. (2018) found a dichotomy in the Eddington-scaled accretion rate distribution out to $z \sim 0.7$, but with considerable overlap in these distributions compared to that found by Best & Heckman (2012) in the local Universe. Therefore, the origin of the differences in these two populations and the precise physical mechanisms that trigger the different modes of AGN in different galaxies are not well understood.

1.3.4.4 Comparison with X-ray black hole binaries

There exists a population of stars, with masses of a few tens of solar masses, that have been found to show X-ray emission. It was found that these sources are in a binary system consisting of a black hole and a companion star, and are hence referred to as 'X-ray black hole binaries' (or X-ray binaries in short). In these systems, the stellar mass black hole appears to be accreting matter, efficiently, from the companion star, which results in the formation of an accretion disc. However, due to the smaller-scale of the accretion disc compared to an AGN, the accretion disc emission peaks at much higher temperatures such that they emit in the X-ray regime. As black holes are believed to be characterised solely based on their mass and spin, it is worth considering the properties of stellar-mass X-ray black hole binaries and the nature of accretion flow around these objects to understand the accretion of matter onto SMBHs.

Observations have shown that there are two different spectral states of X-ray binaries that show distinct properties and appear to be governed by the accretion rate onto the black hole (see reviews by Remillard & McClintock, 2006; Yuan & Narayan, 2014, and reerences therein). The first state is the 'hard' state, where the sources show high-energy ($> 10\,\mathrm{keV}$) X-ray emission and low-power radio jets. In the second, 'soft' state, the X-ray binaries show soft X-ray emission and display signs of a geometrically thin, optically thick accretion disc. The differences between these two spectral states is driven by a switch in the accretion rate onto the black hole, where for the 'hard' state, accretion occurs below a few per cent of the Eddington rate, whereas for the 'soft' state, accretion onto the black hole occurs at above this limit. Moreover, it is found that the SMBH population (i.e. the radiative-mode and jet-mode AGN), and the two states of the X-ray binaries lie on the same relation between radio-luminosity, X-ray luminosity, and black hole mass, the so-called 'fundamental place of BH accretion' (Merloni et al., 2003; Yuan et al., 2009; Gültekin et al., 2019).

Therefore, we can draw a clear analogy between the different modes of the X-ray stellar mass black hole binaries and the two modes of the SMBHs. The jet-mode AGN, found to undergo radiatively inefficient accretion at low Eddington rates (e.g. Best & Heckman, 2012), are analogues of the 'hard' state of the X-ray binaries where the accretion flow in both is described by an ADAF (Narayan & Yi, 1994; Yuan & Narayan, 2014). In contrast, the radiative-mode AGN undergo radiatively efficient accretion at above a few per cent of the Eddington rate, and are analogues of the 'soft' X-ray binary state, where the higher accretion rate results in the formation of a standard accretion disc.

1.4 Feedback from AGN

Having established the properties of AGN and their hosts, we can now consider how the AGN can affect their host galaxies and how this may have set up some of the relationships observed at low redshift. As outlined previously, feedback from AGN is often required by semi-analytical and cosmological numerical simulations (e.g. Bower et al., 2006; Croton et al., 2006; Vogelsberger et al., 2014; Schaye et al., 2015; Volonteri et al., 2016; Davé et al., 2019). In general, feedback from AGN is commonly invoked in two modes. The first, closely associated with the radiative-mode AGN, is thought to be responsible for halting the star-formation and for setting up the observed relationships in the local Universe, for example, the $M_{\rm BH} - \sigma$ relation (Kormendy & Ho, 2013); this is often dubbed as 'quasar-mode' feedback in simulations. The second mode, associated with jet-mode AGN, is in the form of kinetic or mechanical mode of feedback from the powerful radio jets, which are thought to be responsible for maintaining a galaxy as 'red and dead' once it has been quenched, via episodes of recurrent AGN activity; this is often referred to as 'radio-mode' feedback. The feedback from these two modes of AGN is discussed in more detail below.

1.4.1 Feedback from radiative-mode AGN

The radiation output from the AGN resulting from efficient accretion can input significant energy into the surroundings and drive outflows of gas (see review by Fabian 2012). As mentioned previously, the host galaxies also contain significant ongoing star formation, especially near the central regions, for the most luminous AGN. Stellar processes such as starbursts and supernovae have been shown to drive outflows from the nuclear region of the galaxy. Although these gas outflows have typical velocities of a few hundred km s^{-1} and mass outflow rates of tens to a few hundred M$_\odot$ yr^{-1} (see review by Veilleux et al., 2005), only a small fraction of gas completely escapes the potential well of the galaxy. Therefore, the majority of the outflow gas can eventually cool and form stars again and so may not be responsible for quenching star formation (e.g., in ULIRGs: Pereira-Santaella et al. 2018). In the presence of such high star-formation rates, it can be difficult to distinguish between stellar and AGN driven outflows. However, very high velocity outflows (> 500 km s^{-1}) are unlikely to be caused by energy output from stellar processes.

To probe the outflows and their role in affecting the star-formation, typically molecular

outflows are used, as a tracer for neutral atomic gas (which is difficult to detect). Blue-shifted absorption features from the OH molecule in the far-IR spectrum of AGN dominated galaxies have outflows with velocities of 1000 km s^{-1} and mass outflow rates of 1200 $M_{\odot}$ yr^{-1} (e.g. Feruglio et al., 2010; Fischer et al., 2010; Sturm et al., 2011; Veilleux et al., 2013; Runco et al., 2020; Herrera-Camus et al., 2020). Molecular gas outflows have also been mapped using other tracers such as CO line emission at mm/sub-mm wavelengths and other dense gas tracers (e.g. Audibert et al., 2019; Fluetsch et al., 2019; Lutz et al., 2020; Gao et al., 2021). Outflows that act on galactic scales and over a long enough time are required for suppressing galactic scale star formation. Although outflows can contain substantial masses of molecular gas, in some systems, it is found that these are not enough to remove vast amounts of gas from the galaxy with the effects of such outflows being limited to the central $\sim$ parsec regions (e.g. Morganti et al. 2015; Fluetsch et al. 2019, see also reviews by Veilleux et al. 2020; Pérez-Torres et al. 2021).

At higher redshift, radiative-mode AGN are commonly invoked for setting up the M_{BH}-σ relationship seen in the nearby Universe (King, 2003; Fabian, 2012). In addition to molecular outflows, ionised outflows around quasars (radiative-mode AGN) have been found to extend up to few tens of kpc (Harrison et al., 2012; Cresci et al., 2015) at high redshift. However, most of the outflowing gas does not escape the halo of the galaxy. Instead, AGN driven outflows eject gas from central regions but primarily escape from least resistance, low density regions, leaving behind the most dense gas still in the galaxy (e.g., Cano-Díaz et al. 2012; Carniani et al. 2016). Furthermore, as at low redshifts, these high velocity outflows are typically limited to the most luminous and extreme quasars, with little evidence in typical, lower mass, star-forming galaxies. Therefore, the role and importance of radiative-mode AGN in feedback is still debated in the literature.

Recently, other modes of quenching star-formation have been suggested. One such mechanism is 'halo quenching' whereby the main driver for quenching is proposed to be the galaxy halo rather than the stellar component. In such a scenario, as gas falls into the halo and towards the central massive galaxy, gravitational potential energy of the gas is converted to heat. In the most massive galaxies (with halo masses $\gtrsim 10^{12}$ $M_{\odot}$), the cooling time of the gas is larger than the dynamical timescale and, therefore, gas stays in a hot state within the halo. This can prevent gas from cooling and leads to suppression of star formation (Birnboim & Dekel, 2003; Dekel & Birnboim, 2006). However, it is found that this 'halo quenching' process only prevents the gas from cooling and forming stars for a few Gyrs, after which, a cooling flow may develop. Therefore, additional sources of heating such as AGN and environment may play a role in quenching (Gabor

& Davé, 2015).

1.4.2 Feedback from jet-mode AGN

The simplest mode of feedback from the jet-mode AGN comes from the kinetic energy of the jets; the energy output from the jets in these sources can exceed the bolometric radiative luminosity of the AGN by several orders of magnitude. This may also be important in the radio-loud subset of the radiative-mode AGN population. The kinetic (mechanical) energy stored in the jets can provide a significant energetic output in the form of work done on the surrounding environment to inflate the lobes; this so-called mechanical (kinetic) luminosity from the jets can be orders of magnitude larger than the monochromatic radio luminosities for the jet-mode AGN population.

The kinetic energy output from the radio jets can be estimated from the radio luminosity of the source using the minimum energy argument. This argument is based on assuming that the magnetic field strength has a value that minimises the combined energy in the magnetic field and the population of relativistic particles. There are, however, significant uncertainties on this estimate. One of the main sources of uncertainties is the lack of physical understanding of the radio lobes and its composition. Willott et al. (1999a) expressed the combination of all of the uncertainties as a single parameter f_W such that the jet mechanical power is given by

$$P_{\rm mech,sync} = 4 \times 10^{35} (f_{\rm W})^{3/2} \left(\frac{L_{1.4\,\rm GHz}}{10^{25}\,\rm WHz^{-1}} \right)^{0.86} W \tag{1.4}$$

where, $L_{1.4GHz}$ is the radio luminosity measured at 1.4GHz and the values of f_W are typically in the range of 1 - 20 (Heckman & Best, 2014). The relation between the mechanical luminosities ($P_{\rm mech,sync}$) and 1.4 GHz radio luminosities estimated in this way is shown in Fig. 1.6.

An alternative method of estimating the jet kinetic luminosities from radio luminosities has also been developed based on the pV energy needed to inflate the X-ray cavities. As the expanding jets plough through the surrounding material, they create cavities in the hot gas, which can be observed in the X-rays (Boehringer et al., 1993). The jet kinetic luminosity can be estimated from the radio luminosity by considering the energy required to expand the lobes and inflate the X-ray cavities of pressure p with a volume V, $E_{\rm cav} = f_{\rm cav}\, pV$ (e.g. Cavagnolo et al., 2010; Best et al., 2014). It is commonly assumed that this is the sum of the pV of work done in inflating the cavities, and $3pV$ of energy stored in the relativistic plasma of the jets, hence resulting in the commonly

assumed value of $f_{\rm cav} = 4$ (e.g. Cavagnolo et al., 2010; Best et al., 2014; Butler et al., 2019); however, there is a large uncertainty in these assumptions, for example, shock heating of the intergalactic medium from the jets may result in higher values. There is nonetheless a good correlation between the cavity powers and the 1.4 GHz radio luminosities, albeit with a scatter of ~ 0.7 dex. The jet mechanical luminosity, based on the analysis from Cavagnolo et al. (2010) and Heckman & Best (2014) is given as

$$L_{\rm mech,cav} = 7 \times 10^{36} f_{\rm cav} \left(L_{1.4\,\rm GHz} / 10^{25}\,\rm W\,Hz^{-1} \right)^{0.68} \rm W \tag{1.5}$$

where, $L_{1.4\,\rm GHz}$ is the 1.4 GHz radio luminosity. This relation is also illustrated for $f_{\rm cav} = 4$ in Fig. 1.6, and is found to be in good agreement with the mean $L_{\rm radio}$ - $L_{\rm mech,sync}$ relation of Willott et al. (1999b), calculated from the synchrotron properties.

As mentioned in Sect. 1.3.4.2, jet-mode AGN activity is typically triggered by the hot gas halo cooling and accreting as 'cool clouds' onto the black hole. The resulting jets deposit their kinetic energy into the surrounding inter-stellar medium (ISM) which heats up the gas in the halo, preventing further gas cooling and accretion. As the fuel supply to the AGN 'runs out', the radio jets are 'switched off'. After some time, this allows the gas in the hot halo to cool down enough such that accretion and therefore radio jets can resume again. Therefore, feedback from the jet-mode AGN plays an important role in the evolution of massive galaxies as the medium into which the jet energy is deposited also forms the likely fuel source for the AGN, thus providing the conditions for a self-regulating AGN feedback cycle.

The strongest evidence of this feedback mechanism is seen within massive ellipticals in clusters with a cooling flow, where gas inflow and cooling to the central region occurs, resulting in cold gas densities and predicted SFRs of the order of 100 - 1000 $\rm M_\odot\ year^{-1}$. However, as discussed in Sect. 1.2.4, observationally, the SFRs observed are at least an order of magnitude lower than expected (e.g., Böhringer et al. 2002; Peterson et al. 2003). To match the observations, the cooling losses from the gas must be balanced by a heating source, provided by the mechanical energy output from the AGN jets (as discussed earlier in this section; see also Fig. 1.6). Direct observational evidence of this mode of feedback comes from X-ray cavities/bubbles inflated by the jets transferring energy into the inter-galactic medium (IGM), which are observed in most 'cool-core' clusters (Fabian, 2012). It is found that the time-averaged energy output from these radio jets, in the form of mechanical energy, does indeed balance the radiative cooling losses from the hot gas (e.g. Best et al. 2006; McNamara & Nulsen 2007; Heckman & Best 2014; Hardcastle et al. 2019; see also Fig. 1.4). In the most massive cooling flow clusters, the AGN are observed to be switched on almost all of the time to maintain the

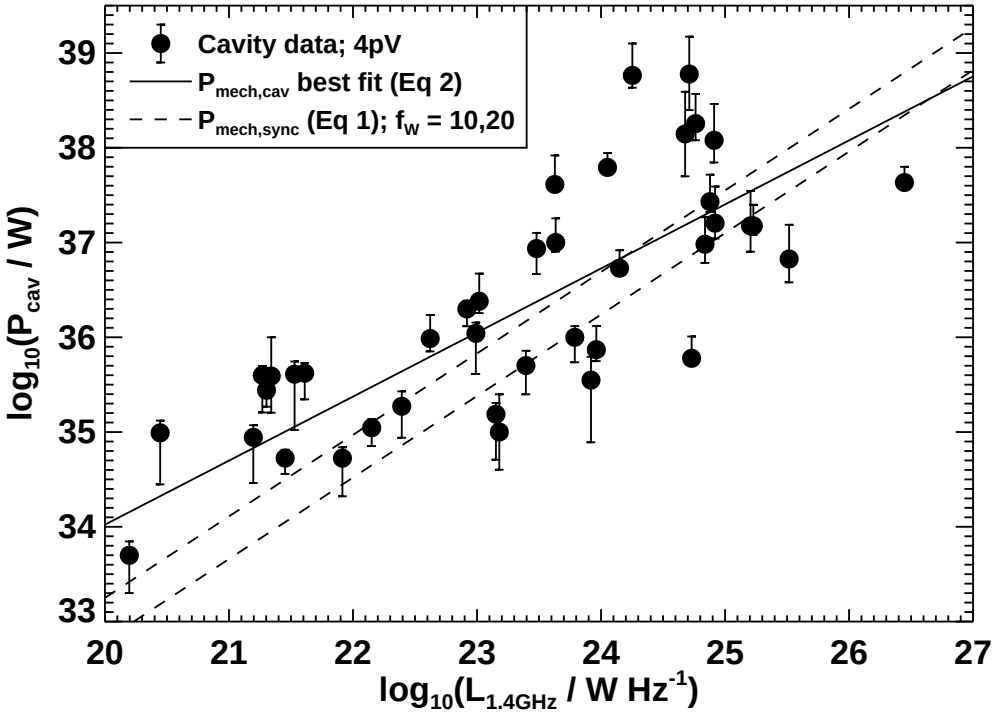

Figure 1.6 *The relation between jet mechanical luminosity and 1.4 GHz radio luminosities. The dashed lines shows the estimate of mechanical luminosities estimated based on the synchrotron properties using the Willott et al. (1999b) relation for* $f_{\rm w}$ = 10 and 20. *The solid line shows the same but estimated using the cavity method based on* $f_{\rm cav}$ = 4. *The mean of both of these estimates shows reasonable agreement. The cavity data were taken from Rafferty et al. (2006); Bîrzan et al. (2008); Cavagnolo et al. (2010). Plot taken from Heckman & Best (2014).*

heating-cooling balance.

On galaxy-scales however, it is found that the instantaneous energy input from radio-AGN is more than an order of magnitude larger than that required to balance the cooling losses (Best et al., 2006, 2007). The prevalence of radio-AGN in lower mass systems is lower, however when the AGN is triggered, it overheats the gas in the galaxy. This additional heating prevents the gas from cooling into 'cool clouds' and effectively shuts off the fuel supply to the AGN and therefore the AGN activity itself. The AGN is 'switched off' for a long time due to this additional heating, such that the time-averaged energy input from the AGN roughly balances the cooling losses. These observations in lower mass systems support the simple model of radio-AGN feedback outlined above.

1.5 Evolution of Radio-Loud AGN

Sects. 1.2- 1.4 present evidence for the different properties of AGN and how they can have an effect on their host galaxies, mainly at low redshift. However, the SMBHs are growing at a much slower rate today than they were in the past and the cosmic SFR is also much lower today (see Sect. 1.2.2). It is therefore useful to study how the properties of the AGN and the hosts in which they reside evolve through cosmic time and how the galaxies are quenched from the peak of star-formation at early times. In this thesis, I am particularly interested in studying the properties of radio-loud AGN (also known as radio-AGN) and the cosmic evolution of their AGN feedback using deep and wide radio continuum surveys. Therefore, in the rest of this chapter, I particularly focus on the properties and evolution of the radio-loud AGN population in the current state of the literature.

Firstly, it is important to note that the radio-selected AGN population contains both the jet-mode AGN and the radio-loud subset of the radiative-mode AGN. In addition, the population of (radio-quiet) radiative-mode AGN may also be detected by radio surveys probing at sufficiently faint luminosities, and in fact form a significant part of the radio-selected samples below ~ 1 mJy at 150 MHz (Wilman et al., 2008), like the dataset being used in this thesis (see Chapter 2). Hereafter, as we are focused on studying the evolution of the radio-loud AGN population in this thesis, the jet-mode AGN are referred to as LERGs, and the radio-loud radiative-mode AGN are referred to as HERGs.

One of the ways to study how the population of radio-loud AGN evolves is by constructing the radio luminosity function (RLF), which represents the space density (i.e. the number of sources per unit comoving volume and per unit log luminosity) as a function of radio luminosity. The evolution of the overall radio-AGN population has been studied in detail (e.g. Dunlop & Peacock, 1990; Rigby et al., 2011, 2015). It is found that the space density of the highest radio luminosity sources increases by a factor of ≈ 100 from $z = 0$ to $z \sim 3$, but at lower luminosities, the population both evolves less strongly, and reaches its peak at lower redshifts ($z \sim 1$; e.g. Rigby et al. 2011, 2015).

This evolution seen in the RLF is likely due to the different evolution of the LERG and HERG populations encompassed by the overall radio-loud AGN populations. Best & Heckman (2012), using a sample of over 18 000 radio-loud AGN, first constructed the RLF of the LERG and HERG populations, separately, in the local Universe. These local RLFs are also shown in Fig. 1.7, taken from Best & Heckman (2012). They found that

although both of these populations are found across all radio luminosities, the LERGs dominate the space densities at low radio luminosities ($L_{1.4\,\mathrm{GHz}} \lesssim 10^{25}\,\mathrm{W\,Hz^{-1}}$), while the HERGs dominate at high luminosities (see Fig. 1.7). They also find that these two populations evolve differently, with the LERGs showing little evolution and the HERGs showing strong evolution at all luminosities; as these dominate the population at high luminosities, this can thus explain the strong evolution seen in the RLF of the overall radio-loud AGN. These results also highlight the importance of studying the cosmic evolution of these two populations separately.

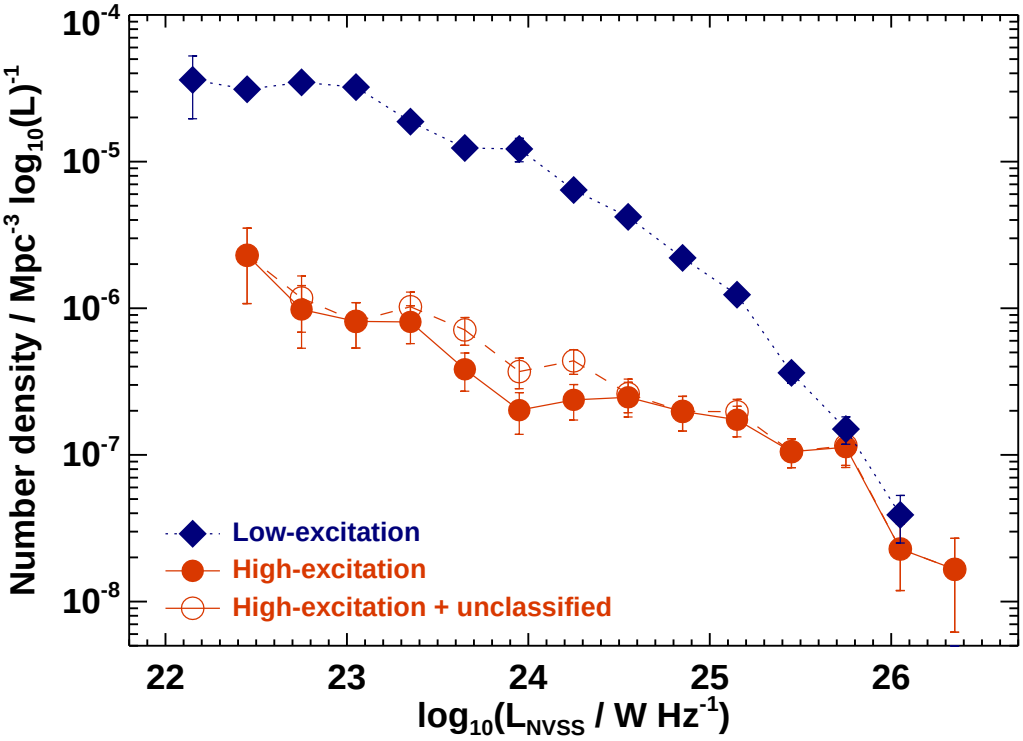

Figure 1.7 *The local radio luminosity function of LERGs and HERGs. The two populations show different shapes, with the LERGs dominating the space densities at low luminosities (L_{NVSS}; from 1.4 GHz observations), while the HERGs dominate at high luminosities. Plot taken from Best & Heckman (2012).*

Best et al. (2014) extended their previous analysis to higher redshifts to provide the first measurement of the evolution of the LERG and HERG populations, separately, out to $z \sim 1$. They compiled a sample of 211 radio-loud AGN from various radio continuum surveys over a wide area, and used emission line diagnostics to separate LERGs and HERGs. They found that at all luminosities, the comoving space density of the HERG population increases by a factor of ~ 7 between $z = 0$ and $z \sim 0.75$. In contrast, the evolution of the LERGs showed a dependence on radio luminosity, with the space

density decreasing mildly with redshift at low-luminosities ($L_{1.4\,\mathrm{GHz}} \lesssim 10^{24}\,\mathrm{W\,Hz^{-1}}$) but increasing slightly with redshift at higher radio luminosities. Similarly, Pracy et al. (2016) constructed the evolving HERG and LERG luminosity functions (at $z < 0.75$) using a sample of around 5000 radio-AGN derived from cross-matching the FIRST (Becker et al., 1995) survey at 1.4 GHz and the SDSS (York et al., 2000b). They also found that the LERG population shows little-to-no evolution over this redshift, whereas the HERGs showed strong evolution with an increase in the characteristic density by a factor of $\approx (1 + z)^3$.

More recently, Butler et al. (2019) extended these analyses to higher redshfits, using a sample of 1729 LERGs from 2.1 GHz observations of the XMM extragalactic survey field (XXL-S) using the Australia Telescope Compact Array (ATCA). They found that the space densities of the LERGs showed weak evolution out to $z \sim 1.3$, largely consistent with previous studies. At low radio frequencies, Williams et al. (2018) studied the evolution of HERGs and LERGs, classified using photometry alone, between $0.5 < z \leq 2$ using 150 MHz LOw Frequency ARray (LOFAR; van Haarlem et al., 2013) observations of the Boötes field. They studied the properties of 243 LERGs and 398 HERGs. They found that HERGs are generally found in star-forming galaxies and that the prevalence of these AGN increases with redshift, consistent with previous results. Similarly, even at high redshift, the prevalence of LERGs with stellar mass obeys the same steep relationship found in the local Universe (see Sect. 1.3.4.2) with a strong preference to be hosted by galaxies with high mass and red colours. Over this redshift range, they find that the LERG population shows a strong decline in space densities with increasing redshift that is found to be consistent with the decline in the space densities of the expected hosts, as in the massive quiescent galaxies.

These results provide further evidence that the two classes of AGN are fuelled by different processes. The HERG population is typically fuelled by cold gas accretion in star-forming galaxies and therefore, the space density of this class of AGN increases with redshift, in line with the increasing cosmic star-formation rate, where the higher gas fractions mean that more 'fuel' was available to induce both the AGN and star-formation activity. Similarly, the LERGs are typically found in massive galaxies with hot gas haloes, and the prevalence of such hosts decreases with increasing redshift.

1.6 The structure of this thesis

Although recent studies of the two modes of radio-loud AGN population have extended analysis out to $z \gtrsim 1$, physical processes underpinning AGN triggering and activity remain poorly understood. Detailed characterisation of how the radio-loud AGN activity depends on various properties (e.g. stellar mass, star-formation rate, radio luminosity, environment, etc.) at earlier times can provide crucial tests for our current models of galaxy formation and evolution. Such analysis is currently lacking, largely as a result of radio surveys that are not deep enough over wide areas to detect large statistical samples of low-luminosity sources out to higher redshifts. The advent of the new generation of radio telescopes such as the Square Kilometre Array (SKA) and its pathfinders will make such detailed statistical studies possible.

In this thesis, I use the deep observations from the LOFAR Two-Metre Sky Survey (LoTSS) Deep Fields (Tasse et al., 2021; Sabater et al., 2021) to study the cosmic evolution of the radio-loud AGN population and feedback from these AGN. The current LoTSS Deep Fields dataset, covering $\sim 26\,\mathrm{deg}^2$ (including multi-wavelength coverage) and reaching an unprecedented rms depth of $\sim 20\,\mu\mathrm{Jy\,beam}^{-1}$ (at 150 MHz), is comparable in depth to the deepest existing radio continuum surveys (e.g. VLA-COSMOS 3 GHz project; Smolčić et al., 2017a) but with more than an order of magnitude larger sky-area coverage. With this combination of deep, high-quality radio and multi-wavelength data over tens of square degrees, and along multiple sight-lines, the LoTSS Deep Fields are now able to probe a cosmological volume large enough to sample all galaxy environments to beyond $z \sim 1$, minimise the effects of cosmic variance (to an estimated level of $\sim 4\%$) for $0.5 < z < 1.0$; Driver & Robotham 2010), and build statistical radio-selected samples of AGN and star-forming galaxies, even when simultaneously split by various physical parameters.

In the following Chapter, I present new forced, matched-aperture multi-wavelength photometry catalogues generated in the LoTSS Deep Fields to provide sufficiently robust catalogues suited to the scientific aims of the LOFAR surveys. Then, in Chapter 3, using these new multi-wavelength catalogues, I detail the identification of the host galaxies of the radio-detected sources in the LoTSS Deep Fields. The identification of the counterparts of radio sources is crucial in order to determine physical properties of the radio-detected sources (e.g. photometric redshifts, stellar masses, star-formation rates, luminosities, etc.). This process is performed using a combination of statistical and visual classification process, resulting in a sample of over 80 000 radio-sources with a host galaxy identification rate of $> 97\%$, owing to the depth of the available ancillary

datasets. The work based on these two Chapters has been published as Kondapally et al. (2021). Then, in Chapter 4, using the datasets generated in Chapters 2 and 3, I study the cosmic evolution of the radio-loud AGN population and their host galaxy properties out to $z \sim 2.5$ using a sample of 11 466 radio-loud AGN and characterise the evolution of the AGN feedback from these sources in the past ~ 10 Gyrs. This work will appear as Kondapally et al. (in prep.). Finally, in Chapter 5, I present the conclusions from this thesis and outline some avenues of future work to extend the analysis presented in the previous Chapters.

Throughout this work, we use a flat ΛCDM cosmology with $\Omega_{\rm M} = 0.3$, $\Omega_{\Lambda} = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}, and a radio spectral index $\alpha = -0.7$ (where $S_{\nu} \propto \nu^{\alpha}$). Where quoted, magnitudes are in the AB system (Oke & Gunn, 1983), unless otherwise stated.

Chapter 2

Creation of multi-wavelength catalogues in LoTSS Deep Fields

2.1 Introduction

This thesis aims to study the cosmic evolution of the properties of radio-AGN and how energetic feedback from these AGN affects galaxy formation and evolution at early cosmic epochs. As discussed in Chapter 1, there has been a considerable amount of work over the past two decades in characterising the nature of the radio-AGN population in the local Universe and on how feedback from these AGN shapes the observed physical galaxy properties.

The cosmic evolution of the radio-AGN at early times, which is not well known, can provide important tests to our current understanding of galaxy formation and evolution models. It is clear that observational studies aiming to characterise the high-redshift radio-AGN population not only require deep radio-continuum observations to extend measurements to fainter radio-luminosities and higher redshifts, but also surveys across large-areas of the sky for deriving robust number statistics, tracing a wide range of galaxy environments to detect the rarer sources, and also in reducing the effects of cosmic variance. The ancillary multi-wavelength information is particularly important for maximising the scientific output of radio-continuum surveys; the identification of the multi-wavelength counterparts of the radio-detected sources is crucial (see Chapter 3) in studying the physical host galaxy properties of the radio sources e.g., via the spectral energy distribution (SED) fitting of their photometry.

As detailed in Chapter 1, in this thesis, we use observations from the LOFAR telescope, which is undertaking deep observations of best studied multi-wavelength, degree scale fields in the northern sky, forming the LOFAR Two-Metre Sky Survey (LoTSS) Deep Fields (Tasse et al. 2021 and Sabater et al. 2021). The first three LoTSS Deep Fields are the European Large-Area ISO Survey-North 1 (ELAIS-N1; Oliver et al. 2000), Lockman Hole, and Boötes (Jannuzi & Dey, 1999); these were chosen to have extensive multi-wavelength coverage from past and ongoing deep, wide-area surveys sampling the X-ray (e.g. Brandt et al. 2001; Hasinger et al. 2001; Manners et al. 2003; Murray et al. 2005), ultra-violet (UV; e.g. Martin et al. 2005; Morrissey et al. 2007) to optical (e.g. Jannuzi & Dey 1999; Cool 2007; Muzzin et al. 2009; Wilson et al. 2009; Chambers et al. 2016; Huber et al. 2017; Aihara et al. 2018) and to infrared (IR; e.g. Lonsdale et al. 2003; Lawrence et al. 2007; Ashby et al. 2009; Whitaker et al. 2011; Mauduit et al. 2012; Oliver et al. 2012) wavelengths; this is ideal for a wide range of our scientific objectives. These fields also benefit from additional radio observations at higher frequencies from the Giant Metrewave Radio Telescope (GMRT; e.g. Garn et al. 2008b,a; Sirothia et al. 2009; Intema et al. 2011; Ocran et al. 2019; Ishwara-Chandra et al. 2020) and the VLA (e.g. Ciliegi et al. 1999; Ibar et al. 2009).

In two of the deep fields, ELAIS-N1 and Lockman Hole, individual catalogues from the various surveys (see Sect. 2.2) detected based on different filters already exist. However, catalogue combination issues, such as when sources are blended in lower resolution catalogues, or only detected in a subset of filters, present significant challenges. Furthermore, the usefulness of existing catalogues for photometric redshifts is limited due to the varying catalogue creation methods. For example, magnitudes in these catalogues were typically measured within different apertures and with different methods of correcting to total magnitudes, leading to colours that are not sufficiently robust. In addition, for the sources that were detected in only a subset of filters, the lack of information or application of a generic limiting magnitude in other filters, would lead to a loss of information on galaxy colours compared to a forced photometry measurement. This can have a significant impact on the accuracy of SED fitting and therefore the photometric redshift estimates. To alleviate these issues in the existing catalogues, we build pixel-matched images and create matched aperture, multi-wavelength catalogues with forced photometry spanning the UV to mid-infrared wavelengths in the ELAIS-N1 and Lockman Hole fields. In ELAIS-N1, this new catalogue generated also includes deeper optical imaging than those available in literature. This process provides high quality catalogues for SED fitting and photometric redshift estimates, and also for identifying the radio-source counterparts; these will be valuable not only for studying the evolution of the radio-AGN in the LoTSS Deep Fields but also for a wide host of

other scientific objectives.

2.2 Description of the data

2.2.1 Radio data

The details of the LOFAR observations used, along with the calibration and source extraction methods employed, are described in detail in Tasse et al. (2021) and Sabater et al. (2021). Here, we summarise these steps and list the key properties of the radio data released (see Table 2.1).

The LOFAR observations for the LoTSS Deep Fields were taken with the High Band Antenna (HBA) array, with frequencies between 114.9–177.4 MHz. The ELAIS-N1 data were obtained from LOFAR observation cycles 0, 2, and 4, consisting of 22 visits of $\sim$ 8 hour integrations (total $\sim$ 164 hours). The Lockman Hole data were obtained from cycles 3 and 10, with 12 visits of $\sim$ 8 hour integrations (total $\sim$ 112 hours). The Boötes dataset was obtained from cycles 3 and 8 with total integration time of $\sim$ 100 hours. The total exposure times, pointing centres and root mean square (RMS) sensitivities of calibrated data are listed in Table 2.1.

The calibration of interferometric data at these low frequencies is a challenging task, in particular due to direction dependent effects caused by the ionosphere and the station beam (Intema et al., 2009). These direction dependent effects (DDEs) are corrected using a facet based calibration, where the entire field-of-view is divided into small facets and the solutions computed for each facet individually (see Shimwell et al. 2019 and Tasse et al. (2021) for details). The overall calibration pipeline involves solving first for direction-independent (van Weeren et al., 2016; Williams et al., 2016; de Gasperin et al., 2019) and then for direction-dependent effects, as described for the LoTSS Data Release 1 (LoTSS DR1; Shimwell et al. 2017, 2019), but with an updated version of the pipeline applied to the LoTSS Deep Fields (Tasse et al., 2021) that is more robust against un-modelled flux absorption and artefacts around bright sources. Finally, the imaging was carried out using DDFacet (Tasse et al., 2018) to generate a high resolution ($6''$) Stokes I image for all fields, reaching unprecedented RMS depths of $S_{150\mathrm{MHz}} \sim 20$, 22, and 32 μJy beam^{-1} at the field centres in ELAIS-N1, Lockman Hole, and Boötes, respectively (see Table 2.1). The current imaging data released includes data from the Dutch baselines only; international station data are available and will be included in future data releases.

Source extraction is performed on the Stokes I radio image in each field using Python Blob Detector and Source Finder (PyBDSF; Mohan & Rafferty 2015). We refer the reader to Mohan & Rafferty (2015) for a detailed description of the software, and to Tasse et al. (2021) and Sabater et al. (2021) for details of the detection parameters used to generate the PyBDSF radio catalogues. In summary, sources are extracted by first identifying islands of emission (using island and peak detection thresholds of 3 and 5σ, respectively). The islands are then decomposed into Gaussians, which are then grouped together to form a source. An island of emission may contain single or multiple Gaussians and sources may be formed of either only one Gaussian or by grouping multiple Gaussians. For unintentional historic reasons, source extraction in Lockman Hole and Boötes were performed with slightly different parameters than in ELAIS-N1, leading to a higher fraction of PyBDSF sources being split into multiple Gaussians; however, after these are correctly grouped using our visual classification schemes (see Sect. 3.4); this should have little or no effect on the final cross-matched catalogue. We summarise some key properties of the radio data and the PyBDSF catalogues for each field in Table 2.1.

2.2.2 Multi-wavelength data in ELAIS-N1

The ELAIS-N1 field has deep multi-wavelength (0.15 μm - 500 μm) observations taken as part of many different surveys, covering up to 10 deg^2. The ELAIS-N1 footprint illustrated in Fig. 2.1 shows the coverage of some of the key optical-IR surveys used, as well as the region imaged by LOFAR (plot limited to the 30% power of the primary beam). In total, we generate photometry from 20 UV to mid-IR filters, with additional far-IR data from *Spitzer* and *Herschel*. The typical depths and areas covered by the multi-wavelength imaging datasets are listed in Table 2.2.

Table 2.1 *Summary of the radio data properties in the current data release of the LoTSS Deep Fields. The radio data have an angular resolution of 6″ and cover around 68 deg^2 out to the primary beam 30% power point.*

	ELAIS-N1	Lockman Hole	Boötes
Centre RA, DEC [deg]	242.75, 55.00	161.75, 58.083	218.0, 34.50
Central frequency [MHz]	146	144	144
Central RMS [μJy/beam]	20	22	32
Integration Time [hrs]	164	112	80
№ PyBDSF radio sources	84 862	50 112	36 767
Reference	Sabater et al. (2021)	Tasse et al. (2021)	Tasse et al. (2021)

2.2.2.1 UV to mid-infrared data in ELAIS-N1

Optical data for ELAIS-N1 comes from Panoramic Survey Telescope and Rapid Response System (Pan-STARRS-1; Kaiser et al. 2010). Pan-STARRS 1 (PS1) is installed on the peak of Haleakala on the island of Maui in the Hawaiian island chain. The PS1 system uses a 1.8m diameter telescope together with a 1.4 gigapixel CCD camera with a 7 deg^2 field-of-view. A full description of the PS1 system is provided by Kaiser et al. (2010) and the PS1 optical design is described in Hodapp et al. (2004). The PS1 photometry is in the AB system (Oke & Gunn, 1983) and the photometric system is described in detail by Tonry et al. (2012). The PS1 data in ELAIS-N1 consists of broadband optical (*g, r, i, z* and *y*) imaging from the Medium Deep Survey (MDS), one of the PS1 surveys (Chambers et al., 2016). As part of the MDS, ELAIS-N1 (and the other fields) was visited on an almost nightly basis to obtain deep, high cadence images, with each epoch consisting of eight dithered exposures. This PS1 dataset provides the deepest wide-area imaging at redder optical wavelengths across ELAIS-N1.

Additional optical data were taken from the Hyper-Suprime-Cam Subaru Strategic Program (HSC-SSP) survey. ELAIS-N1 is one of the 'deep' fields of the HSC-SSP survey, covering a total of ~ 7.7 deg^2 in optical filters *g, r, i, z, y*, and the narrow-band NB921, taken over four HSC pointings. The images were acquired from the first HSC-SSP data release (Aihara et al., 2018).[1] The HSC data have higher angular resolution than the PS1 data, and are of comparable depths at bluer wavelengths. The use of both HSC and PS1 data allows the advantages of each survey to be present in the catalogues, and in addition, provides complementary photometric data points for SED fitting.

The broadband *u*-band data were obtained from the *Spitzer* Adaptation of the Red-sequence Cluster Survey (SpARCS; Wilson et al. 2009; Muzzin et al. 2009). SpARCS is a follow-up of the *Spitzer* Wide-area Infra-Red Extragalactic (SWIRE) survey fields taken using the MegaCAM instrument on the Canada-France-Hawaii Telescope (CFHT). In ELAIS-N1, the data were taken over 12 CFHT pointings (1 deg^2 each) covering $\sim$ 12 deg^2 in total.

The UV data were obtained from the Release 6 and 7 of the Deep Imaging Survey (DIS) taken with the Galaxy Evolution Explorer (GALEX) space telescope (Martin et al., 2005; Morrissey et al., 2007). GALEX observations were taken in the near-UV (NUV)

[1]We note that DR2 of HSC-SSP was released in May 2019 (Aihara et al., 2019). At this time, our optical catalogues had been finalised and the visual cross-identification process was in progress. Processing the new HSC-SSP DR2 data to modify the optical catalogues would have been unfeasible, leading to delays in the visual identification process. However, we plan on including new HSC-SSP data releases for future deep fields data releases.

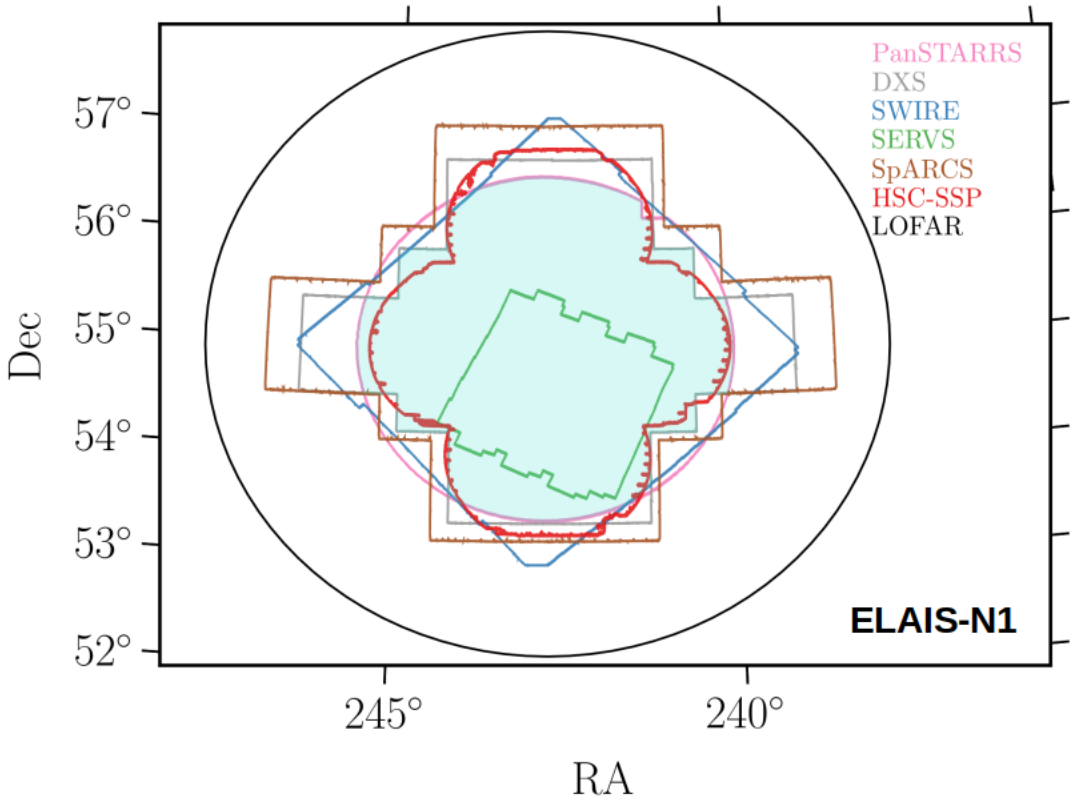

Figure 2.1 *Footprint (north up, east left) for ELAIS-N1 field showing the coverage of multi-wavelength data from various surveys in optical and IR bands described in Sect. 2.2.2. The LOFAR radio coverage is also shown in black. The shaded light blue region shows the selected area of overlap that is used for the radio-optical cross-match in Chapter 3 for ELAIS-N1 (~7.15 deg^2), as described in Sect. 2.2.2.3 with slightly reduced area after bright-star masking.*

and far-UV (FUV) spanning 1350Å - 2800Å and have a field-of-view ≈ 1.5 deg^2 per pointing, covering around 13.5 deg^2 in total.

The near-infrared (NIR) J and K band data come from the UK Infrared Deep Sky Survey (UKIDSS) Deep Extragalactic Survey (DXS) DR10 (Lawrence et al., 2007). Observations were taken using the WFCAM instrument (Casali et al., 2007) on the UK Infrared Telescope (UKIRT) in Hawaii as part of the 7 year DXS survey plan and cover ~ 8.9 deg^2 of the ELAIS-N1 field. The photometric system is described in Hewett et al. (2006).

The mid-infrared (MIR) 3.6 μm, 4.5 μm, 5.8 μm and 8.0 μm data were acquired from the IRAC instrument (Fazio et al., 2004) on board the *Spitzer* Space Telescope (Werner et al., 2004). We use two *Spitzer* surveys that cover the ELAIS-N1 field: the SWIRE (Lonsdale et al., 2003) survey and the *Spitzer* Extragalactic Representative Volume Survey (SERVS; Mauduit et al. 2012). The SWIRE data were taken in January 2004 and cover an area of ~ 10 deg^2 in all four IRAC channels. The SERVS project imaged a small part of the ELAIS-N1 field, covering around 2.4 deg^2 in only two channels (3.6 μm and 4.5 μm) during Spitzer's warm mission but reaching ~ 1 mag deeper than SWIRE.

2.2.2.2 Additional far-infrared data in ELAIS-N1

Longer wavelength data at 24 μm comes from the Multiband Imaging Photometer for *Spitzer* (MIPS; Rieke et al. 2004) instrument on-board Spitzer. Data were also taken from *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012) by the *Herschel* Space Observatory (Pilbratt et al., 2010), using the Spectral and Photometric Imaging Receiver (SPIRE; Griffin et al. 2010) instrument at 250 μm, 360 μm and 520 μm, and Photodetector Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) at 100 μm and 160 μm. The three fields are all part of Level 5 or 6 deep tiers of HerMES, comprising one of the deepest, large-area *Herschel* surveys available. The 70 μm data from MIPS or PACS are not included in our catalogues (and nor within HELP) due to their poorer sensitivity.

In part due to their low angular resolution, these FIR data are not used to generate the forced, matched aperture catalogues. Instead, FIR fluxes are added from existing catalogues from the SPIRE and PACS maps, generated by the *Herschel* Extragalactic Legacy Project (HELP; Shirley et al. 2021). FIR fluxes from HELP were incorporated by performing a cross-match between our multi-wavelength catalogue and HELP

catalogues using a 1.5″ cross-match. If no match was found within the HELP catalogues, FIR fluxes were extracted using the XID+ software (Hurley et al., 2017), incorporating the radio (or optical) positions into the list of priors. The details of the process of generating and adding FIR fluxes is described by McCheyne et al. (subm).

2.2.2.3 Selected survey area in ELAIS-N1

The radio data cover a significantly larger area than the accompanying multi-wavelength data. We therefore define the area used for cross-matching in Chapter 3 for the ELAIS-N1 field as the overlapping area between PanSTARRS, UKIDSS, and SWIRE, covering ~ 7.15 deg^2. This overlap area is indicated by the blue shaded region in the ELAIS-N1 footprint shown in Fig. 2.1. At the largest extent of this selected area from the radio field centre, the radio primary beam correction factor is ~0.65, resulting in a noise level approximately 50% higher than in the centre. There is thus a moderate variation in the depth of the radio data across the survey region.

2.2.3 Multi-wavelength data in Lockman Hole

Lockman Hole also possesses deep multi-wavelength (0.15 μm–500 μm) data and is the field with the largest area of multi-wavelength coverage, as shown by the footprint in Fig. 2.2 . The typical depths and areas covered by the multi-wavelength and radio imaging datasets are listed in Table 2.2.

2.2.3.1 UV to mid-infrared data in Lockman Hole

The optical data in Lockman Hole come from two surveys taken by the CFHT-MegaCam instrument: SpARCS and the Red Cluster Sequence Lensing Survey (RCSLenS; Hildebrandt et al. 2016). The SpARCS data in Lockman Hole consist of broadband *u, g, r, z* filter images taken using 14 pointings of the CFHT, covering around 13.3 deg^2 of the field. The RCSLenS data consist of *g, r, i, z* observations covering around 16 deg^2. The coverage from RCSLenS however, is not contiguous, with gaps between different pointings.

Similar to ELAIS-N1, the NUV and FUV imaging data come from the GALEX DIS Release 6 and 7, and, the NIR data is obtained from the *J* and *K* bands of the UKIDSS-DXS DR10, covering a maximum area of around 8 deg^2. Observations of the Lockman

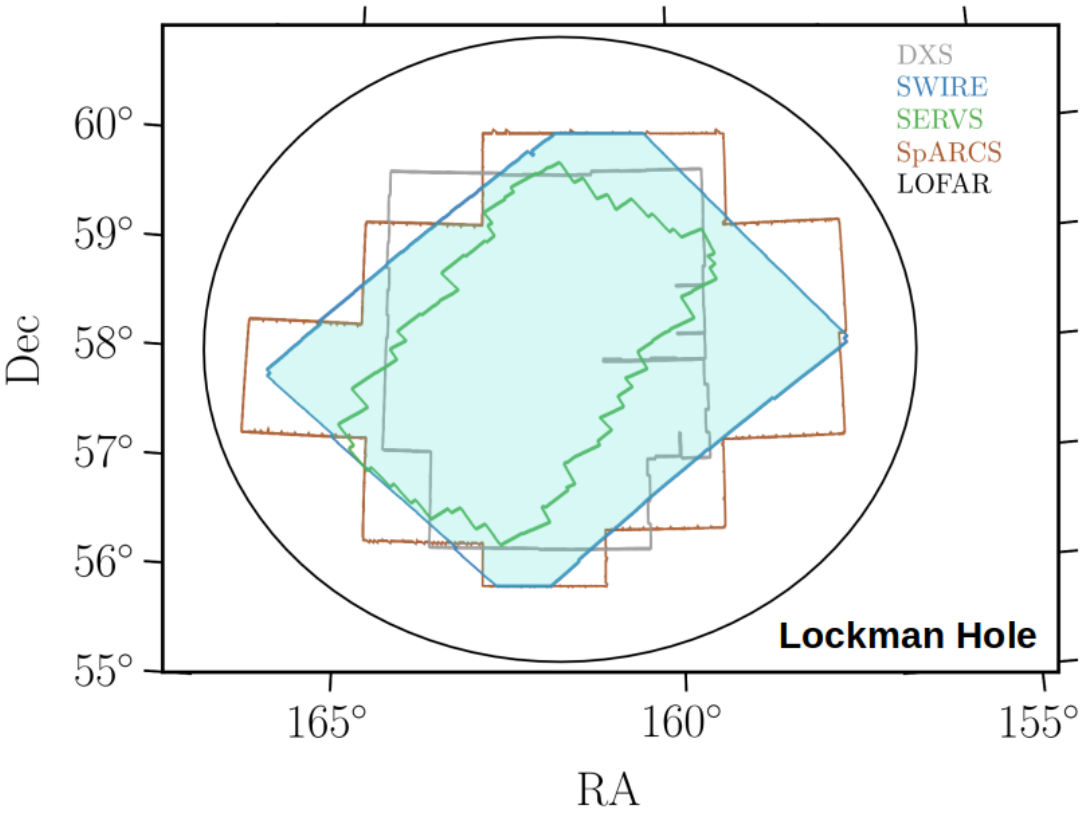

Figure 2.2 *Footprint (north up, east left) for the Lockman Hole field showing the coverage of multi-wavelength data from various surveys in optical and IR bands described in Sect. 2.2.3. The LOFAR radio coverage is also shown in black. The shaded light blue region shows the selected area of overlap that is used for the radio-optical cross-match in Chapter 3 for Lockman Hole (~10.73 deg^2), as described in Sect. 2.2.3.3, with slightly reduced area after bright-star masking.*

Hole field were also taken in IRAC channels as part of SWIRE and SERVS, reaching similar depths as in ELAIS-N1 but over much larger areas. The SWIRE data in all four IRAC channels cover around 11 deg^2 whereas the deeper SERVS data in the two IRAC channels (3.6 μm and 4.5 μm) cover around 5.6 deg^2.

2.2.3.2 Additional far-infrared data in Lockman Hole

Lockman Hole is also covered by both *Spitzer* MIPS and HerMES observations. These FIR fluxes were added using catalogues generated by HELP and by running XID+, following the same method as for ELAIS-N1 (see McCheyne et al. subm).

Table 2.2 *Key properties of the multi-wavelength data in ELAIS-N1 and Lockman Hole. For each filter, we include the Vega-AB conversion factor (if any) used for generating pixel-matched mosaics (see Sect. 2.3.1), the average PSF FWHM, and the approximate area covered by each survey. The 3σ depths (in AB system) in each filter estimated from the variance of empty, source free 3″ apertures, and the filter dependent Galactic extinction values, $A_{band}/E(B-V)$ are listed.*

Field/Survey	Band	Vega-AB [mag]	PSF [arcsec]	3σ depth [mag]	$A_{band}/E(B-V)$	Area [deg^2]
ELAIS-N1						
GALEX-DIS						13.53
	FUV	-	4.2	24.94	28.637	
	NUV	-	5.3	25.26	8.675	
SpARCS						11.81
	u	-	0.9	25.4	4.595	
PanSTARRS						8.05
	g	-	1.2	25.5	3.612	
	r	-	1.1	25.2	2.569	
	i	-	1.0	25.0	1.897	
	z	-	0.9	24.6	1.495	
	y	-	1.0	23.4	1.248	
HSC						7.70
	G	-	0.5	25.6	3.659	
	R	-	0.7	25.0	2.574	

continued . . .

. . . continued

Field/Survey	Band	Vega-AB [mag]	PSF [arcsec]	3σ depth [mag]	$A_{band}/E(B-V)$	Area [deg^2]
	I	-	0.5	24.6	1.840	
	Z	-	0.7	24.2	1.428	
	Y	-	0.6	23.4	1.213	
	NB921	-	0.6	24.3	1.345	
UKIDSS-DXS						8.87
	J	0.938	0.8	23.2	0.797	
	K	1.900	0.9	22.7	0.340	
SWIRE						9.32
	3.6 μm	2.788	1.66	23.4	0.184	
	4.5 μm	3.255	1.72	22.9	0.139	
	5.8 μm	3.743	1.88	21.2	0.106	
	8 μm	4.372	1.98	21.3	0.075	
SERVS						2.39
	3.6 μm	2.788	1.66	24.1	0.184	
	4.5 μm	3.255	1.72	24.1	0.139	
Lockman Hole						
GALEX-DIS						15.53
	FUV	-	4.2	25.43	28.637	
	NUV	-	5.3	25.75	8.675	

continued . . .

. . . continued

Field/Survey	Band	Vega-AB [mag]	PSF [arcsec]	3σ depth [mag]	$A_{band}/E(B-V)$	Area [deg^2]
SpARCS						13.32
	u	-	1.06	25.5	4.595	
	g	-	1.13	25.8	3.619	
	r	-	0.76	25.1	2.540	
	z	-	0.69	23.5	1.444	
RCSLenS						16.63
	g	-	0.78	25.1	3.619	
	r	-	0.68	24.8	2.540	
	i	-	0.60	23.8	1.898	
	z	-	0.65	22.4	1.444	
UKIDSS						8.16
	J	0.938	0.76	23.4	0.797	
	K	1.900	0.88	22.8	0.340	
SWIRE						10.95
	3.6 μm	2.788	1.66	23.4	0.184	
	4.5 μm	3.255	1.72	22.9	0.139	
	5.8 μm	3.743	1.88	21.2	0.106	
	8 μm	4.372	1.98	21.2	0.075	
SERVS						5.58

continued . . .

. . . continued

Field/Survey	Band	Vega-AB [mag]	PSF [arcsec]	3σ depth [mag]	$A_{band}/E(B-V)$	Area [deg^2]
	3.6 μm	2.788	1.66	24.1	0.184	
	4.5 μm	3.255	1.72	24.0	0.139	

2.2.3.3 Selected survey area for Lockman Hole

In this thesis (see Chapter 3), for radio-optical cross-matching, we use the overlapping area between the SpARCS r-band and the SWIRE survey which covers ≈ 10.73 deg^2. As such, Lockman Hole is the largest deep field released with respect to the accompanying multi-wavelength data. This overlap area in Lockman Hole is also illustrated by the blue shaded region in the footprint in Fig. 2.2. At the largest extent of this selected area from the radio field centre, the radio primary beam correction factor is ~0.42.

2.2.4 Multi-wavelength data in Boötes

In Boötes, we make use of existing PSF matched I-band and 4.5 μm band catalogues (Brown et al., 2007, 2008) built using imaging data from the NOAO Deep Wide Field Survey (NDWFS; Jannuzi & Dey 1999) and follow-up imaging campaigns in other filters. This catalogue contains 15 multi-wavelength bands (0.14 μm–24 μm) from different surveys. Fig. 2.3 shows the footprint of the key surveys covering the Boötes field. Typical 3σ depths estimated using variance from random apertures for each filter are listed in Table 2.3.

In summary, deep optical photometry in the B_W, R, and I filters comes from NDWFS (Jannuzi & Dey, 1999). Photometry in the NUV and FUV comes from GALEX surveys. Additional z-band data covering the full NDWFS field comes from the zBoötes survey (Cool, 2007) taken with the Bok 90Prime imager, and additional data from the Subaru z-band (PI: Yen-Ting, Lin). Additional optical imaging in the $U_{\rm spec}$ and the Y bands comes from the Large Binocular Telescope (Bian et al., 2013). NIR data in J, H, and K_s comes from Gonzalez et al. (2010). In the MIR, *Spitzer* surveyed ~10 deg^2 of the NDWFS field at 3.6, 4.5, 5.8 and 8.0 μm across 5 epochs. Primarily, the data consist of 4 epochs from the *Spitzer* Deep Wide Field Survey (SDWFS; Ashby et al. 2009), a subset of which is the IRAC Shallow Survey (Eisenhardt et al., 2004) and, the fifth epoch from the Decadal IRAC Boötes Survey (M.L.N. Ashby PI, PID 10088).

The full details of the data used and the catalogue generation process are provided in Brown et al. (2007, 2008). In summary, images in all filters were first moved on to a common pixel scale and then sources detected using SExtractor (Bertin & Arnouts, 1996). Forced photometry was then performed on optical-NIR filters smoothed to a common PSF. The common PSF was chosen to be a Moffat profile with $\beta = 2.5$ and a FWHM of 1.35″ (B_W, R, I, Y, H and K), a FWHM of 1.6″ (u, z, J), and a FWHM of 0.68″ for the Subaru z-band. Aperture corrections based on the chosen Moffat profile

Table 2.3 *Filter dependent extinction correction factors per unit reddening,* $A_{band}/E(B-V)$ *and* 3σ *depths from* $3''$ *apertures for Boötes.*

Band	$A_{band}/E(B-V)$	3σ depth [mag]
FUV	28.637	26.3
NUV	8.675	26.7
u	4.828	25.9
Bw	4.216	26.2
R	2.376	25.2
I	1.697	24.6
z	1.423	23.4
z_Subaru	1.379	24.3
y	1.185	23.4
J	0.811	23.1
H	0.515	22.5
K	0.338	20.2
Ks	0.348	21.8
3.6 μm	0.184	23.3
4.5 μm	0.139	23.1
5.8 μm	0.105	21.6
8 μm	0.074	21.6

were then applied to account for the different FWHM choices in PSF smoothing.

In Boötes, the FIR data from HerMES and MIPS were obtained by a similar method to ELAIS-N1 and Lockman Hole (see McCheyne et al. subm), and form a new addition to the existing catalogues of Brown et al. (2007, 2008).

In this chapter, subsequent analysis is performed for the overlap of the NDWFS and SDWFS datasets, covering ~ 9.5 deg^2, as shown in Fig. 2.3. This area was chosen as the largest area with coverage in most of the optical-IR bands. At the largest extent of the selected area from the radio field centre, the primary beam correction factor is $\sim$0.39.

2.3 Creation of multi-wavelength catalogues

This section describes the creation of the pixel-matched images and the generation of the new multi-wavelength catalogues in both ELAIS-N1 and Lockman Hole. The Boötes field already possesses PSF-matched forced photometry catalogues created using an I-band and a 4.5 μm band detected catalogue (Brown et al., 2007, 2008). To generate a similar multi-wavelength catalogue in Boötes as the other two fields for radio-optical

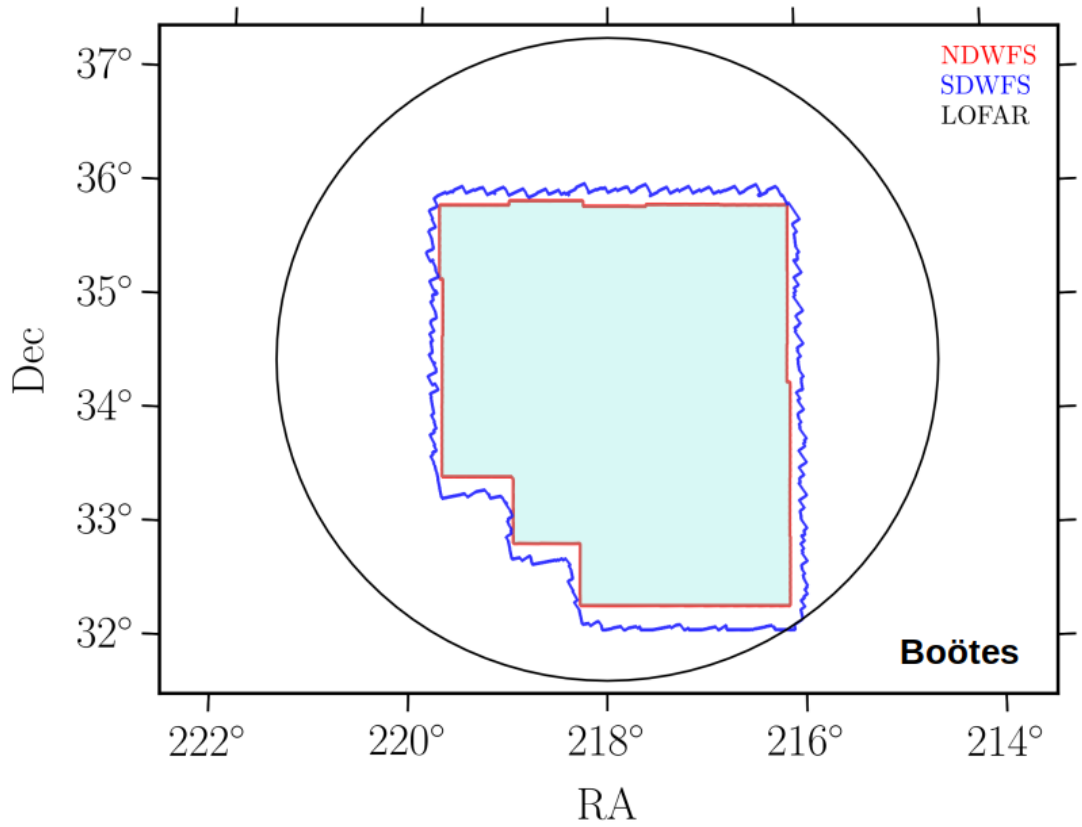

Figure 2.3 *Footprint (north up, east left) for the Boötes field showing the coverage of multi-wavelength data from various surveys in optical and IR bands described in Sect. 2.2.4. The LOFAR radio coverage is also shown in black. The shaded light blue region shows the selected area of overlap that is used for the radio-optical cross-match in Chapter 3 for Boötes (~9.5deg^2), as described in Sect. 2.2.4, with slightly reduced area after bright-star masking.*

cross-matching, we apply only the final steps of our catalogue generation process, namely, the masking around stars (see Sect. 2.3.4.3), the merging of the I-band and 4.5 μm detected catalogues (see Sect. 2.3.4.4), and the Galactic extinction corrections.

2.3.1 Creation of the pixel-matched images

The images from different instruments had different pixel scales and therefore all of the images needed to be re-gridded (resampled) onto the same pixel scale to perform matched aperture photometry across all filters. Observations in most filters consisted of many overlapping exposures of the total area. We obtained reduced images from survey archives for all filters and used SWarp (Bertin et al., 2002) to both resample the individual images in each filter to a common pixel scale of 0.2″ per pixel and then to combine (co-add) these resampled images to make a single large mosaic in each filter. We make no attempt to perform point-spread function (PSF) homogenisation of these observations; instead, we account for the varying PSF in each filter by performing aperture corrections (see Sect. 2.3.3.1).

Changes to the astrometric projection or the photometric calibration were performed during the resampling process by SWarp. During this step, the contribution to the flux from the background/sky is subtracted before the resampling and co-addition process to avoid artefacts resulting from image combination. The flux scale of the images was also adjusted using each input frame's zero-point magnitude, exposure time and any Vega-AB conversion factors (see Table 2.2) to shift the zero-point magnitude of all the images to 30 mag (in the AB system). The resampled images in each filter were then co-added in a 'weighted' manner to take into account the relative exposure time/noise per pixel in multiple input frames and in overlapping frames. Table 2.2 also lists the typical PSF full-width half-maximum (FWHM) for each filter in ELAIS-N1 and Lockman Hole. We compared photometry in fixed apertures for given sources in both the resampled frames and the final mosaics to ensure that the photometry is consistent with the original images.

2.3.2 Source detection

Source detection is performed using SExtractor (Bertin & Arnouts, 1996). We ran SExtractor in 'dual-mode', using a deep image for detecting sources and then performing photometry using these detections on all of the filters. To produce as

complete a catalogue as possible, we built our deep detection image by using SWarp to create deep χ^2 images (Szalay et al., 1999) by combining observations from multiple filters; this is done by computing the χ^2 statistic (using the pixel values and pixel noise) over multiple filters in a pixel matched manner. As stated by Szalay et al. (1999), the χ^2 image represents the probability that a pixel is drawn from the background sky distribution; pixels above a sufficiently high threshold on the χ^2 image are thus unlikely to be drawn from the sky, and hence can be considered as a source.

In practice, due to the significantly worse angular resolution of the *Spitzer* data, we built two χ^2 images, one using optical and NIR filters and a separate χ^2 image using only the *Spitzer*-IRAC data. In ELAIS-N1, the optical χ^2 image was created using SpARCS-*u*, PS1-*griz*, and UKIDSS-DXS-*JK* filters (the PS1 *y*-band is not included due to its shallower depth and lower sensitivity relative to the adjacent filters). In Lockman Hole, we used SpARCS-*ugrz*, RCSLenS-*i*, and UKIDSS-DXS-*JK* filters. The *Spitzer* χ^2 images in both ELAIS-N1 and Lockman Hole are built from the IRAC 3.6 μm and 4.5 μm bands from both SWIRE and SERVS. The longer wavelength *Spitzer* data are not included in the χ^2 detection images due to a further decrease in angular resolution.

The key detection parameters in SExtractor are the ones concerning deblending, the detection threshold, and minimum detection area. These key parameters are listed in Table 2.4 for the optical-NIR and *Spitzer* χ^2 images. We fine-tuned these parameters for each χ^2 image by adjusting their values and inspecting the resulting catalogue overlaid on the χ^2 images.

Although more sophisticated tools exist for performing multi-band photometry that allow model-fitting of detections (e.g. The Tractor; Lang et al. 2016; Nyland et al. 2017 and T-PHOT; Merlin et al. 2015, 2016), SExtractor is a flexible and easily scalable tool that allows both source detection and forced, matched aperture photometry to be performed in a practical and robust manner over ~30 bands across > 18 deg^2 for the two fields. Moreover, the use of SExtractor for ELAIS-N1 and Lockman Hole also provides consistency with the method that was adopted to generate the existing Boötes catalogues.

2.3.3 Photometric measurements

Running SExtractor in dual mode, we measure fluxes in all of the filters using both the optical-NIR and *Spitzer* χ^2 images; this includes *Spitzer* fluxes from sources detected on the optical-NIR χ^2 images and vice versa. We extract fluxes from a wide variety

Table 2.4 *Key* SExtractor *detection and deblending parameters used for the optical-NIR and Spitzer χ^2 detection images in ELAIS-N1 and Lockman Hole.*

Parameter	Value	
	Optical-NIR χ^2	*Spitzer* χ^2
DETECT_MINAREA	8	10
DETECT_THRESH	1	2.5
DEBLEND_NTHRESH	64	64
DEBLEND_MINCONT	0.0001	0.001

of aperture sizes, specifically 1″ – 7″ diameter (in 1″ steps) and also 10″ diameter apertures in each filter.

2.3.3.1 Aperture and Galactic extinction corrections

Fluxes from fixed apertures were corrected to total fluxes using aperture corrections based on the curve of growth estimated from our full range of aperture measurements (assuming all of the flux from a source is contained within the 10″ aperture). We compute median correction factors for each aperture size using relatively isolated (>5″ from nearest neighbour) sources of moderate magnitude (e.g. i-band $19 < i < 20.5$), chosen to have high sky density but also sufficient signal-to-noise ratio (S/N) even in the larger apertures. These sources are typical of moderately distant galaxies, with this selection driven by the primary scientific aims of the LOFAR surveys. It is important to note that the resulting correction factors are found to be not sensitive to the exact choice of the magnitude employed in selecting sources used for calibrating the aperture corrections. The full list of aperture corrections are provided in Table 2.5. In addition, we also provide in Table 2.5, a list of aperture corrections calibrated based on stars with $18 < \mathrm{Gmag} < 20$ in GAIA Data Release 2 (GAIA DR2; Gaia Collaboration et al., 2016, 2018; Riello et al., 2018; Evans et al., 2018). In deriving the aperture corrections, we assume that the PSF variations between different images of a given filter are insignificant compared to the PSF variation across different filters (see discussion in Sect. 2.3.5).

Galactic extinction corrections are computed at the position of each object using the map of Schlegel et al. (1998) [2]. We provide a column of $E(B-V)$ reddening values computed from Schlegel et al. (1998) for each source, which is then multiplied by the filter dependent factor (listed in Table 2.2 for ELAIS-N1 and Lockman Hole; 2.3 provides the corresponding values for Boötes) derived from the filter transmission curve

[2]Performed using the dustmaps package (Green, 2018) for Python.

and the Milky Way extinction curve (Fitzpatrick, 1999). For all fields, we recommend using the 'band_mag_corr' columns which are corrected for aperture and extinction using the 3″ for optical-NIR bands and the 4″ for the *Spitzer* IRAC bands (see discussion in Sect. 2.4). The raw photometry in any aperture can be corrected for both aperture and extinction as follows

$$\begin{aligned} \text{band\_mag\_corr} = \text{MAG\_APER\_band\_ap} - 2.5 \log F_{\text{band,ap}} \\ - \text{EBV} \times A_{\text{band}}/E(B-V) \end{aligned} \tag{2.1}$$

where band is the filter, ap is the aperture size, $F_{\text{band,ap}}$ is the aperture correction factor for a given band and aperture. We apply the values listed in Table 2.5 derived using based on sources typical of moderately distant galaxies, as described above; the stellar based aperture corrections, which are not applied to the catalogues, are nonetheless provided in Table 2.5. EBV is the reddening value computed based on source position and the Schlegel et al. (1998) dust map, and, $A_{\text{band}}/\text{E(B}-\text{V)}$ are the filter dependent extinction factors listed in Table 2.2 and Table 2.3. Equation 2.1 can also be used for any other band or aperture size combination to derive an aperture and Galactic extinction corrected magnitude.

2.3.3.2 Computation of photometric errors

We find that the flux uncertainties reported by SExtractor typically underestimate the total uncertainties. This is a well-known issue and occurs as SExtractor only takes into account photon and detector noise, and does not account for background subtraction errors or correlated noise arising from image combination. We estimate the additional flux error term using the same method as that used by Bielby et al. (2012) and Laigle et al. (2016). Firstly, fluxes were measured in random isolated apertures (with the same sized apertures as our flux measurements) on a background-subtracted image. Then, to remove the contribution from sources to the flux in the random apertures, an iterative sigma clipping of the measured flux distribution is performed. Finally, the standard deviation of the clipped distribution is taken to be the additional contribution to the flux uncertainties from correlated noise and background subtraction errors, and is then added in quadrature with the uncertainties reported by SExtractor on a source-by-source basis to compute the total photometric errors. The magnitude errors were then updated accordingly. The 3σ magnitude depths estimated from the variance of empty, source free 3″ apertures in each filter are listed in Table 2.2.

Table 2.5 *List of aperture corrections derived for each filter in ELAIS-N1 and Lockman Hole, using the method described in Sect. 2.3.3.1. The raw aperture fluxes released in the catalogues should be divided by the values listed to correct the fluxes for aperture effects. The galaxies-based aperture corrections are appropriate for moderately distant galaxies and are applied to both the value-added catalogues and multi-wavelength catalogues in our recommended aperture sizes. Stellar-based corrections (not applied to catalogues) are derived using stars in GAIA-DR2 with 18 < Gmag < 20.*

Survey-Filter	Galaxies-based corrections							Stellar-based corrections						
	Fraction of flux in aperture (arcsec)							Fraction of flux in aperture (arcsec)						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
ELAIS-N1														
SpARCS-u	0.39	0.76	0.9	0.95	0.97	0.98	0.99	0.41	0.79	0.91	0.96	0.98	0.99	1.0
PS1-g	0.25	0.6	0.78	0.87	0.92	0.95	0.97	0.27	0.62	0.79	0.88	0.93	0.96	0.98
PS1-r	0.25	0.6	0.77	0.85	0.91	0.94	0.97	0.3	0.64	0.8	0.88	0.92	0.95	0.97
PS1-i	0.23	0.55	0.74	0.83	0.89	0.93	0.96	0.31	0.65	0.8	0.87	0.92	0.95	0.97
PS1-z	0.23	0.54	0.73	0.84	0.9	0.94	0.96	0.33	0.66	0.81	0.88	0.92	0.95	0.97
PS1-y	0.2	0.49	0.68	0.79	0.86	0.91	0.95	0.29	0.59	0.75	0.84	0.89	0.93	0.95
HSC-g	0.55	0.83	0.91	0.95	0.97	0.98	0.99	0.56	0.84	0.92	0.95	0.97	0.98	0.99
HSC-r	0.46	0.8	0.9	0.94	0.96	0.97	0.98	0.49	0.82	0.91	0.95	0.97	0.98	0.99
HSC-i	0.61	0.87	0.92	0.95	0.97	0.98	0.99	0.64	0.9	0.94	0.96	0.97	0.98	0.99
HSC-z	0.37	0.71	0.85	0.91	0.95	0.97	0.98	0.47	0.8	0.91	0.95	0.97	0.98	0.99
HSC-y	0.49	0.78	0.87	0.92	0.95	0.97	0.98	0.63	0.86	0.92	0.94	0.96	0.97	0.98
HSC-NB921	0.35	0.69	0.84	0.91	0.94	0.97	0.98	0.63	0.88	0.94	0.97	0.98	0.99	0.99
J	0.39	0.78	0.92	0.96	0.98	0.99	1.0	0.45	0.82	0.93	0.97	0.99	0.99	1.0

continued . . .

continued . . .

K	0.32	0.7	0.88	0.95	0.98	0.99	1.0	0.42	0.79	0.92	0.96	0.98	1.0	1.0
IRAC-3.6 μm	0.1	0.34	0.56	0.73	0.83	0.89	0.93	0.13	0.4	0.64	0.79	0.88	0.93	0.95
IRAC-4.5 μm	0.11	0.34	0.56	0.72	0.84	0.9	0.94	0.13	0.4	0.63	0.78	0.88	0.94	0.96
IRAC-5.8 μm	0.08	0.26	0.45	0.62	0.74	0.84	0.91	0.11	0.33	0.53	0.68	0.8	0.89	0.95
IRAC-8.0 μm	0.07	0.25	0.43	0.58	0.68	0.77	0.85	0.09	0.3	0.5	0.64	0.73	0.82	0.9
Lockman Hole														
SpARCS-u	0.31	0.69	0.87	0.94	0.97	0.98	0.99	0.35	0.73	0.88	0.95	0.97	0.99	0.99
SpARCS-g	0.21	0.54	0.76	0.87	0.93	0.96	0.98	0.33	0.72	0.88	0.94	0.97	0.98	0.99
SpARCS-r	0.24	0.57	0.78	0.89	0.94	0.97	0.98	0.49	0.82	0.92	0.96	0.97	0.98	0.99
SpARCS-z	0.29	0.64	0.83	0.91	0.96	0.98	0.99	0.52	0.82	0.91	0.95	0.97	0.98	0.99
RCSLenS-g	0.23	0.57	0.78	0.88	0.93	0.96	0.98	0.48	0.83	0.93	0.96	0.98	0.99	0.99
RCSLenS-r	0.25	0.58	0.78	0.89	0.94	0.97	0.98	0.55	0.86	0.94	0.97	0.98	0.99	0.99
RCSLenS-i	0.33	0.68	0.85	0.93	0.97	0.99	0.99	0.61	0.89	0.95	0.97	0.99	0.99	1.0
RCSLenS-z	0.28	0.62	0.8	0.89	0.94	0.97	0.99	0.56	0.84	0.92	0.96	0.97	0.98	0.99
J	0.29	0.67	0.86	0.95	0.98	1.0	1.0	0.46	0.82	0.93	0.97	0.98	0.99	1.0
K	0.32	0.7	0.88	0.96	0.99	1.01	1.01	0.47	0.82	0.92	0.96	0.98	0.99	1.0
IRAC-3.6 μm	0.1	0.33	0.56	0.73	0.83	0.89	0.93	0.13	0.4	0.63	0.79	0.88	0.92	0.95
IRAC-4.5 μm	0.11	0.34	0.56	0.72	0.83	0.9	0.94	0.13	0.4	0.63	0.77	0.87	0.93	0.96
IRAC-5.8 μm	0.08	0.27	0.46	0.61	0.74	0.84	0.91	0.1	0.32	0.52	0.67	0.78	0.88	0.94
IRAC-8.0 μm	0.07	0.25	0.44	0.58	0.68	0.77	0.85	0.09	0.29	0.49	0.62	0.7	0.79	0.88

2.3.4 Catalogue cleaning and merging

In this section, we describe the key steps used to clean the ELAIS-N1 and Lockman Hole catalogues of spurious sources and low-significance detections. We then discuss masking around bright stars and merging of the optical-NIR and *Spitzer* detected catalogues in all three fields.

2.3.4.1 Cross-talk removal

Cross-talks are non-astronomical artefacts that appear on the UKIDSS (J or K) images at fixed offsets from bright stars due to readout patterns; these may appear in the χ^2 detection image. The cross-talks appear at 128, 256, and 384 pixels in both the RA and Dec directions and in the native pixel scale of UKIDSS images, this corresponds to 51″, 102″, and 153″, respectively. Cross-talks may have extreme colours due to their non-astrophysical nature and, therefore, we use the flux measurements (or lack thereof) in the optical and NIR filters to identify and remove cross-talks from the catalogue. Specifically, we searched for catalogued detections within 2″ of the expected cross-talk positions to identify (and remove) detections that have either extreme optical-NIR colours (i.e. (i - K) > 4) or, have low significance (S/N < 3) measurements in multiple optical bands and a NIR magnitude that is more than 6 mag fainter than the 'host' star. These criteria were confirmed by visual inspection of detections that were removed and retained (i.e. sources present at the expected cross-talk positions but not satisfying other criteria above).

Confidence in the selection method applied comes from investigating the distribution of sources removed and retained around bright stars. Fig. 2.4 shows the radial distribution of sources detected around bright stars ($m_K < 17$ mag) out to $\approx 120''$ in the un-cleaned catalogue (blue) along with the distribution of sources after the above cross-talk removal criteria is applied to the catalogue (red). As evident from this plot, the narrow peaks observed at the radii expected of the cross-talk artefacts (i.e. 51″ and 102″) were eliminated, without the removal of genuine astrophysical sources.

2.3.4.2 Cleaning low significance detections

In the final cleaning step, we removed any sources which have a S/N less than 3 in all apertures of all filters. Such low significance detections may have a S/N < 3 in each of the single band images but could end up in the catalogue due to the use of χ^2

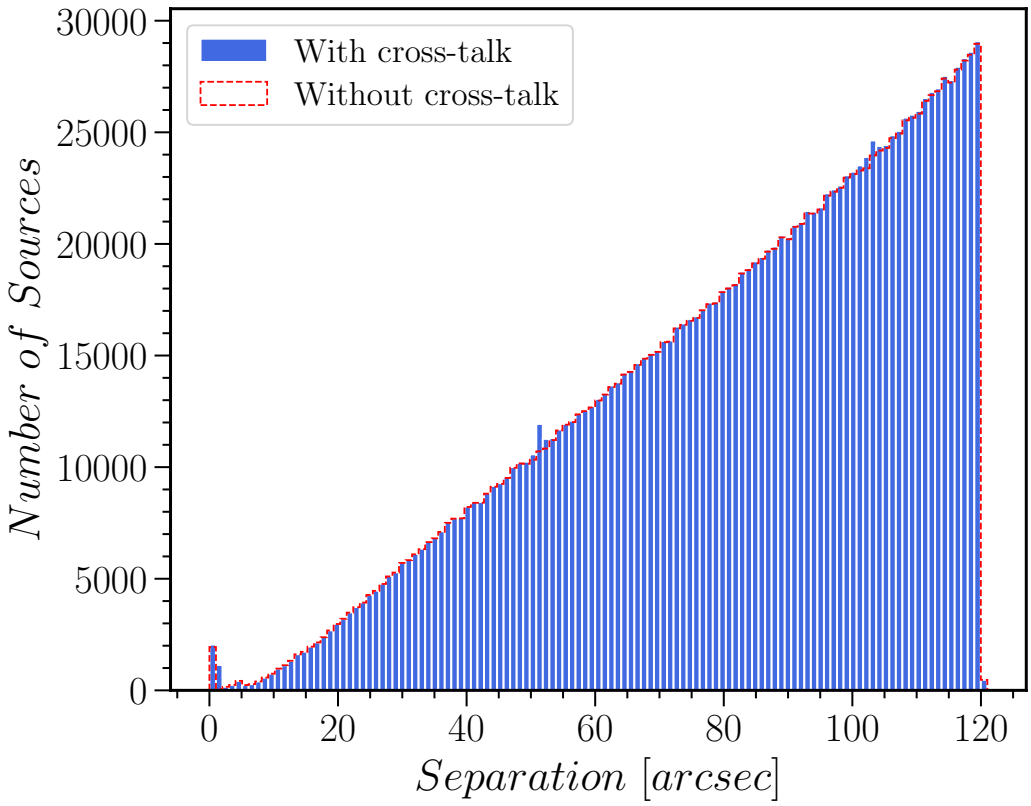

Figure 2.4 *Radial plot showing the number of sources detected around bright stars ($m_K < 17$ mag) as a function of separation, out to $\approx 120''$. The blue bars show the distribution of sources before we cleaned for cross-talks. The presence of these spurious cross-talks is evident in the spikes in the distribution at $\sim 51''$ and $\sim 102''$. The red dashed line corresponds to the distribution of sources after we applied our cross-talk removal criteria, which results in the removal of the artefacts at the expected positions, without the removal of genuine sources.*

detection images which combines the signal from multiple bands. Although probably genuine, such sources are of limited scientific value as none of their flux measurements are sufficiently reliable. This step removes ~15% and 27% of the sources from the ELAIS-N1 and Lockman Hole catalogues, respectively. The higher fraction of low-significance sources removed in Lockman Hole are largely located near the edge of the field where the χ^2 image contains few filters with variable relative depth: ELAIS-N1 possesses both deeper optical data, and also coverage from most filters across a higher fraction of the total area of the field.

Table 2.6 *Radii masked around bright stars (in arcsec) as a function of their GAIA DR2 G-band magnitude using both the optical-NIR and Spitzer detected catalogues.*

G-magnitude [mag]	ELAIS-N1 [arcsec]		Lockman Hole [arcsec]		Boötes [arcsec]	
	Optical-NIR	*Spitzer*	Optical-NIR	*Spitzer*	Optical-NIR	*Spitzer*
$16.0 < \text{Gmag} \leq 16.5$	15	12	13	10	25	21
$15.0 < \text{Gmag} \leq 16.0$	18	14	20	15	31	26
$14.0 < \text{Gmag} \leq 15.0$	23	16	23	18	39	33
$13.0 < \text{Gmag} \leq 14.0$	27	22	35	20	52	44
$12.0 < \text{Gmag} \leq 13.0$	33	27	40	25	62	52
$11.0 < \text{Gmag} \leq 12.0$	40	31	55	40	83	70
$10.0 < \text{Gmag} \leq 11.0$	55	50	80	45	95	80
$\text{Gmag} \leq 10.0$	65	60	130	55	101	85

Table 2.7 *Area flagged by the star masks. Two star mask images are generated using the Spitzer- and optical-detected catalogues (see Sect 2.3.4.3).*

Field	Spitzer-mask area [deg^2]	Optical-mask area [deg^2]
ELAIS-N1	0.40	0.61
Lockman Hole	0.31	0.85
Boötes	0.87	1.18

2.3.4.3 Masking sources near bright stars

Next, we created a mask image by masking regions around bright stars and flagging sources within these regions in our catalogue, in each of the ELAIS-N1, Lockman Hole, and Boötes fields. The rationale behind this is twofold. Firstly, in regions around stars, SExtractor may detect additional spurious sources or miss other sources nearby or behind the star in sky projection. Secondly, the photometry of objects near bright stars will not be reliable. Masking such regions therefore allows scientific analysis to be restricted to areas where there is reliable coverage. This is crucial for some science cases, for example, clustering analysis.

To select the stars around which regions must be masked, we cross-matched our catalogue to stars with Gmag < 16.5 mag in GAIA DR2. Then, we split the stars into narrow magnitude bins and select the radius to mask around stars in each bin by using a plot of the sky density of the sources as a function of the radius from the star. An appropriate radius was chosen where neither the 'holes' in the detections nor a 'ring' of additional spurious sources near the star were affecting the detections (e.g. Coupon et al. 2018). We validated the choice of the magnitude dependent radii using careful visual inspection, with the values chosen listed in Table 2.6.

Detections around stars are affected less by this issue in the *Spitzer* χ^2 image (and catalogue), allowing us to mask a smaller area. Therefore, in practice, we create two such masks, one for the optical-NIR χ^2 image (a conservative mask) and, one for the *Spitzer* χ^2 image (an optimistic mask), with both masks being applied to both the optical-NIR and the *Spitzer* detected catalogues. Using detections from the *Spitzer*-detected catalogue masks a smaller area around stars, recovering some genuine sources detected in the *Spitzer* image that are not affected by source extraction biases. However, photometry of these sources in the optical-NIR images may be less reliable due to stellar emission, and moreover, any optical-only detected sources may be missing from this extra recovered area. The area masked in each field using both the optical-NIR (conservative) and the *Spitzer* (optimistic) mask is given in Table 2.7. For convenience, we include a flag (FLAG_CLEAN) column in both the multi-wavelength catalogues and the radio cross-match catalogues which indicates if a source is within the two masked areas. For readers requiring a clean homogeneous catalogue, we recommend using FLAG_CLEAN = 1 to select sources that are not in either the optical or *Spitzer* star mask region. Instead, if the largest sample of sources is required, with photometry not critical, we recommend using FLAG_CLEAN $\neq$ 3 to exclude only sources in the smaller *Spitzer* star mask. We note that this should be used in conjunction with FLAG_OVERLAP to select sources with reliable

photometry in the majority of the bands (see Sect. 2.4 and Table 2.8).

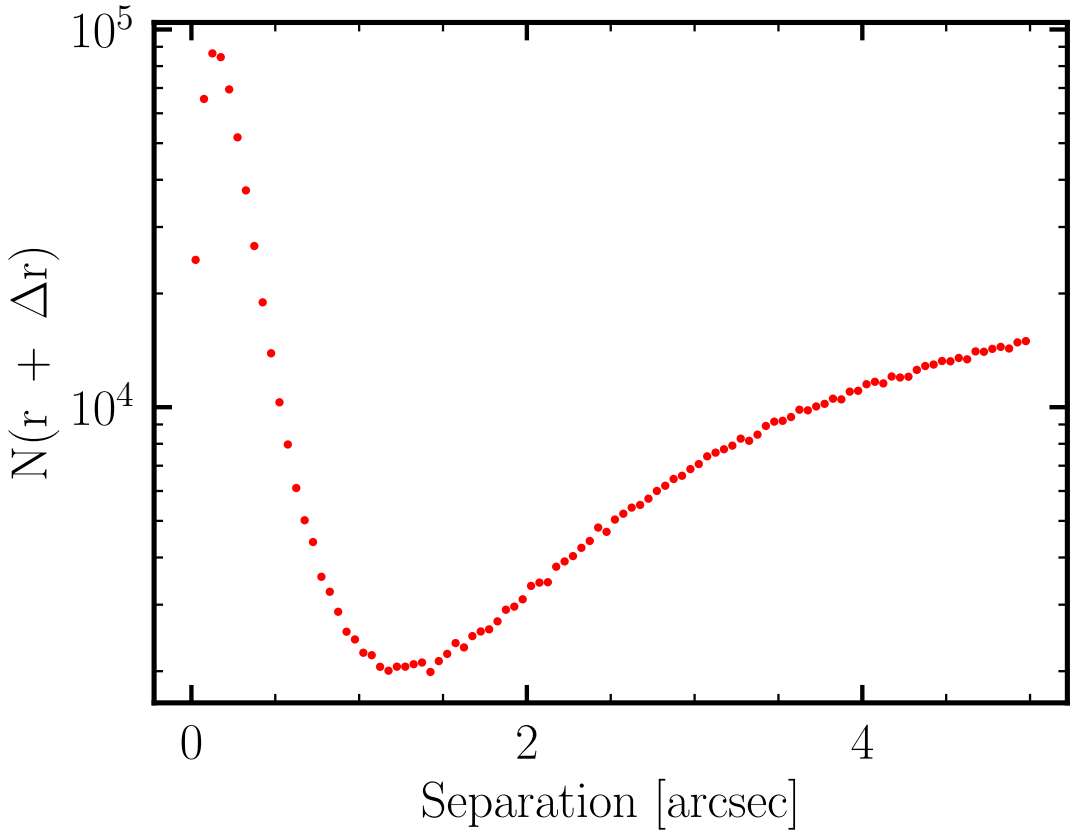

Figure 2.5 *Selection of the cross-match radius to merge the optical-NIR and Spitzer-detected χ^2 catalogues in ELAIS-N1 and Lockman Hole. The plot shows the number of cross-matches between r and r + Δr as the search radius is increased. The chosen cross-match radius of 1.5″ is based on the point where the number of genuine cross-matches decreases from its peak and the number of random matches starts to increase; this was further validated by visual inspection.*

2.3.4.4 Merging optical and *Spitzer* catalogues

After applying our cleaning steps, the optical-NIR detected catalogue was merged with the *Spitzer* detected catalogue in each of the three fields. Many of the *Spitzer*-detected sources, especially those with blue colours, will already be present in the optical-NIR catalogue. We therefore merge the two catalogues by appending '*Spitzer*-only' sources to the optical-NIR catalogue. We define a source as '*Spitzer*-only' if its nearest neighbour in the optical-NIR catalogue is more than 1.5″ away. Fig. 2.5 shows the number of cross-matches within r and $r + \Delta r$ between the optical-NIR and *Spitzer* based χ^2 catalogues for increasing search radii. As evident from the plot, the adopted

cross-match radius was chosen to be the point where the number of genuine matches decreases rapidly after its peak at low separations, and the number of random matches starts to increase. The choice of this cross-match radius was also validated via visual inspection of the resulting '*Spitzer*-only' sources. In ELAIS-N1, we find that $\sim 15\%$ of *Spitzer* detected sources are '*Spitzer*-only' sources, which make up 4.6% of the total number of sources in the final multi-wavelength merged catalogue.

2.3.5 Catalogue validation

To validate the catalogues generated, we compare our astrometry and photometry to publicly available catalogues in ELAIS-N1 and Lockman Hole.

To estimate the astrometric accuracy of our mosaics, we compared the median scatter in the RA and Declination between catalogues derived from individual mosaics. We find median astrometric offsets between $0.07'' - 0.13''$, all of which occur at scales smaller than the pixel size of $0.2''$.

In Fig. 2.6, we plot the K-band selected source counts (per square arcsecond per 0.5 magnitude) from the ELAIS-N1 and Lockman Hole χ^2 catalogues, along with the Ks-band selected source counts from the merged catalogue in Boötes. Number counts from the COSMOS deep area of Laigle et al. (2016) in the Ks-band are also shown. The number counts from the UKIDSS DR10 catalogues in both ELAIS-N1 and Lockman Hole are also shown for comparison with each field. Vertical dashed lines indicate the 3- and 5-σ limiting magnitudes in ELAIS-N1. In all cases, contribution from foreground stars are removed by performing a cross-match to GAIA DR2 stars with Gmag < 19 mag. This plot shows that there is excellent agreement between our χ^2 and the UKIDSS catalogue within each field. We also note that the difference between the ELAIS-N1 and the Lockman Hole number counts seen in our χ^2 catalogues, especially at bright magnitudes, is also seen in the UKIDSS DR10 catalogues, suggesting that this is likely due to large scale structure between the two fields. This difference is also seen with the Laigle et al. (2016) data, which agrees well with the ELAIS-N1 data (both our χ^2 and UKIDSS DR10 catalogues) but not with the other fields at $K < 20$ mag, which is likely due to large scale structure. The plot also shows that our catalogues, especially in ELAIS-N1, reach a slightly higher completeness than the UKIDSS DR10 catalogue at S/N of $3 - 5$ due to the use of χ^2 detection images.

For the optical filters, we have compared our aperture corrected magnitudes with model magnitudes from SDSS DR13 (Albareti et al., 2017) where the coverage overlaps and

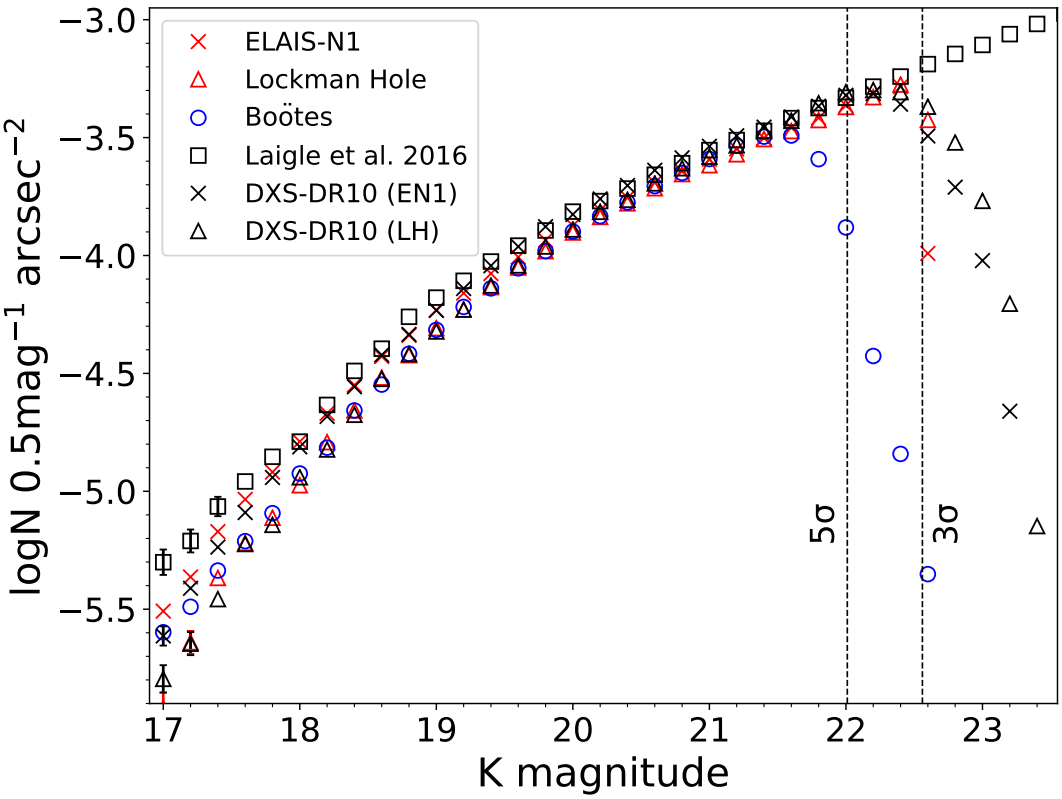

Figure 2.6 *K-band (3σ) selected source counts (per square arcsecond per 0.5 magnitude) in ELAIS-N1 (red crosses) and Lockman Hole (red triangles) from the χ^2 catalogue. In Boötes, we show source counts from the merged catalogue (blue circles). We also show the galaxy counts from the COSMOS deep area taken from Laigle et al. (2016) for comparison (black squares). Additionally, the galaxy counts from the UKIDSS-DXS DR10 catalogues (Lawrence et al., 2007) in both ELAIS-N1 (black crosses) and Lockman Hole (black triangles) are shown. In all cases, we have attempted to remove the contribution from foreground stars via a cross-match to GAIA DR2 catalogues in each field. Vertical lines show the 3- and 5-σ magnitude depths in ELAIS-N1 (dashed lines) estimated from random, source free 3″ diameter apertures. Poissonian error bars are shown only where they are larger than the symbol size, but there may be other cosmic-variance related errors that dominate the uncertainties.*

find very good agreement to a few percent level, well below the typical photometric uncertainties. We show a typical example for the PS1 *r*-band in ELAIS-N1 in Fig. 2.7 (*top left*) which illustrates the median magnitude difference in cells of 0.06 deg^2. This is calculated by comparing our photometry for relatively bright ($r < 21$ mag) sources with SDSS model magnitudes, accounting for the small differences in the PS1 and SDSS filters using colour terms estimated from Finkbeiner et al. (2016). There is good agreement with SDSS for most of the PS1 footprint, however, the PS1 *r*-band

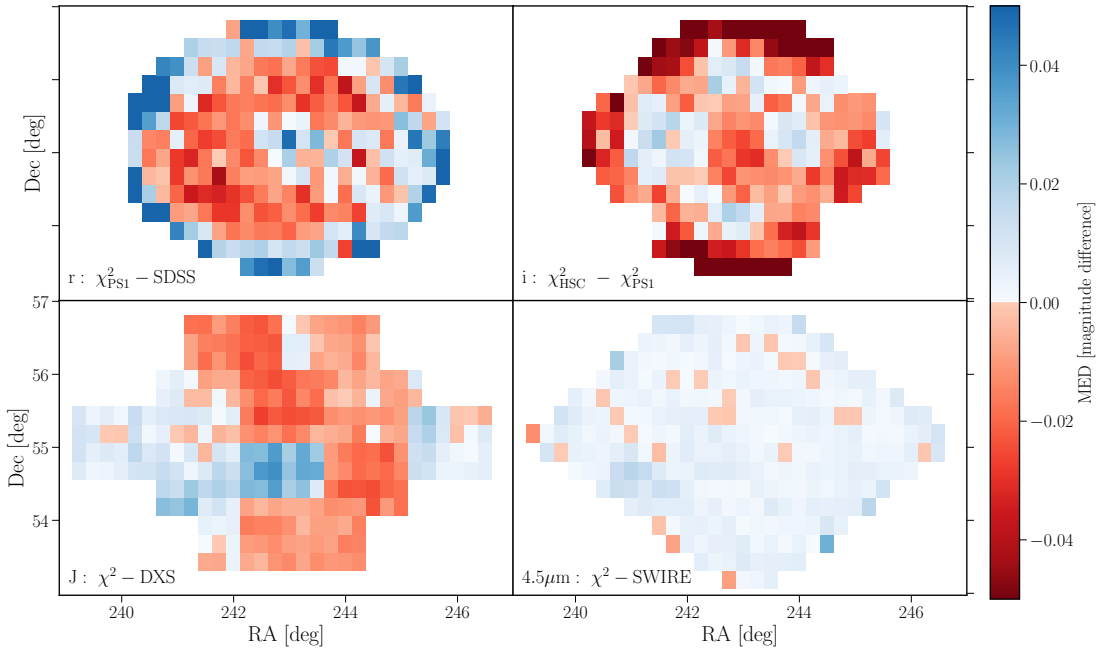

Figure 2.7 *Comparison of photometry in the r, i, J and 4.5 µm bands in ELAIS-N1. The colour-map shows the median magnitude difference computed over cells of 0.06 deg^2 between our χ^2 and publicly available catalogues from SDSS DR12, DXS DR10, and the SWIRE survey for the r, J and 4.5 µm bands, respectively. For the i-band, we compare the photometry between PS1 and HSC within our χ^2 catalogue. We use aperture corrected magnitudes based on the 3 arcsec aperture for optical-NIR bands and 4 arcsec for the 4.5 µm band. There is excellent agreement between our χ^2 and publicly released catalogues, with differences in optical bands likely driven by zero-point calibration of individual PS1 chips.*

magnitudes are too faint by 5-10% near the edge of the PS1 footprint. We find that this trend, which is observed across all PS1 filters (albeit sometimes with a smaller offset, or larger scatter), is likely driven by the zero-point calibration of the individual chips in the PS1, which gets fainter by up to $\lesssim 10\%$ by ~ 1.5 deg from the field centre. In Fig. 2.7 (top right) we also show a comparison in the *i*-band between HSC-*i* and PS1-*i* ($\mathrm{HSC_i}$ - $\mathrm{PS_i}$), both taken from our χ^2 catalogue, which shows good agreement across the field. The median magnitude difference gets more negative (i.e. PS1 is too faint compared to HSC) near the edges of the field by $\sim 8\%$, which is also consistent with the trend in zero-point variation discussed above. This suggests that the PanSTARRS photometry near the edge of the field can typically become more uncertain by $\lesssim 10\%$. However, it is important to note that this effect is comparable to the additional 10% flux error typically added to the photometric uncertainties before SED fitting and moreover, as this effect occurs near the edges of the PanSTARRS footprint, some of these regions will be outside our recommended multi-wavelength area, where photometry is the most reliable.

For the NIR *J* and *K* bands, we compare our aperture corrected fluxes to the UKIDSS-DXS DR10 catalogues in both of these fields and to that of 2MASS, finding excellent agreement to within 2-3%. As a typical example, the comparison between the ELAIS-N1 *J*-band and UKIDSS DR10 across the full field is shown in Fig. 2.7 (*bottom left*). There are some small systematic offsets with position across the field, driven by the varying PSF across the field between different exposures. Therefore, our assumption of a constant PSF per filter is not entirely accurate for this band; the resulting photometry is, however, affected at the $\lesssim 5\%$ level, which is much smaller than the typical additional photometric uncertainties used for photometric redshifts and SED fitting.

In the *Spitzer*-IRAC bands, we compare photometry to the public SWIRE and SERVS catalogues, finding a remarkably good agreement to within a $\pm$ 1% level (e.g. Fig. 2.7, *bottom right*).

2.4 Final multi-wavelength catalogues

The resulting multi-wavelength catalogue in ELAIS-N1 contains more than 2.1 million sources with over 1.5 million sources in the overlapping region of Pan-STARRS, UKIDSS-DXS, and *Spitzer*-SWIRE surveys that are used for the cross-match with the radio catalogue. Similarly, the multi-wavelength catalogue in Lockman Hole consists of over 3 million sources with over 1.9 million sources in the overlapping region of

SpARCS *r*-band and *Spitzer*-SWIRE coverage. Finally, the merged Boötes catalogue consists of over 2.2 million sources, with around 1.9 million sources in the coverage of the original NDWFS area. Some of the key properties of the multi-wavelength and initial PyBDSF radio catalogues are listed in Table 2.8.

For each field, we release the multi-wavelength catalogue over the full field coverage. For convenience, we include a FLAG_OVERLAP bit value for each source in both the multi-wavelength catalogues and the radio cross-matched catalogues released, which indicates which survey footprint a source falls within. In Table 2.8, we list the recommended FLAG_OVERLAP value to use for each field, to select sources that are within our selected multi-wavelength overlap area.

For ELAIS-N1 and Lockman Hole, we release the raw (uncorrected for any aperture effects or Galactic extinction) aperture fluxes and magnitudes in each filter and in addition, provide, for each filter, a flux and magnitude corrected for aperture (in our recommended aperture) and Galactic extinction. We choose the 3″ aperture fluxes for all optical-NIR bands and the 4″ aperture for all *Spitzer* IRAC bands as our recommended apertures. While the 3″ aperture may have a lower S/N than the 2″ aperture for compact objects, the fluxes will be less sensitive to PSF variations or astrometric uncertainties, resulting in more robust colours. The 4″ aperture corresponds to roughly twice the PSF FWHM of the IRAC bands, and was found by Lonsdale et al. (2003) to reduce scatter in colour magnitude diagrams for stars. These aperture sizes are therefore used in our radio-optical cross-matching and for the photometric redshift estimates (described by Duncan et al. 2021) and for the SED fitting process (described by Best et al. in prep.).

The existing Boötes catalogues have already been aperture corrected. We therefore apply Galactic extinction corrections to the 3 (for optical-NIR bands) and 4 (for IRAC bands) arcsec aperture fluxes and magnitudes, in the same way as for ELAIS-N1 and Lockman Hole, and only provide these recommended fluxes and magnitudes in the catalogues released for this field. The $E(B-V)$ values used for each source are also provided in an additional column; the filter dependent extinction factors are listed in Table 2.3. We refer readers who require photometry in other apertures to Brown et al. (2007, 2008).

Table 2.8 *Properties of the initial* PyBDSF *catalogues and the final multi-wavelength catalogues in ELAIS-N1, Lockman Hole, and Boötes. We also list here the overlapping multi-wavelength coverage area, the number of radio and multi-wavelength sources within this region, and the overlap bit flag,* FLAG_OVERLAP *for each field, which can be used to select both radio and multi-wavelength sources within our chosen area.*

	ELAIS-N1	Lockman Hole	Boötes
Nº PyBDSF radio sources	84 862	50 112	36 767
Nº multi-wavelength sources	2 106 293	3 041 956	2 214 358
FLAG_OVERLAP[a]	7	3	1
Overlap Area [deg^2][b]	6.74	10.28	8.63
Nº PyBDSF radio sources overlap[c]	31 059	29 784	18 766
Nº optical sources overlap[c]	1 470 968	1 906 317	1 911 929
Multi-wavelength catalogue sky density [arcsec^{-2}]	0.0168	0.0143	0.0171
PyBDSF radio catalogue sky density [arcsec^{-2}]	3.6×10^{-4}	2.2×10^{-4}	1.7×10^{-4}

[a] Overlap bit flag (FLAG_OVERLAP) provided in the full multi-wavelength catalogues and the radio cross-match catalogues indicating the coverage of each source. The overlap flag value in this table should be used to select sources in the overlapping multi-wavelength area defined in Sect. 2.2.

[b] The overlap area listed covers the overlapping multi-wavelength coverage (based on FLAG_OVERLAP) and excludes the region masked based on the *Spitzer* star mask. Radio-optical cross-matching is only performed for sources in this overlap area.

[c] Number of radio (in initial PyBDSF list) and optical sources in the overlap area above can be selected using the flag combination of FLAG_CLEAN $\neq 3$ and the respective FLAG_OVERLAP listed above.

It is worth re-iterating the key differences between the construction of the existing Boötes catalogues and the new catalogues generated for ELAIS-N1 and Lockman Hole. First, unlike in Boötes, where sources are detected in the I- and 4.5 μm bands, source detection in the other two fields is performed using χ^2 images which incorporates information from a wider range of wavelengths; as such, the resultant multi-wavelength catalogue would be expected to be more complete. Second, in generating the matched-aperture photometry in ELAIS-N1 and Lockman Hole, we do not smooth the PSFs unlike in Boötes; the variation of the PSFs is instead accounted for by computing different aperture corrections for each filter. In Boötes, aperture corrections are computed based on the Moffat profile PSF smoothing. Nevertheless, despite these differences, in both cases, the catalogues are built using both optical and IR data, extracted using SExtractor in dual-mode, and magnitudes are aperture corrected; thus, the catalogues are expected to be broadly comparable.

We provide here an itemised description of the key properties of the multi-wavelength catalogues released. Some of the properties (e.g. raw aperture fluxes) are only released for ELAIS-N1 and Lockman Hole.

- Unique source identifier for the catalogue ("ID")
- Multi-wavelength source position ("ALPHA_J2000", "DELTA_J2000")
- Aperture and extinction corrected flux (and flux errors) from our recommended aperture size (<band>_flux_corr and <band>_fluxerr_corr) in μJy
- Aperture and extinction corrected magnitude (and magnitude errors) from our recommended aperture size in the AB system (<band>_mag_corr and <band>_magerr_corr)
- Raw aperture flux (and flux errors) in 8 aperture sizes in μJy (FLUX_APER_<band>_ap and FLUXERR_APER_<band>_ap; excluding Boötes)
- Raw aperture magnitude (and magnitude errors) in 8 aperture sizes in the AB system (MAG_APER_<band>_ap and MAGERR_APER_<band>_ap; excluding Boötes)
- Overlap bit flag indicating the coverage of source across overlapping multi-wavelength surveys ("FLAG_OVERLAP"). See Table 2.8 for the recommended flag values.

- Bright star masking flag indicating masked and un-masked regions in the *Spitzer*- and optical-based bright star mask ("FLAG_CLEAN")
- Position based E(B − V) reddening values from Schlegel et al. (1998) dust map ("EBV").
- Manual masking and duplicate source flag from Brown et al. (2007, 2008) *I*-band catalog ("FLAG_DEEP"; for Boötes only).

We also refer the reader to the accompanying documentation for full description of all of the columns provided in the multi-wavelength catalogues[3]. Additional value-added columns regarding photometric redshifts, rest-frame colours, absolute magnitudes and stellar masses are described by Duncan et al. (2021), while the far-infrared fluxes are described by McCheyne et al. (subm). A full description of all columns provided in the multi-wavelength catalogues (both those described here, and in the value-added catalogue) can be found in the documentation accompanying the data release.

Now that the multi-wavelength catalogues have been constructed, we proceed to use these to identify and characterise the properties of the radio sources for the LoTSS Deep Fields in Chapter 3.

[3] See https://lofar-surveys.org/deepfields.html

Chapter 3

Host-galaxy identification in LoTSS Deep Fields

3.1 Introduction

The identification of the multi-wavelength counterparts to the radio-detected sources is crucial in maximising the scientific output from radio continuum surveys. This allows the characterisation of the host-galaxy properties, via SED fitting and photometric redshift estimates, for example. In this Chapter, we use the new forced, matched aperture multi-wavelength catalogues generated in Chapter 2 for cross-identification of radio sources detected in the LoTSS Deep Fields.

The identification of genuine counterparts to radio sources as opposed to random background objects is a challenging task. Emission from radio sources can be extended and the typically lower resolution of the radio data can lead to poor positional accuracy (and large, asymmetric positional uncertainties). This is compounded by the high source density of deep optical and infrared (IR) surveys, meaning that the genuine counterpart could lie anywhere within a large region around the radio source, with multiple potential counterparts within this region. For this reason, a simple nearest neighbour (NN) search is not always reliable, producing significant numbers of false identifications.

Moreover, radio surveys detect many classes of sources (e.g. star-forming galaxies, radio quiet quasars, radio-loud AGN, etc.) with a wide variety of morphologies which complicates this effort. For example, source extraction algorithms may split extended radio sources into multiple components, and sources nearby in sky-projection may be

blended together. Automatic association of the components and the identification of the genuine counterpart for such complex sources is difficult. While PyBDSF, the source extractor used for the LoTSS Deep Fields, is a very useful tool for source detection and measurement, the association of islands of radio emission into distinct radio sources is not expected to be perfect for all sources in all fields. Such incorrect associations in the PyBDSF catalogue can occur in a few ways, as noted by Williams et al. (2019; hereafter W19). Firstly, radio emission from physically distinct nearby sources could be associated as one PyBDSF source (blended sources). Such blends are much more common in these deep LOFAR data than the LoTSS-DR1. Secondly, sources with multiple components could be incorrectly grouped into separate PyBDSF sources due to a lack of contiguous emission between the components. For example, this can occur for sources with double radio lobes, or with large extended or diffuse radio emission.

In this Chapter, we aim to form correct associations of the sources and components generated by PyBDSF, and to cross-match (identify) the multi-wavelenth counterparts of the radio sources. In Sect. 3.2, we describe how we utilised the properties of a radio source and its neighbours to develop a decision tree to identify radio sources that are correctly associated, with secure radio positions and are hence suitable for an automated, statistical approach of cross-identification. For these sources, we use the Likelihood Ratio (LR) method (de Ruiter et al., 1977; Sutherland & Saunders, 1992), which is a commonly used statistical technique to identify real counterparts of sources detected at different wavelengths (e.g. Smith et al., 2011; McAlpine et al., 2012; Fleuren et al., 2012). In particular, we use the colour-based adaptation of the LR method, developed by Nisbet (2018) and used in the LoTSS-DR1 (Williams et al., 2019). This method incorporates positional uncertainties of the radio sources along with the magnitude and colour information of potential counterparts to generate a highly reliable and complete sample of cross-identifications. This process is described in Sect. 3.3. For sources where the decision tree indicates that the LR method is not suitable, we make use of a visual classification scheme to identify counterparts and perform accurate source association using a combination of the LOFAR Galaxy Zoo (LGZ), where source association and counterpart identification is performed using a group consensus, and, a separate workflow for specialised cases which are classified by a single expert. This process is detailed in Sect. 3.4. Sect. 3.5 then describes the properties of the final radio cross-matched and source associated catalogues released, and Sect. 3.6 presents magnitude distributions of the host galaxies of the radio sources in the LoTSS Deep Fields.

Both the radio-optical cross-match and the source associations for the LOFAR Deep

Fields adapt the techniques developed and presented for LoTSS-DR1 by W19. We refer the reader to that paper for details of the process; here, we summarise these methods and in particular, describe our specific adaptation and implementation of these methods to the LoTSS Deep Fields. We note that our radio cross-matching techniques and the cross-matched catalogues released are performed for sources within the overlapping multi-wavelength coverage defined in Sect. 2.2, less the region of the *Spitzer*-based bright star mask for each field. These areas are quoted in Table 2.8.

3.2 Decision tree

The decisions for how a source would be identified and/or classified are shown pictorially in Fig. 3.1, with numbers and fractions on the plot tracking the 31059 PyBDSF sources in ELAIS-N1 (see also Table 2.8). The decisions used (many of which are the same as those of W19) are listed in Table 3.1, and were based on both the radio source properties (e.g. size, source density, etc.) and the LR cross-matches (if any) of both the PyBDSF source and the Gaussian component catalogues in each field. Compared to the decision tree in LoTSS-DR1 (W19), some of the decision blocks could be simplified by sending sources directly for visual classification (without compromising feasibility) as the LR cross-match rate in the deep fields is significantly higher than in LoTSS-DR1, and the number of sources reaching these visual classification end-points is also much smaller (with very few extremely large/extended sources in these smaller areas). Furthermore, given the very high LR identification rates of up to 97% (see Sect. 3.3.5), it is feasible to send any source for which a counterpart cannot be determined using the LR method to visual inspection for confirmation that there is no possible counterpart.

We now describe in detail the key decision blocks and end-points of the decision tree. To select 'simple' and 'complex' sources in the decision tree, we use the S_Code parameter from the PyBDSF catalogues. We define a 'simple' source ('S' in Table 3.1) to only include sources that were fitted by a single Gaussian and are also the only source in the island (S_Code = S). Sources that were instead fitted with either multiple Gaussians (S_Code = M) or were fitted with a single Gaussian but were in the same island as other sources (S_Code = C) are defined as being 'complex'. Throughout this Chapter, we define a source as having a 'LR identification' (or LR-ID), if the LR value of the cross-match is above the LR threshold chosen (see Sect. 3.3 and Sect. 3.3.4).

In the decision tree, we first consider the size of the radio source. Radio sources with large sizes are typically complex or have poor positional accuracy; statistical methods

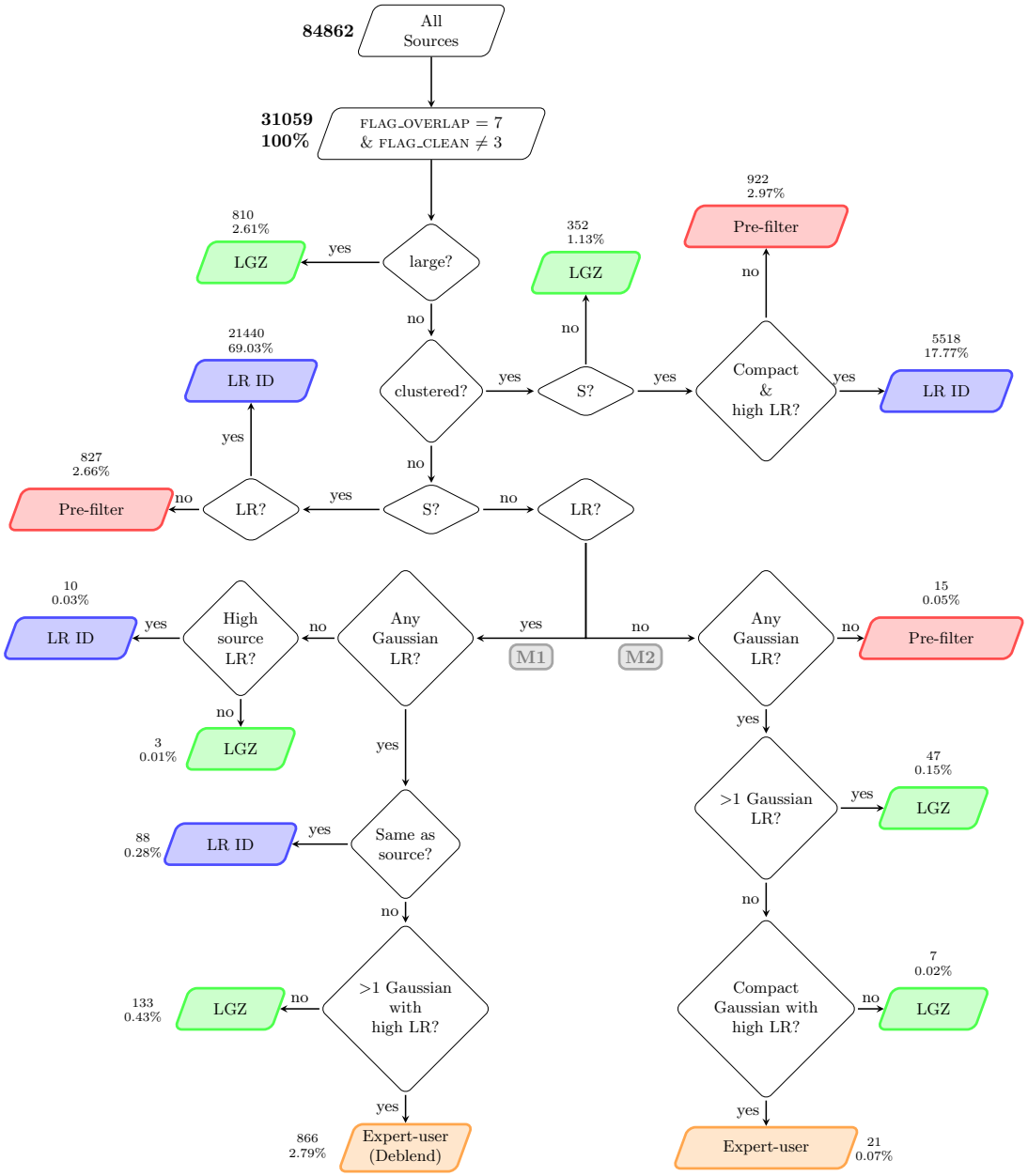

Figure 3.1 *(Caption on the following page).*

Figure 3.1 *Flowchart developed for the deep fields to select the most appropriate method (end-point) for identification (LR or visual) based on properties of the radio source and LR-identification (if any). The 'LR-ID' end-point indicates that the LR cross-match is accepted (see Sect. 3.3). Sources with the end-point of 'LGZ', 'pre-filter', and 'expert-user workflow' are sent to their respective visual classification and identification workflows (see Sect. 3.4). The numbers and percentages of sources at each end-point relate to the total number of sources in the* PyBDSF *catalogue within our defined multi-wavelength area (*FLAG_OVERLAP *= 7; see Table 2.8) and not in the Spitzer-masked region (*FLAG_CLEAN *≠ 3; see Table 2.7), corresponding to 31059* PyBDSF *sources. Table 3.1 lists the definitions used for each decision block. The decision tree is described in detail in Sect. 3.2. In ELAIS-N1, 27056 (87.1%) sources were selected as suitable for analysis by the statistical LR method, with the rest (4003 sources = 12.9%) selected as requiring some form of visual classification.*

of cross-identification for these sources are not accurate. Moreover, large PyBDSF sources may be part of even larger physical sources that are not correctly associated; these sources would need to be associated visually before the correct multi-wavelength ID can be selected. We therefore directly sent all large (major axis size > 15″) sources (810 sources = 2.6% in ELAIS-N1) in the PyBDSF catalogue to LGZ (see Sect. 3.4.1 for details of LGZ).

Next, for sources that are not large, we then test if they are in a region of high source density (referred to by W19 as 'clustered'); sources in high source density regions are more likely to be a part of some larger or complex source (although in some cases, could be just a chance occurrence due to sky projection). We define a source as 'clustered' if the separation to the fourth nearest neighbour (NN) is < 45″, the same criteria as used in LoTSS-DR1. All 'clustered' sources that are 'complex' were sent to LGZ since the complex nature of the sources would probably make them unsuitable for LR. Where instead these sources were 'simple', we checked if the source was compact and if the LR identification found was highly secure (see Table 3.1); if so, we accepted the secure LR identification found. Otherwise, the source was sent to the pre-filter workflow where one expert would quickly inspect the source and decide whether the LR cross-match found (or the lack of a LR-ID) is correct and should be accepted as the identification (or lack of), or if the source is complex and requires additional association and identification via visual methods (see Sect. 3.4). Section 3.4.3 provides a description of the pre-filter workflow.

Table 3.1 *Key criteria and definitions used in the decision blocks of the decision tree in Fig. 3.1. LR_{th} is the LR threshold corresponding to the intersection of the completeness and reliability. This is a scalar value that varies for each field.*

Parameter	Definition
Large	PyBDSF major axis > 15″
Clustered	Distance to fourth nearest neighbour < 45″
S	'Simple' source: single Gaussian PyBDSF source (and only source in the island)
LR	$LR > LR_{th}$
High source LR	$LR_{source} > 10 \times LR_{th}$
Compact & high LR	PyBDSF source major axis < 10″ and $LR_{source} > 10 \times LR_{th}$
Same as source	The ID(s) for the Gaussian component(s) is identical to ID for the source
Any Gaussian LR	At least one Gaussian component with $LR_{gauss} > LR_{th}$
> 1 Gaussian LR	More than 1 Gaussian with $LR_{gauss} > LR_{th}$
> 1 Gaussian with high LR	More than 1 Gaussian with $LR_{gauss} > 10 \times LR_{th}$
Compact Gaussian with high LR	Gaussian major axis < 10″ and $LR_{gauss} > 10 \times LR_{th}$

The largest branch of the decision tree was formed of the remaining small, non-clustered sources (23457, 75.5%). The LR analysis is most suitable for non-clustered, 'simple' sources with compact radio emission so, if sources at this stage had a LR value above the LR threshold (LR_{th}), we accept the multi-wavelength ID as found by the LR analysis (i.e. LR-ID). In ELAIS-N1, 21440 (69.0%) sources were identified at this end-point. If instead a match is not found by the LR analysis (i.e. the LR is lower than the threshold), the source was sent to the pre-filter workflow to either confirm that there is no acceptable LR match, or to send for visual classification in the case that the multi-wavelength ID is missed by the LR analysis. In ELAIS-N1, 827 (2.66%) sources were sent to the pre-filter workflow from this branch.

Next, the small, non-clustered sources that are 'complex' instead, were treated in two separate branches based on whether the source LR value is above ('M1' branch) or below ('M2' branch) the threshold. For these sources, we considered both the LR identification of the source (LR_{source}) and the LR identification of the Gaussian components of the source (LR_{gauss}).

If the 'complex' source had a LR above the threshold, we decided the end-point of the source by considering the LR value and LR-ID found by the source and by the individual Gaussian components that form the source (see 'M1' branch of Fig. 3.1). We do not simply accept the source LR identification for such sources as this branch may include sources that have complex emission fitted by multiple Gaussians, or cases where PyBDSF has incorrectly grouped Gaussians associated with multiple physical sources into a single PyBDSF catalogue source (i.e. blends). If the source LR-ID and all of its constituent Gaussian LR-IDs are the same, or if the PyBDSF source has a highly secure LR-ID and with no individual Gaussians having a LR-ID, we accepted the source LR-ID. If multiple Gaussians have secure LR-IDs, these are likely to be blended sources. We therefore sent these to the 'expert user workflow' (see Sect. 3.4.2 for details) to perform de-blending. Sources at all other end-points were sent to LGZ in the 'M1' branch, as detailed in Fig. 3.1.

The non-clustered, 'complex' sources that don't have a source LR match above the threshold were considered in the 'M2' branch (see Fig. 3.1). In this branch, sources were sent to the pre-filter workflow if none of the Gaussians have a LR match, with the main aim of confirming the lack of a multi-wavelength counterpart. If only one of the constituent Gaussians had a LR match (which is also highly secure) and a compact size, we sent the source to the 'expert user workflow' to confirm the LR match or to change this (and the PyBDSF Gaussian grouping; if necessary). Sources at all other end-points of the 'M2' branch (comprising $< 0.5\%$ of the total PyBDSF catalogued sources) were

then sent for visual classification via LGZ.

Of the 31059 PyBDSF sources in ELAIS-N1, 27056 (87.11%) sources were selected as suitable for the statistical LR analysis. 1352 (4.35%) sources were sent directly to LGZ and 887 (2.86%) sources were sent to the 'expert-user workflow', with the majority of these selected as being potential blends. Finally, 1764 (5.68%) sources were sent to the pre-filter workflow; these were appropriately flagged and then sent to the expert-user, LGZ, and LR workflows (if required). We note that the number of sources that actually underwent de-blending was different to the number of potential blends listed above, as some sources that were initially selected as blends turned out not to be genuine blends, while additional sources were input from both LGZ and pre-filter that were flagged as blends. In the rest of this section, we describe in detail how we classify and identify the host galaxies of sources that are in each of the four distinct end-points of the decision tree.

3.3 The Likelihood Ratio method

The statistical Likelihood Ratio (LR) method (de Ruiter et al., 1977; Sutherland & Saunders, 1992) is commonly used to identify counterparts to radio and milli-metre sources (e.g. Smith et al. 2011; Fleuren et al. 2012; McAlpine et al. 2012). Defined simply, the LR is the ratio of the probability that a galaxy with a given set of properties is a genuine counterpart as opposed to the probability that it is an unrelated background object. In this Chapter, we use the magnitude m, and the colour c information to compute the LR of a source. Nisbet (2018) and W19 show that incorporating colour information greatly benefits the LR analysis, finding that redder galaxies are more likely to host a radio source. The LR is given by

$$LR = \frac{q(m,c)f(r)}{n(m,c)}, \tag{3.1}$$

where $q(m,c)$ gives the a priori probability that a source with magnitude m and colour c is a counterpart to the radio (LOFAR) source. $n(m,c)$ represents the sky density of all galaxies of magnitude m and colour c. $f(r)$ is the probability distribution of the offset between the radio source and the possible counterpart, while accounting for the positional uncertainties of both of the sources. A full description of the theoretical background and the method of the LR technique is given in W19 and is not reproduced here. Instead, we focus mainly on the specific application of the LR technique to the LOFAR Deep Fields dataset.

Table 3.2 *Q_0 values in the optical (i-band for ELAIS-N1 and Boötes and r-band for Lockman Hole; see text) and 4.5 μm bands for the magnitude only LR run in each field.*

Field	$Q_{0,\mathrm{opt}}$	$Q_{0,4.5}$
ELAIS-N1	0.85	0.95
Boötes	0.75	0.84
Lockman Hole	0.78	0.95

3.3.1 Calculating $n(m)$ and $n(m, c)$

The $n(m)$ corresponds to the number of objects in the multi-wavelength catalogue at a given magnitude per unit area of the sky. This is computed simply by counting the number of sources within a large representative area (typically > 3.5 deg^2 in our case) in each of the three fields. We adopt a Gaussian kernel density estimator (KDE) of width 0.5 mag to smooth the $n(m)$ distribution and provide a more robust estimate when interpolated at a given magnitude. The $n(m, c)$ is then simply given by computing the $n(m)$ separately for different colour bins (see Sect. 3.3.3).

3.3.2 Calculating $f(r)$

The $f(r)$ term accounts for the positional difference between the radio source and a potential multi-wavelength counterpart. The form of the distribution is given as a 2D Gaussian with

$$f(r) = \frac{1}{2\pi\sigma_{maj}\sigma_{min}} \exp\left(\frac{-r^2}{2\sigma_{dir}^2}\right) \tag{3.2}$$

where, σ_{maj} and σ_{min} are the combined positional uncertainties along the major and minor axes, respectively, and σ_{dir} is the combined positional uncertainty, projected along the direction between the radio source and the potential counterpart. The σ_{maj} and σ_{min} terms are a combination of the uncertainties in both the radio and the potential multi-wavelength counterpart positions, and the uncertainties in the relative astrometry of the two catalogues, calculated using the method of Condon (1997). For the potential multi-wavelength counterparts, as the positional uncertainties from a χ^2 detection image are unreliable, we adopt a circular positional uncertainty of $\sigma_{\mathrm{opt}} = 0.35''$. Similar to W19, an additional astrometric uncertainty between the radio and multi-wavelength catalogues of $\sigma_{\mathrm{ast}} = 0.6''$ was adopted. These terms were then added in quadrature for radio source and potential counterparts to derive σ_{maj} and σ_{min}.

3.3.3 Calculating $q(m)$ and $q(m, c)$

$q(m)$ (and $q(m, c)$) is the a priori probability distribution that a radio source has a genuine counterpart with magnitude m (and colour c). The integral of $q(m)$ to the survey detection limit gives Q_0, the fraction of radio sources that have a genuine counterpart up to the magnitude limit of the survey.

The LR analysis is not suitable for large or complex radio sources, and to reduce the bias introduced by such sources on the LR analysis, we initially performed the LR analysis only for radio sources with a major axis size smaller than 10″. In each field, this subset of radio sources was used initially to calibrate the $q(m, c)$ distributions (using the two stage method, as described below in this section). These calibrated $q(m, c)$ distributions were then used to compute the LRs for all radio sources within the multi-wavelength coverage area listed in Table 2.8. The decision tree described in Sect. 3.2 was then used to re-select radio sources that were more suitable for the LR analysis. For this purpose, we choose to calibrate on all 'simple' sources that reach the LR-ID or the pre-filter end-points of the decision tree. The $q(m, c)$ distributions were re-calibrated on this sample and then used to re-compute the LRs for all the radio sources in the field to derive the final counterparts. We found that further iterations of the decision tree made insignificant changes ($\lesssim 1\%$) to the number of sources selected for visual analysis or LR, suggesting that the calibration was being performed on sources most suitable for the LR analysis.

It is well known that radio-loud AGN preferentially lie in over-dense regions such as galaxy groups or clusters (e.g. Prestage & Peacock, 1988; Best, 2004). Various methods have been developed to estimate $q(m)$ and Q_0, using the data itself, in a manner that is unbiased by galaxy clustering (e.g. Smith et al., 2011; McAlpine et al., 2012; Fleuren et al., 2012); here, we follow the approach detailed by Fleuren et al. (2012). In summary, $q(m)$ is determined by first calculating the magnitude distribution of all multi-wavelength sources within a search radius of $r_{\rm max}$, total(m), that is then corrected for the background galaxy counts using the $n(m)$ distribution; this results in the magnitude distribution of sources associated with the radio sources (i.e. a combination of the true radio-source hosts and associated galaxies due to galaxy clustering; see below), given as

$$\mathrm{real}(m) = \mathrm{total}(m) - n(m) N_{\rm radio} \pi r_{\rm max}^2, \tag{3.3}$$

where $N_{\rm radio}$ is the number of radio sources in the catalogue, and $r_{\rm max} = 10''$ (chosen to be comparable to the resolution of the radio data). The $q(m)$ distribution is then

determined by normalising the real(m) distribution and scaling it by Q_0, such that

$$q(m) = \frac{\text{real}(m)}{\sum_{m_i} \text{real}(m)} Q_0. \tag{3.4}$$

Here, the $q(m)$ distribution is corrected for galaxy clustering by calculating the number radio sources without a potential multi-wavelength counterpart out to a search radius, (known as 'blank fields'), and comparing this to the number of black fields around random positions. This method takes into account any biases associated with galaxy clustering as it is based on counting the number of blank fields, which are unaffected by the presence of associated objects around radio sources.

However, as explained by W19, these methods cannot be used to estimate $q(m, c)$ and $Q_{0,c}$ in different colour bins. Instead, we use the iterative approach developed by Nisbet (2018) and applied to the LoTSS DR1 by W19 for estimating $q(m, c)$ in two stages. Briefly, the first stage of this approach involves identifying an initial estimate of the host galaxies using well established magnitude-only LR techniques (e.g. Fleuren et al. 2012). In the second stage, this initial set of host galaxies is split into various colour bins, to allow a starting estimate of $q(m, c)$ to be obtained, which is then used to recompute the LRs, incorporating colour information. This then provides a new set of host galaxy matches, and hence an improved estimate of $q(m, c)$, with this process iterated until the $q(m, c)$ distribution converges.

In practice, in the first stage, we generated a set of initial counterparts to the radio sources using only the magnitude information by cross-matching the radio sources to both the 4.5 μm detected and optical detected [1] sources (separately). For the optical dataset, we use the PS1 i-band in ELAIS-N1, the NDWFS I-band in Boötes, and the SpARCS r-band in Lockman Hole. While there exist i-band data from RCSLenS in Lockman Hole, the survey coverage had gaps in the field between different pointings due to the survey strategy employed. Therefore, we compromise slightly on the choice of optical filter for LR analysis in favour of area coverage. The method of Fleuren et al. (2012) was then used to compute Q_0 in each filter. We list the Q_0 values in the optical and 4.5 μm bands for each field from this first stage in Table 3.2. The differences in these Q_0 values are largely driven by the relative depths of the optical and *Spitzer* surveys between the three fields.

The Q_0 values were then used to derive the corresponding $q(m)$ distributions following the method of Fleuren et al. (2012). The final part of the first stage then involves

[1] We define a source as being detected in a given filter if the S/N > 3 inside the 2″ aperture in that filter.

Table 3.3 *Table of iterated $Q_0(c)$ values for ELAIS-N1, Lockman Hole, and Boötes. The colour* c *is derived using optical - 4.5 µm magnitude where we use the i- (or I-) band in ELAIS-N1 and Boötes, and the r-band in Lockman Hole. The LR thresholds (LR_{th}) derived from the intersection of the completeness and reliability function (see Sect. 3.3.4) and used for selecting genuine cross-matches are also listed. Scaled (by excluding sources in region of around stars, i.e.* FLAG_CLEAN$\neq$ *3) total $Q_0(c)$ values are also listed.*

	$Q_0(c)$		
Colour Bin	ELAIS-N1	Lockman Hole	Boötes
$c \leq -0.5$	0.0031	0.0013	0.0016
$-0.5 < c \leq -0.25$	0.0034	0.0012	0.003
$-0.25 < c \leq 0.0$	0.0081	0.0041	0.0103
$0.0 < c \leq 0.25$	0.0177	0.0086	0.0204
$0.25 < c \leq 0.5$	0.0301	0.0154	0.0316
$0.5 < c \leq 0.75$	0.0468	0.022	0.0465
$0.75 < c \leq 1.0$	0.0562	0.0302	0.059
$1.0 < c \leq 1.25$	0.0606	0.0393	0.0608
$1.25 < c \leq 1.5$	0.0566	0.044	0.0589
$1.5 < c \leq 1.75$	0.0523	0.0457	0.0555
$1.75 < c \leq 2.0$	0.0557	0.045	0.0519
$2.0 < c \leq 2.25$	0.0486	0.0491	0.0504
$2.25 < c \leq 2.5$	0.0489	0.0486	0.0477
$2.5 < c \leq 2.75$	0.0467	0.0481	0.0448
$2.75 < c \leq 3.0$	0.0478	0.0498	0.0472
$3.0 < c \leq 3.25$	0.0481	0.0484	0.0473
$3.25 < c \leq 3.5$	0.0456	0.0493	0.0433
$3.5 < c \leq 3.75$	0.0422	0.0496	0.0357
$3.75 < c \leq 4.0$	0.0388	0.0463	0.0325
$c > 4.0$	0.076	0.1359	0.0546
optical-only	0.0044	0.0068	0.001
4.5-only	0.1129	0.1771	0.1107
no-magnitude	0.0019	0.0061	0.001
Total $Q_0(c)$	95.3%	97.2%	91.6%
Total $Q_0(c)$ with FLAG_CLEAN$\neq$3	96.2%	97.4%	94.2%
LR threshold	0.056	0.055	0.22

computing the LRs for all optical and 4.5 μm detected sources (separately) within 10″ of a radio source. An optical or 4.5 μm detected source was accepted as a cross-match if the LR was above the threshold in that particular filter; in this first stage, the LR threshold is simply estimated as the value for which a fraction Q_0 of cross-matches were accepted in that band. If multiple sources within 10″ were above the LR threshold, the source with the highest LR (in either the optical or the 4.5 μm band) was retained as the most-probable cross-match. The main output of the first stage generates a first-pass set of multi-wavelength counterparts.

In the second stage, the counterparts generated from the first stage were divided into colour categories to provide an initial estimate of $Q_0(c)$ ($= N_c/N_{LOFAR}$) and $q(m, c)$. Colour bins were derived from the (optical - 4.5 μm) colour, provided the source is detected with S/N > 3 in both bands. These sources were then split into 20 colour bins, as listed in Table 3.3. In addition to these, some sources are only detected (S/N > 3) in either the optical or the 4.5 μm band. For these sources, we define two additional colour categories: optical-only and 4.5-only sources. Finally, as mentioned earlier, due to the nature of the detection method using χ^2 images, there are sources that have a low S/N in both the i (or r) and 4.5 μm filters but appear in the catalogue due to detections in other bands. These sources were placed in a final colour category, the 'no-magnitude' category, for which we manually set a first-pass value for the cross-match fraction of $Q_0 = 0.001$ and use the corresponding sky density of all sources in this bin to compute the LRs.

After the division into the colour categories, $n(m, c)$ and $q(m, c)$ can be determined trivially. We again smooth these distributions using a Gaussian KDE of width 0.5 mag. The LR analysis was then repeated in the same manner as stage one where for each source in the multi-wavelength catalogue that is within 10″ of a radio source, the $n(m)$ and $q(m)$ distribution corresponding to the colour bin of that source is used to compute new LRs. A new LR threshold was determined using the completeness and reliability of the cross-matches (see Sect. 3.3.4 for a detailed description), improving upon the estimate from the first stage, with the highest LR match above the threshold retained in each case to produce a new set of cross-matches. The process in the second stage was iterated until the cross-matches converged (i.e. no changes in the sources cross-matched between two consecutive iterations), which was typically within 5 iterations. The total Q_0 is then simply given by summing over the contributions from each colour category, that is, $Q_0 = \sum_c Q_0(c)$. Iteration of the LRs can progressively drive down the $Q_0(c)$ values to zero in the rarest bins. To avoid this, we set a minimum $Q_0(c) = 0.001$ for any colour bin.

3.3.4 Likelihood Ratio thresholds

The LR value for each potential counterpart to be the genuine counterpart of a radio source is computed using equation 3.1 as described in Sect. 3.3. We determine the LR threshold ($LR_{\rm th}$) above which to accept a match as being the genuine counterpart as follows. For a given LR threshold $LR_{\rm th}$, one can compute the completeness $C(LR_{\rm th})$ and reliability $R(LR_{\rm th})$ as

$$C(LR_{\rm th}) = 1 - \frac{1}{Q_0 N_{\rm radio}} \sum_{LR_i < LR_{\rm th}} \frac{Q_0\, LR_i}{Q_0\, LR_i + (1 - Q_0)}, \tag{3.5}$$

$$R(LR_{\rm th}) = 1 - \frac{1}{Q_0 N_{\rm radio}} \sum_{LR_i \geq LR_{\rm th}} \frac{1 - Q_0}{Q_0\, LR_i + (1 - Q_0)}, \tag{3.6}$$

where $N_{\rm radio}$ is the number of radio sources in the catalogue and LR_i is the LR of the radio source i (de Ruiter et al., 1977; Best et al., 2003). The completeness sums over the lower LR values and is defined as the fraction of real identifications that are accepted. The reliability sums over the LR values above the threshold and is defined as the fraction of accepted identifications that are correct.

We determine an appropriate threshold by using the cross-over point between the completeness and reliability curves as shown in Fig. 3.2 (*top*) for ELAIS-N1 (see W19 for further discussion). In ELAIS-N1, the $\rm LR_{th} = 0.056$ chosen, returns a cross-matching completeness and reliability in excess of 99.7%. The division of the radio sources into colour bins drives down the LR values and hence the LR thresholds compared to the magnitude-only run, resulting in LR thresholds below unity. We inspect the sources with LR values near the $\rm LR_{th}$ by visual examination and find that the $\rm LR_{th}$ chosen results in genuine counterparts. The choice of the LR threshold can also be validated by considering the additional LR-matches, along with the change in completeness and reliability as the LR threshold is lowered, following the method described in detail by Nisbet (2018). In summary, using equations 3.5 and 3.6, the number of genuine matches above a threshold T, is given by $N_{radio}\ Q_0\ C(T)$, and the total number of matches above the threshold T is given by $N_{radio}\ Q_0\ C(T)/R(T)$. Then, if the threshold is lowered from T to $T - \Delta T$, this will result in a set of additional matches, some of which will be genuine counterparts. The probability of the additional matches added being genuine

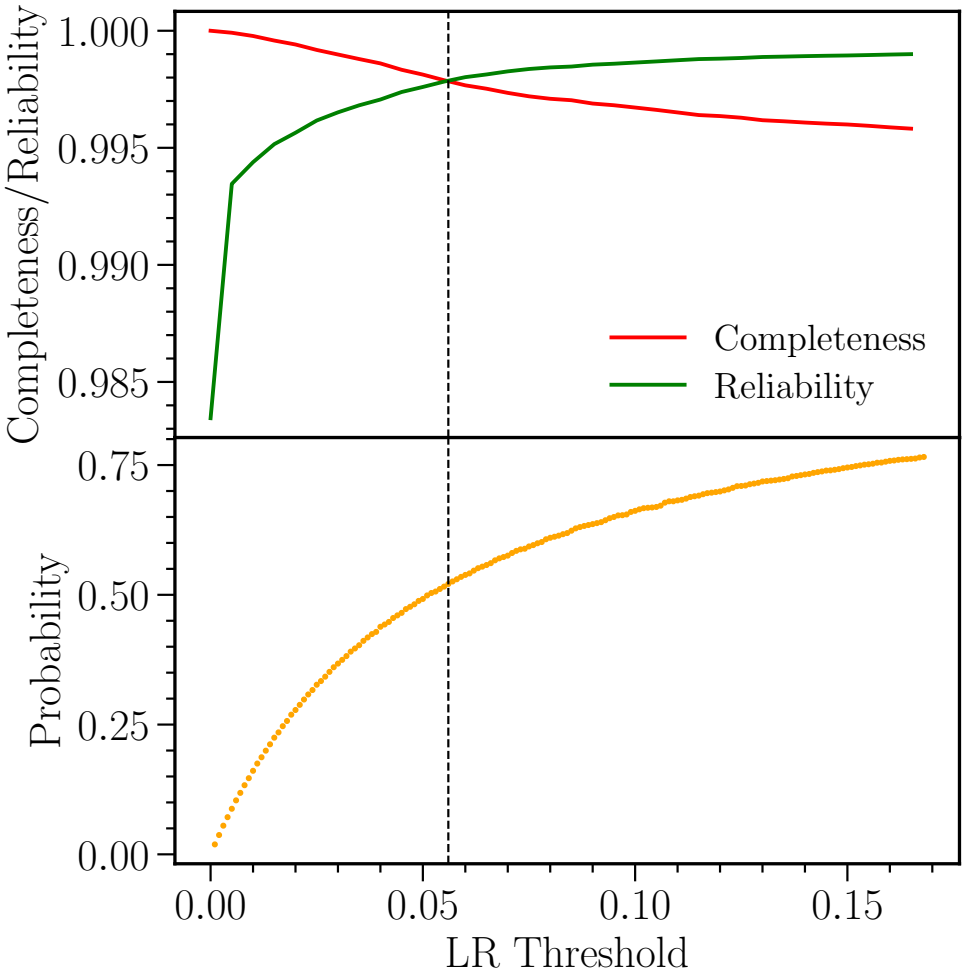

Figure 3.2 *Top: Completeness and Reliability curves as a function of the LR threshold (LR_{th}) in ELAIS-N1. The LR_{th} (= 0.056; vertical line) is chosen as the cross-over point of the two curves, achieving both completeness and reliability > 99.7%. Bottom: The probability of a LR-match being a genuine cross-match as a function of the LR threshold, as it is lowered (see text and Nisbet 2018 for more detail). At the adopted LR_{th}, the LR matches have a probability of just over 50% of being genuine counterparts, confirming that this threshold maximised completeness with limited loss of reliability.*

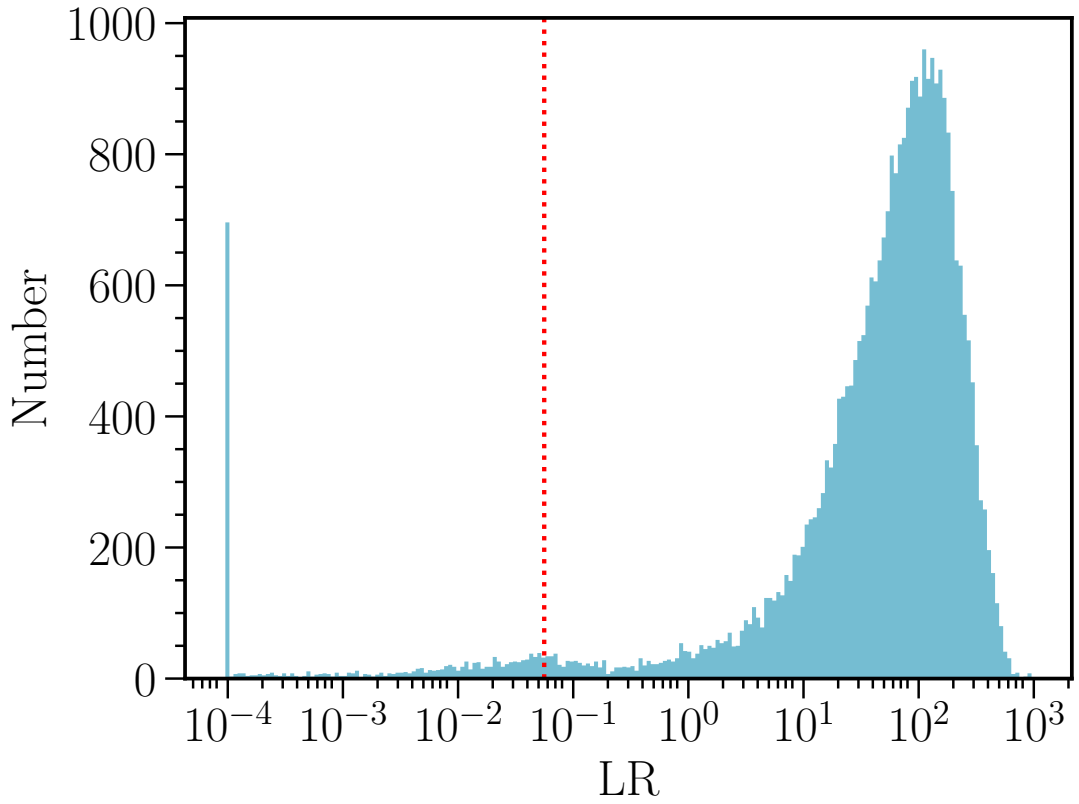

Figure 3.3 *Histogram of the LR values in ELAIS-N1 with equal spaced bins in log-space. The vertical red line shows the LR threshold chosen (0.056). Some sources can have exceptionally low LR values, hence, sources with LR* $\lesssim 10^{-4}$ *are placed in the first bin to aid visual inspection.*

when the threshold is changed from T to $T - \Delta T$ is given as

$$P_{genuine}(T \rightarrow T - \Delta T) = \frac{C(T) - C(T - \Delta T)}{\frac{C(T)}{R(T)} - \frac{C(T - \Delta T)}{R(T - \Delta T)}}. \tag{3.7}$$

This probability $P_{genuine}(T \rightarrow T - \Delta T)$, as a function of the LR threshold is also shown in Fig. 3.2 (*bottom*) for ELAIS-N1. The plot shows that at the LR threshold value chosen, the probability of the LR match being a genuine counterpart is > 52%, suggesting that the counterparts with LR values below unity (but above the threshold) are more likely to be genuine matches than false identifications. A similar analysis in Lockman Hole and Boötes yields probabilities of ∼65% and ∼70%, respectively, at the respective LR thresholds.

An appropriate choice for the LR threshold can also be visualised using a histogram of the LR values, as shown in Fig. 3.3 for ELAIS-N1 considering the PyBDSF catalogue

sources. There are ≈ 1700 sources with LRs < LR_{th}, which were sent to visual inspection. This corresponds to a fraction of ∼ 0.05, consistent with the final iterated $Q_0 \sim 0.95$ obtained using the cross-over point of the completeness and reliability curves (Fig. 3.2). The plot also shows that most of the sources have LR values significantly higher than LR_{th}. Moreover, > 99% of the PyBDSF sources with a LR > LR_{th} have LR values above $2 \times LR_{th}$ (= 0.11); the probability of the cross-match identified being genuine at this point is ∼ 70% (using Fig. 3.2).

3.3.5 LR method results

The colour bins and the corresponding final iterated $Q_0(c)$ values are provided in Table 3.3. The colour c is the same optical - mid-IR colour that was used for the LR analysis. Table 3.3 also lists the iterated LR threshold values derived from the intersection of the completeness and reliability plots (see Sect. 3.3.4). The full sample in all fields achieves a completeness and reliability > 99.7% (see Fig. 3.2 for ELAIS-N1). Visual inspection of low LR matches, and an analysis of the completeness and reliability of sources with LRs close to the LR threshold, gives confidence that the LR thresholds chosen result in genuine cross-matches (see Sect. 3.3.4 for full details).

The total $Q_0(c)$, given by summing the contribution from each colour category, gives an identification fraction of ∼ 95%, 97% and 92% for ELAIS-N1, Lockman Hole, and Boötes, respectively. Interestingly, Lockman Hole has a higher total $Q_0(c)$ than ELAIS-N1, which contains IR data to a similar depth, but much deeper optical data. This difference can be understood by considering the Q_0 of the 4.5 μm band and its coverage, in particular, that of the deeper SERVS data between the two fields. Although the optical data in ELAIS-N1 is much deeper than in Lockman Hole, the 4.5 μm data dominates the identification fraction (see Table 3.2). The SERVS 4.5 μm data in both ELAIS-N1 and Lockman Hole reach a similar depth (as listed in Table 2.2) and achieve the same $Q_{0,4.5}$. However, Lockman Hole benefits from having SERVS coverage (and therefore this high identification rate) over ∼5.6 deg^2, compared to only ∼2.4 deg^2 in ELAIS-N1, resulting in the difference in the total Q_0. The overall Q_0 values shown in Table 3.3 are significantly higher than the total cross-identification fraction of 71% achieved in the shallower LoTSS DR1 (W19).

The iterated $Q_0(c)$ values show remarkable agreement across the three fields, especially between ELAIS-N1 and Boötes which both use the similar optical filters. This agreement can be visualised using the iterated (and KDE smoothed) $\log q(m, c)/n(m, c)$ ratio distributions, which are shown in Fig. 3.4 as a function of magnitude. We show

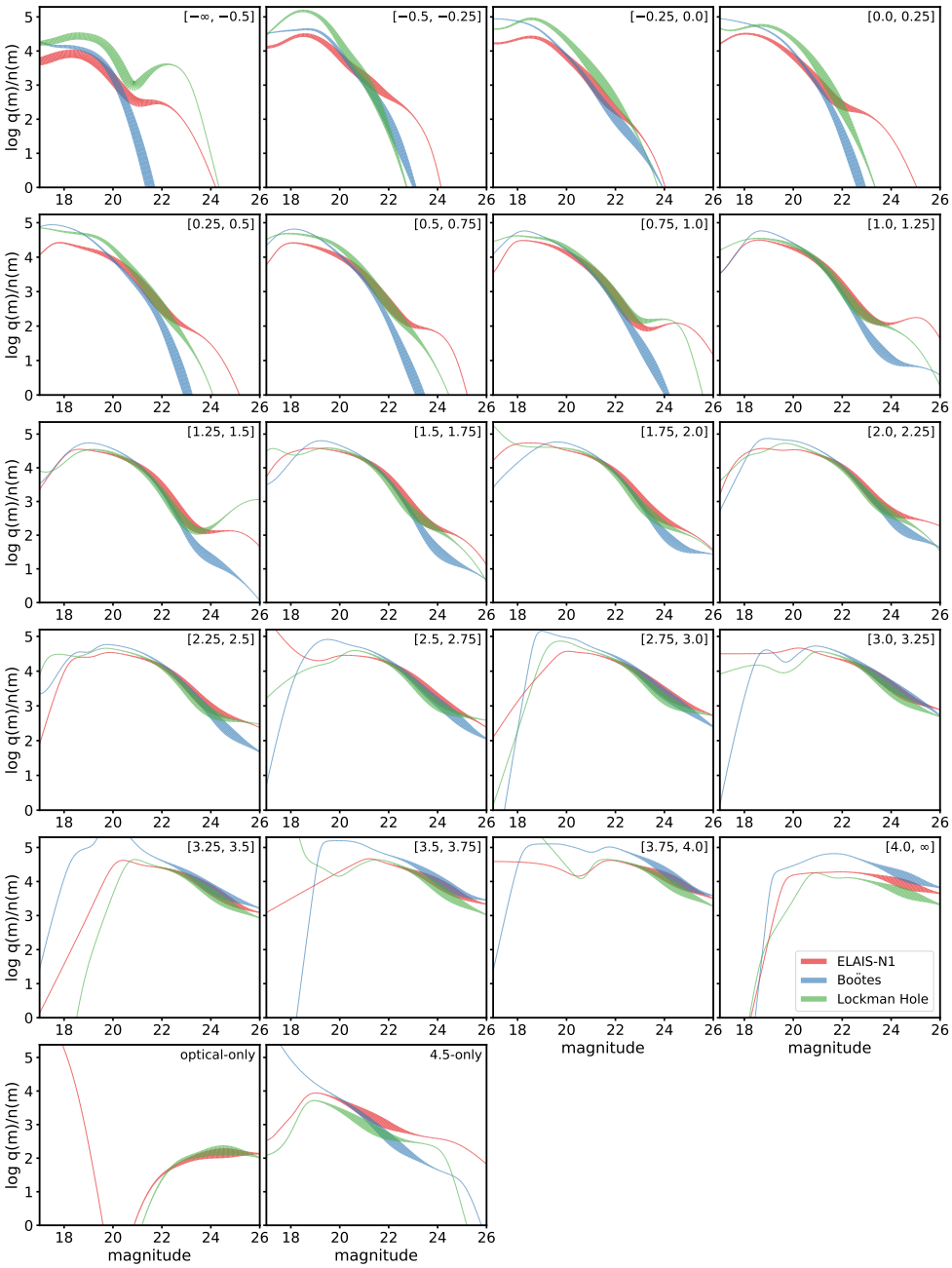

Figure 3.4 q(m, c)/n(m, c) *ratio distributions versus magnitude, smoothed using a KDE. The x-axis displays the optical magnitude in each colour bin, except for the* 4.5*-only category where the 4.5 µm magnitude is used. The width of the lines corresponds to the number of radio sources within that magnitude bin (hence, the thicker lines indicate well-constrained regions of parameter space). The optical magnitudes plotted are the same as those chosen for the LR analysis (i-band for ELAIS-N1, I-band for Boötes and r-band for Lockman Hole). Although these filters are different, no attempt at filter or colour transformation is made (see Sect. 3.3.5). Even without these corrections, the distributions agree well between the three fields, especially considering the log scaling of the y-axis.*

the distributions for ELAIS-N1 (red), Boötes (blue), and Lockman Hole (green) across all the colour bins. The x-axis for all colour bins is the optical magnitude (i.e. i (and I) band for ELAIS-N1 and Boötes, r-band for Lockman Hole), except for the 4.5-only bin where the 4.5 μm magnitude is used. The bin edges for the (optical - 4.5) colour bins (as in Table 3.3) are shown at the top right corner in each panel. The thickness of the curves corresponds to the number of sources within that magnitude bin, such that the distributions and statistics are reliable where lines are thick, and with thin lines corresponding to poorly constrained regions of parameter space, often influenced by the tails of the KDE smoothing. We note that for Lockman Hole, the comparison to the other two fields is not exactly like-for-like due to the different filters. This relates not only to an x-axis shift in colours, but also the selection of sources in each colour bin; for example, sources in Lockman Hole with $3.5 < (i - 4.5) < 3.75$ have a typical colour of $(r - i) \sim 1$, and hence $(r - 4.5) > 4$, so would appear in the c > 4.0 colour category instead. The key note of importance here is that even without the filter transformation for Lockman Hole, the distributions agree well between the three fields, across the colour bins. This agreement is expected as the $q(m, c)$ distribution represents the genuine host galaxy population of radio sources in magnitude and colour space, which should be consistent between the three fields with similar radio survey properties.

In Fig. 3.5, we again show the iterated (and KDE smoothed) $q(m, c)/n(m, c)$ ratio distribution for all colour bins in ELAIS-N1, all on one plot. The numbers in the legend show bin edges in the (i - 4.5) colour space, same as in Fig. 3.4. The evolution of the curves going from blue to redder bins indicates that redder galaxies are more likely to host radio sources, especially at faint magnitudes.

This colour dependence on the identification rate can also be visualised by considering the fraction (percentage) of all multi-wavelength sources that host a LOFAR source as a function of the (i - 4.5) colour, as shown in Fig. 3.6 for ELAIS-N1. The size of the markers indicates the number of LOFAR sources within that colour bin. The sharp rise in the fraction of matches with colour again shows that redder galaxies are more likely to host a LOFAR source as compared to the general galaxy population.

Compared to the shallower radio data available in LoTSS-DR1 (e.g. see Fig. 3 of W19), we note a rise in the fraction of LOFAR sources at blue ($1 < (i - 4.5) < 2$) colours in these deep fields (as seen in Fig. 3.6), and an increase in the $q(m, c)/n(m, c)$ ratios for blue bins (as shown in Fig. 3.5). These trends compared to W19 are probably due to the significant increase in depth of the radio data, where the faint radio source population ($\lesssim 1$ mJy at 150 MHz) starts to be dominated by radio quiet quasars and star-forming galaxies (see fig. 4 of Wilman et al. 2008), which are typically found in bluer galaxies.

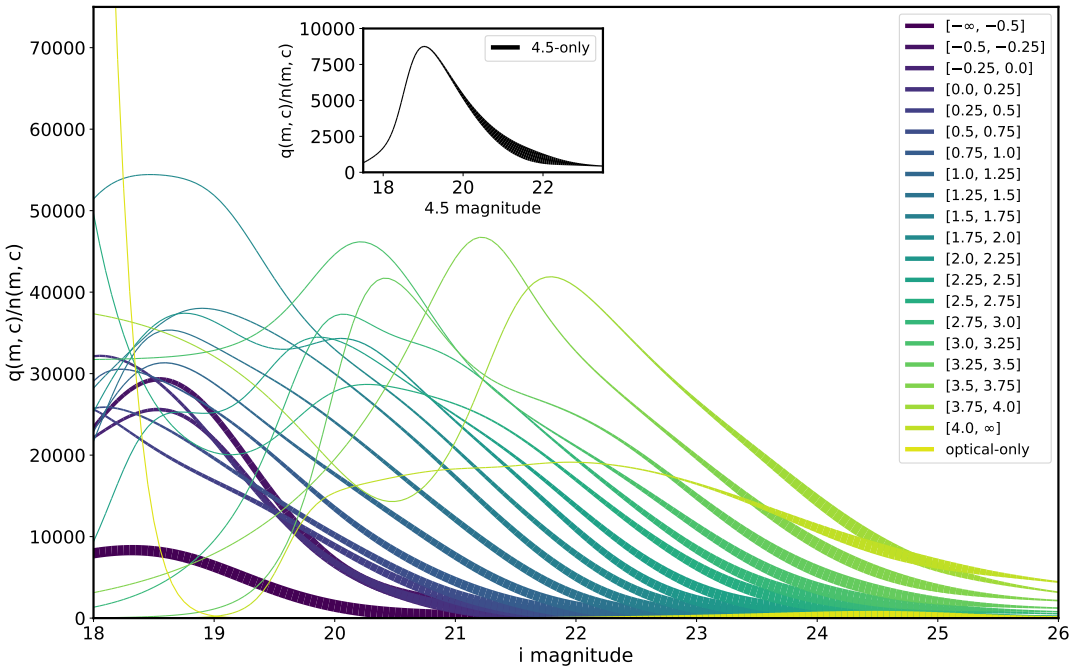

Figure 3.5 *Plot of the* q(m, c)/n(m, c) *ratio distributions versus magnitude across the (i - 4.5) colour bins in ELAIS-N1. The inset shows the same for the 4.5-only bin. The ratios are computed in bins of 0.5 mag (and smoothed using a KDE), with the thickness of the lines corresponding to the number of sources within a given magnitude bin (i.e. thicker lines represent better constrained regions of parameter space). The numbers in the legend correspond to the bin edges in (i - 4.5) colour space. The evolution of the peak and thickness of the curve across the colour bins indicate that radio galaxies are more likely to be hosted by redder galaxies, especially at faint magnitudes.*

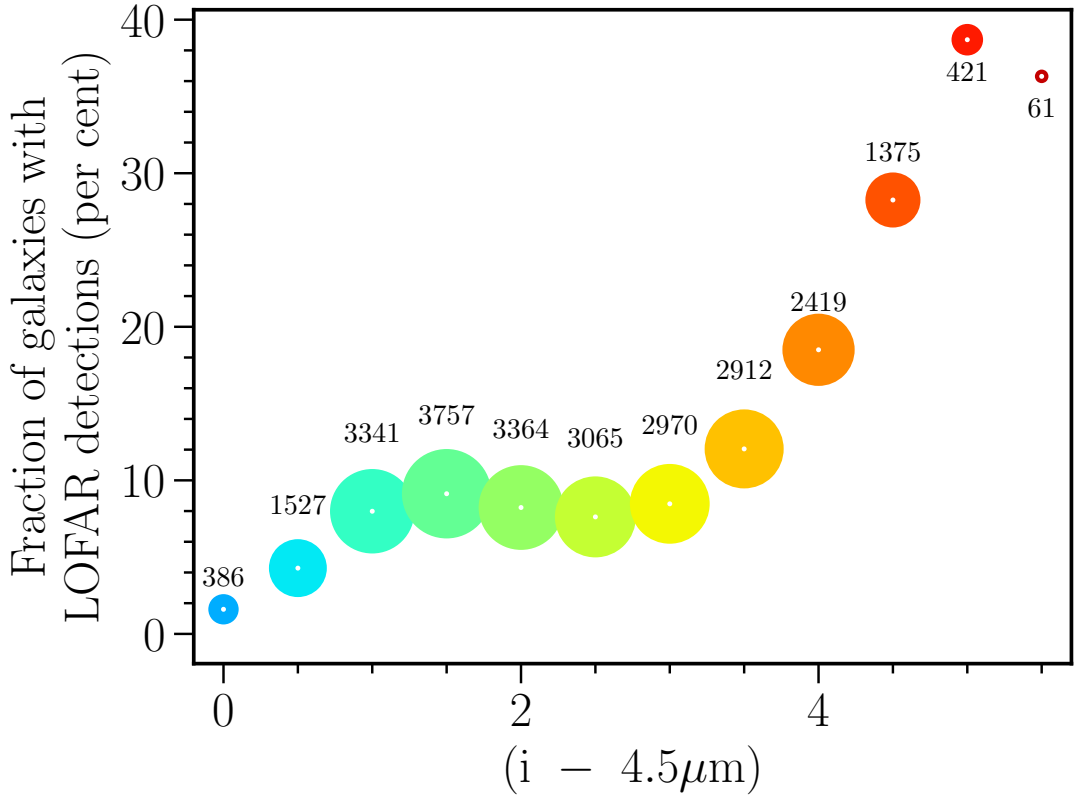

Figure 3.6 *Fraction of all multi-wavelength sources that host a LOFAR source in ELAIS-N1 as a function of (i - 4.5) colour. The size of the data points corresponds to number of LOFAR sources within that colour bin (indicated by the adjacent number), and the colour of the points is a proxy for counterpart colour. The reddest galaxies are more than an order of magnitude more likely to host a radio source than the bluest of galaxies.*

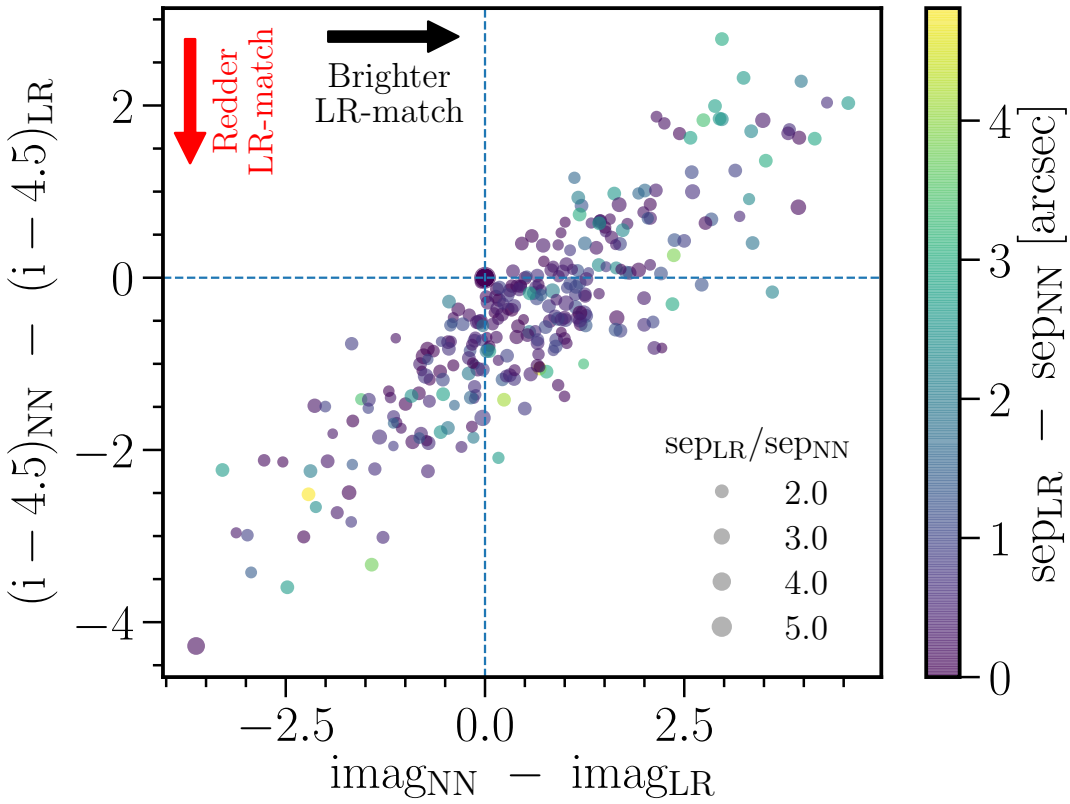

Figure 3.7 *Comparison between the LR method and a simple NN cross-match for radio sources selected by the decision tree to accept the LR-ID in ELAIS-N1. The plot shows colour (i - 4.5μm) difference versus i-band magnitude difference between a NN search and the LR method. Negative y-axis values correspond to a redder LR match compared to the NN match, and positive x-axis values correspond to a brighter LR match, as indicated by the arrows. For the vast majority (≈98%) of the radio sources used for this comparison (see Sect. 3.3.5), the LR match and the NN match are the same (indicated by the large point at (0,0)). The offset from the origin shows that, where these differ, the LR method preferentially selects sources which are either redder or brighter (or both) than the NN match. The colour of the points corresponds to the difference in separation between the LR match and the NN match* ($\mathrm{sep_{LR}} - \mathrm{sep_{NN}}$)*. The size of the points corresponds to the ratio of the separations between the LR method and the NN match (i.e. larger points indicate larger ratios; see plot legend). For radio sources where the LR match is different to the NN match, the separation to the two sources are similar.*

In Fig. 3.7, we compare the counterparts identified by the LR method with those that would be selected by a simple NN match for those radio sources that are selected by the decision tree as being suitable for the LR method in ELAIS-N1 (27056 sources). The plot shows the difference in the (i - 4.5μm) colour versus i-band magnitude difference between the NN match and the LR match. For > 98% of the sources chosen for this analysis, the LR match is also the NN match, as indicated by the cluster of points at x,y = (0,0). For 500 sources, the selected LR-ID is not the same as the NN match: the deviation of these sources from the origin illustrates the role of the LR method. In all of these cases, the LR is either redder or brighter (or both) than the NN match. In some cases where the LR match is bluer, it is always brighter with typically larger counterpart separations than the NN matches (see Fig. 3.7).

3.4 Visual classification and source association

The LR technique is not suitable for cross-identification of sources with significantly extended (large) or complex radio emission. For such sources, visual classification must be used to identify the multi-wavelength counterparts. In addition, it is more likely that for large and complex sources, the individual radio components of a given physical source may not be grouped together correctly by PyBDSF - whether that be extended emission (e.g. from radio lobes) not being grouped as a single source or multiple physical sources being grouped (blended) into a single radio source. To perform correct associations and then identifications for these sources, we use a combination of LOFAR Galaxy Zoo (described in Sect. 3.4.1) and an expert-user workflow (described in Sect. 3.4.2) based on a source's end-point from the decision tree (Sect. 3.2).

3.4.1 LOFAR Galaxy Zoo

For this task of visual classification, we use the Zooniverse framework that was adapted for the LoTSS DR1: LOFAR Galaxy Zoo (LGZ; W19). LGZ is a web based interface for performing source association and host galaxy identification by visually inspecting a given radio source using the radio data and corresponding multi-wavelength images. The user can then perform identification and association by selecting appropriate radio or optical sources on the images and answering questions about the source. The details of the LGZ interface and the choice of images and options provided to the user are almost identical to the LoTSS DR1, and are described by W19, and hence not reproduced here.

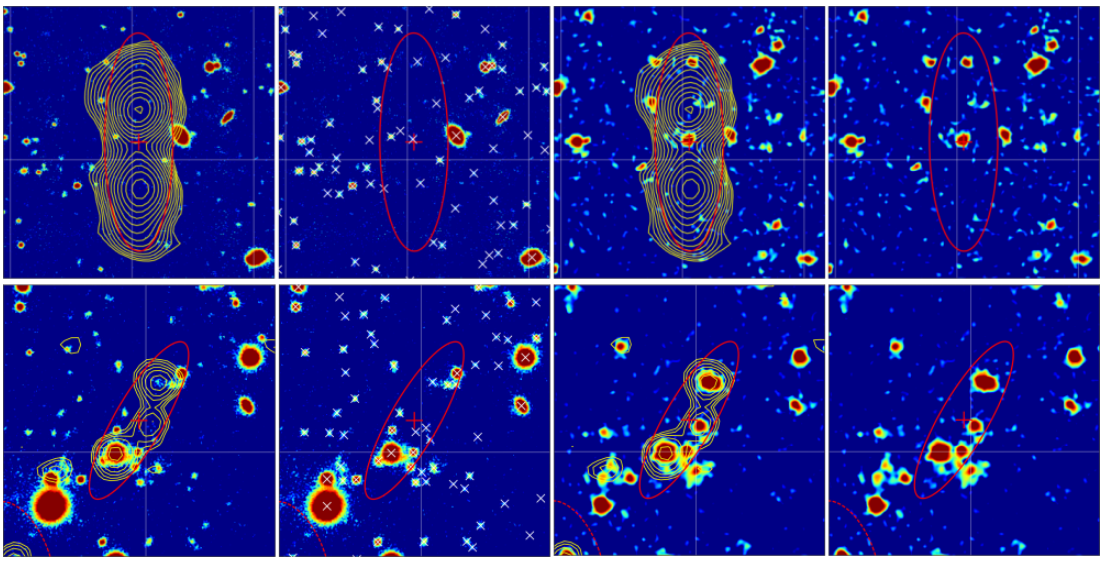

Figure 3.8 *Example set of images used for visual classification of two sources (in rows) using LOFAR Galaxy Zoo (LGZ). The radio source to be classified is in the red ellipse with its* PyBDSF *radio position marked by a red cross. The first frame shows the optical image with radio contours overlaid. The second frame shows the same optical image now without the radio contours, but with white crosses to mark multi-wavelength catalogue detections. The third frame shows the 4.5 µm image with radio contours overlaid. The fourth frame shows the same 4.5 µm image but without radio contours to aid in host galaxy identification.* ***Top:*** *An example of a large radio source.* ***Bottom:*** *An example of a blended radio source initially sent to LGZ, where the radio emission (contours) arises from three distinct physical sources that have been incorrectly grouped together into one* PyBDSF *source (red ellipse). This source was flagged as a blend during the LGZ process by the majority of volunteers and was appropriately sent to the expert-user workflow for de-blending.*

We briefly summarise the interface and the capabilities of LGZ, highlighting differences from the LoTSS DR1 approach. As with LoTSS DR1, the LGZ sample was only made available to members of the LOFAR consortium.

The user was presented with four sets of images when classifying a source. An example of the images presented for two radio sources are shown in Fig. 3.8. The first frame shows an optical image with contours of radio emission. The second frame shows the same optical image without the radio contours but with white crosses to indicate a detection in the multi-wavelength catalogue. The third frame shows the radio contours overlaid on the *Spitzer* 4.5 μm image. For the deep fields, we introduce an additional fourth frame, which is the same as the third frame but without the radio contours to aid in visual inspection. On all four frames, the PyBDSF source in question is marked

with a solid red ellipse and a red cross, while other PyBDSF sources are marked with a dashed red ellipse. The instructions given to the user for the task remain the same as LoTSS DR1. Using these four images, the user must first select any additional radio source components (i.e. dashed red ellipses) that are associated with the radio source in question. Then, the user must select all plausible multi-wavelength identifications (if any). Finally, the user must answer the following questions: Is this an artefact? Is this a radio source blend? Is the image too zoomed in? Are any of the images missing?

Each radio source sent to LGZ was classified by at least five astronomers and the output from LGZ was converted into a set of quality flags for the association and identification steps; the consensus from these classifications and flags was used to form the source associations and identifications. The details of the flags used to decide the associations are as described in W19. The questions in the final step of LGZ were asked to enable the selection of sources for which source association and/or identification could not be fully carried out and therefore may require further inspection. Sources flagged as artefacts by a majority (more than 50%) of the users were removed from the PyBDSF catalogue. Sources flagged as 'image too zoomed in' or, as 'blends' by more than 40% of users were associated separately by a single expert in the expert-user workflow (see Section 3.4.2).

If the LGZ consensus was for source association, a new source was generated by combining its constituent PyBDSF sources and the constituent PyBDSF sources were then removed from the final catalogue. We generate other radio source properties, similar to the ones in the PyBDSF catalogue (e.g. total flux, size, position, etc.) for this new source. We refer the reader to W19 for the details of this process of source association. We also note here that the LGZ association and identification takes precedence over LR identification. For example, consider a radio-AGN split by PyBDSF into three sources, one PyBDSF source consisting of only the compact core and a PyBDSF source for each of the two lobes. In such a case, it is likely that the LR method would have identified the genuine host galaxy belonging to the compact core, but the extended lobes would have been sent to LGZ, where the three components would be associated together and the host galaxy identified for the new source; this over-rides the LR identification.

3.4.2 Expert-user workflow

While testing sources that went to visual classification from initial versions of the decision tree, it was immediately apparent that there was a significant increase in the

occurrence of blends of radio sources compared to LoTSS DR1, due to the deeper radio data. It would be very inefficient to simply send such sources to LGZ. We therefore first attempt to select sources (see Fig. 3.1) that could potentially be 'blends' and send them directly to the expert-user workflow, which has de-blending functionality. Other potential blends were sent to this workflow as an output from either the pre-filter workflow (see Sect. 3.4.3) or from LGZ.

In the expert-user workflow, non-static LGZ style images were provided to a single expert, but also with information from the PyBDSF Gaussian component catalogue displayed. The expert user has the ability to split each PyBDSF source into its constituent Gaussians, which can then be associated (if needed) to generate multiple new sources. Then, multi-wavelength identification (or lack thereof) can be performed for the newly generated sources. For these de-blended sources, the final catalogue contains other radio source properties as in the PyBDSF catalogue, in this case generated from the PyBDSF Gaussian catalogue (see W19). We note that not all sources sent as potential blends to the expert-user workflow were genuine blends; for such sources, no de-blending was performed but the host galaxy identification was still carried out as part of this workflow. For a small number of cases, there are more potential distinct physical sources of emission than fitted PyBDSF Gaussians. In such cases, we only de-blend the PyBDSF source to the number of Gaussians available, selecting the most appropriate host galaxies that contributed the majority of the flux to the available Gaussians.

In addition, the expert-user workflow also has a zoom in or out functionality, and so the sources flagged as 'image too zoomed in' in LGZ (160, 96, and 60 sources in ELAIS-N1, Lockman Hole, and Boötes, respectively) were re-classified by a single expert with re-generated images using the initial LGZ classification as a starting point. The expert-user workflow was also used to identify radio source host galaxies that were missing from the multi-wavelength catalogues, and in addition, used to perform a final inspection of some large-offset LR-IDs, and all radio sources without an identification (hereafter; no-IDs), as detailed in Sect. 3.4.4 – 3.4.6. The expert-user workflow is adapted from the 'too zoomed in' and 'deblend' workflows developed for LoTSS DR1 and we refer the reader to W19 for a full description of this workflow.

Table 3.4 *Output of pre-filter workflow. Percentages are calculated based on the number of* PyBDSF *catalogue sources in the multi-wavelength overlap area (listed in Table 2.8). Sources flagged as 'Blend', 'Too zoomed in' or 'Uncatalogued host' are sent to the expert-user workflow for classification.*

Outcomes	ELAIS-N1		Lockman Hole		Boötes	
	Number	Fraction	Number	Fraction	Number	Fraction
LGZ	346	1.11%	121	0.41%	94	0.5%
Accept LR match	410	1.32%	110	0.37%	43	0.23%
No plausible match	739	2.38%	555	1.86%	320	1.71%
Too zoomed in	23	0.07%	6	0.02%	8	0.04%
Artefact	72	0.23%	15	0.05%	7	0.04%
Uncatalogued host[a]	97	0.31%	77	0.26%	102	0.54%
Blend[b]	77	0.25%	18	0.06%	4	0.02%
Total	1764	5.68%	902	3.03%	578	3.08%

[a] Uncatalogued Host: sources where the host galaxy was not detected in the multi-wavelength catalogue. These were later manually added using the expert-user workflow and forced photometry (see Sect. 3.4.4).
[b] Slightly different PyBDSF parameters adopted (accidentally) for ELAIS-N1 compared to Lockman Hole and Boötes, result in more sources being initially separated into different PyBDSF components in Lockman Hole and Boötes, and hence fewer pre-filter 'Blends' (but a higher proportion of sources needing the expert-user workflow; see Table 3.6).

3.4.3 Pre-filter workflow

In some cases, the radio source or the LR identification properties alone were not sufficient to decide if a source should be sent to LR, LGZ, or the expert user workflow for identification. Rather than send all such sources to LGZ, which is by far the most time consuming process as it requires classification of each source by five volunteers, we instead perform quick visual sorting (pre-filtering) of some stages of the decision tree prior to deciding the most appropriate workflow for counterpart identification. The aim of this pre-filtering step was to quickly assess whether: (i) the best candidate ID selected by LR is unambiguously correct (regardless of whether the LR is above or below the LR_{th}); (ii) the source needs to be sent to LGZ (this was the option used in case of any doubt, to enable a consensus decision to then be taken); (iii) the source is correctly associated but has no plausible multi-wavelength counterpart; (iv) the source is a blend, to be sent to the expert-user workflow; (v) the source is an artefact; (vi) the host galaxy detection is missing in the multi-wavelength catalogue (these sources are also sent to the expert-user workflow); or (vii) the image is too zoomed in (also sent to the expert-user workflow). In practice, for sources sent to the pre-filter workflow, static optical and 4.5 μm images showing the radio contours and the current best LR match and LR value (if any) were generated and categorised by a single expert for all three fields using a PYTHON based interface. The categorised sources were then sent to the appropriate workflows, as shown in Table 3.4. In cases where the host galaxy is missing from the multi-wavelength catalogue, we manually added these to the multi-wavelength catalogue using the process described in Sect. 3.4.4.

3.4.4 Missing host galaxies in multi-wavelength catalogues

During the visual classification steps, we noticed that the host galaxies of a small but non-negligible fraction of radio sources were present in our optical or IR mosaics but missing from our multi-wavelength catalogues (hereafter, 'uncatalogued hosts'; see Table 3.4). There were a few key reasons for this lack of detections; for example, the host galaxy being too close to bright stars where detections were typically missing (especially sources within the optical mask region), and missed '*Spitzer*-only' sources that were blended in the lower resolution *Spitzer* data. We therefore attempt to select the missing host galaxies (uncatalogued hosts) and manually add them to our multi-wavelength catalogues in each field as follows.

These sources with uncatalogued hosts were selected from each of the pre-filter, expert-

user, and LGZ workflows. In the pre-filter workflow, this was one of the options available (see Sect. 3.4.3). For LGZ, one of the outputs is the Badclick flag which indicates the number of volunteers who have clicked on an host galaxy position that is not in the multi-wavelength catalogue. Through visual inspection, we found that radio sources with Badclick > 2 typically correspond to a host galaxy which was missing in the multi-wavelength catalogue but which was sufficiently visible in the LGZ images to be identified by the volunteers. These radio sources with uncatalogued hosts were then sent to the expert-user workflow, where a single expert performed the identification and generated the coordinates of the uncatalogued hosts. Similarly, for sources that were directly sent to the expert-user workflow (e.g. as potential blends), the clicked position of the host galaxy (if uncatalogued) was also generated at the same time.

In processing host galaxy click positions from the expert-user workflow, we define uncatalogued hosts as those where the separation between the host galaxy click position and the multi-wavelength catalogue is more than $1''$. These uncatalogued hosts are then added by either searching in the full *Spitzer*-detected catalogue (which picks up *Spitzer*-only sources that were not added to the merged catalogue) or, if they are not found there, by performing forced photometry (in all filters) at the positions of the uncatalogued hosts.

3.4.5 Cleaning and inspection of large-offset LR matches

A small number of sources (140, 101, and 27 in ELAIS-N1, Lockman Hole, and Boötes, respectively) had a counterpart identified by the LR method that was significantly offset ($> 3''$) from its radio source. Such a large offset is surprising, casting doubt on whether the LR-ID is accurate; we therefore visually inspected all of these sources via the expert-user workflow to either confirm that the multi-wavelength ID found by LR method is the genuine host, or to assign the correct host galaxy (where possible). Roughly half of these sources were confirmed to be reliable; the other half were typically associated with extended sources which should not have been selected for statistical cross-matching. For these, the correct counterpart (or lack thereof) was assigned by the expert user.

3.4.6 Investigation of sources without an identification

The outputs from all of the various identification methods were joined to generate a cross-matched radio catalogue, with the correct source associations. Sources without an

identification were then visually inspected in an expert-user workflow to confirm that the lack of an ID was correct and to indicate the reason for the lack of identification. For a small fraction of sources, this was found to be in error, typically due to the source being either an artefact, a blend, or a potential host galaxy missing from the catalogue, but which had not satisfied the criteria in LGZ output for selection. Such sources were then sent to the expert-user workflow (except artefacts, which were removed) to resolve the association and identification.

For sources genuinely without an ID, a flag ("NoID") was assigned to indicate the reason for the lack of an identification. The flag values and their definitions are listed in Table 3.5 along with the numbers in each category per field. In addition to studying the nature of these sources, an advantage of assigning the NoID flag is that with upcoming spectroscopic surveys (e.g. WEAVE-LOFAR; Smith et al. 2016), a fibre could well be positioned at the position of those radio sources with secure positions to obtain spectra of (and of any emission lines from) the host galaxies where existing optical to MIR imaging data is too faint.

A large fraction (typically > 70%) of the radio sources without an identification were un-resolved (or barely resolved) sources with a secure radio source position. The host galaxy, however, was below the survey depths of our multi-wavelength dataset (albeit in some cases, low significance emission may be present). The second biggest fraction consisted of extended radio sources, with large positional uncertainties and poorly defined positions; some of these had no plausible ID whereas others had one or more plausible IDs, but none reliable enough to be chosen.

Table 3.5 *Description of the "NoID" flag values. Flag = 0 indicates that an identification is present, and the higher flag values indicate the reason for the lack of an identification.*

Flag	Description	Field		
		ELAIS-N1	Lockman Hole	Boötes
0	Source has an ID	30839	30402	18579
1	Radio source position accurate	407	392	217
2	Radio position accurate; possible faint ID but below catalogue limit	164	158	213
3	Extended (radio position may be inaccurate); no plausible ID	100	103	54
4	Radio position lies under another un-associated object	34	12	33
5	Extended source, one or more potential IDs, but none unambiguous	66	95	83

Table 3.6 *The number of radio sources in the source-associated radio-optical cross-matched catalogue and the number and fraction of sources that have an identification (or lack thereof), split by the identification method. ID fractions are calculated based on the total number of radio sources (listed at the bottom of the table) in the source-associated and cross-matched radio catalogue.*

	ELAIS-N1		Lockman Hole		Boötes	
	Number	Fraction	Number	Fraction	Number	Fraction
LR	26701	84.5%	24851	79.7%	16151	84.2%
LGZ	1966	6.2%	2395	7.7%	1058	5.5%
Expert-user	2172	6.9%	3156	10.1%	1370	7.1%
Total-ID	30839	97.6%	30402	97.6%	18579	96.9%
No-ID	771	2.4%	760	2.4%	600	3.1%
Total	31610		31162		19179	

3.5 Final cross-matched catalogues

The final cross-matched and associated catalogue in ELAIS-N1 contains 31610 radio sources, with host galaxies identified for 97.6% of these. Similarly, there are 31162 sources in Lockman Hole with host galaxies found for 97.6%, and 19179 sources in Boötes with host galaxies identified for 96.9%. These properties, along with the number of sources (and the fraction) identified by each method, are listed in Table 3.6.

Compared to similar cross-matching efforts in the literature, for example in the ELAIS-N1 field by Ocran et al. (2019) using 610 MHz GMRT observations, we find a higher cross-identification rate by > 5%. We also note the larger (by ∼3%) fraction of sources requiring visual classification (expert-user and LGZ) in Lockman Hole, and a similar decrease in the fraction of sources where the LR identification was accepted compared to the other two fields. This is likely due to the slight difference in the PyBDSF source extraction parameters used for Lockman Hole (and for Boötes), where a significantly larger fraction of the sources were fitted with multiple Gaussian components, resulting in ambiguity in the decision tree and requiring source association or de-blending. The effect of the difference in the PyBDSF source extraction parameters is less prominent in Boötes, likely due to the shallower radio data depth. It is important to note that these differences should not affect the final source-associated, cross-matched catalogues, but simply result in a difference in the method of the identification: visual classifications were more often used to form the correct source associations for sources split into multiple Gaussians.

The LOFAR Deep Fields value-added catalogue released contains properties of the correctly associated radio sources, and the multi-wavelength counterpart identifications and properties (where available).

The radio source properties are as follows:

- The IAU source identification ("Source_Name") based on source position.
- Radio source position and uncertainties ("RA", "E_RA", "Dec", and "E_Dec").
- Radio source peak and total flux densities and corresponding uncertainties ("Peak_flux", "E_Peak_flux", "Total_flux", "E_Total_flux").
- Ellipse shape parameters and corresponding uncertainties ("Maj", "Min", "PA", "E_Maj", "E_Min", "E_PA"). These are blank for associated sources; see below for properties of associated sources. For de-blended sources, these are taken from the PyBDSF Gaussian catalogue.
- A code to define the source structure ("S_Code"; 'S' = single-Gaussian, 'M' = multi-Gaussian, 'Z' = associated/compound source)
- Overlap bit flag indicating the coverage of the multi-wavelength surveys at the radio source position ("FLAG_OVERLAP_RADIO"). See Table 2.8 for the recommended flag values.
- Bright star masking flag indicating masked and un-masked regions in the *Spitzer*- and optical-based bright star mask ("FLAG_CLEAN_RADIO"), based on radio position.

Associated sources have additional radio source properties given by:

- Ellipse shape parameters for associated sources ("LGZ_Size", "LGZ_Width", "LGZ_PA")
- Gaussian de-convolved shape parameters ("DC_Maj", "DC_Min", "DC_PA")
- Number of PyBDSF source components associated ("Assoc")
- Quality flag of the association ("Assoc_Qual")

The multi-wavelength identification (if any) and host galaxy properties are as follows:

- Unique identifier of the ID to the multi-wavelength catalogue ("ID")

- multi-wavelength ID source position ("ALPHA_J2000", "DELTA_J2000")
- Aperture and extinction corrected fluxes (and flux errors) from our recommended aperture size (<band>_flux_corr and <band>_fluxerr_corr) in μJy.
- Aperture and extinction corrected magnitude (and magnitude errors) from our recommended aperture size (<band>_mag_corr and <band>_magerr_corr) in the AB system.
- Overlap bit flag indicating the coverage of the multi-wavelength surveys at the counterpart source position ("FLAG_OVERLAP"). See Table 2.8 for the recommended flag values.
- Bright star masking flag indicating masked and un-masked regions in the *Spitzer*- (= 3) and optical- (= 1) based bright star mask ("FLAG_CLEAN")
- The maximum LR match (if an ID is present, and if the ID is obtained from the LR method; "lr_fin")
- multi-wavelength ID position based E(B − V) reddening values from Schlegel et al. (1998) dust map ("EBV").
- Flag indicating reason for lack of identification ("NoID"; '0' = an identification exists). Flag definitions are listed in Table 3.5.

Additional columns pertaining to the photometric redshifts, rest-frame colours, absolute magnitudes, and stellar masses are described in Duncan et al. (2021), and columns relating to the far-infrared data are described in McCheyne et al. (subm). For full details of all columns presented, please see the accompanying data release documentation.

Fig. 3.9 shows the number of sources (top panel) and the identification fraction (bottom panel) as a function of the radio flux density and the identification method. The fraction of sources requiring LGZ for identification decreases from 100% at the brightest fluxes down to well below 5% at the faintest fluxes. There is a transition of the dominant method of identification at ∼10 mJy from LGZ to the LR method, which accounts for ≳90% of the sources at the faintest fluxes. This higher identification rate by the statistical method showcases the power of the ancillary data available in these deep fields compared to the shallower LoTSS DR1. The expert-user method plays a sub-dominant role across most of the flux density range but, as a result of the depth of the radio data and consequently increasing number of blends, begins to dominate the identification rates achieved from the visual methods at the faintest fluxes, and therefore corresponds

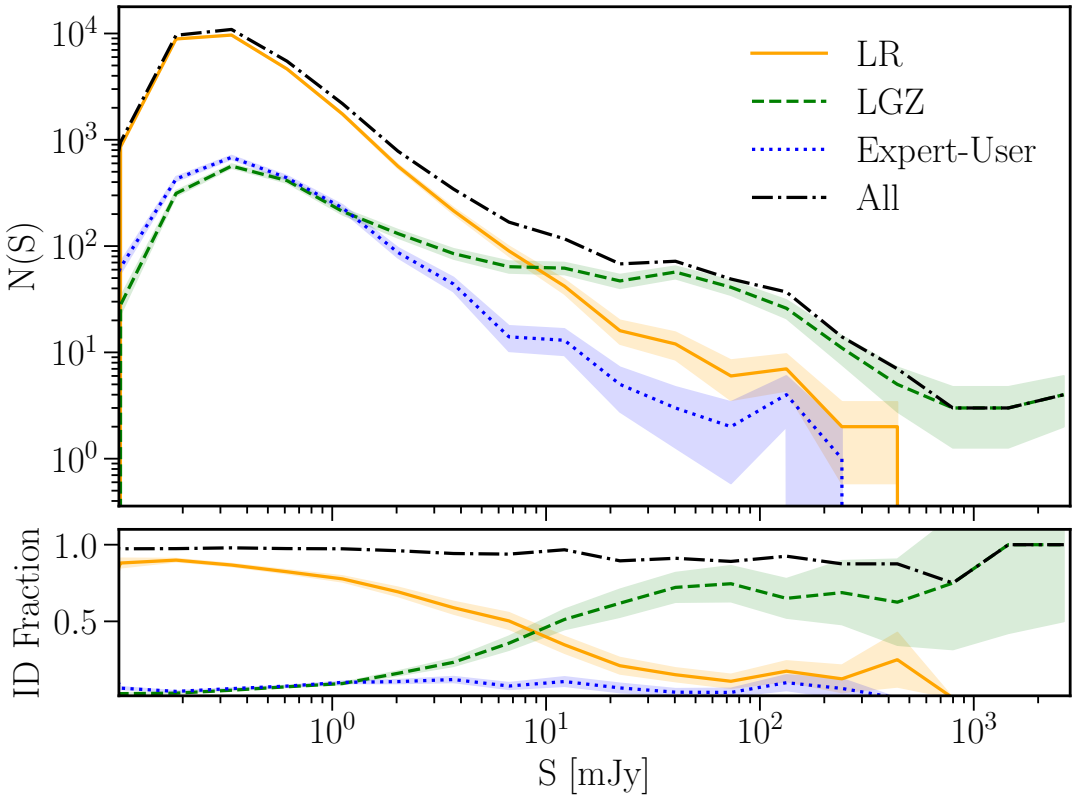

Figure 3.9 *Top: The number of radio sources with identifications as a function of radio flux density and the identification method used (LR, LGZ or 'expert user workflow'). The flux density distribution of all sources with identification is shown by the dot-dashed black line. Bottom: The identification fraction as a function of the flux density, also split by the identification method. The identification fraction is computed based on the total number of radio sources (with or without an identification). The LGZ method dominates the identification rate above ~10mJy, with the LR method dominating below this. The filled regions show Poisson error estimates.*

to a significant number of sources within our sample (as shown in Table 3.6). The trend of decreasing overall identification rate (see Fig. 3.9) with decreasing radio flux densities noted by W19 (down to ∼ 1 mJy in W19) in LoTSS-DR1 (see fig. 8 of W19) is not observed in these deep fields. This is expected as this decrease in identification rate was attributed to the shallow PanSTARSS and WISE data available for cross-matching by W19. In the LoTSS Deep Fields, although the typical redshift of sources probed increases with decreasing radio flux density, the significantly deeper multi-wavelength data available allows us to effectively identify counterparts down to lower radio flux densities than LoTSS-DR1, where the LR method starts to dominate the identification rates.

3.6 Properties of host galaxies

In Fig. 3.10, we show the magnitude distribution of all counterparts identified in ELAIS-N1 for a range of optical (SpARCS and PS1) to IR bands (UKIDSS and SWIRE). Also shown in shaded regions are the distributions for the subset of radio sources with 150 MHz radio flux densities > 1mJy (pink) and > 10 mJy (blue). The top value listed in each panel is the percentage of all radio sources in ELAIS-N1 that have a counterpart detected within that given band. The second and third values provide the corresponding percentages for the number of radio sources with radio flux densities > 1mJy and > 10mJy, respectively, that have a counterpart detected in that band.

For the 3.6 and 4.5 μm channels, the magnitude distribution of the counterparts is clearly peaked at a magnitude of 19–20, and declines towards fainter magnitudes. This is well within the detection limit of the *Spitzer* survey, illustrating that the vast majority of radio counterparts are detected by SWIRE and SERVS, as indicated by the high (96%) identification rates in these two channels. In the NIR filters, the distributions show a broad peak around 20th–21st magnitude, turning over close to the magnitude limit of the UKIDSS survey. In contrast the distributions in the bluer (optical) filters show no signs of turning down at the faintest magnitudes probed, consistent with the lower identification rates achieved in these filters being limited by the depth probed by the available optical surveys. This trend in the shape of the magnitude distributions can also be seen by the increase in identification rate achieved with wavelength, increasing from ∼ 56% to 96% from the u-band to the 4.5 μm band. Even in the bluer filters, however, the faint end of the distributions flatten, as compared to the well-known monotonic increase of the number counts of all galaxies towards fainter magnitudes, indicating that radio galaxies are preferentially hosted in brighter galaxies (a result which motivates

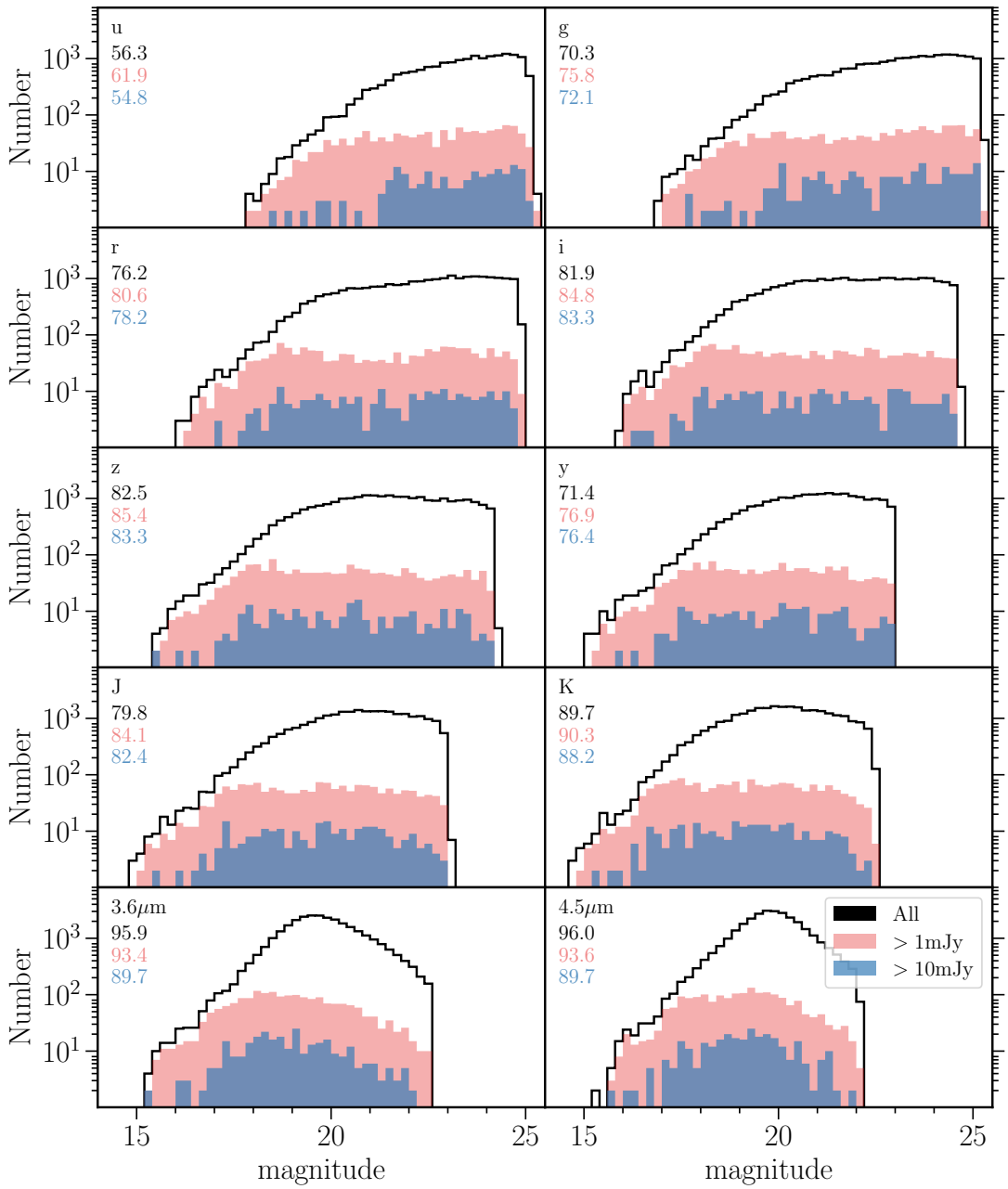

Figure 3.10 *Magnitude distributions of the host galaxies of radio sources in optical (SpARCS and PanSTARRS) to IR bands in ELAIS-N1. Each panel also shows the same for the subset of radio sources with radio flux densities >1mJy (pink) and >10mJy (blue). The numbers in each panel corresponds to the fraction of all radio sources with a counterpart detected in that band, for all sources, and for >1 mJy and >10 mJy sources. The identification fraction increases with wavelength from ~56% in u-band up to 96% in 4.5 μm.*

the LR approach).

Interestingly, the distributions for the >1 mJy sources peak at brighter magnitudes than those of 'All' sources, suggesting that higher flux density radio sources even more strongly favour brighter host galaxies. Comparing the difference between the distributions of 'All' sources and > 1mJy sources, it is clear that at faint optical and IR magnitudes, the radio sources with flux densities below 1 mJy dominate the population. In contrast, at bright optical and IR magnitudes, the majority of the radio population has flux densities above 1 mJy. This trend may be driven by a shift in the dominant radio source population below ~1 mJy, where we expect a significant fraction of both nearby star-forming galaxies and high redshift (obscured) radio quiet quasars (Wilman et al., 2008), which are likely hosted by fainter optical galaxies.

In conclusion, in this Chapter, I have used the new multi-wavelength catalogues generated in Chapter 2 to identify the counterparts and form correct source associations of the radio-detected sources in the three LoTSS Deep Fields. This was performed using a combination of the statistical colour-based Likelihood Ratio method and an extensive visual classification scheme, with a decision tree developed to identify the most suitable method of identification for each source. This process results in a catalogue of 81 951 radio-sources with counterparts identified for over 97% of these. These catalogues are then used for the analysis of radio-AGN in the following Chapter.

Chapter 4

Evolution of Radio-Loud AGN in LOFAR Deep Fields

4.1 Introduction

In this Chapter, we present the first robust measurement of the LERG luminosity functions out to $z \sim 2.5$ and study the cosmic evolution of their host galaxy properties using a sample of over 10 000 LERGs constructed by using the value-added catalogues from the LoTSS Deep Fields generated in the previous two chapters of this thesis. Photometric redshifts from Duncan et al. (2021) and SED fitting from Best et al. (in prep) were used for this analysis. Our combination of deep and wide radio and multi-wavelength datasets is ideal for probing much fainter luminosities out to higher redshifts than previous studies (e.g. Best et al., 2014; Pracy et al., 2016; Williams et al., 2018; Butler et al., 2019) and also allows a better sampling of the bright end of the luminosity function compared to previous deep observations over small areas (e.g. Smolčić et al., 2017a,b), while limiting the effects of cosmic variance. In this study, we focus on the evolution of the LERG population as our deep radio dataset is particularly well-suited to sample the low-luminosity AGN population, which are dominated by LERGs; this allows us to characterise the evolution of this population in unprecedented detail.

4.2 Description of data

In this Chapter, I use the radio data from the LoTSS Deep Fields, covering the ELAIS-N1, Lockman Hole, and Boötes fields, as presented by Tasse et al. (2021) and Sabater et al. (2021), and described in Chapter 2. These data, together with the new multi-wavelength catalogues generated in Chapter 2 and the radio cross-matched value added catalogues generated in Chapter 3 form the basis of the dataset used for analysis in this Chapter.

4.2.1 Photometric redshifts and spectral energy distribution fitting

Building on the robust multi-wavelength photometric catalogues in each field, photometric redshifts for the full multi-wavelength catalogues are presented by Duncan et al. (2021). These are generated using a hybrid approach, combining template fitting and machine learning methods, developed for the next generation of radio surveys (see Duncan et al., 2018a,b). In this chapter, we use spectroscopic redshifts where available and reliable for a small fraction of the radio-sources (5%, 5%, and 22% of all radio-sources in ELAIS-N1, Lockman Hole, and Boötes, respectively), otherwise the photometric redshifts are used.

Spectral energy distribution (SED) fitting was performed using four different SED fitting codes for all LoTSS Deep Fields sources with a counterpart that satisfy the quality cuts defined in Chapters 2 and 3 and Duncan et al. (2021); this process is described in detail by Best et al. (in prep.).

The SED fitting was performed using this same input catalogue using each of AGNFitter (Calistro Rivera et al., 2016), Bayesian Analysis of Galaxies for Physical Inference and Parameter Estimation (bagpipes; Carnall et al., 2018), Code Investigating Galaxy Evolution (cigale; Burgarella et al., 2005; Noll et al., 2009; Boquien et al., 2019), and Multi-wavelength Analysis of Galaxy Physical Properties (magphys; da Cunha et al., 2008). One of the key differences between the SED fitting codes employed here is that AGNFitter and cigale, unlike the other two routines, are also able to model emission from AGN which imprint features in the mid-infrared regime in particular. This is done by incorporating models for the accretion disc and dusty torus surrounding the AGN in both SED fitting codes; for cigale, this includes a run based on the Fritz et al. (2006) AGN models, and a separate run based on the skirtor (Stalevski et al.,

2012, 2016) prescription which assumes a clumpy two-medium torus model (rather than a smooth torus structure; see Boquien et al., 2019; Yang et al., 2020). The other key difference between the different codes is that BAGPIPES, MAGPHYS, and CIGALE all enforce an energy balance between the UV-optical emission from starlight absorbed by dust and the re-radiated emission in the infrared; this ensures that the spectral distributions of galaxies are physically consistent. AGNFITTER does not enforce this, for example to account for cases where the UV emission is spatially offset from the dust emission (e.g. Calistro Rivera et al., 2018). The implications of this when comparing our results with previous AGNFITTER-based work from Williams et al. (2018) are discussed in Sect. 4.4.

The output from the four different SED fitting routines were combined to generate *consensus* estimates of physical galaxy properties; the key parameters relevant for analysis in this study are the stellar masses and star-formation rates (SFRs). This process is described in detail by Best et al. (in prep.) but in summary, for sources that showed no signatures of a radiative-mode AGN, the stellar masses and SFRs were estimated by taking the average of the MAGPHYS and BAGPIPES results (accounting for their goodness of fits); both stellar masses and SFRs agree very well with each other other, with a small scatter (see Best et al. in prep.). For sources that showed signs of a radiative-mode AGN (also referred to as 'optical-AGN' hereafter; see below), the stellar masses and SFRs were taken by averaging the two CIGALE runs with the Fritz et al. (2006) and SKIRTOR AGN models, provided a good fit was found; AGNFITTER was found to result in less reliable fits and was hence excluded from this step. We use these *consensus* values determined by Best et al. (in prep.) throughout this chapter unless otherwise stated.

The 'optical-AGN' (i.e. AGN showing emission from the accretion disc or torus), were identified using the outputs of the four SED fitting routines. Firstly, Best et al. (in prep.) defined a diagnostic based on the $f_{\rm AGN}$ parameter, which corresponds to the ratio of MIR luminosity that arises from AGN components compared to that from the stellar galaxy component, as fitted by both AGNFITTER and CIGALE. In particular, Best et al. used the $16^{\rm th}$ percentile of this value, to avoid bias from objects with large uncertainties on $f_{\rm AGN}$. Secondly, Best et al. also considered the goodness of fit estimates from BAGPIPES and MAGPHYS (which do not include AGN components) compared to that from CIGALE and AGNFITTER; the latter two codes should find a better goodness of fit for sources with significant AGN contribution. Based on the combination of the $f_{\rm AGN,16^{th}}$ parameters and the relative goodness of fit values, Best et al. identified the likely 'optical' (radiative-mode) AGN. Finally, a small subset of the LOFAR-detected sources have either optical spectroscopy indicating the presence of a typical AGN or

X-ray observations (see Duncan et al., 2021); most of these were already identified as 'optical-AGN' by the method above, but any additional sources were also added to this sample.

4.2.2 Selection of LERGs and HERGs

Radio continuum surveys detect both synchrotron emission associated with jets from AGN, and also emission from supernovae indicative of star-formation activity. The faint radio source population, especially at $S_{150\,\mathrm{MHz}} \lesssim 1$ mJy, is expected to be dominated by star-forming galaxies (Wilman et al., 2008). Source classification of the LOFAR-detected sources was therefore performed to separate star-forming galaxies from different classes of AGN using the outputs from the SED fitting process. This process is described in detail by Best et al. (in prep.) and summarised below.

As low-frequency radio observations trace cosmic ray electrons from supernovae from massive (recently formed) stars, there is a well-known relation between the radio luminosity and SFR for star-forming galaxies (e.g. Calistro Rivera et al., 2016; Gürkan et al., 2018; Smith et al., 2021), with a correlation also seen between the far-infrared and radio luminosities (far-infrared radio correlation; FIRC). We selected radio-AGN (also known as 'radio-excess AGN') as sources that show excess radio emission (> 0.7 dex, $\approx 3\sigma$) compared with that expected from star-formation processes alone using a ridgeline analysis by Best et al. (in prep.). To this sample, we added the small fraction of sources that are resolved and show extended radio emission that is associated with jets from the AGN even if they show a radio-excess < 0.7 dex. The other sources without such radio-excess are largely star-forming galaxies (SFGs) but can also include 'radio-quiet AGN' which are known to broadly follow the same FIRC (e.g. Bonzini et al., 2015). Both of these groups are excluded from this analysis as we are focused on the radio-loud AGN population in this study.

As discussed in the introduction, the total radio-AGN population consists of two classes of AGN: high-excitation radio galaxies (HERGs) and low-excitation radio galaxies (LERGs). HERGs display strong emission lines indicating the presence of an accretion disc and dusty obscuring structure; these sources are consistent with being the radio-loud subset of the radiative-mode AGN population. The LERGs show powerful radio emission from the AGN but little-to-no evidence of having an accretion disc or obscuring structure, and as such not identified as AGN at other wavelengths. Using the above definitions of the two classes of AGN, in this study, we define LERGs as sources that host a 'radio-excess AGN' but not an 'optical-AGN' (based on the SED fitting;

see Sect. 4.2.1). Likewise, HERGs are defined as sources that are classed as both 'radio-excess AGN' and 'optical-AGN'.

The final sample of LERGs and HERGs used in this chapter and for the construction of the luminosity functions is limited to sources that have a radio detection with peak flux density at $> 5\sigma$ level based on the local RMS, and excludes sources that might be masked in optical bright-star masks (i.e. FLAG_CLEAN = 1; see Chapter 2 for details) to ensure a clean and robust sample. In addition, we limit our analysis of the luminosity functions to $0.03 < z \leq 2.5$; at lower redshifts the volume sampled by LoTSS-Deep is small and there may be incompleteness due to the larger angular size of the nearby sources, and at higher redshifts, this reaches the limits of the multi-wavelength datasets available in these fields, beyond which the source classifications of Best et al. from the SED fitting process become less secure. The above criteria result in a sample of 10 429 LERGs (of which 4979 are hosted by quiescent galaxies and 5450 are hosted by star-forming galaxies; see Sect. 4.6), and 1037 HERGs, across the deep fields.

Fig. 4.1 (*top*) shows the 150 MHz radio luminosity as a function of redshift for the sample of LERGs from the LoTSS Deep Fields. The red line corresponds to the 5σ detection limit calculated based on the noise in the central region of ELAIS-N1, $\sigma = 20\,\mu$Jy, and a radio spectral index $\alpha = -0.7$. Fig. 4.1 (*bottom*) shows the redshift distribution for the sample of LERGs (split in to those hosted by quiescent and star-forming galaxies; see Sect. 4.5.1) and HERGs.

4.3 Total Radio-AGN luminosity functions

4.3.1 Building the luminosity functions

We calculate the radio luminosity of our sources by assuming that the radio spectrum is described by a simple power law in frequency, ν, with $S_\nu \propto \nu^\alpha$, where S_ν is the radio flux density at frequency ν and $\alpha = -0.7$ is the assumed spectral index throughout this study. The radio luminosity L_ν can be computed using the radio flux density S_ν as

$$L_\nu = \frac{4\pi D_{\rm L}^2(z)}{(1+z)^{1+\alpha}} S_\nu \tag{4.1}$$

where $D_{\rm L}$ is the luminosity distance to the source and the $(1+z)^{-(1+\alpha)}$ term accounts for the radio spectrum k-correction.

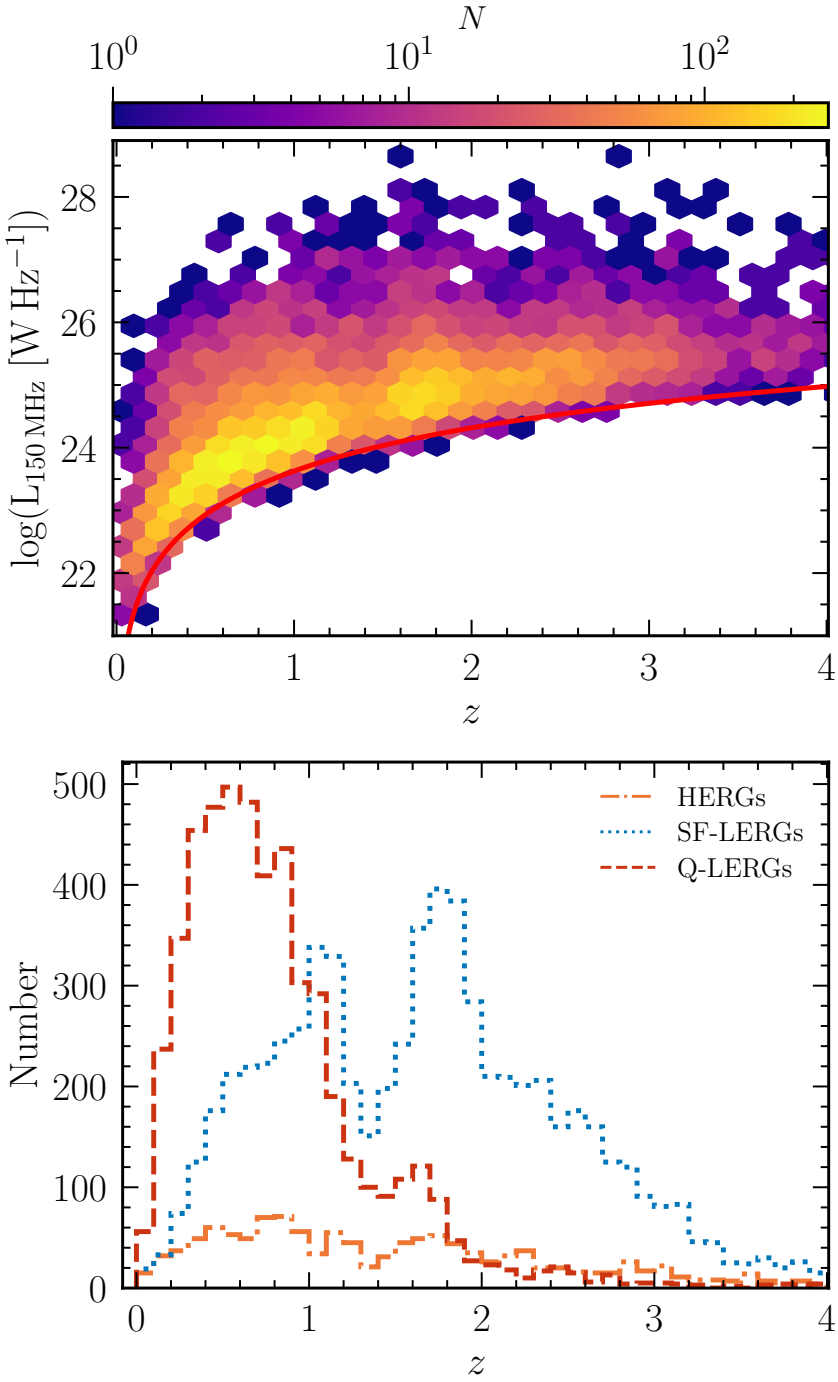

Figure 4.1 *Top: Radio luminosity as a function of redshift for the LERGs detected in the LoTSS Deep Fields. The colourbar with a logarithmic scale shows the number of sources across the parameter space. The red line shows the 5σ detection limit using the $\sigma = 20\mu$ Jy (achieved in the central region of ELAIS-N1) and assuming a spectral index $\alpha = -0.7$. Bottom: Histogram of redshifts for the sample of LERGs split into those hosted by quiescent (red dashed line) and star-forming (blue dotted line) galaxies (see Sect. 4.5.1), along with the HERGs (orange line), selected in the LoTSS Deep fields.*

The luminosity functions (LFs) were built using the standard $1/V_{\rm max}$ technique (Schmidt, 1968; Condon, 1989), which weights each source in the sample by the maximum volume that the source could be observed in, given the potential redshift range, and still satisfy all selection effects to be included in the sample. This is particularly important for surveys like the LoTSS Deep Fields where the RMS in the radio images (and hence the 5σ flux density limit) varies as a function of the position in the field due to primary beam effects, increased RMS around bright sources, and facet-to-facet variations in calibration. The luminosity function $\rho(L, z)$ gives the number of sources per unit comoving volume observed per unit of log luminosity, and is given by

$$\rho(L, z) = \frac{1}{\Delta \log L} \sum_{i=1}^{N} \frac{1}{V_{max,i}} \tag{4.2}$$

where $V_{max,i}$ is the maximum volume that the source i could be observed in, within the limits of the multi-wavelength surveys, and $\Delta \log L = 0.3$ is the luminosity bin width in log-space and the sum is calculated over all sources in a given luminosity and redshift bin.

The $V_{max,i}$ for a given source is then computed as

$$V_{max,i} = \int_{z_{min}}^{z_{max}} V(z)\, \theta(S, z)\, dz, \tag{4.3}$$

where $V(z)dz$ is the comoving volume across the whole sky between redshift z and $z + dz$, $\theta(S, z)$ is the fractional sky coverage that accounts for the non-uniform radio-map noise and radio flux density incompleteness, and S is the radio flux density that a source i with a given intrinsic luminosity would have at redshift z. In practice, we evaluate this integral numerically, with a step size of $\Delta z = 0.0001$ between $z_{\rm min}$ and $z_{\rm max}$, the minimum and maximum range of the redshift bin. At each redshift, $\theta(S, z)$ for each source i is given as

$$\theta(S, z) = \frac{\Omega[S(z)]}{4\pi} \times C_{\rm radio}[S(z)] \tag{4.4}$$

where $\Omega[S(z)]$ is the solid angle (in units of sr) of the survey area in which a source i with a flux density S can be detected at 5σ given the non-uniform radio image noise, and $C_{\rm radio}[S(z)]$ is the radio flux density completeness correction at flux density S. The $\Omega[S(z)]$ term is computed by performing a linear interpolation using the cumulative area versus noise plot for each field, shown in Fig. 4.2. Completeness corrections are required to account for the faint undetected radio sources close the survey detection limit, otherwise leading to an underestimate of the space density; these are discussed in detail in Sect. 4.3.3.

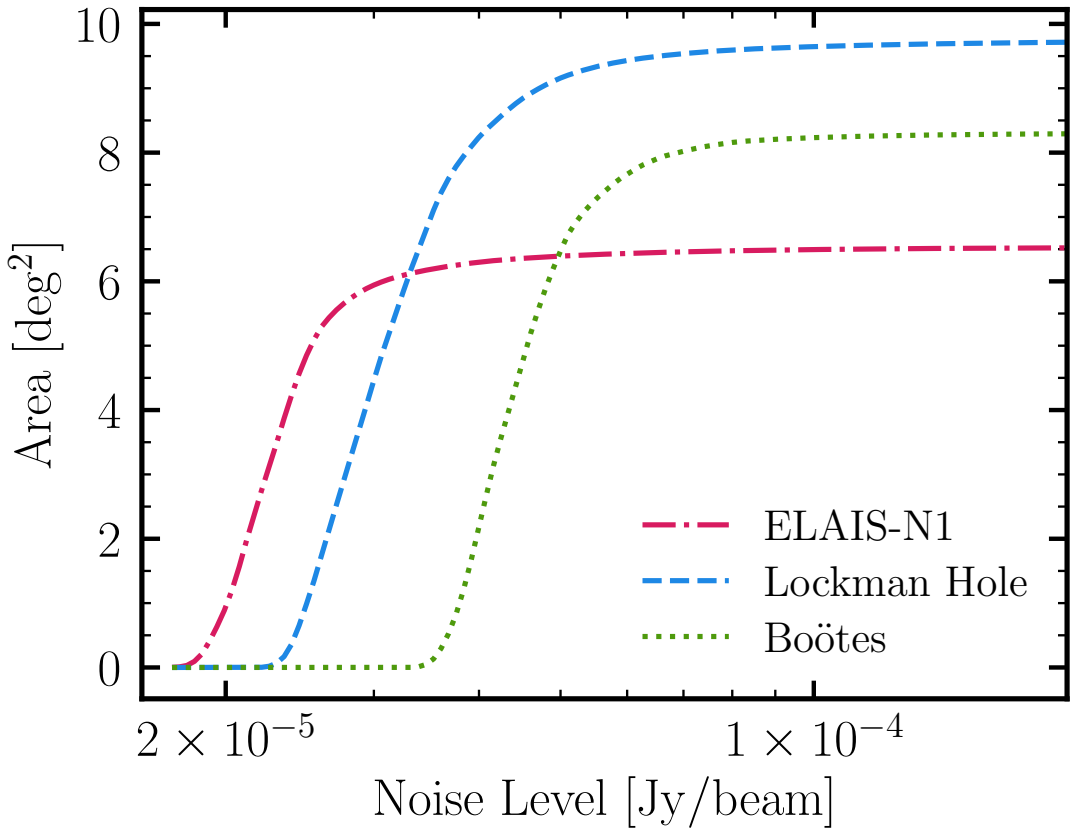

Figure 4.2 *Cumulative area covered down to a given noise level in the three LoTSS Deep Fields. The area is computed after applying the optical-based bright-star masks and only including the regions with coverage from the best available multi-wavelength surveys in each field (see Chapter 2). The total area across the three fields covers ≈ 24.6 deg^2.*

In constructing the LFs, we only consider the radio sources that have a host galaxy counterpart; this corresponds to 97% of the radio sources in the LoTSS Deep Fields (see Chapter 3). For the remaining 3% of the sources, the lack of a multi-wavelength counterpart identification means that no photometric redshift was available, and hence SED fitting and subsequent source classification was not possible. In Chapter 5, I examine the nature of these sources, including an optical to mid-IR stacking analysis of the subset of these sources with secure radio positions (which forms the majority) and find that these sources are likely to be predominantly $z > 3$ radio-AGN. Therefore, we expect that this will have a negligible effect on the derived LFs in this chapter as we only consider sources with $z \leq 2.5$.

To compute the uncertainties on the luminosity functions, we performed bootstrap sampling (random sampling by replacement) of the catalogue to generate a distribution of 1000 realisations of the luminosity function. The lower and upper 1σ uncertainties on our luminosity functions are then determined from the 16^{th} and 84^{th} percentiles of the bootstrap samples. For the faint luminosity bins, where the samples are large and the uncertainties computed from bootstrapping correspondingly small, the uncertainties are likely to be dominated by other factors such as the photometric redshift errors and source classification uncertainties. Therefore, we set a minimum uncertainty of 0.03 dex in the luminosity functions reported, based on the ~ 7 per cent photometric redshift outlier fraction in these fields.

4.3.2 Optical/IR magnitude limits on V_{max}

In the calculation above, we have assumed that the radio flux density sets the limit on the V_{max} for a given source. It is, however, possible that the optical/IR dataset may set the limits on the maximum observable volume for a source. In this section, we consider the effects of applying optical/IR magnitude completeness limits in the calculation of V_{max} on the evolution of the LERG LFs.

The multi-wavelength catalogues, using which the photometric redshifts and SED fitting are determined, are based on source detection from χ^2 images; as such there is no well-defined magnitude limit in the optical or IR bands. Instead, we apply a 5σ limiting magnitude in the IRAC $4.5\,\mu$m band, chosen to be the band in which $> 96\%$ of the radio sources are detected (see Chapter 3), in addition to the radio-based selection criteria in Sect 4.2.2. This effectively corresponds to having a multiplicative term in equation 4.3, $f(m)$ that equals unity if $m < m_{\text{lim}}$ and zero if $m > m_{\text{lim}}$, where m_{lim} is the magnitude limit. To compute this, we first calculate the absolute $4.5\,\mu$m magnitude, $M_{4.5\,\mu\text{m}}$ for

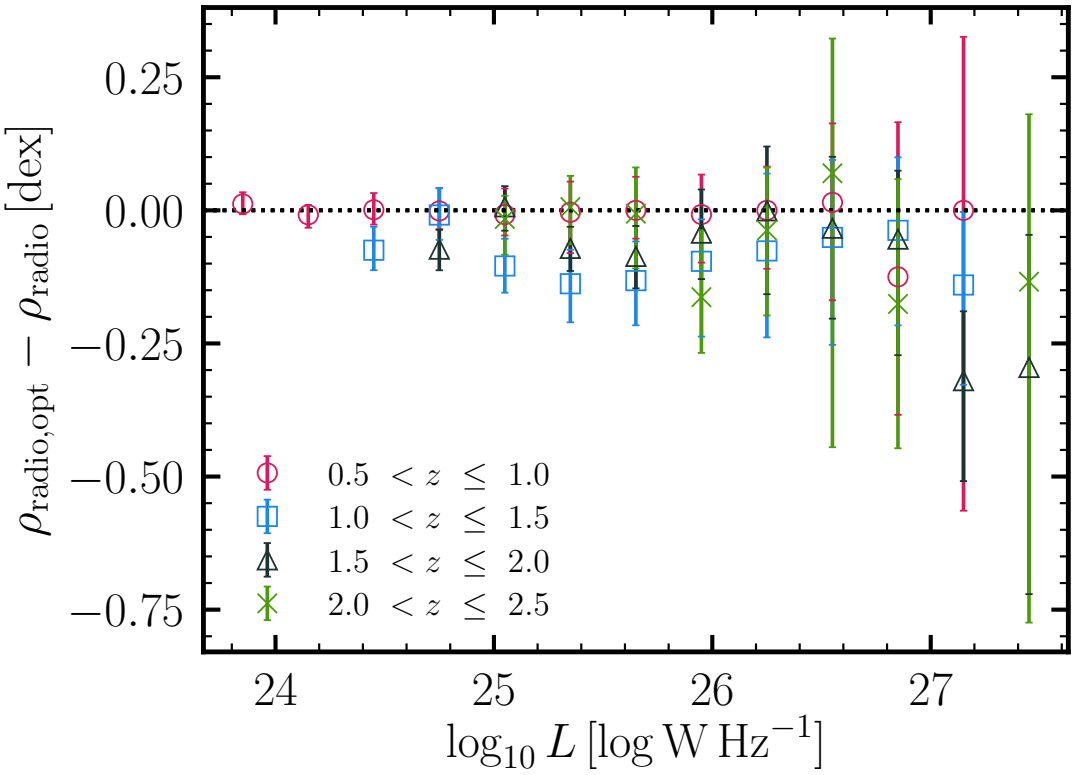

Figure 4.3 *Difference between the LERG LF calculated with ($\rho_{\rm radio,opt}$) and without ($\rho_{\rm radio}$) an optical magnitude limit applied to the $V_{\rm max}$ calculation. We find good agreement between the two determinations of the LFs across the redshifts examind in this study.*

each source at its redshift z, including a k-correction,

$$M_{4.5\,\mu{\rm m}} = m_{4.5\,\mu{\rm m}} - DM - K_{4.5\,\mu{\rm m}}(z) \tag{4.5}$$

where, DM is the distant modulus and $K_{4.5\,\mu{\rm m}}(z)$ is the k-correction. We calculated two sets of k-corrections based on typical star-forming and quiescent galaxies, forming two extremes of the corrections. For the former, we consider a template with a recent burst of star-formation occurring 50 Myrs ago computed using the stellar population models of Bruzual & Charlot (2003), the Chabrier (2003) initial mass function (IMF) and the Calzetti et al. (2000) attenuation curve and an extinction of $A_{\rm V} = 3$. For the latter, we considered a galaxy with an old stellar population (formed at $z = 12$) with an exponentially declining star-formation rate with a characteristic time of 100 Myr.

Then, for a given redshift bin, we evaluate $m_{4.5\,\mu{\rm m}}$ and $\theta(S, z)$ at small intervals ($\delta z = 0.001$; see equation 4.3) between $z_{\rm min}$ and $z_{\rm max}$ until the source i fails to satisfy the 5σ 4.5 μm limit and the radio selection criteria to determine the total comoving volume that

the source i could be observed in, $V_{\rm max,i}$.

We refer to the LFs determined using the above method, incorporating both radio and optical selection criteria, as $\rho_{\rm radio,opt}$, and the LFs determined by considering the radio selection effects alone (i.e. the LFs used throughout this chapter) as $\rho_{\rm radio}$. In Fig. 4.3, we show the difference between these LFs as a function of radio luminosity for the LERGs in found broad redshfit bins that are also used for the analysis of the LERGs in Sect. 4.4. At $z < 1$, we find a remarkably good agreement between the two LFs, and even at higher redshift $z > 2$, where we find median differences of ~ -0.1 dex, the LFs determined are still in good agreement with respect to the uncertainties. Moreover, we found that the maximum observable volume for almost all of the sources is set by the radio dataset, and hence the similarities in the two LFs is not surprising; any differences seen in the LFs here are due to radio sources simply not satisfying the additional 5σ 4.5 μm selection in the first instance. The $\rho_{\rm radio,opt}$ space densities shown in the plot were calculated using k-corrections based on a typical quiescent galaxy, however we find that we get very similar results, within the uncertainties, when using k-corrections based on the star-forming galaxy template.

It is also worth noting that the above calculation of $\rho_{\rm radio,opt}$ assumes that the multi-wavelength catalogues are complete down to the 5σ limit chosen, which is not necessarily accurate. Application of a completeness correction to the multi-wavelength data would only increase the space densities, bringing $\rho_{\rm radio,opt}$ more in line with $\rho_{\rm radio}$. Moreover, a proper treatment of the optical/IR magnitude limits requires consideration on the effects on the photometric redshift accuracy and source classification which is beyond the scope of this chapter. Therefore, in order to retain a cleaner selection function, we only considered the effects of the radio selection in determining the space densities of the AGN population in this chapter.

4.3.3 Completeness Corrections

The radio flux density completeness corrections are generated by performing simulations of inserting mock sources of various intrinsic source-size and flux density distributions in to the radio image, and then recovering them using the same PyBDSF parameters as that used for the real sources (see Sabater et al. 2021 for the PyBDSF parameters used).

We simulated mock sources at fixed total flux density intervals separated by 0.2 dex in the range $0.4 < S < 40$ mJy and, using a finer sampling interval of 0.081 dex in the range $90 < S < 400\,\mu$Jy to better probe the steep part of the curve. For the Lockman

Hole and Boötes fields, we used the same sampling intervals but only simulating sources down to $\sim 110\,\mu$Jy and $\sim 190\,\mu$Jy, respectively, due to the slightly shallower depth of the radio data. The flux intervals used are listed in Table 4.1.

For each field, we simulated 120 000 – 150 000 mock sources to sample the full range in (total) flux density and source-size parameter space, while also obtaining robust statistics for the bright and extended rare sources. Practically, this was done by inserting 1000 mock sources with convolved sizes between 6″ – 30″, where 6″ is the size of the LOFAR beam, for a given (total) flux density value into the radio image and extracting the sources using PyBDSF, with this step repeated many times (see Sec. 4.3.3.1 for details). The injected sources were modelled as Gaussians, although the structure of real sources may be more complex. We ensure that a mock source is placed at least twice its FWHM (along the major axis) away from other mock sources and real radio sources to avoid source overlapping, which can complicate the process of determining if a simulated source has been detected. A mock source is defined as being 'detected' if there is a nearest neighbour match within 2″ in the extracted PyBDSF source catalogue. This angular separation criteria was determined as the point where the number of genuine cross-matches falls and the number of random matches starts to dominate.

4.3.3.1 Source-Size Distributions

Completeness depends not only on the total flux density but also on the size of the source as source detection is performed based on the peak flux density of a source; therefore, for a given total flux density, the peak flux density for a larger source is more likely to fall below the detection threshold than for a smaller source. However, an accurate source-size distribution of the sub-mJy radio source population at low frequencies is lacking and we must therefore make some assumptions in deriving the corrections.

We start by assuming that the observed size distribution of sources, within a flux density range that is unaffected by completeness, is an accurate description of sizes at fainter flux densities. As our work is focused on generating completeness corrections suited for the AGN subset of the radio population, we generate an 'AGN' source-size distribution by selecting all sources classified as radio-excess AGN or optical AGN in the total flux density range 1 mJy $< S_{150,\mathrm{tot}} <$ 10 mJy, where we expect the sample to be largely complete. We only consider sources with sizes in the range 6″ – 30″; larger sizes are not used in our simulations as such sources are poorly represented by a Gaussian surface brightness profile; the small number of sources with larger sizes are all placed at 30″ in our simulations. Within each simulated flux density bin, we weight the simulation

Table 4.1 *Completeness corrections determined for the LoTSS Deep Fields calculated using a source-size distribution derived from the observed sources classified as AGN by Best et al. (in prep.) with* $1\,\mathrm{mJy} < S_{150,\mathrm{tot}} \leq 10\,\mathrm{mJy}$ *in each field. The completeness curves for each field are shown in Fig. 4.5.*

Flux density [mJy]	ELAIS-N1	Lockman Hole	Boötes
0.09	0.1	-	-
0.11	0.214	-	-
0.13	0.343	0.182	-
0.16	0.555	0.306	-
0.19	0.68	0.43	0.18
0.23	0.734	0.57	0.311
0.28	0.78	0.664	0.462
0.33	0.802	0.746	0.619
0.4	0.818	0.781	0.71
0.63	0.85	0.844	0.819
1.01	0.885	0.874	0.87
1.59	0.921	0.912	0.915
2.52	0.947	0.945	0.949
4.0	0.959	0.955	0.968
6.34	0.971	0.969	0.977
10.05	0.972	0.976	0.981
15.92	0.978	0.971	0.984
25.24	0.985	0.978	0.986
40.0	0.987	0.979	0.984

output by this size distribution. To determine the completeness correction for each flux density interval, we then considered the subset of mock sources with total flux density $> 5\sigma$ based on the local rms, and determined the fraction of these that were detected by PyBDSF with a peak flux density above 5σ (the same criterion as that applied to the observed source sample and folded in to the calculation of the V_{max}; thereby avoiding any double-counting). The 'AGN' completeness corrections obtained in this way in ELAIS-N1 are shown by the grey line in Fig. 4.4. Similarly, for comparison, we repeat this process using the size distribution of the star-forming galaxy (SFG) subset of the radio population, deriving the pink completeness curve in Fig. 4.4. The relatively lower completeness for the 'AGN' line compared to the 'SFG' curve is driven by the sizes of (resolved) radio-AGN being significantly larger than the SFG population; this results in a higher fraction of extended sources (with consequently lower peak flux densities for a given total flux density), which leads to a lower completeness as expected.

To confirm the robustness of the above approach, we consider also the flux-dependent

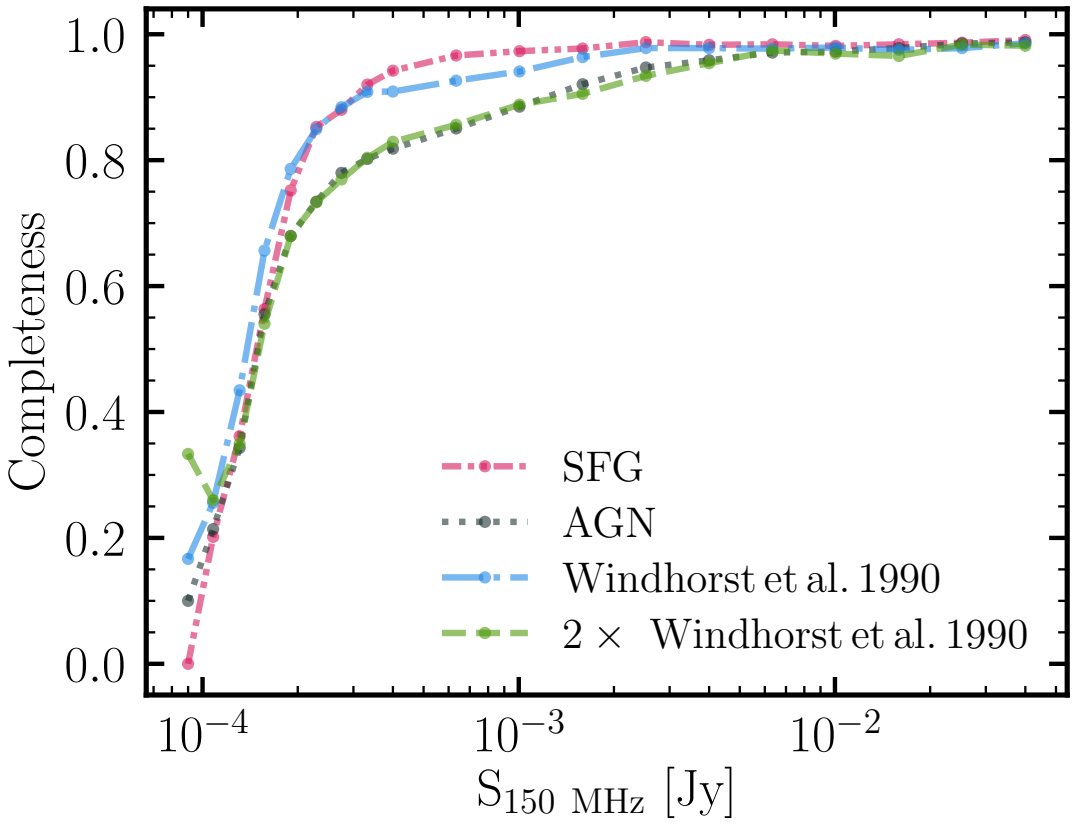

Figure 4.4 *Radio flux density completeness corrections computed using the method outlined in Sect. 4.3.3. Completeness curves in the ELAIS-N1 field as a function of the assumed source-size distribution. Curves show completeness based on applying the size-distribution of moderately bright sources ($S_{tot} = 1-10$ mJy; i.e. where completeness is high) classified as either SFGs or AGN in Best et al. (in prep.). In addition, the completeness curves resulting from the Windhorst et al. (1990) and 2 × Windhorst et al. (1990) integral angular size distribution, often used in literature, are also shown.*

angular size distribution based on GHz surveys commonly used in literature. Windhorst et al. (1990) describe the integral angular size distribution as

$$h\,(>\psi) = \exp\left[-\ln 2\left(\frac{\psi}{\psi_{\rm med}}\right)^{0.62}\right] \tag{4.6}$$

where, $h(>\psi)$ is the integral angular size distribution for sources with angular sizes larger than ψ (in arcsec) at a given flux density and $\psi_{\rm med}$ is the median angular size at the given flux density. Windhorst et al. (1990, 1993) proposed a relationship between $\psi_{\rm med}$ and the radio flux at 1.4 GHz, $S_{1.4\,\rm GHz}$:

$$\psi_{\rm med} = 2\,(S_{1.4\,\rm GHz})^{0.3}\rm arcsec, \tag{4.7}$$

which was converted to 150 MHz using a spectral index $\alpha = -0.7$. They also considered a potential floor in this relationship at a size of 2″, however given that this needs to be convolved with the 6″ LOFAR beam, this makes no significant difference to the observed size distribution. Using equations 4.6 and 4.7, $h(>\psi)$ was computed for each flux density interval and the simulation outputs were weighted by this to determine the completeness corrections. The resulting correction in the ELAIS-N1 field is shown in Fig. 4.4 (blue line). Other low radio frequency studies in literature (e.g. Retana-Montenegro et al., 2018) find good agreement between the LOFAR data and the Windhorst et al. (1990) size distributions if the median angular size relation (equation 4.7) is scaled by a factor of two. We therefore also compute the completeness corrections for this relation, also shown in Fig. 4.4 (green line), with the larger median sizes resulting in lower completeness. We find a good agreement between the 2× Windhorst et al. (1990) and the 'AGN' curves, and similarly between the Windhorst et al. (1990) and the 'SFG' curves.

In subsequent analysis, we use the 'AGN' completeness corrections in the construction of the luminosity functions. The 'AGN' completeness corrections for all three LoTSS Deep Fields are shown in Fig. 4.5. Using 'AGN' based corrections, the observations in LoTSS Deep Fields reach a completeness of 50 per cent (and 90 per cent) at 150 μJy (1.3 mJy), 209 μJy (1.4 mJy), and 289 μJy (1.4 μJy), in ELAIS-N1, Lockman Hole, and Boötes, respectively. We note that in all fields, the completeness does not reach 100 per cent; this is largely due to the source finding algorithm struggling to detect simulated sources placed in the higher noise, lower dynamic range regions near bright genuine sources. We generate a "combined" completeness curve for use in constructing the LFs by performing an area-weighted average of the completeness curves in the three fields (black line in Fig. 4.5).

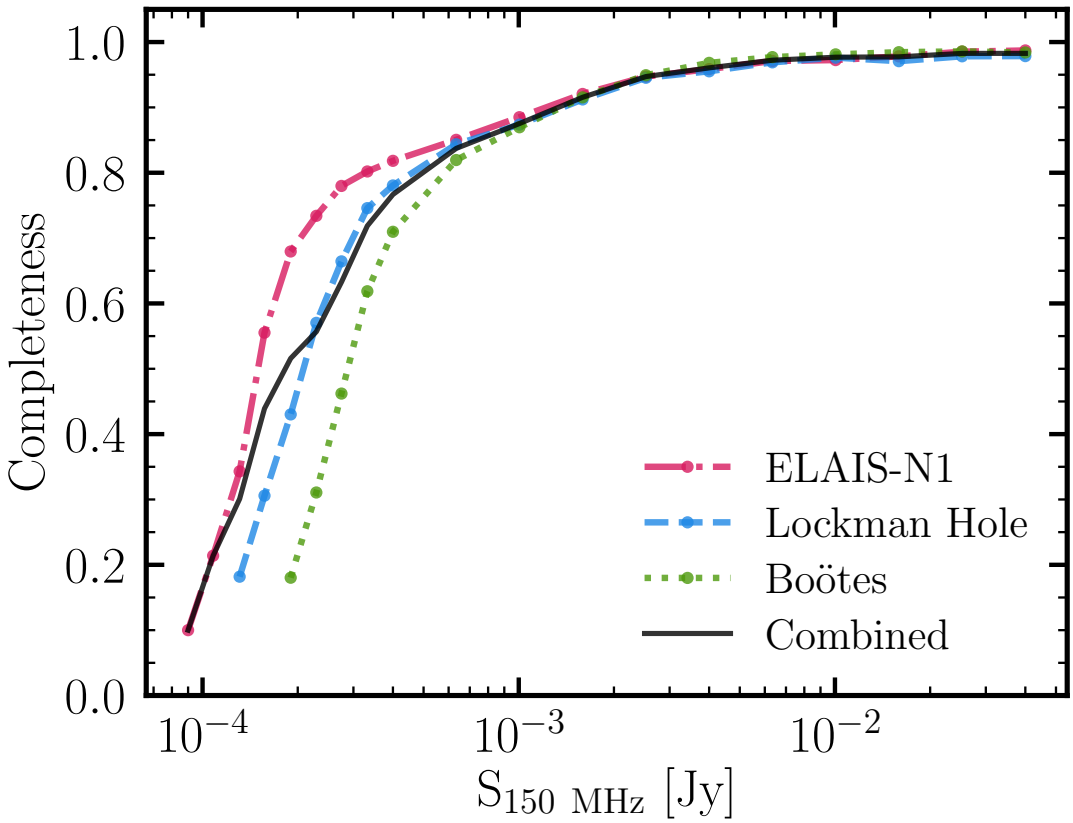

Figure 4.5 *Radio flux density completeness correction curves for each of the three LoTSS Deep Fields assuming an 'AGN' size-distribution; this is used in the construction of the AGN LFs for each field. A table listing these adopted corrections for each field is shown in Table 4.1. The 'combined' curve is the area-weighted average of the completeness curves in the three fields, used for constructing the combined LFs across the three fields.*

We note that the completeness corrections applied throughout are calculated based on the input flux density of the simulated sources, rather than the recovered flux densities from PyBDSF. This may have a particularly important effect at faint fluxes, close to the detection limit, where sources detected on a noise peak may appear at higher flux densities in the catalogue. This so-called Eddington bias causes the flux densities to be boosted resulting in an increase in the probability of detecting a source below the detection limit, and hence artificially boosting the completeness close to the detection limit. We estimate the level flux boosting by comparing the measured flux density with the true flux density of the simulated sources ($S_{\rm obs}/S_{\rm true}$) in various flux density bins. To asses the impact of this bias on the luminosity functions derived in Sect. 4.3.5 and Sect. 4.4, we calculate the median $S_{\rm obs}/S_{\rm true}$ for the faintest luminosity bin that we probe within each redshift bin. We find that for the faintest luminosity bin, the typical flux boosting level is $\sim$ 10% across all of the redshift bins studied, with even lower values for higher luminosity bins. Therefore, given the broad luminosity bin width of 0.3 dex, this bias is negligible over most of the luminosity function; for the faintest bin, this will have a small effect on the estimated completeness correction, leading to slightly higher space densities.

4.3.3.2 Application of the Completeness Corrections

The completeness corrections were applied when determining the luminosity functions as in equation 4.4, where the completeness corrections were linearly interpolated for each flux density, $S(z)$. We applied a maximum completeness correction of a factor of 10 as any larger corrections are likely not reliable. Then, to determine the point where the completeness corrections to the LFs are too large to be reliable, we recalculated the luminosity functions but this time without applying any completeness corrections (i.e. by setting $C_{\rm radio}[S(z)] = 1$ in equation 4.4); in our analysis we do not plot or list the space densities for the luminosity bins where the difference between the data points with and without the corrections is larger than 0.3 dex. To account for uncertainties in the completeness corrections (e.g. the lack of knowledge of the true source-size distribution), we add 25 per cent of the completeness correction in quadrature to the error obtained from bootstrap sampling at each luminosity bin.

We calculate the LF in each of the three LoTSS Deep Fields separately, and to take full advantage of the LoTSS Deep Fields dataset, we build a combined LF across the three fields covering $\sim 25\,{\rm deg}^2$ to obtain more robust number statistics across the full luminosity range and to limit the effects of cosmic variance.

Table 4.2 *Luminosity functions of the local* ($0.03 < z \leq 0.3$) *radio-excess AGN population in the LoTSS Deep Fields. These LFs are also illustrated in Fig. 4.6.*

z	$\log L_{150\,\mathrm{MHz}}$ $\log \mathrm{W\ Hz^{-1}}$	$\log \rho$ $\mathrm{Mpc^{-3}\ log}\, L^{-1}$	N
$0.03 < z \leq 0.3$	21.15	$-3.61^{+0.24}_{-0.37}$	2
	21.45	$-3.32^{+0.15}_{-0.17}$	10
	21.75	$-3.43^{+0.11}_{-0.12}$	22
	22.05	$-3.49^{+0.08}_{-0.09}$	47
	22.35	$-3.47^{+0.06}_{-0.07}$	122
	22.65	$-3.64^{+0.05}_{-0.05}$	177
	22.95	$-3.79^{+0.04}_{-0.04}$	168
	23.25	$-4.00^{+0.04}_{-0.04}$	113
	23.55	$-4.36^{+0.06}_{-0.07}$	52
	23.85	$-4.43^{+0.06}_{-0.06}$	46
	24.15	$-4.84^{+0.11}_{-0.14}$	18
	24.45	$-4.75^{+0.09}_{-0.11}$	22
	24.75	$-4.87^{+0.11}_{-0.12}$	17
	25.05	$-4.99^{+0.12}_{-0.16}$	13
	25.35	$-4.93^{+0.10}_{-0.13}$	15
	25.95	$-5.63^{+0.22}_{-0.48}$	3

4.3.4 The local radio-AGN luminosity function

Although the LoTSS Deep Fields cover a relatively small volume at low redshifts, comparison of the low redshift LF against previous measurements in literature can be useful. Using the methods outlined above, we have built the local $0.03 < z \leq 0.3$ luminosity function for the radio-excess AGN (848 sources) in the LoTSS Deep Fields, which is shown in Fig. 4.6. The combined luminosity function across the three fields is shown by pink filled circles, with the shaded pink region showing the 1σ uncertainties. The luminosity functions for each individual deep field are also shown by pink crosses, triangles, and squares (with their respective error bars) for ELAIS-N1, Lockman Hole, and Boötes, respectively. We note that the ELAIS-N1 data points are consistently higher than the other fields, an effect that is also seen in the K-band number counts (see Fig. 2.6 in Chapter 2), likely due to large-scale structure within the field as the volume probed at these low redshifts is relatively small. The combined LoTSS Deep LFs are also listed in Table 4.2.

Fig. 4.6 also shows comparison between our local radio-AGN luminosity functions and other previous studies. We find good agreement with the local radio-AGN luminosity

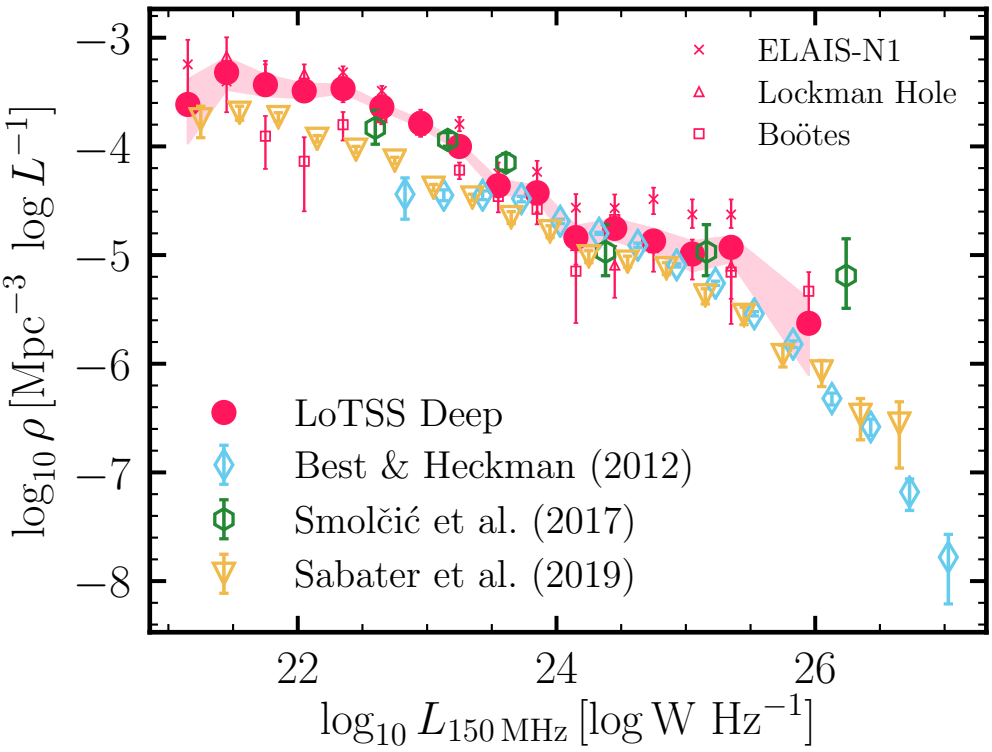

Figure 4.6 *Local* ($0.03 < z \leq 0.3$) *radio-excess AGN luminosity functions for LoTSS Deep Fields built with a bin width of* $\Delta \log L = 0.3$ *dex. The combined luminosity function is shown in pink filled circles (for 848 sources) with shaded regions showing the* 1σ *uncertainties computed from bootstrapping. The luminosity functions for each field individually are also shown. We find fairly good agreement with the Best & Heckman (2012) and Smolčić et al. (2017b) results from GHz surveys, scaled to 150 MHz using a spectral index* $\alpha = -0.7$.

function from Smolčić et al. (2017b) built using deep radio imaging from the VLA-COSMOS 3 GHz Large project (Smolčić et al., 2017a). Their LFs were computed over $0.1 < z < 0.4$ (105 sources), and for illustration in Fig. 4.6 are shifted to 150 MHz using a spectral index $\alpha = -0.7$. Best & Heckman (2012) present a large sample of radio-detected AGN drawn from the combination of FIRST and NVSS with SDSS spectroscopcic sample data. We find good agreement with the luminosity function of Best & Heckman (2012), shifted to $\mathrm{L}_{150\,\mathrm{MHz}}$ using a spectral index $\alpha = -0.7$, in particular at $\mathrm{L}_{150\,\mathrm{MHz}} \sim 10^{23.5} - 10^{25}\,\mathrm{W\,Hz^{-1}}$. We are however not able to sample enough volume to probe significantly above the break in the luminosity function. Best & Heckman (2012) also found their LF to be in good agreement with other determinations (e.g. Mauch & Sadler, 2007; Pracy et al., 2016).

Also shown in Fig. 4.6 (green triangles) is the radio-AGN luminosity function from the

shallower but wider LoTSS Data Release 1 (LoTSS-DR1) from Sabater et al. (2019). Their data, covering $> 400\,\mathrm{deg}^2$, is better suited to sample higher radio luminosities. We find good agreement at moderate luminosities, however our luminosity function is consistently offset to higher space densities by ~ 0.5 dex, especially at $\mathrm{L}_{150\,\mathrm{MHz}} \lesssim 10^{23.5}\,\mathrm{W\,Hz^{-1}}$. We also find that the Best & Heckman (2012) LFs also appear to turn over at around the same point in luminosity. There are a few possible reasons for the difference with both of those studies. Neither Best & Heckman (2012) nor Sabater et al. (2019) apply any completeness corrections; we re-derive the Sabater et al. luminosity function by simply applying a 10σ radio flux density cut, which increases the space densities by ~ 0.2 dex at $\mathrm{L}_{150\,\mathrm{MHz}} \lesssim 10^{22.3}\,\mathrm{W\,Hz^{-1}}$. Secondly, the Best & Heckman (2012) and Sabater et al. (2019) radio-AGN samples were defined by combining their radio data with the SDSS main galaxy sample; source like quasars or radio-quiet quasars are missing from this sample, likely biasing the luminosity function to lower space densities. Finally, the median redshift of the Best & Heckman (2012) sample is $z_{\mathrm{med}} \sim 0.16$, and that of Sabater et al. (2019) is $z_{\mathrm{med}} = 0.14$, whereas the median redshift for the LoTSS Deep Fields sample (in this redshift bin) is $z_{\mathrm{med}} = 0.21$; any cosmic evolution of this population can therefore contribute to the difference in space densities observed. We also note that this difference is not simply due to a mis-classification of sources between radio-AGN and SFGs, as a similar offset is also found by Cochrane et al. (in prep.) when comparing the local SFG luminosity function from the LoTSS Deep Fields with LoTSS-DR1.

4.3.5 Evolution of the radio-AGN LFs

Fig. 4.7 shows the redshift evolution of the total radio-AGN LF in the LoTSS Deep fields in six redshift bins ($0.4 < z \leq 2.5$). The LoTSS Deep LF is shown as pink circles, with the LF from individual fields shown as open pink symbols (using the same symbols as in Fig. 4.6). The LFs are also listed in Table 4.3. For comparison, we show the parametric fit to the local radio-AGN luminosity function by Best et al. (2014) in each panel, shifted to 150 MHz using $\alpha = -0.7$. We again compare our results with the evolution of the total radio-AGN population presented in Smolčić et al. (2017b), shifted to $\mathrm{L}_{150\,\mathrm{MHz}}$ using $\alpha = -0.7$, shown by green symbols, with our redshift bins chosen to match their analysis. As is evident from Fig. 4.7, we find excellent agreement with their results at all radio luminosities, out to $z \sim 2.5$; this gives us further confidence that our source classification method for separating radio-AGN from star-forming galaxies is appropriate out to high redshifts.

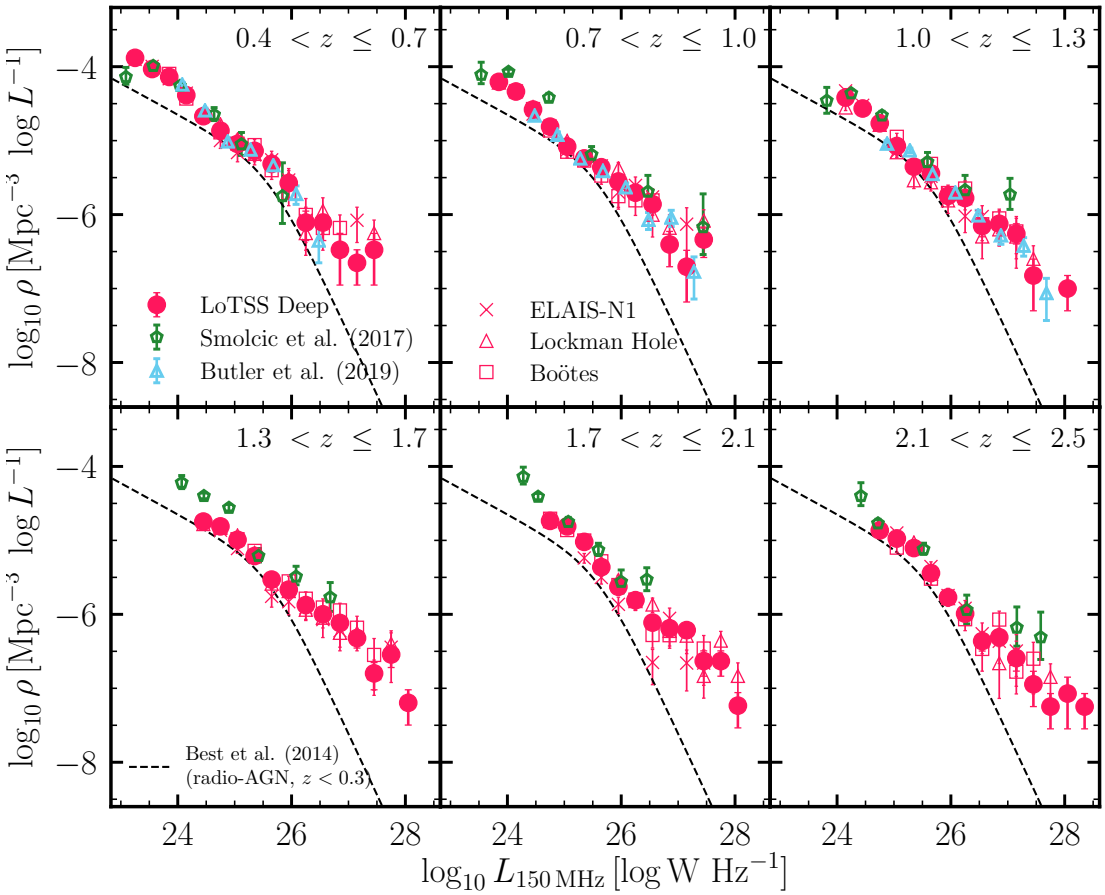

Figure 4.7 *Cosmic evolution (across $0.4 < z \leq 2.5$) of the total radio-AGN luminosity functions at 150 MHz for the combined LoTSS Deep sample (pink circles). The LFs in individual LoTSS Deep fields are shown in open pink symbols (same symbols as in Fig. 4.6). The dashed line in each panel shows the parameterised form of the local radio-AGN LF by Best et al. (2014), converted to 150 MHz using a spectral index $\alpha = -0.7$ to guide the eye. The total radio-AGN luminosity functions from Smolčić et al. (2017b) and Butler et al. (2019) in green and cyan open points, respectively, converted to 150 MHz using $\alpha = -0.7$, are also shown. Overall, we find good agreement with the results from both Butler et al. (2019) and Smolčić et al. (2017b) across redshift. The LFs are also listed in Table 4.3.*

Also shown in Fig. 4.7 are the radio-AGN LFs of Butler et al. (2019) across $0.3 < z \leq 0.6$, $0.6 < z \leq 0.9$, and $0.9 < z \leq 1.3$ in the closest redshift bins to our LFs. These LFs were compiled using data from the Australian Telescope Compact Array (ATCA) 2.1 GHz observations of the XXL-S field (Butler et al., 2018) and the radio-AGN sample was selected based on radio-source luminosity, morphology, spectral indices, and radio-excess emission based on the FIRC. We also find good agreement with their LFs where available, but probing fainter in luminosity.

Table 4.3 *Luminosity functions of the radio-excess AGN in the LoTSS Deep Fields between* $0.5 < z \leq 2.5$*, illustrated in Fig. 4.7. Space densities for bins with* $N < 2$ *are not shown.*

z	$\log L_{150\,\mathrm{MHz}}$ $\log \mathrm{W\,Hz^{-1}}$	$\log \rho$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	N
$0.4 < z \leq 0.7$	23.25	$-3.88^{+0.08}_{-0.08}$	277
	23.55	$-4.03^{+0.05}_{-0.05}$	494
	23.85	$-4.14^{+0.03}_{-0.03}$	528
	24.15	$-4.39^{+0.03}_{-0.03}$	327
	24.45	$-4.67^{+0.04}_{-0.03}$	181
	24.75	$-4.87^{+0.04}_{-0.04}$	118
	25.05	$-5.04^{+0.04}_{-0.05}$	79
	25.35	$-5.14^{+0.05}_{-0.06}$	64
	25.65	$-5.32^{+0.07}_{-0.07}$	43
	25.95	$-5.57^{+0.08}_{-0.10}$	24
	26.25	$-6.11^{+0.15}_{-0.24}$	7
	26.55	$-6.11^{+0.11}_{-0.24}$	7
	26.85	$-6.48^{+0.22}_{-0.48}$	3
	27.15	$-6.65^{+0.18}_{-0.30}$	2
	27.45	$-6.48^{+0.22}_{-0.48}$	3
$0.7 < z \leq 1.0$	23.85	$-4.21^{+0.06}_{-0.06}$	455
	24.15	$-4.34^{+0.03}_{-0.03}$	534
	24.45	$-4.58^{+0.03}_{-0.03}$	345
	24.75	$-4.81^{+0.03}_{-0.03}$	217
	25.05	$-5.08^{+0.04}_{-0.04}$	121
	25.35	$-5.24^{+0.04}_{-0.05}$	85
	25.65	$-5.36^{+0.05}_{-0.05}$	65

continued ...

... continued

z	$\log L_{150\,\mathrm{MHz}}$ $\log \mathrm{W\ Hz^{-1}}$	$\log \rho$ $\mathrm{Mpc^{-3}\ log}\, L^{-1}$	N
	25.95	$-5.55^{+0.06}_{-0.07}$	42
	26.25	$-5.70^{+0.07}_{-0.08}$	30
	26.55	$-5.86^{+0.09}_{-0.09}$	21
	26.85	$-6.40^{+0.18}_{-0.30}$	6
	27.15	$-6.70^{+0.22}_{-0.48}$	3
	27.45	$-6.34^{+0.15}_{-0.24}$	7
$1.0 < z \leq 1.3$	24.15	$-4.42^{+0.06}_{-0.06}$	354
	24.45	$-4.57^{+0.03}_{-0.04}$	419
	24.75	$-4.77^{+0.03}_{-0.03}$	296
	25.05	$-5.07^{+0.04}_{-0.04}$	156
	25.35	$-5.35^{+0.05}_{-0.05}$	85
	25.65	$-5.44^{+0.05}_{-0.05}$	70
	25.95	$-5.75^{+0.07}_{-0.07}$	35
	26.25	$-5.78^{+0.07}_{-0.09}$	33
	26.55	$-6.15^{+0.11}_{-0.15}$	14
	26.85	$-6.12^{+0.10}_{-0.13}$	15
	27.15	$-6.26^{+0.13}_{-0.14}$	11
	27.45	$-6.82^{+0.22}_{-0.48}$	3
	28.05	$-7.00^{+0.18}_{-0.30}$	2
$1.3 < z \leq 1.7$	24.45	$-4.75^{+0.06}_{-0.06}$	298
	24.75	$-4.81^{+0.03}_{-0.03}$	384
	25.05	$-4.99^{+0.03}_{-0.03}$	279
	25.35	$-5.21^{+0.03}_{-0.04}$	180
	25.65	$-5.53^{+0.05}_{-0.05}$	89
	25.95	$-5.67^{+0.05}_{-0.05}$	66
	26.25	$-5.87^{+0.06}_{-0.07}$	41
	26.55	$-6.00^{+0.07}_{-0.09}$	31
	26.85	$-6.12^{+0.08}_{-0.10}$	24
	27.15	$-6.32^{+0.10}_{-0.13}$	15
	27.45	$-6.80^{+0.15}_{-0.22}$	5
	27.75	$-6.54^{+0.12}_{-0.18}$	9

continued ...

... continued

z	$\log L_{150\,\mathrm{MHz}}$ $\log \mathrm{W\,Hz^{-1}}$	$\log \rho$ $\mathrm{Mpc^{-3}\ log}\, L^{-1}$	N
	28.05	$-7.20^{+0.18}_{-0.30}$	2
$1.7 < z \leq 2.1$	24.75	$-4.74^{+0.05}_{-0.05}$	387
	25.05	$-4.81^{+0.03}_{-0.03}$	432
	25.35	$-5.02^{+0.03}_{-0.03}$	289
	25.65	$-5.36^{+0.04}_{-0.04}$	140
	25.95	$-5.63^{+0.04}_{-0.05}$	78
	26.25	$-5.81^{+0.05}_{-0.06}$	52
	26.55	$-6.11^{+0.08}_{-0.09}$	26
	26.85	$-6.19^{+0.09}_{-0.11}$	22
	27.15	$-6.21^{+0.09}_{-0.11}$	21
	27.45	$-6.63^{+0.14}_{-0.20}$	8
	27.75	$-6.63^{+0.14}_{-0.20}$	8
	28.05	$-7.23^{+0.18}_{-0.30}$	2
$2.1 < z \leq 2.5$	24.75	$-4.87^{+0.08}_{-0.08}$	167
	25.05	$-4.98^{+0.04}_{-0.04}$	273
	25.35	$-5.10^{+0.03}_{-0.03}$	236
	25.65	$-5.44^{+0.04}_{-0.04}$	116
	25.95	$-5.77^{+0.06}_{-0.06}$	57
	26.25	$-5.99^{+0.07}_{-0.08}$	35
	26.55	$-6.36^{+0.10}_{-0.13}$	15
	26.85	$-6.31^{+0.09}_{-0.12}$	17
	27.15	$-6.59^{+0.13}_{-0.18}$	9
	27.45	$-6.95^{+0.18}_{-0.30}$	4
	27.75	$-7.25^{+0.18}_{-0.30}$	2
	28.05	$-7.07^{+0.22}_{-0.48}$	3
	28.35	$-7.25^{+0.18}_{-0.30}$	2

As mentioned in Appendix 4.A, our selection of radio-excess AGN, based on radio-excess $> 3\sigma$, may miss low-luminosity radio-AGN, particularly in massive star-forming galaxies where the radio-emission is dominated by star-formation processes rather than the jets; this can lead to missing radio-AGN in the derived luminosity functions for the radio-excess AGN (Fig. 4.7) and of the LERGs (Fig. 4.8). To determine the

luminosity down to which our luminosity functions are robust against this selection effect, we calculated the typical radio-luminosities due to star-formation (based on their median star-formation rates and the radio-star-formation rate relation e.g. Smith et al. 2021), for the sources in each of the six redshift bins in Fig. 4.7. We find that over all redshift bins, the typical star-forming radio luminosity is lower than the faintest luminosity bin that we consider by at least 0.7 dex (corresponding to a radio excess of $\approx 3\sigma$). We therefore conclude that although some individual high star-formation rate sources (most commonly at high stellar mass) may be affected, the global impact of this selection effect on our luminosity functions is negligible. We repeated this analysis for the evolving LERG luminosity functions shown in Fig. 4.8, and find a similar result.

4.4 Cosmic evolution of the LERG luminosity functions

The total radio-AGN luminosity functions presented in Sect. 4.3.5 contain a mixture of both the LERG and the HERG populations. These two populations are expected to evolve differently with the LERGs dominating the space densities at low luminosities whereas the HERGs dominate at high radio luminosities (e.g. Best & Heckman, 2012); the LERG population is also particularly interesting for radio-AGN feedback considerations. Previous studies (e.g. Best et al., 2014; Pracy et al., 2016; Williams et al., 2018; Butler et al., 2019) have attempted to model the evolution of the LERGs, however the current LoTSS Deep Fields dataset with vastly greater numbers of LERGs resulting from the combination of deep radio and multi-wavelength datasets over $\sim 25\,\mathrm{deg}^2$ allows us to study the cosmic evolution of this population in unprecedented detail.

In Fig. 4.8, we show the evolution of the LERG LF for the LoTSS Deep Fields in four redshift bins ($0.5 < z \leq 1.0$, $1.0 < z \leq 1.5$, $1.5 < z \leq 2.0$, and $2.0 < z \leq 2.5$), each spanning $3 - 4$ decades in luminosity. The combined LoTSS Deep Fields LFs, also tabulated in Table 4.4, are shown as pink circles. The LFs for individual fields are also shown in pink, using the same symbols for each field as in Fig. 4.6. In each panel, the black dotted line shows the parametrised form of the local ($z < 0.3$) LERG luminosity function determined by Best et al. (2014), scaled to 150 MHz using a spectral index $\alpha = -0.7$. We find that the LERG population shows modest evolution between between $0 < z < 0.75$, especially at high luminosities, although the details of this depend on the assumed spectral index. Beyond $z = 0.75$, we see a relatively mild evolution in the LFs with redshift.

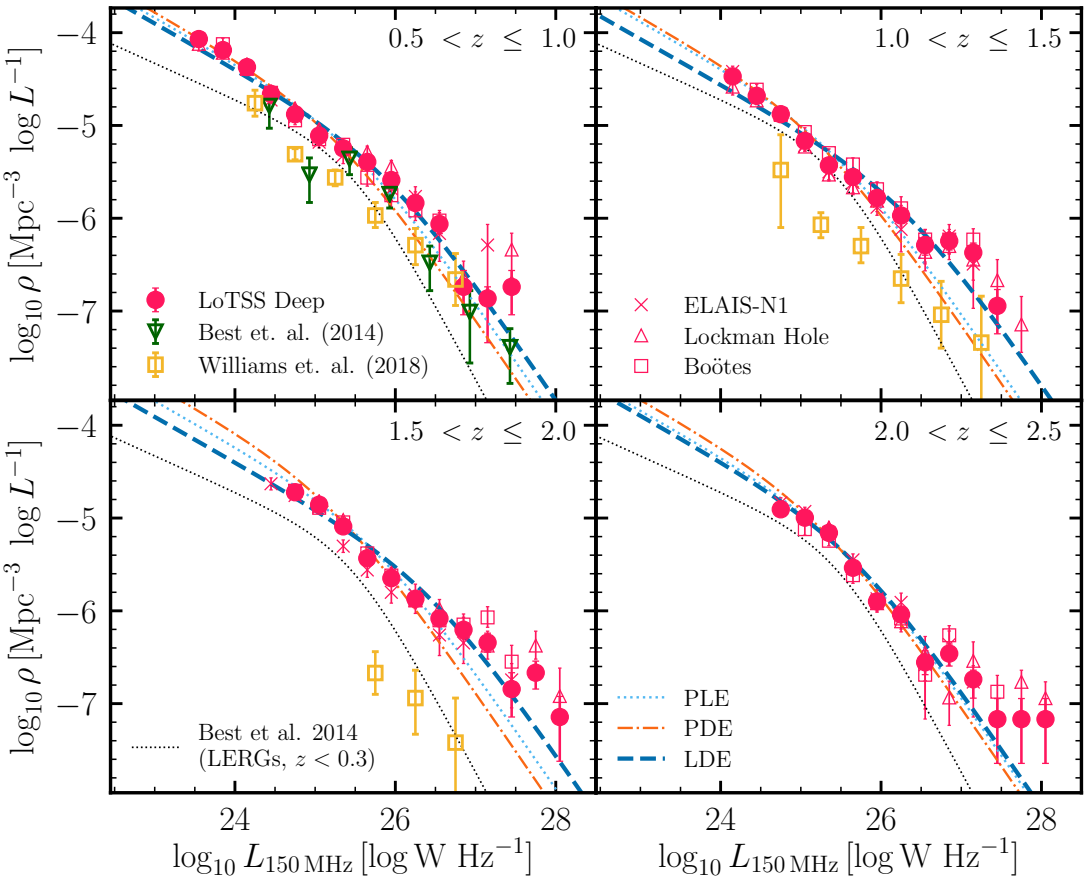

Figure 4.8 *Cosmic evolution (across* $0.5 < z \leq 2.5$*) of the LERG luminosity functions at 150 MHz for the combined LoTSS Deep sample (pink circles), showing relatively mild evolution across the redshift bins examined. Also shown are the LFs in individual LoTSS Deep fields (in pink; same symbols as in Fig. 4.6) and the parametric fit to the local LERG AGN LF from Best et al. (2014) (black dotted line), scaled to 150 MHz assuming a spectral index* $\alpha = -0.7$*. The pure density evolution (PDE), pure luminosity evolution (PLE), and combined luminosity and density evolution (LDE) models from fits to individual redshift bins are shown as orange (dash-dotted) and blue (dashed), dark blue (dotted) lines, respectively (see Sect. 4.4.1 for details). For* $0.5 \leq 1.0$*, the LERG LF from Best et al. (2014, green triangles), scaled to 150 MHz using a spectral index* $\alpha = -0.7$*, shows good agreement with our results. The LFs from Williams et al. (2018), reaching* $z \sim 2$ *(yellow squares in the three lowest redshift bins), have systematically lower space densities compared to our results. Further tests show that the differences are largely driven by our improvements in the source classification criteria used.*

We compare our results to the LERG LF derived by Best et al. (2014) at $0.5 < z \leq 1$ in Fig. 4.8, shifted to $L_{150\,\mathrm{MHz}}$ using a spectral index $\alpha = -0.7$. Best et al. (2014) compiled catalogues of radio-detected AGN from eight surveys, covering different radio-depths and areas, to obtain a sample of 211 radio-loud AGN. They used spectroscopic information to classify their AGN sample into LERGs and HERGs, representing the first study on the evolution of the two modes of AGN, separately, out to $z \sim 1$. The LoTSS-Deep LF shows good agreement with their resulting luminosity function (green triangles), with our data probing fainter luminosities.

Also in Fig. 4.8, we show the LERG LFs computed by Williams et al. (2018), who studied the cosmic evolution ($0.5 < z \leq 2$) of a sample of 1224 LOFAR-detected sources within the Boötes field (albeit with shallower radio data). They used the combination of the FIR-radio correlation of Calistro Rivera et al. (2016) and results from SED fitting via AGNFitter to perform source classifications. Radio sources that lie 2σ away from the FIR-radio correlation (Calistro Rivera et al., 2017) were identified as radio-loud AGN. Then, the various AGN and galaxy models fitted by AGNFitter were used to quantify the fraction of MIR emission arising from the AGN compared to the galaxy component (f_{AGN}); sources with $f_{\mathrm{AGN}} > 0.25$, indicative of significant emission from the torus, were classified as HERGs, as would be expected, whereas the remainder are classified as LERGs, which are expected to show little to no torus emission. Their final sample contains 243 LERGs (and 398 HERGs) within $0.5 < z \leq 2$, with their resulting LF shown as yellow squares in Fig. 4.8.

We note that the Williams et al. LFs appear systematically offset to lower space densities than our dataset in all redshift bins, and also appear offset by ~ 0.4 dex at moderate to faint luminosities ($L_{150\,\mathrm{MHz}} < 10^{26}\,\mathrm{W\,Hz^{-1}}$) compared to Best et al. (2014). We have performed tests to investigate the source of this discrepancy, and find this to be largely driven by the different source classification schemes adopted. In principle, the source classification method of Williams et al. (2018) detailed above is similar to the one used by Best et al. (in prep.), however, in practise, the exact criteria differ for both the selection of radio-excess AGN and for separating HERGs and LERGs (see Sect. 4.2.2). Therefore we reconstructed the LERG LFs with the LoTSS-Deep data using the classification scheme adopted by Williams et al. (2018) to test the effect this has on the LFs and the comparison with the Williams et al. (2018) results. Specifically, we identified radio-excess sources making use of the FIR-radio correlation of Calistro Rivera et al. (2017) using physical galaxy properties derived from AGNFitter only instead of the *consensus* values, and we classified these sources as HERGs or LERGs using the Williams et al. (2018) f_{AGN} parameter.

Table 4.4 *Luminosity functions of the LERGs in the LoTSS Deep Fields between* $0.5 < z \leq 2.5$*, illustrated in Fig. 4.8. Also shown are the LFs in the same redshift bins, for the subsets of the LERG population that are hosted by quiescent galaxies and by star-forming galaxies, respectively; these resulting LFs are discussed in more detail in Sect. 4.6, with the separation of quiescent and star-forming host galaxies of the LERGs described in Sect. 4.5.1. Space densities for bins with* $N < 2$ *are not shown.*

z	$\log_{10} L_{150\,\mathrm{MHz}}$	$\log_{10} \rho_{\mathrm{LERG}}$	N_{LERG}	$\log_{10} \rho_{\mathrm{LERG,Q}}$	$N_{\mathrm{LERG,Q}}$	$\log_{10} \rho_{\mathrm{LERG,SF}}$	$N_{\mathrm{LERG,SF}}$
	$\log_{10}\mathrm{W\,Hz^{-1}}$	$\mathrm{Mpc^{-3}\ log}\,L^{-1}$		$\mathrm{Mpc^{-3}\ log}\,L^{-1}$		$\mathrm{Mpc^{-3}\ log}\,L^{-1}$	
$0.5 < z \leq 1.0$	23.55	$-4.07^{+0.07}_{-0.07}$	428	$-4.22^{+0.07}_{-0.07}$	303	$-4.60^{+0.08}_{-0.08}$	125
	23.85	$-4.19^{+0.05}_{-0.05}$	808	$-4.39^{+0.05}_{-0.05}$	510	$-4.62^{+0.05}_{-0.05}$	298
	24.15	$-4.37^{+0.03}_{-0.03}$	743	$-4.61^{+0.03}_{-0.03}$	431	$-4.75^{+0.03}_{-0.04}$	312
	24.45	$-4.66^{+0.03}_{-0.03}$	424	$-4.91^{+0.03}_{-0.03}$	238	$-5.02^{+0.03}_{-0.03}$	186
	24.75	$-4.88^{+0.03}_{-0.03}$	271	$-5.08^{+0.03}_{-0.04}$	172	$-5.32^{+0.04}_{-0.04}$	99
	25.05	$-5.11^{+0.03}_{-0.03}$	164	$-5.26^{+0.04}_{-0.04}$	117	$-5.65^{+0.05}_{-0.07}$	47
	25.35	$-5.25^{+0.04}_{-0.04}$	121	$-5.34^{+0.04}_{-0.05}$	97	$-5.95^{+0.08}_{-0.10}$	24
	25.65	$-5.39^{+0.04}_{-0.05}$	87	$-5.48^{+0.05}_{-0.06}$	71	$-6.13^{+0.10}_{-0.12}$	16
	25.95	$-5.59^{+0.05}_{-0.06}$	56	$-5.71^{+0.06}_{-0.07}$	42	$-6.19^{+0.11}_{-0.15}$	14
	26.25	$-5.84^{+0.09}_{-0.08}$	32	$-5.88^{+0.07}_{-0.08}$	29	$-6.86^{+0.22}_{-0.48}$	3
	26.55	$-6.06^{+0.08}_{-0.10}$	19	$-6.20^{+0.11}_{-0.15}$	14	$-6.64^{+0.15}_{-0.22}$	5
	26.85	$-6.74^{+0.18}_{-0.30}$	4	$-6.86^{+0.22}_{-0.48}$	3		
	27.15	$-6.86^{+0.12}_{-0.48}$	3	$-7.04^{+0.18}_{-0.30}$	2		
	27.45	$-6.74^{+0.18}_{-0.30}$	4				

continued ...

. . . continued

z	$\log_{10} L_{150\,\mathrm{MHz}}$ $\log_{10}\mathrm{W\ Hz^{-1}}$	$\log_{10}\rho_{\mathrm{LERG}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	N_{LERG}	$\log_{10}\rho_{\mathrm{LERG,Q}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,Q}}$	$\log_{10}\rho_{\mathrm{LERG,SF}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,SF}}$
$1.0 < z \leq 1.5$	24.15	$-4.47^{+0.07}_{-0.07}$	401	$-4.79^{+0.08}_{-0.08}$	189	$-4.76^{+0.07}_{-0.07}$	212
	24.45	$-4.68^{+0.04}_{-0.04}$	516	$-5.13^{+0.05}_{-0.05}$	183	$-4.87^{+0.04}_{-0.04}$	333
	24.75	$-4.88^{+0.03}_{-0.03}$	392	$-5.40^{+0.04}_{-0.04}$	117	$-5.03^{+0.03}_{-0.03}$	275
	25.05	$-5.17^{+0.03}_{-0.03}$	218	$-5.61^{+0.05}_{-0.05}$	78	$-5.36^{+0.03}_{-0.04}$	140
	25.35	$-5.43^{+0.04}_{-0.04}$	124	$-5.78^{+0.06}_{-0.06}$	55	$-5.68^{+0.05}_{-0.05}$	69
	25.65	$-5.55^{+0.05}_{-0.05}$	96	$-5.98^{+0.07}_{-0.08}$	36	$-5.76^{+0.05}_{-0.05}$	60
	25.95	$-5.78^{+0.05}_{-0.07}$	57	$-6.10^{+0.07}_{-0.09}$	27	$-6.06^{+0.07}_{-0.10}$	30
	26.25	$-5.97^{+0.06}_{-0.08}$	37	$-6.16^{+0.08}_{-0.10}$	24	$-6.42^{+0.09}_{-0.16}$	13
	26.55	$-6.29^{+0.09}_{-0.11}$	18	$-6.59^{+0.12}_{-0.18}$	9	$-6.59^{+0.12}_{-0.18}$	9
	26.85	$-6.25^{+0.08}_{-0.10}$	20	$-6.70^{+0.11}_{-0.24}$	7	$-6.43^{+0.12}_{-0.11}$	13
	27.15	$-6.37^{+0.10}_{-0.13}$	15	$-6.94^{+0.18}_{-0.30}$	4	$-6.50^{+0.11}_{-0.14}$	11
	27.45	$-6.94^{+0.18}_{-0.30}$	4	$-7.25^{+0.18}_{-0.30}$	2	$-7.25^{+0.18}_{-0.30}$	2
$1.5 < z \leq 2.0$	24.75	$-4.72^{+0.04}_{-0.04}$	549	$-5.47^{+0.06}_{-0.06}$	96	$-4.81^{+0.04}_{-0.04}$	453
	25.05	$-4.86^{+0.03}_{-0.03}$	488	$-5.71^{+0.05}_{-0.06}$	68	$-4.92^{+0.03}_{-0.03}$	420
	25.35	$-5.09^{+0.03}_{-0.03}$	306	$-5.91^{+0.05}_{-0.07}$	46	$-5.16^{+0.03}_{-0.03}$	260
	25.65	$-5.43^{+0.04}_{-0.04}$	145	$-6.26^{+0.09}_{-0.09}$	22	$-5.50^{+0.04}_{-0.04}$	123
	25.95	$-5.65^{+0.05}_{-0.05}$	91	$-6.26^{+0.09}_{-0.11}$	22	$-5.77^{+0.05}_{-0.06}$	69
	26.25	$-5.87^{+0.06}_{-0.06}$	55	$-6.46^{+0.11}_{-0.15}$	14	$-6.00^{+0.06}_{-0.07}$	41

continued . . .

. . . continued

z	$\log_{10} L_{150\,\mathrm{MHz}}$ $\log_{10}\mathrm{W\ Hz^{-1}}$	$\log_{10}\rho_{\mathrm{LERG}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	N_{LERG}	$\log_{10}\rho_{\mathrm{LERG,Q}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,Q}}$	$\log_{10}\rho_{\mathrm{LERG,SF}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,SF}}$
	26.55	$-6.08^{+0.08}_{-0.07}$	34	$-6.57^{+0.10}_{-0.14}$	11	$-6.25^{+0.09}_{-0.11}$	23
	26.85	$-6.20^{+0.06}_{-0.10}$	26	$-6.72^{+0.14}_{-0.20}$	8	$-6.36^{+0.09}_{-0.11}$	18
	27.15	$-6.34^{+0.10}_{-0.10}$	19	$-6.84^{+0.18}_{-0.18}$	6	$-6.51^{+0.09}_{-0.11}$	13
	27.45	$-6.84^{+0.18}_{-0.30}$	6			$-6.84^{+0.12}_{-0.30}$	6
	27.75	$-6.67^{+0.12}_{-0.18}$	9			$-6.78^{+0.12}_{-0.24}$	7
	28.05	$-7.14^{+0.22}_{-0.48}$	3			$-7.14^{+0.22}_{-0.48}$	3
$2.0 < z \leq 2.5$	24.75	$-4.91^{+0.08}_{-0.08}$	206	$-5.75^{+0.11}_{-0.12}$	27	$-4.97^{+0.08}_{-0.08}$	179
	25.05	$-5.00^{+0.04}_{-0.04}$	329	$-6.07^{+0.08}_{-0.09}$	27	$-5.03^{+0.04}_{-0.04}$	302
	25.35	$-5.16^{+0.03}_{-0.03}$	260	$-6.62^{+0.13}_{-0.18}$	9	$-5.17^{+0.03}_{-0.03}$	251
	25.65	$-5.54^{+0.04}_{-0.04}$	117	$-6.83^{+0.13}_{-0.30}$	6	$-5.56^{+0.04}_{-0.04}$	111
	25.95	$-5.90^{+0.05}_{-0.06}$	53	$-7.02^{+0.18}_{-0.30}$	4	$-5.93^{+0.06}_{-0.06}$	49
	26.25	$-6.04^{+0.06}_{-0.07}$	39	$-6.79^{+0.15}_{-0.24}$	7	$-6.13^{+0.07}_{-0.09}$	32
	26.55	$-6.55^{+0.12}_{-0.13}$	12	$-7.33^{+0.18}_{-0.30}$	2	$-6.63^{+0.11}_{-0.15}$	10
	26.85	$-6.46^{+0.10}_{-0.13}$	15			$-6.49^{+0.11}_{-0.10}$	14
	27.15	$-6.74^{+0.10}_{-0.20}$	8			$-6.80^{+0.15}_{-0.24}$	7
	27.45	$-7.17^{+0.22}_{-0.48}$	3			$-7.17^{+0.22}_{-0.48}$	3
	27.75	$-7.17^{+0.22}_{-0.48}$	3			$-7.17^{+0.22}_{-0.48}$	3

continued . . .

. . . continued

z	$\log_{10} L_{150\,\mathrm{MHz}}$ $\log_{10}\mathrm{W\ Hz^{-1}}$	$\log_{10}\rho_{\mathrm{LERG}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	N_{LERG}	$\log_{10}\rho_{\mathrm{LERG,Q}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,Q}}$	$\log_{10}\rho_{\mathrm{LERG,SF}}$ $\mathrm{Mpc^{-3}\ log}\,L^{-1}$	$N_{\mathrm{LERG,SF}}$
	28.05	$-7.17^{+0.22}_{-0.48}$	3			$-7.17^{+0.22}_{-0.48}$	3

Fig. 4.9 shows comparison with Williams et al. (2018) when applying their source classification scheme to our LoTSS Deep dataset (cyan squares; hereafter LoTSS-Deep W18). We find that the derived LFs now show better agreement at all redshifts with the LFs derived in this work (within 2σ) based on Best et al. (in prep.) classifications, than with the Williams et al. (2018) results, especially at high radio luminosities and higher redshifts. This appears to be largely because we used an updated version of AGNFitter compared to the one used in Williams et al. with improved models being used in the fitting process resulting in better fits in general. This has a significant effect on the classification of sources as HERGs or LERGs.

The reconstructed LoTSS-Deep W18 LFs do not agree well with ours at faint radio luminosities, at all redshifts. We find that this is due to a significant number of sources being classified as radio-excess sources (and subsequently as LERGs) by our criteria, but as SFGs under the Williams et al. (2018) criteria. The discrepancy in the source classification arises from the two different estimates of SFRs in the two procedures: *consensus* SFRs used in this study, and AGNFitter based IR luminosities (and hence SFRs) used by Williams et al. (2018). Best et al. (in prep.) show that the SFRs reported by AGNFitter are found to be systematically offset to higher values than those reported by the other three SED fitting routines used. Therefore, if using the AGNFitter estimated SFRs, we find these sources (for which the classification disagrees) to lie less than 2σ above the FIR-radio correlation and hence be identified as SFGs, whereas when the *consensus* SFRs are used, the sources lie well above the relation for star-forming galaxies and hence are selected as radio-loud AGN.

This is likely a result of one of the key differences between AGNFitter and the three other SED fitting codes used here. Unlike the others, AGNFitter does not assume an "energy balance" between the stellar emission in the UV and optical, and the re-radiated dust emission in the infrared; instead both of these components are allowed to be fitted independently, with a prior that the infrared energy must be at least equal to the energy absorbed in the UV-optical regime. This is also discussed in more detail by Best et al. (in prep.) but in conclusion, the *consensus* SFRs computed are therefore expected to be more robust and hence used for source classification by Best et al. (in prep.) and for other analysis in this work. We also note that the good agreement in the evolution of the total radio-AGN LFs seen in Sect. 4.3.5 and also for the star-forming galaxy LFs (Cochrane et al. in prep.) gives us more confidence in our adopted classification scheme.

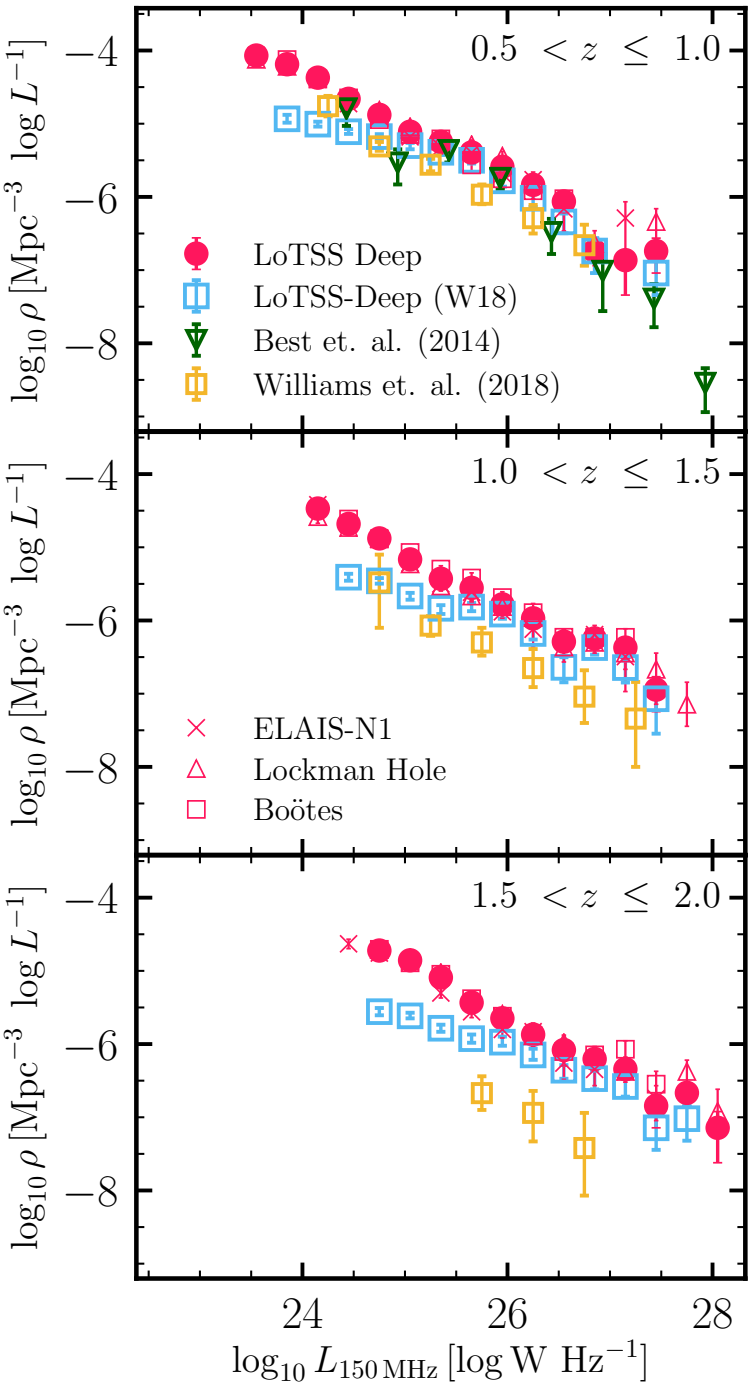

Figure 4.9 *Comparison of the cosmic evolution of LERG LFs with Williams et al. (2018) using different source classification schemes. We show the LFs resulting from the LoTSS-Deep classification scheme (pink circles; as in Fig. 4.8), and the resultant LoTSS-Deep LF when the source classification method of Williams et al. (2018) is applied to our LoTSS-Deep sample (cyan squares; see also Sec. 4.4). The literature results from Williams et al. (2018) (yellow squares) and Best et al. (2014) (green triangles) are also shown.*

4.4.1 Modelling the cosmic evolution of the LFs

The radio luminosity function of AGN is generally modelled as a broken power-law of the form

$$\rho(L) = \frac{\rho^\star}{(L^\star/L)^\beta + (L^\star/L)^\gamma}, \tag{4.8}$$

where $\rho^\star$ is the characteristic space density, $L^\star$ is the characteristic luminosity, and, β and γ are the bright and faint end slopes, respectively (e.g. Dunlop & Peacock, 1990). As evident from comparison with the $0.5 < z \leq 1$ Best et al. (2014) LF in Fig. 4.8, the area covered by the first data release of the LoTSS Deep Fields is not sufficient to probe bright enough luminosities to constrain the bright-end slope of the LERG LF, at all redshifts. Therefore, we considered modelling the evolution of the LERG population as the luminosity evolution and density evolution of the local LF such that in equation 4.8, $\rho^\star(z)$ accounts for the redshift evolution of the normalisation and $L^\star(z)$ is the redshift evolution of the characteristic luminosity. For this process, we assumed that the shape of the LF remains constant by fixing the bright-end slope $\beta = -1.27$, and the faint-end slope $\gamma = -0.49$, as found by a broken power-law fit to the local radio-AGN LF by Mauch & Sadler (2007). Although the Mauch & Sadler LF includes both the LERGs and HERGs, the faint-end slope, which is the key parameter to constrain as our data does not probe the bright end well, will be dominated by the LERGs; the Mauch & Sadler (2007) faint-end slope provides a better match to our dataset than, for example, the jet-mode AGN LF of Best et al. (2014). We also find that a broken power law fit to the radio-AGN LF of Sabater et al. (2019) gives a faint-end slope consistent with the Mauch & Sadler (2007) value, and hence gives similar results in modelling the evolution below.

We first considered a luminosity and density evolution (LDE) model where we allowed $\rho^\star(z)$ and $L_\star(z)$ to be free parameters: both of these parameters were then fitted for each redshift bin using the EMCEE Markov Chain Monte Carlo (MCMC) fitting routine (Foreman-Mackey et al., 2013), with the resulting fitted $\rho^\star(z)$ and $L_\star(z)$ shown in Table 4.5. This LDE model can also be expressed in terms of a $(1+z)$ parameterisation such that the evolution of the normalisation is expressed as $\rho^\star(z) = \rho^\star(0) \times (1+z)^{\alpha_\mathrm{D}}$, and, the redshift evolution of the characteristic luminosity, $L^\star(z)$, is given as $L^\star(z) = L^\star(0) \times (1+z)^{\alpha_\mathrm{L}}$. Here, α_D and α_L correspond to the density and luminosity evolution parameters, respectively. This can be used to both illustrate how strong the evolution is, and if it follows a simple trend with redshift. For these cases, we used the Mauch & Sadler normalisation $\rho^\star(0) = \frac{1}{0.4} 10^{-5.5} \log \mathrm{L}^{-1}\,\mathrm{Mpc}^{-3}$ (converted to per log luminosity) and the characteristic luminosity $L^\star(0) = 10^{25.27}\,\mathrm{W\,Hz^{-1}}$ (scaled to 150 MHz assuming

$\alpha = -0.7$) along with their two slopes. The corresponding values for the evolution parameters for this LDE model, along with the reduced chi-squared goodness of fit values, χ^2_ν, are also listed in Table 4.5.

For each redshift bin, we also considered a pure luminosity evolution (PLE) model (i.e. we set $\alpha_D = 0$) and a pure density evolution (PDE) model (i.e. we set $\alpha_L = 0$). The resulting best-fitting α_D and α_L parameters (as appropriate for each model) for each redshift bin are listed in Table 4.5, along with the χ^2_ν values. The resulting best-fitting LFs for all three models at each redshift bin are also shown in Fig. 4.8. We find that both the PLE and PDE models are not sufficient at describing the evolution of the LERGs, in particular at high radio luminosities, as also noted by the resulting χ^2_ν values in Table 4.5. The LDE model is better able match the high-luminosity end of the LF. The LDE fits indicate that at least out to $z \sim 2$, there is a mild increase in the characteristic luminosity and a corresponding mild decrease in the characteristic space density with increasing redshift. This appears to reverse in the highest redshift bin, however it should be noted that there is some degeneracy between the fitted $L_\star(z)$ and $\rho_\star(z)$, especially due to the lack of a prominent break in the LFs.

Finally, we considered how uncertainties in the source classification criteria employed affects our results on the LERG luminosity functions; this is discussed in detail in Appendix 4.A. In summary, given the good agreement between our LFs and literature results of the total radio-AGN LFs across redshift, we studied the effects of varying the selection of 'optical-AGN' on the LERG to HERG classification. We find that the evolution of the LERG population is largely insensitive (at the $\lesssim 0.2$ dex level) to the exact threshold used for separating the two modes of AGN, even under a very conservative definition of HERGs; this gives us confidence in our source classification criteria adopted and on the resulting evolution of the LERG LFs.

To understand the origin of the mild evolution seen in the luminosity functions of the LERGs, we investigate the incidence of the LERGs as a function of stellar mass and redshift in different types of host galaxies in Sect. 4.5.

Table 4.5 *Results of modelling the evolution of the LERG LFs using a pure luminosity evolution (PLE), pure density evolution (PDE), and luminosity and density evolution (LDE) models, relative to the local relation from Mauch & Sadler (2007), as detailed in Sect. 4.4.1. The best-fitting LFs for each model are also shown in Fig. 4.8.*

z	PLE			PDE			LDE				
	$\log_{10} L_\star(z)$	α_L	χ^2_ν	$\log_{10} \rho^\star(z)$	α_D	χ^2_ν	$\log_{10} \rho^\star(z)$	$\log_{10} L_\star(z)$	α_D	α_L	χ^2_ν
$0.5 < z \leq 1.0$	25.59	$1.30^{+0.07}_{-0.07}$	5.46	-4.89	$0.88^{+0.05}_{-0.05}$	8.53	-5.34	25.95	$-0.99^{+0.28}_{-0.29}$	$2.78^{+0.44}_{-0.47}$	4.72
$1.0 < z \leq 1.5$	25.50	$0.65^{+0.05}_{-0.05}$	11.85	-4.94	$0.46^{+0.04}_{-0.04}$	15.83	-5.72	26.38	$-1.76^{+0.23}_{-0.25}$	$3.15^{+0.37}_{-0.41}$	5.84
$1.5 < z \leq 2.0$	25.79	$1.18^{+0.04}_{-0.04}$	8.74	-4.67	$0.98^{+0.03}_{-0.03}$	16.16	-5.62	26.49	$-1.17^{+0.17}_{-0.17}$	$2.78^{+0.26}_{-0.27}$	6.05
$2.0 < z \leq 2.5$	25.58	$0.60^{+0.04}_{-0.04}$	2.78	-4.83	$0.53^{+0.04}_{-0.04}$	3.62	-5.23	25.73	$-0.25^{+0.22}_{-0.22}$	$0.90^{+0.27}_{-0.28}$	2.95

4.5 Prevalence of AGN activity with stellar mass and star-formation rate

It is well-known that the prevalence of radio-AGN increases strongly with stellar mass in the local Universe (e.g. Best et al., 2005b; Smolčić et al., 2009; Janssen et al., 2012; Sabater et al., 2019) and early studies out to $z \sim 1$ suggest that this trend holds at earlier cosmic time (e.g. Tasse et al., 2008; Simpson et al., 2013). Here, we study the redshift evolution of LERG activity by measuring how the fraction of galaxies that host a LERG varies with stellar mass and redshift.

For the LERG sample, we used the *consensus* stellar masses derived from SED fitting (see Sect. 4.2.1). For the underlying galaxy population, we used the stellar masses (50th percentile of the posterior distribution) computed by Duncan et al. (2021) using a grid-based SED fitting method (see also Duncan et al., 2014, 2019) for the full multi-wavelength catalogue in each field. Duncan et al. (2021) validated their stellar masses for the population as a whole using comparison with literature galaxy stellar mass functions, and also showed them to be in good agreement on a source-by-source basis (in ELAIS-N1, where comparison was possible); Best et al. (in prep.) also find no systematic offset between the Duncan et al. and the *consensus* stellar masses. As detailed by Duncan et al. (2021), the photometric redshifts, and hence the derived stellar masses, are found to be reliable at $z < 1.5$ for host-galaxy dominated sources; we therefore restrict our analysis of the LERG mass fractions to $z < 1.5$ in this section. In deriving the LERG mass fractions, to avoid biasing our results by stellar mass incompleteness, we restrict analysis to masses above the 90 per cent stellar mass completeness limits estimated by Duncan et al. (2021), following the empirical Pozzetti et al. (2010) method. In summary, within narrow redshift slices, 20% of the faintest sources above the 90% magnitude completeness limit (in the K-band) were selected. Then, the limiting stellar mass was computed for this subset of sources, defined as the mass of a galaxy at its redshift if its apparent magnitude was equal to the limiting magnitude. The magnitude limit is then determined to be the 95th percentile of the limiting stellar mass distribution[1]. We estimated the stellar mass completeness limit in each of the five redshift bins as the stellar mass above which a source would be detected over the full redshift bin, and simply removed sources with stellar masses below this completeness limit from the

[1]The completeness limits estimated in this manner are expected to be appropriate for the galaxy population, however, we note that for AGN in the star-forming population (see Sect. 4.5.1), we may be complete to lower masses, albeit only by one or two mass bins; we therefore do not expect this to have a significant effect on the derived results.

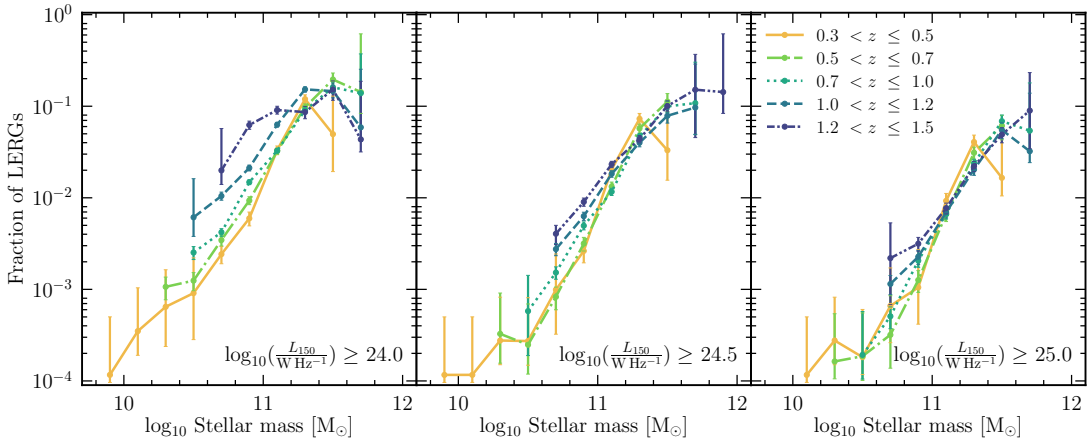

Figure 4.10 *Fraction of galaxies hosting a LERG as a function of stellar mass in five redshift bins between* $0.3 < z \leq 1.5$ *in the LoTSS Deep Fields. Panels from left-to-right show this for increasing radio luminosity limits of* $L_{150\,\mathrm{MHz}} \geq 10^{24}\,\mathrm{W\,Hz^{-1}}$, $L_{150\,\mathrm{MHz}} \geq 10^{24.5}\,\mathrm{W\,Hz^{-1}}$, *and* $L_{150\,\mathrm{MHz}} \geq 10^{25}\,\mathrm{W\,Hz^{-1}}$. *The error bars represent Poisson uncertainties, following Gehrels (1986) for* $\mathrm{N} < 10$. *The fraction of galaxies that host a LERG shows a steep stellar mass dependence, in particular at high radio luminosity limits; this relation remains at least out to* $z \sim 1.5$. *We also find evidence of this relation flattening at lower radio luminosity limits suggesting the presence of an additional fuelling mechanism which is investigated in Sect. 4.5.1.*

analysis.

In Fig. 4.10, we present the fraction of all galaxies that host a LERG as a function of stellar mass in five redshift bins ($0.3 < z \leq 0.5$, $0.5 < z \leq 0.7$, $0.7 < z \leq 1.0$, $1.0 < z \leq 1.2$, $1.2 < z \leq 1.5$), shown by different coloured lines, for three equally-spaced radio luminosity limits (shown in different panels) from $L_{150\,\mathrm{MHz}} \geq 10^{24}\,\mathrm{W\,Hz^{-1}}$ to $L_{150\,\mathrm{MHz}} \geq 10^{25}\,\mathrm{W\,Hz^{-1}}$. The lowest radio luminosity limit chosen corresponds to a 5σ detection (based on noise in the deepest regions) at $z \sim 1.5$, and is broadly comparable to the 1.4 GHz luminosity limit ($L_{1.4\,\mathrm{GHz}} = 10^{23}\,\mathrm{W\,Hz^{-1}}$, assuming $\alpha = -0.7$) typically used in similar studies in the local Universe. To account for the lack of a volume-limited sample, we weight each source by a $1/\mathrm{V_{max}}$ factor based on the maximum volume that a source could be detected in at above 5σ based on the local radio RMS level. The error bars show Poisson uncertainties (following Gehrels (1986) for stellar mass bins with $\mathrm{N} < 10$). In total, this results in 2722, 1776, and 835 LERGs (within $0.3 < z \leq 1.5$) with $L_{150\,\mathrm{MHz}} \geq 10^{24}\,\mathrm{W\,Hz^{-1}}$, $L_{150\,\mathrm{MHz}} \geq 10^{24.5}\,\mathrm{W\,Hz^{-1}}$, and $L_{150\,\mathrm{MHz}} \geq 10^{25}\,\mathrm{W\,Hz^{-1}}$, respectively.

For the high radio-power LERGs, corresponding to a radio luminosity limit of

10^{25} W Hz^{-1} (right panel of Fig. 4.10), we find that the fraction of galaxies that host a LERG has a steep dependence on the stellar mass, approaching ~ 10 per cent at the highest stellar masses; these results agree with observations in the local Universe (e.g. Best et al., 2005a; Janssen et al., 2012) and show no signs of redshift evolution. As the radio luminosity limit is lowered (i.e. going from the right panels to left in Fig. 4.10), we observe an increase in the LERG fraction at lower stellar masses at all redshifts, and in particular at higher redshifts, resulting in a shallower dependence of the overall AGN fraction with stellar mass. Physically, the strong dependence with stellar mass is expected from arguments based on the fuelling mechanism of the LERGs (e.g. Best et al., 2006; Hardcastle et al., 2007). The flattening of this relation, in particular at lower radio luminosities and stellar masses where star-forming galaxies are expected to dominate the galaxy stellar mass function, suggests the presence of a star-forming LERG population with a potentially different fuelling mechanism. We investigate this in more detail in Sect. 4.5.1 by considering the dependence of AGN activity with both stellar mass and star-formation rate.

4.5.1 Prevalence of LERGs in star-forming and quiescent galaxies

It is well known that star-forming galaxies occupy a well-defined sequence in the SFR-$M_\star$ plane, known as the 'main-sequence' of star-formation (e.g. Whitaker et al., 2012; Speagle et al., 2014a; Schreiber et al., 2015; Tasca et al., 2015), with quiescent galaxies lying below this sequence. Therefore, the ratio of the star-formation rate and the stellar mass of a galaxy, known as the *specific* star-formation rate (sSFR) can be used to identify quiescent galaxies (e.g. Fontana et al., 2009; Pacifici et al., 2016; Merlin et al., 2018). In this study, using the *consensus* SFRs and stellar masses derived from SED fitting (see Sect. 4.2.1), we select sources as quiescent galaxies if they satisfy the condition

$$\mathrm{sSFR} < \mathrm{f_{sSFR}}/\mathrm{t_H(z)}, \tag{4.9}$$

where, $t_\mathrm{H}(z)$ is the age of the Universe at redshift z of the source, and $f_\mathrm{sSFR} = 1/5$ defines the threshold in sSFR. Galaxy samples selected by this criteria have been found to show good agreement with traditional UVJ rest-frame colour selected samples (e.g. Pacifici et al., 2016; Carnall et al., 2018, 2020).

Fig. 4.11 shows the SFR-$M_\star$ relation for the LERG sample in ELAIS-N1 in the same five redshift bins as in Fig. 4.10 ($0.3 < z \leq 1.5$), split into those hosted by quiescent and

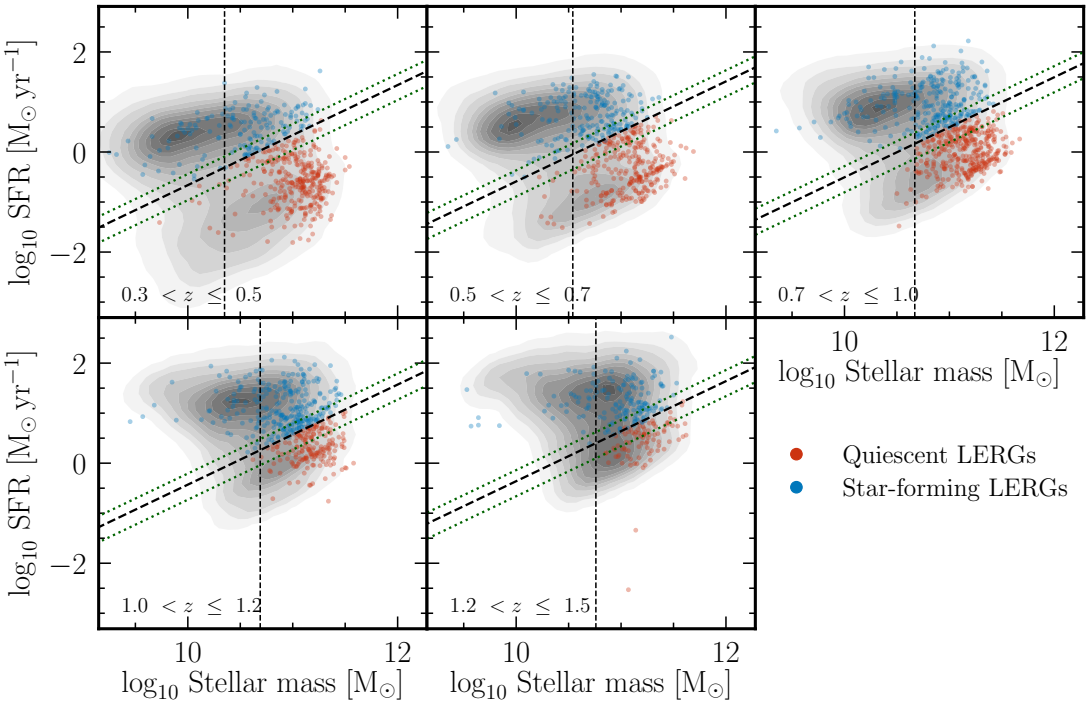

Figure 4.11 *The location of LERGs in the ELAIS-N1 field on the SFR-stellar mass ($M_\star$) plane in five redshift bins ($0.3 < z \leq 1.5$). The red and blue points correspond to quiescent and star-forming host galaxies of LERGs, respectively. The shaded contours show the SFR-$M_\star$ distribution (from* MAGPHYS*) for the underlying IRAC flux-selected sample from Smith et al. (2021) in each redshift bin. The dashed black line corresponds to the sSFR selection threshold adopted for identifying quiescent galaxies (i.e. $f_{\rm sSFR} = 1/5$ in equation 4.9), as described in Sect. 4.5.1. Similarly, the green dotted lines, above and below this line correspond to sSFR thresholds based on $f_{\rm sSFR} = 1/3$ and $f_{\rm sSFR} = 1/10$, respectively. The vertical lines correspond to the 90 per cent stellar mass completeness limits for each redshift bin in ELAIS-N1; sources below this are removed from the analysis of LERG fractions in Fig. 4.12.*

star-forming galaxies (red and blue points, respectively), separated using the criteria from equation 4.9. The black dashed line corresponds to the sSFR $< 0.2/t_H(z)$ threshold used to separate the two populations; this results in 1238 quiescent and 1100 star-forming host galaxies of LERGs within $0.3 < z < 1.5$. We also note here that the vast majority of the star-forming LERGs lie on or below the main-sequence. The green dotted lines above and below this dividing line correspond to $f_{sSFR} = 1/3$ and $f_{sSFR} = 1/10$, respectively, in equation 4.9; we use these variations on the standard sSFR threshold to test the robustness of our results on the prevalence of LERGs (see below). Vertical dashed lines show the stellar mass completeness limits in ELAIS-N1 in each redshift bin that are applied when generating the LERG fraction plot. The grey shaded contours in each redshift bin show the SFR-$M_\star$ relation for the underlying parent galaxy population in ELAIS-N1; these are drawn such that the outermost contour encompasses 99 per cent of all the underlying sources. For this underlying population, we use the $F_{3.6\mu m} > 10\,\mu$Jy IRAC selected sample of Smith et al. (2021) in ELAIS-N1, consisting of 183 399 sources for which SED fitting was performed using MAGPHYS. We limit the following analysis to the ELAIS-N1 field as a similar MAGPHYS SED fitting output catalogue for the underlying population is not available in the other two fields. Limiting to the sources with the best available multi-wavelength coverage, good χ^2 SED fits, and within the chosen redshift range (i.e. $0.3 < z \leq 1.5$) results in a final sample of 140 754 sources in the underlying population that are used in this analysis. We separate this underlying population into star-forming and quiescent galaxies using the same sSFR threshold as that used for the LERG population. As evident from the panels in Fig. 4.11, the $f_{sSFR} = 1/5$ threshold is found to be an appropriate division between star-forming and quiescent galaxies at all redshifts.

Fig. 4.12 shows the fraction of quiescent galaxies (of a given mass) that host a LERG (*left*), and similarly, the fraction of star-forming galaxies hosting a LERG (*right*), for the same five redshift bins as in Fig. 4.10, with a radio luminosity limit of $L_{150\,MHz} \geq 10^{24}\,W\,Hz^{-1}$. The black dashed line also plotted in each panel shows the $f_{LERG} \approx 0.01(M_\star/10^{11}\,M_\odot)^{2.5}$ relationship found for the prevalence of jet-mode AGN with stellar mass in the local Universe (Best et al., 2005a). The exact value of the normalisation in this relationship depends on the luminosity limit and hence on the spectral index adopted; the line drawn in Fig. 4.12 is for illustration purposes only. The fraction of quiescent galaxies hosting a LERG agrees well with this steep stellar mass dependence found in the local Universe, showing essentially no evolution with redshift. In contrast, the LERG fractions for the star-forming host galaxies show a weaker dependence on stellar mass, with this relation evolving with redshift such that the LERG fraction increases with increasing redshift at a fixed stellar mass, with the prevalence of these

LERGs in star-forming hosts reaching comparable levels to the quiescent hosts at high redshifts.

As discussed in Sect. 4.4 and Appendix 4.A, our selection of radio-excess AGN based on radio-excess $> 3\sigma$ may miss low-luminosity radio-AGN, hosted in massive star-forming galaxies, in particular, where the radio-emission is dominated by star-formation processes rather than the jets; this may impact the results determined in Fig. 4.12 given our adoption of an absolute radio-luminosity limit, as sources of a given jet power would be harder to identify in such star-forming galaxies. Missing low-luminosity radio-AGN at high stellar masses may have led to an artificial flattening of the relation with stellar mass for the star-forming LERGs. We have investigated applying a higher radio-luminosity limit, $L_{150\,\mathrm{MHz}} \geq 10^{25}\,\mathrm{W\,Hz^{-1}}$; this ensures a cleaner sample of radio-AGN as sources with such high radio powers are likely dominated by the jets rather than star-formation processes. Using this higher radio luminosity limit results in a steeper relation with stellar mass for the star-forming LERGs. However, such a high radio luminosity limit removes a large fraction of the LERGs from the sample, leading to poorer statistics. Moreover, this limit also means that our investigation is appropriate for LERGs with the most powerful jets, rather than the "typical" LERG population; these results are hence also affected by the selection effects. More robust measurements of the star-forming LERG population will be enabled by upcoming datasets from sub-arcsecond LOFAR imaging (Sweijen et al. submitted) and the WEAVE-LOFAR survey (see Chapter 5). The high-resolution radio dataset will allow us to identify the presence of jets associated with AGN, even in star-forming galaxies, allowing a more robust selection of the star-forming LERGs, whereas spectroscopy from the WEAVE-LOFAR survey will complement the separation of radio sources into star-forming galaxies and AGN using emission line diagnostics.

To quantitatively investigate the trends in the evolution of the LERG mass fractions seen in Fig. 4.12, we parametrise the LERG mass fractions as a power law of the form

$$f_{\mathrm{LERG}}(M_\star) = c\left(\frac{M_\star}{10^{11}\,\mathrm{M}_\odot}\right)^{\beta} \tag{4.10}$$

such that $f_{\mathrm{LERG}}(M_\star)$ is the LERG fraction at mass $M_\star$, c is the normalisation at $M_\star = 10^{11}\,\mathrm{M}_\odot$, and β is the power law slope. We then fitted this power law form to the LERG fractions shown in Fig. 4.12 at each redshift individually, for both the star-forming and quiescent LERG populations.

Fig. 4.13 shows the results of this power law fitting process for the quiescent (red) and star-forming (blue) LERG populations in the same five redshift bins as the ones used in

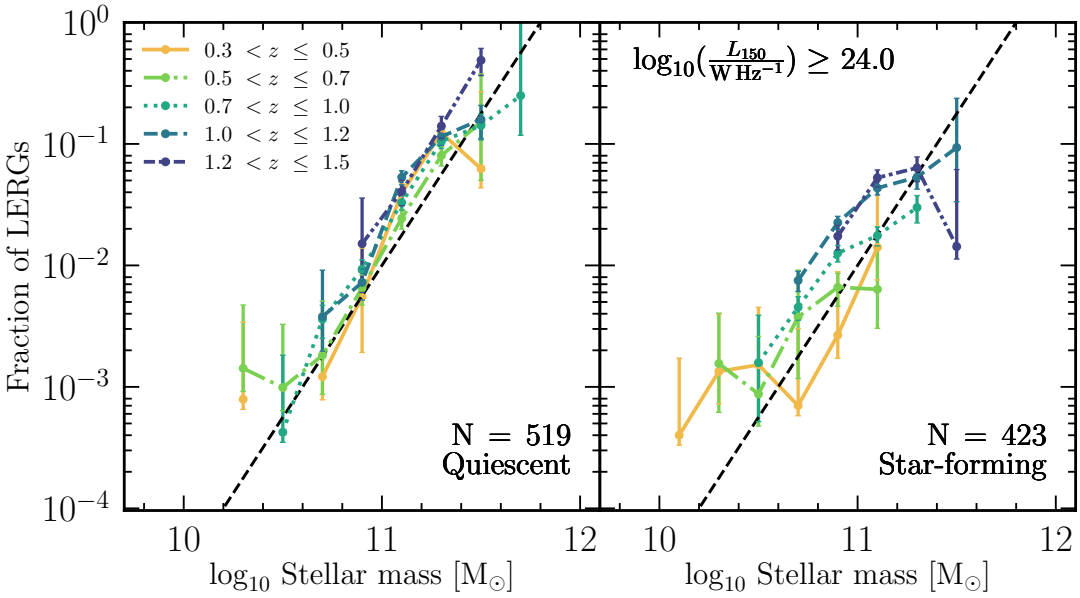

Figure 4.12 *Fraction of quiescent LERGs with $L_{150\,\mathrm{MHz}} \geq 10^{24}\,\mathrm{W\,Hz^{-1}}$ hosted in quiescent galaxies (left), and the fraction of star-forming LERGs hosted in star-forming galaxies (right) as a function of stellar mass in ELAIS-N1. The number in each panel lists the number of quiescent and star-forming LERGs, respectively, within $0.3 < z \leq 1.5$, for the given radio luminosity limit. The error bars represent Poisson uncertainties, following Gehrels (1986) for $\mathrm{N} < 10$. The black dashed line shows the $f_{\mathrm{LERG}} \approx 0.01(M_\star/10^{11}\,\mathrm{M}_\odot)^{2.5}$ relation found from studies of LERGs in the local Universe; the normalisation of this relation however depends on the radio luminosity limit, and hence the assumed spectral index. We find that the quiescent LERG fractions show a steep stellar mass dependence with essentially no evolution with redshift out to $z \sim 1.5$. The star-forming LERGs show a much shallower dependence on stellar mass, with an increase in the normalisation with redshifts.*

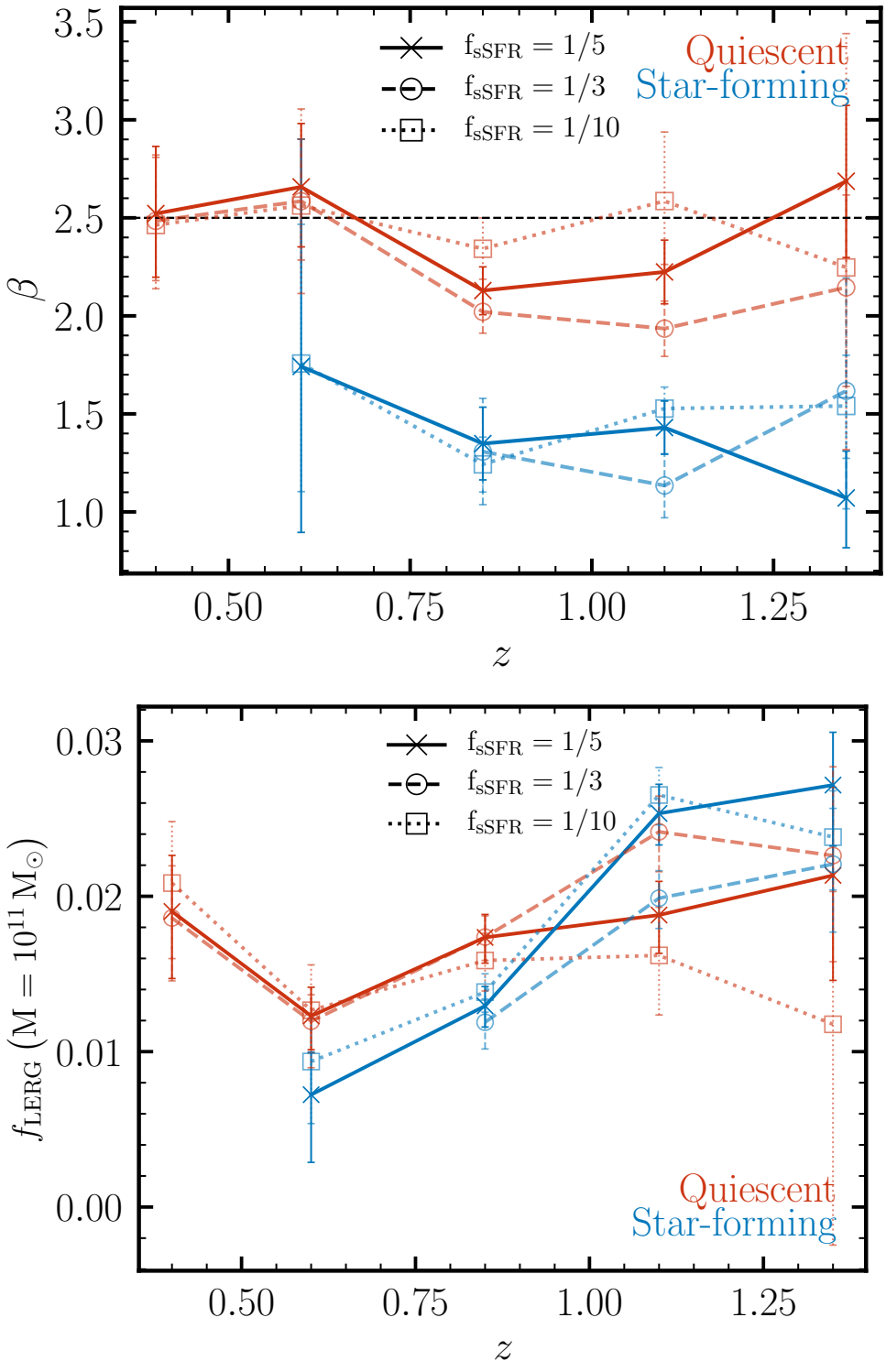

Figure 4.13 *(Caption on the following page).*

Figure 4.13 *Results of power-law fits to the star-forming and quiescent LERG fractions with a functional form given by equation 4.10. top panel shows the evolution of the power law slope for quiescent (red lines) and star-forming (blue lines) hosts and the bottom panel similarly shows the evolution in the normalisation of the LERG fractions at $M_\star = 10^{11}\,\mathrm{M}_\odot$. The solid line (with crosses) corresponds to the results obtained from using the $f_{\mathrm{sSFR}} = 1/5$ threshold adopted throughout this chapter for separating star-forming and quiescent galaxies (see Sect. 4.5.1). The dashed and dotted lines correspond to results obtained from using variations of the f_{sSFR} threshold adopted. The black horizontal line in the top panel corresponds to the slope obtained in the local Universe (e.g. Best et al., 2005a; Janssen et al., 2012). We find a clear difference in the slope of LERG mass fractions between the quiescent and star-forming hosts of LERGs. The steep power law slope of the quiescent LERGs agrees well with the local Universe measurement at all redshifts, consistent with these AGN being fuelled from the hot gas haloes in massive galaxies; the flatter relation for the star-forming LERGs suggests that a different fuelling mechanism associated with the accretion at low levels from the cold gas within these galaxies may be present. Confidence for this link with the cold gas also comes from the increase in the normalisation of the LERG fractions for this sub-group at earlier times when gas fractions were higher.*

Fig. 4.12. The *top* panel of Fig. 4.13 shows the redshift evolution of the power law slope, β, and *bottom* panel shows the redshift evolution of the normalisation at $M_\star = 10^{11}\,\mathrm{M}_\odot$ for both populations. The error bars show the 1σ uncertainties in the fitted parameters. We omit data points where the uncertainties are of order or larger than the magnitude of the fitted parameters due to a poor fit; this in particular affected the lowest redshift bin of the star-forming subset due to large uncertainties in the AGN fractions. In each panel, the solid lines correspond to $f_{\mathrm{sSFR}} = 1/5$, our adopted threshold (see Sect. 4.5.1 for details) for separating star-forming (blue) and quiescent (blue) host-galaxies used throughout this study. Also shown are the results from fitting the AGN fractions that were derived based on variations of the f_{sSFR} selection thresholds equal to $1/3$ and $1/10$ (see also Fig. 4.11). The black dashed line in the *top* panel corresponds to the power law slope ($\beta = 2.5$) found in studies of jet-mode AGN population in the local Universe (e.g. Best et al., 2005a; Janssen et al., 2012).

We find that the quiescent and star-forming LERG populations show distinct power law slopes, with the quiescent LERG mass fractions having a steep slope, close to the local redshift relation, that remains roughly constant with redshift. Moreover, if a stricter definition of 'quiescence' is used (i.e. $f_{\mathrm{sSFR}} = 1/10$; dotted line), we find

that the slope of our quiescent LERG population at higher redshifts agrees even better with the local relation. The star-forming LERG fractions show a much shallower slope, which is interestingly similar to the relation of the radiative-mode AGN population in the local Universe (e.g. Janssen et al., 2012). As is evident from Fig. 4.13, the LERG fractions resulting from the two alternative selection thresholds for quiescent/star-forming galaxies agree well with the values derived from our adopted 'quiescence' criteria (within 1σ based on our uncertainties); this demonstrates that our results are robust to changes in how quiescent and star-forming host-galaxies are selected.

Looking at the evolution of normalisation of the power law at $M = 10^{11}\,\mathrm{M}_{\odot}$ in Fig. 4.13 (*bottom*), we find that for the quiescent LERGs, the normalisation stays roughly constant out to $z \sim 1.5$. In contrast, the star-forming subset shows a strong increase in the normalisation of the LERG fraction with increasing redshift, showing an increase by a factor of ~ 4 by $z \sim 1$, and showing a higher prevalence of being hosted in star-forming galaxies than quiescent galaxies at these redshifts.

4.5.2 The nature of star-forming and quiescent LERGs

Physically, the differences observed between the quiescent and star-forming hosts of LERGs suggests that the LERGs in these galaxies are fuelled by different mechanisms. For the quiescent LERGs, we observe a steep stellar mass dependence on the LERG fraction, similar to that found in the local Universe (e.g. Best et al., 2005a; Janssen et al., 2012), where the LERG population is dominated by red, quiescent host galaxies. The quiescent LERGs are understood to be associated with fuelling from the hot phase (e.g. Allen et al., 2006; Best et al., 2006; Hardcastle et al., 2007), which requires a supply of hot gas and a massive black hole; given the correlation between galaxy stellar mass and both the black hole mass and galaxy halo mass, we expect a strong correlation between LERG activity and both stellar and black hole mass. Best et al. (2005a) and Best et al. (2006) have indeed shown that radio-loud AGN activity $f_{\mathrm{radio-loud}}$, depends strongly on both stellar mass ($f_{\mathrm{radio-loud}} \propto M_{\star}^{2.5}$) and black hole mass ($f_{\mathrm{radio-loud}} \propto M_{\mathrm{BH}}^{1.6}$). Best et al. (2006) also showed that the black hole mass dependence of radio-loud AGN activity is very similar to the black hole mass dependence on the black hole accretion rate under standard Bondi accretion. These results have interesting implications for the AGN-feedback cycle of LERGs: they show that the time-averaged mechanical energy from the radio-AGN is able to balance the cooling losses from their hot host haloes, providing evidence for an AGN-feedback controlled fuelling of the black hole as the likely fuel source of the AGN is the same medium into which energy from the

AGN is deposited (see also Heckman & Best 2014). Therefore, the steep stellar mass dependence observed for the quiescent LERGs and the lack of redshift evolution seen in Fig. 4.12 suggests that these systems are fuelled by cooling of hot gas from haloes, and that the AGN-feedback cycle may be in place in these galaxies since at least $z \sim 1.5$ (see also Sect. 4.7).

In contrast, the LERG activity in star-forming galaxies shows a much flatter dependence with stellar mass, suggesting that a different physical mechanism may be responsible for fuelling the AGN in these galaxies. Given that this population is undergoing a recent episode of star-formation, likely fuelled by the availability of cold gas (e.g. via mergers or secular processes), we may also expect this cold gas to provide a fuel source for the black hole as well. Cold gas accretion is often thought to lead to a radiative-mode AGN (which are widely associated with star-forming galaxies), but these star-forming LERGs are understood to be fuelled by a radiatively-inefficient accretion flow; hence, the cold gas associated with the on-going star-formation must also be capable of fuelling the black hole in a different manner. This is consistent with the idea that it is the Eddington-scaled accretion rate onto the black hole that determines the accretion mode (i.e. radiatively efficient/inefficient mode) and not necessarily the fuel source, such that it is possible to achieve accretion of cold material at low Eddington rates, resulting in a radiatively inefficient mode of AGN (e.g. see discussion by Best et al., 2014; Hardcastle, 2018; Hardcastle & Croston, 2020).

Other evidence in support of this picture comes from the observation of enhanced LERG activity in interacting galaxies (e.g. Sabater et al., 2013; Gordon et al., 2019) and that in lower-mass galaxies ($M_\star \lesssim 10^{11}\,\mathrm{M}_\odot$), the merger fraction of LERGs is higher than that of a control sample (by ~ 20 per cent; Gao et al. 2020); these results could be due to the supply of cold gas brought in to the galaxy due to the interaction/merger events. Indeed there is evidence for large reservoirs of cold molecular gas ($\sim 10^8 - 10^{10}\,\mathrm{M}_\odot$) in the host galaxies of LERGs in particular (e.g. Okuda et al., 2005; Ocaña Flaquer et al., 2010; Ruffa et al., 2019a), along with evidence for the cold gas playing a role in the fuelling of the AGN in some systems (e.g. Tremblay et al., 2016; Maccagni et al., 2018; Ruffa et al., 2019b). However, it is also worth noting that the presence of cold gas within the host galaxy does not necessarily imply that this gas reaches and accretes on to the central engine; the cold gas in some systems is mostly found in a relaxed rotating disc (e.g. North et al., 2019) and as such would be expected to lead to only relatively low accretion rates.

In the above scenario, the HERGs may then be expected to be the high accretion rate (or low black hole mass) extension of this population, such that stochastic variations

in the fuelling of the black hole may lead to the different nuclear properties, while it is also possible that fundamental differences in the nature and origin of the accreting gas may lead to the observed differences. The present data are not sufficient for us to determine the physical processes underpinning the AGN fuelling and the direct link with the cold gas; instead detailed characterisation of the multi-phase medium and its kinematics for matched samples of the three sub-groups of the radio-excess AGN are required. Nonetheless, the trends of shallower stellar mass dependence of the LERG fractions and the increase in their normalisation at earlier times is consistent with the idea that a different mechanism is responsible for fuelling the LERG activity in these more star-forming galaxies, likely associated with the cold gas component.

4.6 The luminosity functions of the quiescent LERGs

Given that we see this significant population of LERGs being hosted in star-forming galaxies at higher redshifts in Sect. 4.5, it is informative to study how the luminosity functions of both of these populations evolve with redshift in order to understand the evolution of the LERG population.

Fig. 4.14 shows the evolution of the total LERG LFs (black points, as shown in Fig. 4.8), but also split into the subsets that are hosted by quiescent (red triangles) and star-forming (blue squares) galaxies. The error bars represent the 1σ uncertainties with the LFs also tabulated in Table 4.4. Across the four redshift bins, the LFs are constructed in total for 3406 quiescent and 5024 star-forming LERGs. We find that LERGs hosted by star-forming and quiescent galaxies, are both found across all radio luminosities. At $z < 1$, the total LERG population is primarily hosted in quiescent galaxies at almost all radio luminosities, with the star-forming population reaching comparable space densities only at the faint end. At higher redshifts, we find a steady decline in the space density of the quiescent LERGs; simultaneously, the space density of the star-forming subset increases steadily with redshift, with a switch-over in the dominant population occurring at $1 < z \leq 1.5$. This star-forming LERG population dominates over the quiescent LERGs across all radio luminosities at $z > 1.5$. We note that this trend is not simply due to an increase in the number of star-forming galaxies at early epochs as our analysis of the AGN mass fractions (see Sect. 4.5.1), which account for this effect, show an increase in the prevalence of this population too. The combination of the evolution of these two sub-groups can thus explain the source of the little evolution seen in the

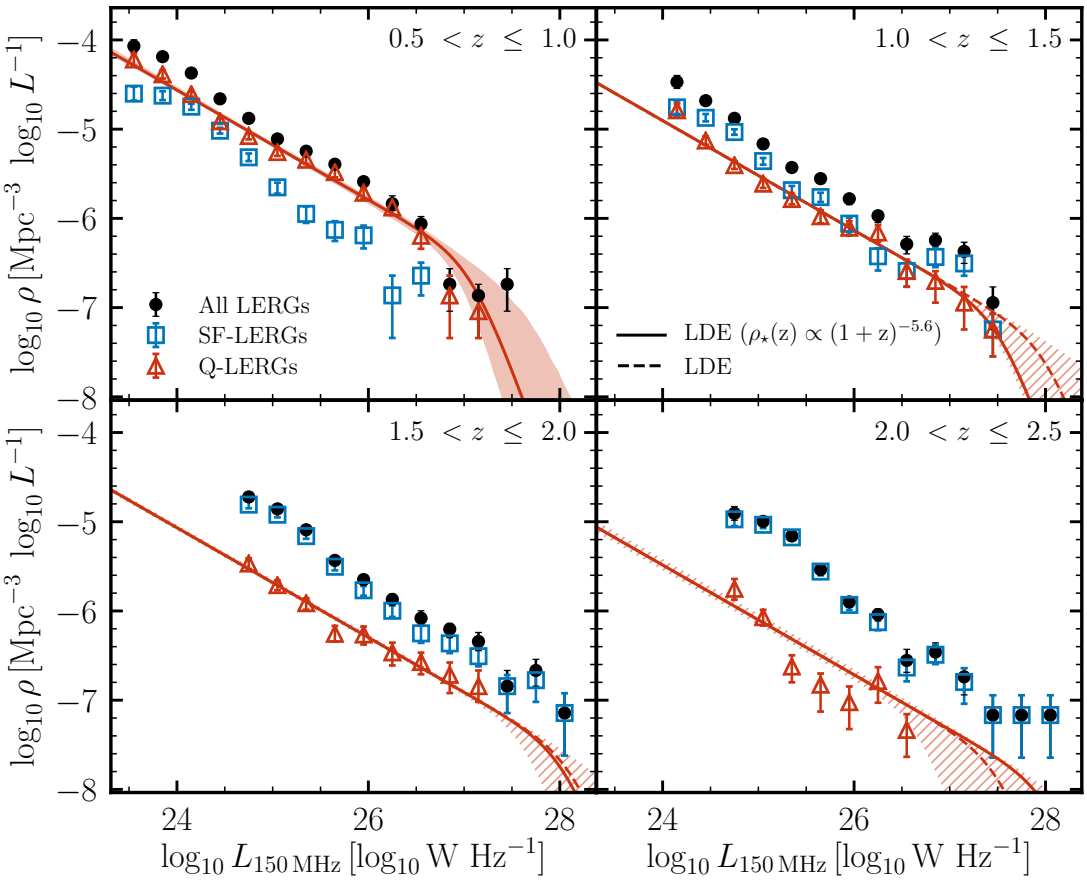

Figure 4.14 *The evolving LFs for the LERGs in LoTSS-Deep split into the subset hosted by quiescent and star-forming galaxies for the same redshift bins as in Fig. 4.8. The space densities of the quiescent LERGs dominate the total LERG population and decrease steadily with redshift, however the star-forming LERGs begin to prevail beyond* $z \sim 1.5$*; the combination of the evolution of these two sub-groups explains the evolution of the total LERG population seen in Sect. 4.4. For the quiescent LERGs, we show a broken power law fit to the data in the* $0.5 < z \leq 1$ *bin, with the shaded region forming the* 1σ *uncertainty on the best fit. At higher redshifts, the dashed line and hatched regions represent the best-fit and* 1σ *uncertainties from a 2-parameter LDE fit, whereas, the solid line shows an LDE fit where the density evolution is fixed based on the evolution of the expected host galaxies, and can be seen to provide a good match to the data within the* 1σ *region at each redshift.*

total LERG LFs (see Sect. 4.4).

We also considered how uncertainties in the separation of LERGs and HERGs affect the LFs of the quiescent and star-forming LERGs; this analysis is detailed in Appendix 4.A. In summary, we modified and applied additional criteria for selecting 'optical-AGN' (i.e. those that show signs of MIR emission from an AGN, typical for HERGs) and reconstructed the LFs of the HERGs and the two sub-groups of LERGs. We found that at all redshifts, the quiescent LERGs are a robust population, largely unaffected by the exact threshold used to separate HERGs from the LERGs. Similarly, the star-forming LERGs also show little change in their LFs, even for quite extreme changes in classification criteria, with changes being largely consistent within 2σ of the errors.

Here, we focus on modelling the evolution of the quiescent LERG LFs as this population is particularly important for radio-mode feedback considerations. To do this, we first fitted the $0.5 < z \leq 1.0$ quiescent LERG LFs as a broken power law (with all four free parameters), with the best-fitting model shown as a solid red line in the $0.5 < z \leq 1$ panel of Fig. 4.14 and the shaded region representing the 1σ uncertainties on the best fit[2]. The fitted parameters are also tabulated in Table 4.6.

Fitting all parameters for a broken power law independently at higher redshifts for the quiescent LERGs is not possible due to the lack of any strong break in the LFs. Instead, we fixed the two power law slopes at the values derived from the fit at $0.5 < z \leq 1$, and adopted an LDE model by fitting directly for $\rho_\star$ and $L_\star$ at each redshift[3]. The resulting fitted model and 1σ uncertainties are shown by the dashed lines and hatched regions in Fig. 4.14. Due to the lack of an obvious break in the LFs, there is a large degeneracy between the evolution of $\rho_\star$ and $L_\star$, as can be visualised in Fig. 4.15 that shows a 2D contour plot of the posterior distribution of the LDE model fit in $L_\star(z)$ and $\rho_\star(z)$ parameter space for different redshift bins. For the $0.5 < z \leq 1$ redshift bin, the contours shown are from the broken power law fit described above.

[2]Interestingly, in Fig. 4.14, it can be seen that all redshifts, the LF could potentially be fit better with a shallower faint-end slope, with then an upturn in the space densities of the quiescent LERGs at the faintest luminosities. At the lowest redshift bin ($0.5 < z \leq 1$), the upturn starts to appear at $L_{150\,\mathrm{MHz}} \approx 10^{24.5}\,\mathrm{W\,Hz^{-1}}$; this is typically below the limit probed by previous surveys. If this upturn is real, then it might suggest that a double Schechter function, which is also used to model the stellar mass functions, may be more appropriate than a broken power law for modelling the luminosity functions of the quiescent LERGs in particular. However, the present dataset does not have enough data points at the faint end beyond the upturn to provide constraints on such a model; deeper radio data, for example from subsequent data releases of the LoTSS Deep Fields, will allow us to better sample the faint end of the quiescent LERG population.

[3]We also examined the PLE and PDE models, but neither is able to match the data well: the PLE model cannot explain the strong evolution in space densities seen at low luminosities, whereas the PDE model fails to match the evolution at high luminosities.

Table 4.6 *Results of modelling the evolution of the quiescent LERGs. We fitted a broken power law to the* $0.5 < z \leq 1$ *LF. Using this broken power law fit (i.e. constant slopes), the LF at higher redshifts is modelled with a luminosity and density evolution. To reduce the degeneracy between parameters, the evolution of* $\rho_\star(z)$ *was fixed based on the evolution of the expected host galaxies of the quiescent LERGs (i.e.* $\rho_\star(z) \propto (1+z)^{-5.6}$*, see Fig. 4.16), with* $L_\star(z)$ *fitted as a free parameter. The best-fitting models are also shown in Fig. 4.14.*

z	$\log_{10}\rho^\star(z)$	$\log_{10} L_\star(z)$	β	γ	χ^2_ν
$0.5 < z \leq 1.0$	$-6.46^{+0.16}_{-0.30}$	$27.10^{+0.47}_{-0.21}$	$-2.98^{+1.05}_{-1.24}$	$-0.61^{+0.02}_{-0.02}$	2.88
$1.0 < z \leq 1.5$	$-7.08^{+0.16}_{-0.30}$	$27.54^{+0.03}_{-0.03}$	...	...	1.30
$1.5 < z \leq 2.0$	$-7.57^{+0.16}_{-0.30}$	$28.07^{+0.05}_{-0.05}$	...	...	0.76
$2.0 < z \leq 2.5$	$-7.98^{+0.16}_{-0.30}$	$28.06^{+0.10}_{-0.10}$	...	...	1.96

Under a simple model, we can consider the evolution of the quiescent LERG LFs to be dictated by the availability of their expected host galaxies (i.e. the massive quiescent galaxies; e.g. Best et al. 2014), which can help reduce the degeneracy of the two-parameter LDE fit. To determine the evolution of the space densities of the hosts of the quiescent LERGs, we follow the method outlined in Best et al. (2014). Best et al. used various determinations of the quiescent stellar mass functions from the literature (from Domínguez Sánchez et al., 2011; Baldry et al., 2012; Moustakas et al., 2013; Ilbert et al., 2013; Muzzin et al., 2013) and combined this with the prevalence of LERGs as a function of stellar mass to predict the total space density of the expected LERG hosts. We use these data compiled by Best et al. (2014), but scale these by a factor of 0.018/0.01 based on the normalisation of quiescent LERG fractions found in this study (where 0.018 is the mean normalisation at $10^{11}\,M_\odot$ across $0.3 < z \leq 1.5$, and 0.01 is the normalisation derived by Best et al. 2005a, at their 1.4 GHz radio luminosity limit). We also remove the small scale-factors that Best et al. (2014) introduced to line up the different datasets. To this, we add additional points based on a more recent measurement of the quiescent stellar mass function by McLeod et al. (2021).

Fig. 4.16 shows the results of this analysis, where we find that the space density of the available hosts remains relatively flat out to $z \sim 0.75$, beyond which there is a sharp decline. Given that we are interested in modelling the evolution of the LFs at $z > 0.75$, and that the flattening is observed at $z \lesssim 0.75$, we only model the evolution of the data points at $z > 0.75$ in Fig. 4.16. We find that this evolution can be reasonably approximated as $\rho \propto (1+z)^{-5.6}$, illustrated by the red line in Fig. 4.16. This predicted evolution of the expected host galaxies, and hence of $\rho_\star(z)$, can then be used to constrain

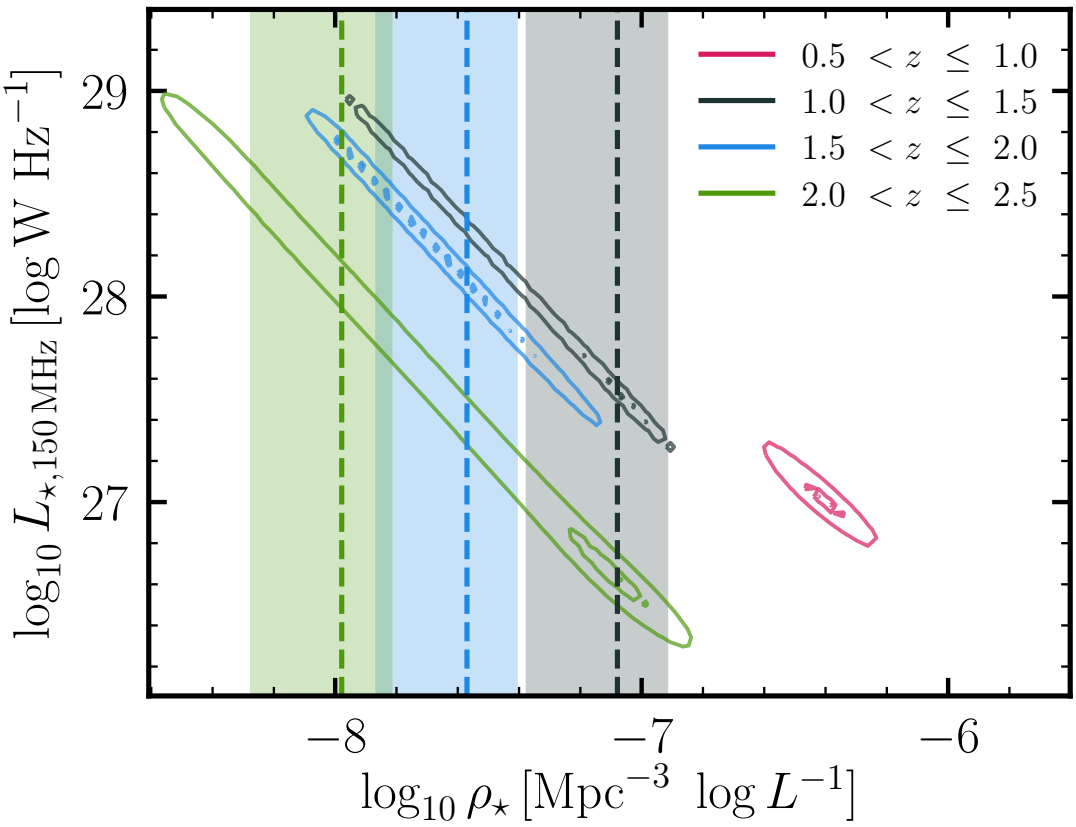

Figure 4.15 *2D contour plot showing the posterior distribution of $\rho_\star(z)$ and $L_\star(z)$ for the quiescent LERGs. For $0.5 < z \leq 1$, the contours are based on a broken power law fit, while at higher redshifts, the contours show the posterior distributions from a 2-parameter LDE fit to the LFs. The large degeneracy in these parameters can be better constrained when the evolution of $\rho_\star(z)$ is fixed based on the evolution of the expected host galaxies of quiescent LERGs as shown by the dashed lines (see Fig. 4.16); based on this estimate, the shaded vertical regions correspond to the 1σ uncertainties on the evolved $\rho_\star(z)$.*

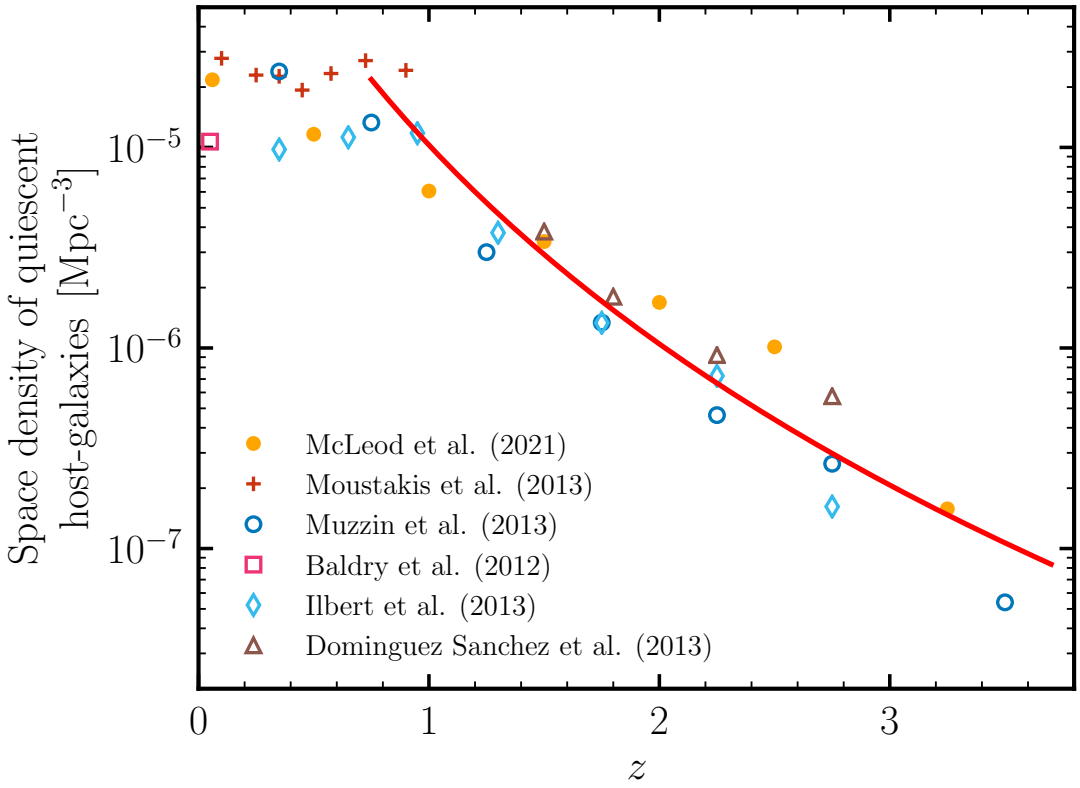

Figure 4.16 *The evolution of the space density of the expected host-galaxies of the quiescent LERGs from literature datasets at $z < 4$. The McLeod et al. (2021) data points were determined by combining the stellar mass function of quiescent galaxies with the prevalence of quiescent LERGs with stellar mass ($f_{\mathrm{AGN(M_\star)}} = 0.018(M_\star/10^{11}\,\mathrm{M_\odot})$; see text) and integrating in stellar mass to obtain the space density of the expected host galaxies. The other data points (from Domínguez Sánchez et al., 2011; Baldry et al., 2012; Moustakas et al., 2013; Ilbert et al., 2013; Muzzin et al., 2013) were complied by Best et al. (2014, taken from their fig. 6) and scaled appropriately to match the normalisation in $f_{\mathrm{AGN}}(\mathrm{M_\star})$ found in this study (see text).*

the LDE model, as shown by the dashed lines in the 2D contour plot in Fig. 4.15 for each redshift bin, with the shaded regions representing the uncertainties in $\rho_\star(z = 0.75)$ from the broken power law fit. We find that these shaded regions agree reasonably well with the peaks of probability distribution functions at $z < 2$, with only the highest redshift bin being poorly constrained. To test this explicitly, we fixed the evolution of $\rho_\star(z)$ at higher redshifts based on the value from the broken power law fit at $0.5 < z \leq 1$ and a $(1+z)^{-5.6}$ evolution, and fitted directly for $L_\star(z)$ only for each of the three $1.0 < z \leq 2.5$ bins. The resulting best-fitting models are shown as red solid lines in each panel; these are found to show a good match to the data, lying within the 1σ hatched region at each redshift.

These results agree well with studies of the total LERG population out to $z \lesssim 1$ that find a mild evolution (e.g. Best et al., 2014; Pracy et al., 2016); at these redshifts, the total LERG population is dominated by the quiescent hosts which show little evolution (see Fig. 4.16). At higher redshifts, while the total LERG population still shows weak evolution, as also found by Butler et al. (2019, out to $z < 1.3$), the quiescent LERG population declines strongly with increasing redshift in line with the expected host galaxies, while the characteristic luminosity of the sources increases; this is counter-balanced by an increasing star-forming LERG population at higher redshifts.

4.7 Implications for cosmic feedback from LERGs

Feedback from LERGs has been incorporated into semi-analytical models to solve the cooling flow problem in clusters and to recover the observed properties of massive galaxies in the local Universe. In this section, we consider the implications of our LERG luminosity functions on the evolution of feedback from radio-AGN across cosmic time.

4.7.1 Kinetic luminosities of LERGs

Although the LERG population have low radio-luminosities, the kinetic energy stored in the jets can provide a significant energetic output in the form of work done on the surrounding environment to inflate the lobes; this so-called mechanical (kinetic) luminosity from the jets can be orders of magnitude larger than the monochromatic radio luminosities for the LERGs. As the expanding jets plough through the surrounding material, they create cavities in the hot gas, which can be observed in the X-rays (Boehringer et al., 1993). As discussed in Chapter 1, the jet kinetic luminosity can

be estimated from the radio luminosity by considering the energy required to expand the lobes and inflate the X-ray cavities of pressure p with a volume V, $E_{\rm cav} = f_{\rm cav}\, pV$, where $f_{\rm cav} = 4$ is the commonly adopted value (e.g. Cavagnolo et al., 2010; Best et al., 2014; Pracy et al., 2016; Butler et al., 2019), with a good correlation found between the cavity powers and 1.4 GHz radio luminosities (e.g. Bîrzan et al., 2004; Cavagnolo et al., 2010). In this study, we have used the jet kinetic luminosity relation from Heckman & Best (2014), based on the analysis from Cavagnolo et al. (2010), given as

$$L_{\rm kin} = 7 \times 10^{36} f_{\rm cav} \left(L_{1.4\,\rm GHz}/10^{25}\,\rm W\,Hz^{-1}\right)^{0.68}\,\rm W \tag{4.11}$$

where, $L_{1.4\,\rm GHz}$ is the 1.4 GHz radio luminosity. This is found to be in good agreement with the mean $L_{\rm radio}$ - $L_{\rm kin}$ relation of Willott et al. (1999b), calculated from the synchrotron properties, also commonly used in literature (see Chapter 1 and the review by Heckman & Best, 2014).

It is important to note that the above relation has a large scatter that is dominated by the systematic effects and assumptions about the unknown physics of the radio sources. Specifically, a simple conversion between radio luminosity and kinetic luminosity is likely not appropriate as the radio luminosity varies over the lifetime of a radio source, depends on the energy density and magnetic field of the radio lobes and hence the environment into which the radio lobes are expanding, and also on the assumed spectral index. However, while a relation based on radio luminosity alone may not be appropriate for inferring jet kinetic luminosities for individual sources, we are interested in the heating output from the AGN at a population level; the use of this scaling relationship should therefore provide a reasonable mean value. However, an additional issue is that the above scaling relation has been determined from characterisation of sources at low-redshifts and it is possible that the radio luminosity of the sources of a given jet power (and hence the above relation) may evolve with redshift (e.g. due to cosmic evolution of magnetic field strengths, or increasing inverse-Compton losses); a well-constrained relation as a function of redshift is, however, still lacking. In the analysis that follows, we assume that the local relation is applicable over the full redshift range studied, but note that this might lead to systematic errors.

4.7.2 Evolution of the kinetic luminosity density

The evolving radio luminosity function of LERGs determined in this study can be combined with equation 4.11 and integrated to estimate the total kinetic luminosity

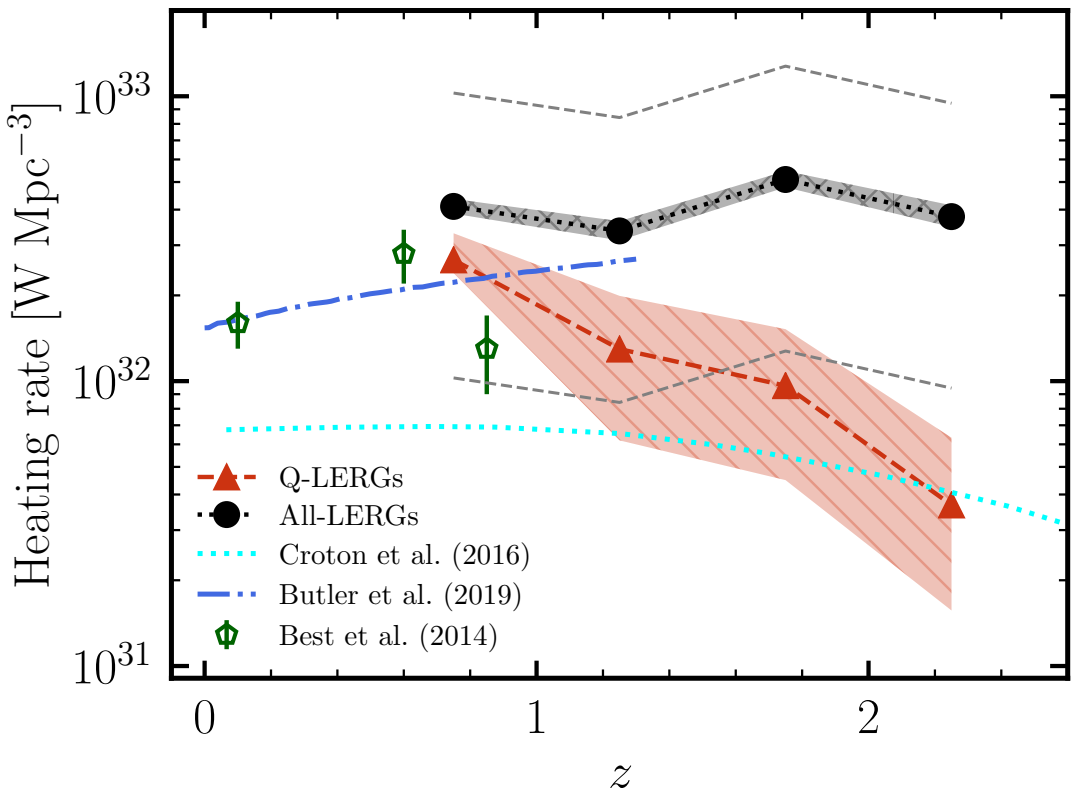

Figure 4.17 *The evolution of the kinetic luminosity density (also known as heating rate), $\Omega_{\rm kin}(z)$ for all LERGs (black points) and quiescent LERGs (red triangles) in the LoTSS Deep Fields. The shaded regions in each case represent the 1σ uncertainties on the kinetic luminosity density based on uncertainties in fitting the evolution of the LFs. The kinetic luminosity density was calculated using the Cavagnolo et al. (2010) relation with $f_{\rm cav} = 4$ in equation 4.11 using the best fit LDE models for both the total and quiescent LERG LFs (see text); these LFs were shifted to 1.4 GHz using a spectral index of $\alpha = -0.7$ and integrated in luminosity as given by equation 4.12 to obtain $\Omega_{\rm kin}(z)$. The grey lines show the systematic shift that would be obtained for the total LERG heating rate using the values $f_{\rm cav} = 1$ and $f_{\rm cav} = 10$ that account for the systematic uncertainties in the estimate of kinetic luminosities. Observational results for the LERG population from Best et al. (2014) and Butler et al. (2019), along with predictions from the SAGE model (Croton et al., 2016) are also shown.*

density, $\Omega_{\rm kin}(z)$ (also referred to as the kinetic heating rate). This is given, in units of W Mpc^{-3}, as a function of 1.4 GHz radio luminosity as

$$\Omega_{\rm kin}(z) = \int L_{\rm kin}(L_{1.4\,\rm GHz}) \times \rho\,(L_{1.4\,\rm GHz}, z)\ {\rm d}\log {\rm L}_{1.4\,\rm GHz}. \tag{4.12}$$

We calculated the kinetic luminosity density for the total LERG and the quiescent LERG populations, separately. For the total LERGs, we used the LDE model fits for each redshift bin based on the evolution of the Mauch & Sadler (2007) local LF (see Table 4.5). For the quiescent LERG population, we used the LDE model where the evolution of the characteristic space density is determined based on the evolution of the expected host galaxies, as detailed in Sect. 4.6 (see also Table 4.6). For both populations, the best-fitting LF models were shifted to 1.4 GHz using the standard spectral index $\alpha = -0.7$ to obtain $\rho(L_{1.4\,\rm GHz}, z)$. Then, for each redshift bin, we calculated $L_{\rm kin}(L_{1.4\,\rm GHz})$ using equation 4.11 and then integrated the LF down to $10^{-4}\ L^{\star}(z)$ to obtain $\Omega_{\rm kin}(z)$.

The resulting evolution of the kinetic luminosity density for each of the total LERG population (black points) and the quiescent LERG population (red triangles) is shown in Fig. 4.17. In each case, the shaded region corresponds to the uncertainties in the heating rate based on uncertainties in fitting the evolution of the LFs. The dashed grey lines show the systematic shift that would be produced for the total LERG heating rate based on the extreme values of $f_{\rm cav} = 1$ and $f_{\rm cav} = 10$; this illustrates the effect of uncertainties in the calibration of the radio-luminosity to kinetic luminosity relationship on the absolute values obtained. The kinetic heating rate for the total LERG population remains roughly constant between $0.5 < z \leq 2.5$ with $\Omega_{\rm kin} \approx 4 \times 10^{32}$ W Mpc^{-3} at $z = 0.75$. In contrast, for the quiescent-LERG population, the heating rate decreases steadily with increasing redshift, from $\Omega_{\rm kin} \approx 2.7 \times 10^{32}$ W Mpc^{-3} at $z = 0.75$ to nearly an order of magnitude lower by $z \sim 2.5$.

Also in Fig. 4.17, we show the estimates of heating rate obtained for the jet-mode AGN population by Best et al. (2014) for $z < 1$ (green symbols), also derived using the Cavagnolo et al. (2010) kinetic luminosity relation with $f_{\rm cav} = 4$. The dash-dotted line illustrates the kinetic luminosity density for the LERG population as determined by Butler et al. (2019) out to $z \sim 1.3$. The results from Butler et al. (2019) show a steady increase in the heating rate with increasing redshift and agree well with our estimates for the total LERG population, even when extrapolating their results to higher redshifts. The results from Best et al. (2014) show an increase out to $z \sim 0.6$, agreeing well with the Butler et al. (2019) results, but at higher redshifts, show a decrease in the heating rates, resulting in better agreement with our quiescent LERG population, within

errors, than the total LERG population. One potential reason for the better agreement with the quiescent-LERGs could be due to the method by which jet-mode AGN were selected by Best et al. (2014); these were identified using emission line ratio diagnostics to select sources with relatively low emission line fluxes or equivalent widths from [OII] or [OIII] lines. Jet-mode AGN hosts with considerable star-forming activity will result in higher equivalent widths of these spectral lines and therefore the spectroscopic classification used by Best et al. (2014) may result in a sample similar to the quiescent LERG population identified in this study. We note that even if the Best et al. (2014) selection is similar to our quiescent LERGs, the good agreement between the Best et al. (2014) and Butler et al. (2019) results is expected at $z < 0.6$, where the total LERG population is dominated by quiescent LERGs.

We also compare our results with predictions from the Semi-Analytic Galaxy Evolution (SAGE) model (Croton et al., 2016), providing an enhancement to the previous model by Croton et al. (2006). Croton et al. (2006) considered the prescription of SMBH growth and AGN feedback in two classes: "quasar mode" and "radio mode" in their terminology; we consider predictions of feedback from their "radio mode" class of AGN which is most closely linked with the LERG population in his study. To do this, we used the predictions of the black hole accretion rate density ($\dot{m}_{\rm BH,R}$) over cosmic time for the "radio-mode" population from Croton et al. (2006, see their fig. 3). The $\dot{m}_{\rm BH,R}(z)$ were then translated into a luminosity of the black hole assuming $L_{\rm BH} = \eta \dot{m}_{\rm BH,R} c^2$, where $\eta = 0.1$ is the standard black-hole accretion efficiency and c is the speed of light. This luminosity generated from the accretion process is considered equivalent to the kinetic luminosity density output from the LERGs derived in this study. As part of an update to this model, Croton et al. (2016) introduced a "radio mode efficiency" parameter, $\kappa_{\rm R}$ within SAGE to provide a more realistic treatment of AGN feedback cycle by attempting to couple the heating provided by the AGN with the gas cooling. As a result, Croton et al. (2016) scale their $\dot{m}_{\rm BH,R}(z)$ and hence the black hole luminosities by $\kappa_{\rm R} = 0.08$. The resulting prediction for heating rate from the SAGE model is shown by the dotted cyan line in Fig. 4.17. The total LERGs show a similar shape to the SAGE model prediction, but offset to higher values by nearly an order of magnitude. This could be due to calibration errors in the scaling relation, or due to some fraction of the heating from LERGs (in particular from very extended sources) being deposited on scales larger than that useful to offset cooling losses. Moreover, the radio-mode feedback in the SAGE model is scaled to a level required to provide AGN heating that can offset radiative cooling, however, at higher redshifts in particular, the LERG activity occurs predominantly in star-forming LERGs (see Sect. 4.6) which may not be part of the feedback cycle. The kinetic luminosity density of the quiescent LERG population is

also systematically higher than the SAGE model at all but the highest redshift bin; these results therefore suggest that the energy output by the quiescent LERG population alone is sufficient to offset the radiative cooling losses in these systems at least out to $z \sim 2.5$, thereby providing evidence for a self-regulating AGN feedback cycle in these systems.

However, as discussed above, the uncertainties in the scaling relation between radio and kinetic luminosities results in a large scatter in this relation, moreover the validity of this relation at earlier cosmic times is not well characterised; in the absence of a better constrained physical relation, especially at high redshifts, a detailed comparison between different estimates of the kinetic luminosity density requires caution.

In summary, in this Chapter, I have used a sample of 10 429 LERGs from the first data release of the LoTSS Deep Fields, covering $\sim 25\,\mathrm{deg}^2$, to present the first measurement of the evolution of the radio luminosity function (LF) of LERGs out to $z \sim 2.5$; this shows relatively mild evolution. We split the LERG population into those hosted by quiescent galaxies and those hosted by star-forming galaxies. The prevalence of LERGs in quiescent galaxies shows a steep dependence on stellar-mass out to $z \sim 1.5$, consistent with local Universe measurements and of accretion occurring from cooling of hot gas haloes. The quiescent-LERGs dominate the LFs at $z < 1$, showing a strong decline in space densities with redshift, tracing that of the available host galaxies, while there is an increase in the characteristic luminosity. The energy dissipated into the surrounding gas by these quiescent-LERGs is sufficient to balance cooling losses at least out to $z \sim 2.5$. The star-forming LERG LF increases with redshift, such that it begins to dominate the LFs at all radio-luminosities at $z \gtrsim 1.5$. The prevalence of LERGs in star-forming galaxies shows a much weaker stellar-mass dependence, and increases with redshift, suggesting a different fuelling mechanism associated with the cold gas.

4.A Effects of varying the source classification criteria

In this Appendix, we explore how deviations from our adopted criteria for source classification affect our main results on the evolution of the LERG LFs presented in Sects. 4.4 – 4.7.

As discussed in Sects. 4.3.4 – 4.3.5, the local and evolving LFs for the total radio-AGN population derived in LoTSS-Deep agree well with literature results (as do those of the SFGs by Cochrane et al. in prep.), giving us confidence in our separation of star-forming galaxies from radio-excess AGN across redshift. Furthermore, we have investigated increasing and decreasing the radio-excess criteria by 0.1 dex and find that the LERG LFs are largely unchanged, with typical changes in space densities of < 0.15 dex. Therefore, in this section, we focus on investigating the modification of the selection of the so-called 'optical-AGN', which will result in a change between sources classified as HERGs and LERGs. We note that it can be challenging to identify (optical) AGN of moderate to low luminosities using photometric data; this may be particularly important in star-forming galaxies as both AGN activity and star-formation contribute to the MIR emission, making it difficult to disentangle contributions associated with the different emission processes. If such sources also show radio-loud emission and the optical-AGN component is missed, then based on our source classification criteria, these would be mis-classified as LERGs (and likely as star-forming LERGs) instead of HERGs. Although this issue is inevitable in the absence of emission line diagnostics that allow a more robust selection of optical AGN, we consider the impact of this potential contamination on our results.

Fig. 4.18 shows the cosmic evolution of the quiescent LERGs, star-forming LERGs, and HERGs between $0.5 < z \leq 2.5$ based on our adopted source classification criteria from Best et al. (in prep.), plotted with triangle (red), square (blue), and cross (black) symbols, respectively. The shaded regions correspond to the 1σ uncertainties on the LFs. Also shown in Fig. 4.18 are the LFs for these three classes of AGN obtained using three variations of the 'optical-AGN' selection criteria from Best et al. (2021, in prep.): SC-1, SC-2, and SC-3. As outlined in Sect. 4.2.2, based on the output from each of AGNFitter and cigale, Best et al. (in prep.) defined a value $f_{\rm AGN}$ for each code, as the fraction of the mid-infrared luminosity associated with AGN emission compared to the host galaxy. This diagnostic (in particular the $16^{\rm th}$ percentile of $f_{\rm AGN}$) forms a key part in the selection of 'optical-AGN', and hence the separation of 'radio-excess AGN' into

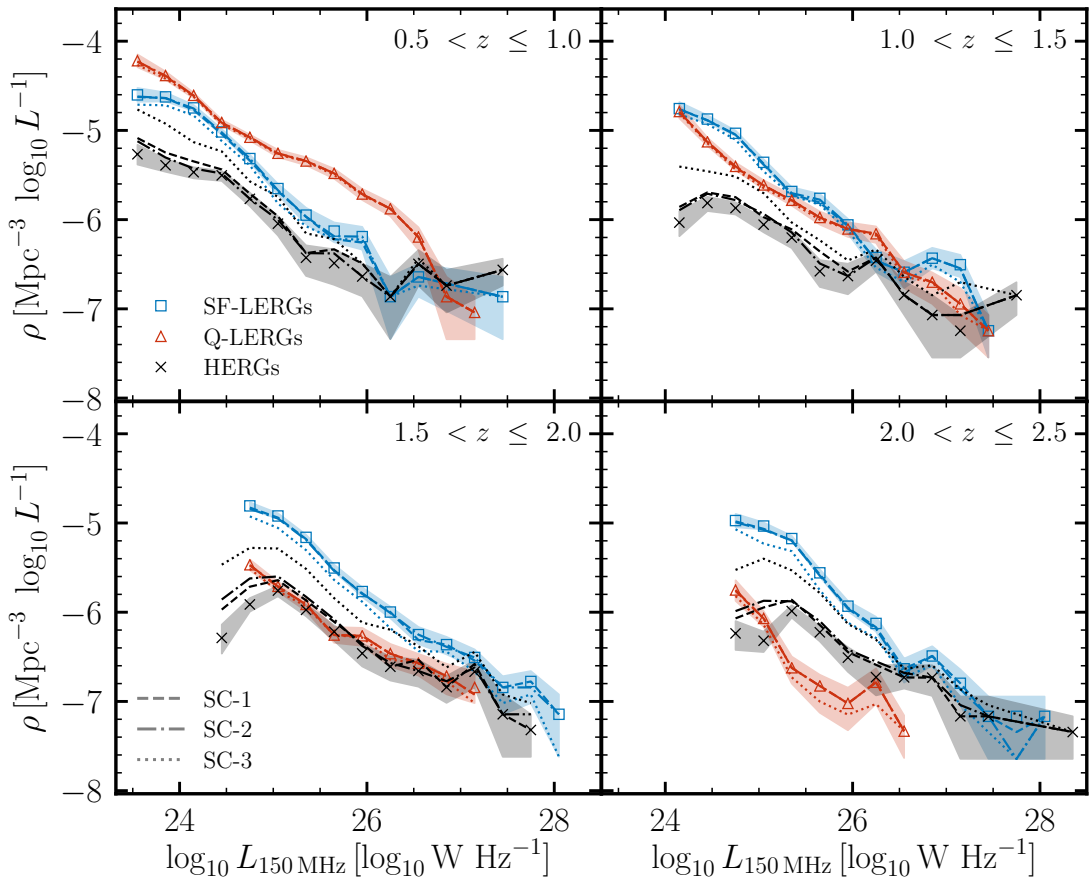

Figure 4.18 *Evolution of the luminosity functions of star-forming and quiescent LERGs, and the HERGs in the LoTSS-Deep fields using different source classification criteria. The separation of quiescent and star-forming LERGs is performed based on the sSFR criteria defined in Sect. 4.5.1. The markers and associated shaded regions correspond to the LFs and* 1σ *uncertainties resulting from the adopted source classification criteria from Best et al. (in prep.), whereas the lines show the LFs obtained under three variations of the source classification criteria (see text). In general, we find that the adopted quiescent and star-forming LERGs, which are the focus of this chapter, show little change in space densities with changes in source classification selection, suggesting that the results in the main paper are robust against uncertainties in the source classification scheme adopted.*

LERGs and HERGs. We arbitrarily reduced the thresholds for these parameters from both AGNFitter and CIGALE by a factor of two (i.e. leading to an increase in the number of HERGs) and reconstructed the LFs, as shown by the SC-1 (dotted) lines.

We also implemented an additional criteria to that adopted by Best et al. (2021, in prep.) in the selection of 'optical-AGN', using the best-fitting $f_{\rm AGN}$ values from AGNFitter and CIGALE in addition to the 16$^{\rm th}$ percentile values. We used $f_{\rm AGN}$ as we wanted to select objects with high $f_{\rm AGN}$ values, but with large uncertainties (e.g. due to low S/N), that may be missed by the Best et al. (in prep.) criteria; this allowed us to test a more complete but possibly contaminated sample, compared to the clean but possibly incomplete 'optical-AGN' sample of Best et al. (in prep.). In order to find an appropriate threshold in $f_{\rm AGN}$, we considered first the distribution of sources in the $f_{\rm AGN}$(AGNFitter) versus $f_{\rm AGN}$(Cigale) 2D parameter space. Fig. 4.19 shows a scatter plot in this 2D space in two broad redshift bins ($0 < z \leq 1.5$ and $1.5 < z \leq 2.5$), with the red, blue, and black points corresponding to the quiescent LERGs, star-forming LERGs, and HERGs, respectively, as classified based on the Best et al. (in prep.) criteria. It is reassuring to see that although the best-fitting $f_{\rm AGN}$ parameters were not used in the selection of the 'optical-AGN', the HERGs occupy a well-defined region of this 2D plane, corresponding to high $f_{\rm AGN}$ values from both the SED fitting codes, albeit with some overlap with the star-forming LERG population. Moreover, the quiescent and star-forming LERGs also show very similar distributions in best-fitting $f_{\rm AGN}$ values (from both AGNFitter and CIGALE) – and distinct from the HERGs – suggesting that these are both drawn from the same parent population and that the star-forming LERG population is not significantly affected by contamination from HERGs. Based on the location of the HERGs, and that of X-ray detected AGN (indicative of radiative-mode AGN) from the X-Boötes survey in the 2D scatter plot in Fig. 4.19, we defined an extra criteria of 'optical-AGN' as sources with $f_{\rm AGN}$(Cigale) > 0.15 and $f_{\rm AGN}$(AGNFitter) > 0.25. We note that this selection criteria is applied in addition to the Best et al. criteria, such that there will always be an increase in the number of sources classified as HERGs; in the case adopted here, this leads to 258 more sources being classified as HERGs. The LFs resulting from this criteria (SC-2) are shown as dashed lines in Fig. 4.18.

Finally, we considered adopting a more extreme approach such that sources with either $f_{\rm AGN}$(Cigale) > 0.15 or $f_{\rm AGN}$(AGNFitter) > 0.25 are classed as 'optical-AGN'. As seen from Fig. 4.19, while this selection may result in a higher completeness of 'optical-AGN', it will suffer from even more contamination; this is particularly the case at higher redshifts. This is due to a significant population of sources (mostly star-forming LERGs based on the Best et al. classification) with $f_{\rm AGN}$(AGNFitter) > 0.25 but

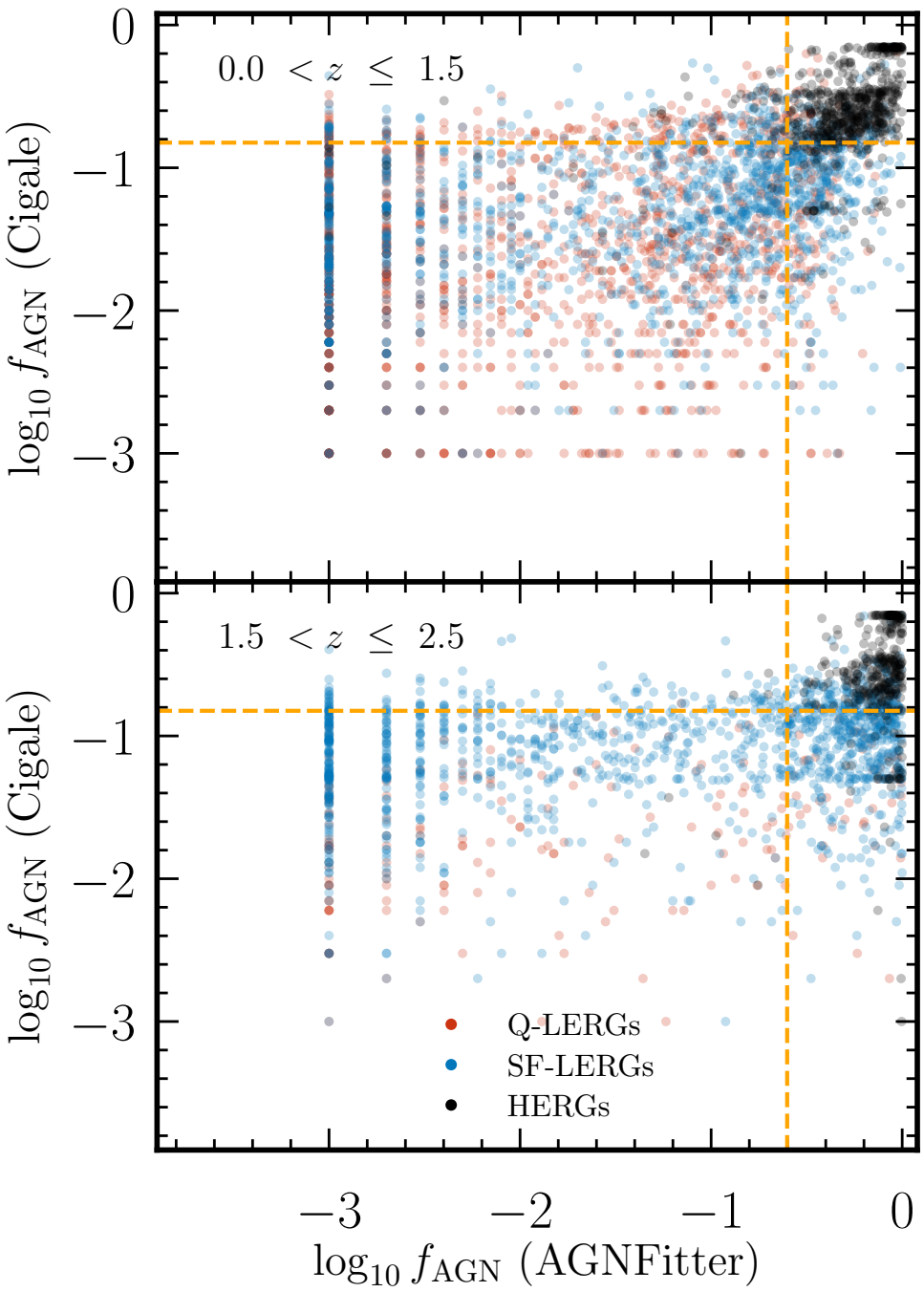

Figure 4.19 *2D scatter plot showing the location of the quiescent LERGs, star-forming LERGs, and HERGs (in ELAIS-N1), in AGN fraction* (CIGALE) *- AGN fraction* (AGNFITTER) *parameter space in two broad redshift bins* ($0 < z \leq 1.5$, $1.5 < z \leq 2.5$)*. The dashed lines along the x and y-axes show the* $f_{\rm AGN}$(AGNFITTER) $= 0.25$ *and* $f_{\rm AGN}$(CIGALE) $= 0.15$ *thresholds, respectively, used to selection additional 'optical-AGN' for testing the source classification criteria.*

significantly lower (up to a factor of 10 lower) values of $f_{\rm AGN}$(CIGALE), which is a result of the significantly higher uncertainties on the AGNFitter values. Nevertheless, the SC-3 criteria allows us to examine the extent to which even an extreme classification could potentially affect the LERG LFs (dash-dotted lines in Fig. 4.18).

As evident from Fig. 4.18, given both the small number of HERGs in the LoTSS Deep Fields (see also Fig. 4.1) and the small numbers of sources with a change in their classification based on SC-1 and SC-2 criteria, there is a relatively small effect on the HERG LFs across all redshifts, but much less so for the LERGs. The SC-1 and SC-2 lines agree with the originally determined Best et al. HERG LFs within $1-2\sigma$ of our best estimate LFs at all but the faintest few luminosity bins; at the faint end, the change in the LFs is comparable to the field-to-field variation found in the LFs. For the quiescent and star-forming LERGs, both SC-1 and SC-2 variations result in very similar LFs across redshift. The quiescent LERG population in particular remains robust against changes to the source classification method adopted within the uncertainties across all of the redshift bins; this also holds even when they become the sub-dominant population of all LERGs at $z > 1.5$. The star-forming LERG LFs at $z \leq 1$ based on SC-1 and SC-2 are lower by $\sim 0.1-0.2$ dex, however this has little to no effect on the total LERG LFs given that the quiescent LERGs dominate the space densities over most of the radio luminosities studied here; at higher redshifts, the differences with the SC-1 and SC-2 based LFs are consistent within $1-2\sigma$ of the uncertainties.

For SC-3, the space density of the HERGs (2734 sources) now increases substantially across all luminosities, and especially at the faint end at $z > 1.5$, as might be expected. For the LERG population however, which is the focus of this chapter, the reconstructed LFs at $z \leq 1.5$ show good agreement with our best estimate of the LERG space densities; at higher redshifts, the space densities decrease, in particular at faint luminosities, reaching ~ 0.2 dex lower in the faintest bins.

We therefore conclude that the LERG LFs presented in the main sections of the Chapter based on the Best et al. (in prep.) source classification criteria – and subsequent results on the AGN mass fractions – are robust to uncertainties in the source classification criteria used.

Chapter 5

Conclusions and Future Prospects

In Chapters 2 and 3, I have presented the value-added catalogue of multi-wavelength counterparts to the radio sources detected in the first LoTSS Deep Fields Data Release by Tasse et al. (2021) and Sabater et al. (2021), covering the ELAIS-N1, Lockman Hole, and Boötes fields and forming the deepest radio continuum survey to date at low frequencies.

To achieve this, in Chapter 2, I first built new multi-wavelength catalogues in both ELAIS-N1 and Lockman Hole, consisting of forced, matched aperture photometry across 20 and 16 bands, respectively, from UV to mid-IR. These catalogues were built using deep χ^2 detection images, using information from the optical and IR bands, to maximise the catalogue completeness and to generate clean, robust photometry and colours; this provides a significant improvement to catalogues existing in literature for photometric redshifts and SED fitting.

Chapter 3 describes the process of how the counterparts to the radio sources were identified using the multi-wavelength catalogues generated in Chapter 2. This was performed using a combination of the statistical Likelihood Ratio method and visual classification schemes. We use the LR method that incorporates both magnitude and colour information, as described in Nisbet (2018) and Williams et al. (2019), to maximise the identification rate and increase the robustness of the cross-matching. The deep ancillary data available in these fields allows us to achieve an identification rate of up to 97% using the LR method alone. The LR method, however, is not suitable for large or complex radio sources; such sources require visual classification instead, which is mainly performed using the LOFAR Galaxy Zoo framework developed for LoTSS-DR1 (Williams et al., 2019). To determine sources that can be identified using the LR

method and those that require visual classification, we adapted and further developed the decision tree used in LoTSS-DR1. The high LR identification rates allowed us to require any source without a LR identification to undergo visual inspection.

The cross-matching effort leads to multi-wavelength identifications for 97.6%, 97.6%, and 96.9% of sources in ELAIS-N1, Lockman Hole, and Boötes, respectively. The colour properties of host galaxies show that (at a fixed faint magnitude) the reddest of galaxies are more than an order of magnitude more likely to host a LOFAR source than the bluest of galaxies. This is also visualised by the magnitude distributions of the host galaxies in different filters, which show that we are able to identify most of the LOFAR sources in the mid-IR. In contrast, deeper optical data are required to achieve higher identification rates and probe the optically faint counterparts to beyond the peak of host-galaxy magnitude distributions in these optical filters. In total, the value-added radio-optical cross-matched catalogues presented contain 81 951 radio sources, with counterparts identified and matched aperture optical to infrared properties presented for 79 820 sources (> 97%), covering ~ 26 deg^2 in total, across multiple sight-lines.

The scientific potential of the catalogues presented is further increased by the availability of photometric redshifts. Rest-frame colours and photometric redshifts for both the multi-wavelength catalogues and the radio-optical cross-matched catalogues in the three fields are presented in Duncan et al. (2021). This enables one to determine physical properties of host galaxies (such as luminosities, stellar masses, star-formation rates, etc.), which are used to perform source classification as presented by Best et al. (in prep.).

In Chapter 4, I have used the the datasets generated in the previous chapters to study the evolution of the radio-AGN population out to $z \sim 2.5$. The depth of the radio and multi-wavelength dataset in combination with the wide area coverage makes this an ideal survey to constrain the evolution of the low-luminosity radio-AGN out to high redshifts.

The total radio-AGN luminosity functions were constructed out to $z \leq 2.5$ using a sample of 11 466 radio-AGN; this shows good agreement across all redshifts with the literature LFs determined from GHz surveys. Using the outputs from the SED fitting process by Best et al. (in prep.), the radio-AGN were separated into LERGs and HERGs via the identification of accretion disc and torus features in galaxy SEDs; the presence of such signatures is indicative of a HERG. This results in a sample of 10 429 LERGs and 1037 HERGs across the deep fields (within $z \leq 2.5$). Using this sample, we have then presented the first robust measurement of the cosmic evolution of the LERG luminosity

functions out to $z \sim 2.5$. We find relatively mild evolution within the redshift bins ($0.5 < z \leq 2.5$) examined in this study; this evolution is found to be driven by a very different evolution of those LERGs hosted by quiescent galaxies and those hosted by star-forming galaxies. We therefore investigated these two populations separately.

The main results of our study for the quiescent LERGs are as follows:

- The quiescent LERG LFs dominate the space densities of LERGs across essentially all luminosities at $z \lesssim 0.75$.
- The incidence of LERG activity in quiescent galaxies shows a strong stellar mass dependence, which can be well-described at all redshifts as $f_{\rm LERG,Q} = c(M_\star/10^{11}\,{\rm M}_\odot)^\beta$, with a power-law slope of $\beta \approx 2.5$ and a normalisation of $c \approx 0.018$ (for a luminosity limit of $L_{150\,\rm MHz} \geq 10^{24}\,{\rm W\,Hz^{-1}}$). This relation for the quiescent LERGs also shows no evolution with redshift (between $0.3 < z \leq 1.5$).
- The steep dependence of LERG activity with stellar mass, and the lack of a redshift evolution suggest that the LERGs in these quiescent galaxies are fuelled by cooling of hot gas from haloes, consistent with local Universe observations, at least over the past ~ 9 Gyrs.
- The quiescent LERGs show a strong negative evolution in their LFs beyond $z \sim 0.75$, which is consistent with the characteristic space density evolving in accordance with the availability of the expected host galaxies, while there is also an increase in the characteristic luminosity with increasing redshift.

The main results of our study for the star-forming LERGs are as follows:

- The incidence of LERG activity in star-forming galaxies shows a much weaker stellar mass dependence ($\beta \approx 1.5$ consistently across redshift) than that found for the quiescent LERGs in this study, and for LERGs in the local Universe, and instead is comparable to that found for the HERGs locally.
- The incidence of LERGs in star-forming galaxies (at fixed stellar mass; i.e. the normalisation c) increases by nearly a factor of four between $0.3 < z \leq 1.5$, reaching a higher prevalence than the quiescent LERGs at the highest redshifts.
- The star-forming LERGs are a minority of the overall LERG population at $z < 1$, but space densities of the star-forming LERGs increase with increasing redshift, such that they become the dominant population at $z \gtrsim 1.5$ over all radio luminosities.

The flatter stellar mass dependence and the increase in the prevalence of LERGs in star-forming galaxies at earlier cosmic times, when gas fractions were higher, suggests that a different fuelling mechanism, likely associated with the cold gas, is responsible for triggering LERG activity within these star-forming galaxies.

We then considered the implications of the observed evolution of the LERG LFs on cosmic radio-AGN feedback. A scaling relation between radio and jet-kinetic luminosity, along with the best-fitting models for the LERG LFs, were combined to compute the cosmic evolution of the kinetic luminosity density. We find that the heating rate for the total LERG population remains roughly constant across $0.5 < z \leq 2.5$, whereas this decreases by an order of magnitude for the quiescent LERG subset. A comparison with the recent SAGE model finds that the quiescent LERGs alone may deposit sufficient energy into their surrounding environment to balance the radiative cooling losses, at least out to $z \sim 2.5$, providing evidence for an AGN feedback cycle in these systems.

5.1 Ongoing and future work

We are continuing to acquire more radio observations for all three LoTSS Deep Fields, with the current allocated time bringing the ELAIS-N1 observations up to 500 hours (expected to reach rms noise levels of $\sim 10\,\mu$Jy), and up to at least 300 hours in the other two fields. Moreover, since the first data release of the LoTSS Deep Fields, radio observations for a fourth deep field, the North Ecliptic Pole (NEP) field have been taken and processed, with ~ 300 more hours allocated in upcoming LOFAR observing cycles.

Future data releases will present these deeper and wider radio data that will provide significantly better constraints on both the low luminosity and high luminosity ends of the luminosity functions. In addition, the current radio imaging dataset is calibrated based on the Dutch baselines only (~ 100 kilometre scales), offering $6''$ resolution; the addition of international LOFAR baselines, spanning many hundreds of kilometres, will provide $0.3''$ resolution radio imaging, rivalling that of the optical and NIR datasets, and assist greatly in source identification (e.g. due to deblending) and classification (due to resolved morphologies), and will also enable resolved studies of samples across multiple wavelengths.

In addition to advances in the radio datasets, ELAIS-N1 will continue to obtain deeper and higher resolution optical imaging from future releases of the HSC-SSP survey. The NEP field, in particular, will be one of the prime deep fields in the coming decade;

it will be one of the *Euclid* deep fields (*Euclid* Deep Field North) with incredibly deep NIR observations at 0.3″ resolution reaching $H \sim 26$ mag over 10 deg^2 that will allow studies of mass-complete samples out to $z = 3$. Furthermore, the NEP field also possesses deep *Spitzer* observations over 10 deg^2 that reach ~ 0.8 mag deeper than SERVS, and will soon fully obtain deep optical imaging from the HSC telescope as part of the Hawaii-Two-0 project. The location of this field in the continuous viewing zone (CVZ) also makes it an ideal target for space-based observations from the *James Webb Space Telescope (JWST)* and *e-ROSITA*. Finally, future multi-wavelength catalogues will also benefit from PSF-matched photometry which will allow the calculation of model-fit total magnitudes.

Further to the improvements in the radio and multi-wavelength datasets, there are other ongoing projects that represent future work; in the sections below, I highlight potential future projects based on these.

5.1.1 Radio sources without an identification

The identification of the host-galaxies of the radio-detected sources in the LoTSS Deep Fields (see Chapter 3) resulted in counterparts identified for over 97% of the radio-detected sources, with no multi-wavelength counterpart found for $\approx 3\%$ of the radio sources. We investigated the potential nature of the host galaxies of these radio sources without an identification (hereafter, 'No-ID' sources), focusing on ELAIS-N1, by stacking the optical to mid-IR images. To do this, we first selected LOFAR sources without an identification that have a secure radio position (571 sources). Most ($> 90\%$) of these radio sources have radio flux densities of $S_{150\mathrm{MHz}} < 1$mJy; at these flux densities, the radio population is a mix of different source types (see Wilman et al. 2008 and discussion in Sect. 3.6). A median stack based on the radio positions was then performed, with photometry extracted in the optical-NIR (3″ aperture) and Spitzer (4″ aperture) filters, corrected to total magnitudes to match the catalogues. For the 24 μm MIPS band, we extracted photometry from the 10″ aperture and perform aperture corrections based on the values computed by Engelbracht et al. (2007).

The resulting stacked SED for all the unidentified sources (filled black circles) is shown in Fig. 5.1. Also shown in blue, orange and green solid lines are templates of galaxies dominated by a recent burst of star-formation occurring 100 Myrs ago, at $z = 2, 3$ and 4, respectively, computed using the stellar population models of Bruzual & Charlot (2003), the Chabrier (2003) initial mass function (IMF) and the Calzetti et al. (2000) attenuation curve. The star-formation rate (SFR) normalisation of these templates is

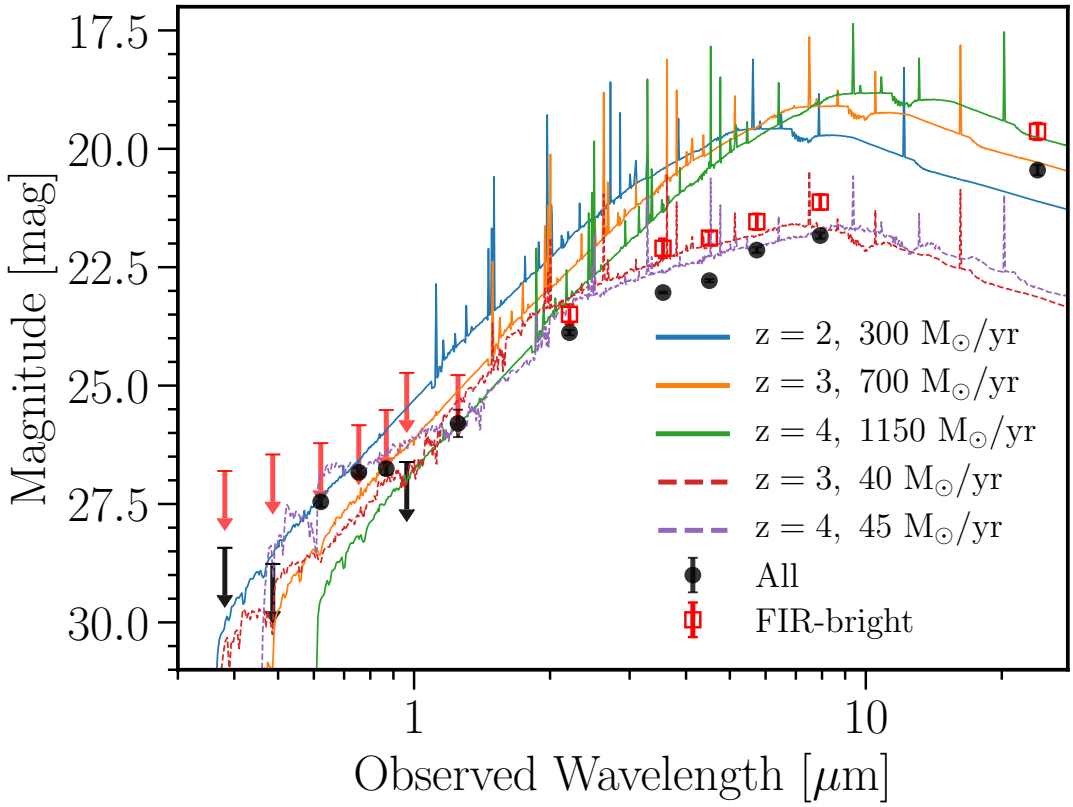

Figure 5.1 *Optical to mid-IR stacked SED of the unidentified radio sources (see Chapter 3 for details) with secure radio positions in ELAIS-N1 (571 sources; black points). The stack for the subset of the 'FIR-bright' sources (50 sources) is shown as red squares. Overlaid on top are typical star-forming galaxy templates, computed using models of Bruzual & Charlot (2003) at $z = 2, 3$ and 4 (blue, orange and green solid lines, respectively), scaled to the median SFR detectable for these sources at the depth of the LOFAR data in ELAIS-N1. The dashed lines show two-component templates of galaxies with an old stellar population undergoing a recent burst of star-formation (see text) at $z = 3$ and 4 (red and purple, respectively).*

fixed to the SFR required to produce the median radio flux density observed of these sources, assuming all of the radio emission is associated with star-formation. This is calculated using the relation between the 150 MHz luminosity (L_{150}) and SFR derived by Gürkan et al. (2018), assuming a spectral index $\alpha = 0.7$, and requires a SFR of $300\,M_\odot\,yr^{-1}$ at $z = 2$, $700\,M_\odot\,yr^{-1}$ at $z = 3$ and $1150\,M_\odot\,yr^{-1}$ at $z = 4$. As evident from the stacked SED, if the radio emission is entirely due to star formation, we do not observe enough emission, particularly at near- and mid-infrared wavelengths, to match the star-forming galaxy templates and the difference in flux by over an order of magnitude cannot simply be explained by extinction (especially in the IRAC bands).

To further investigate the potential role of the star-forming galaxies within this radio source population, we examine the FIR emission from these sources. 50 of these 571 No-ID sources (~7%) have significant FIR emission (hereafter 'FIR-bright' sources); these are defined as sources with a 250μm flux density ($F_{250\,\mu m}$) > 15 mJy. When stacked separately (red squares), these FIR-bright sources are found to be around 1-1.5 mag brighter in the mid-IR bands, however still nearly an order of magnitude fainter in flux than the star-forming templates at $z = 2 - 4$. Thus, even though the FIR measurements suggest that some star-formation may be on-going in these sources, we conclude that the radio emission in the majority of these sources is not dominated by star-formation.

Instead, to provide an illustration of the SED that a typical radiative-mode (or high-excitation radio galaxy; HERG) AGN might have, we consider a two component model, including both an old stellar population and a period of recent star-formation (since high-redshift radiative-mode AGN are found to lie close to the star-forming main sequence; e.g. Mainieri et al. 2011; Bonzini et al. 2015; Suh et al. 2019). Specifically, the red and purple dashed lines on Fig. 5.1 show a 10^{10} $M_\odot$ old stellar population (formed at $z = 12$ with an exponentially declining star-formation rate with a characteristic time of 150 Myr) with a burst of star-formation within the past 50 Myr with a SFR of 40 and 45 $M_\odot\,yr^{-1}$ (consistent with being on the star-forming main sequence; Speagle et al. 2014b) at $z = 3$ and 4, respectively. These illustrative SEDs broadly trace the stacked data points, suggesting that a significant fraction of the unidentified radio source population is likely dominated by high-redshift obscured/radiative-mode AGN, which are expected to contribute significantly to source counts at $S_{150\,MHz} < 1$ mJy (Wilman et al., 2008). We also observe a significant detection at 24 μm for the 'All' stack, indicative of hot dust emission; this emission could potentially arise from the hot torus surrounding an AGN (i.e. a HERG-like AGN) that is expected to peak at rest-frame wavelength of $\sim$ 10 μm (e.g. Silva et al., 2004). The median 150 MHz luminosity for the 'All' sources, assuming a spectra index $\alpha = 0.7$, is $L_{150\,MHz} \approx 10^{25}\,W\,Hz^{-1}$ if at $z = 3$, and

$L_{150\,\mathrm{MHz}} \approx 2 \times 10^{25}$ W H^{-1} if at $z = 4$. These are significantly fainter than the radio galaxies known at these redshifts (e.g. Jarvis et al. 2009; Saxena et al. 2018).

5.1.1.1 Sub-millimetre follow-up of candidate z > 3 galaxies

Although the majority of the optically-dark radio-detected sources are likely to be comprised of AGN, a small proportion of these sources, in particular those that are bright in the FIR wavelengths, are potentially high-redshift SFGs or AGN-starburst composite systems. Characterising the nature of these can provide interesting constraints on cosmic star-formation, early galaxy formation and early AGN activity in these extreme systems.

It is believed that rare, dusty star-forming galaxies (DSFGs) may dominate the star-formation rate density (SFRD) at the earliest epochs (see the review by Casey et al. 2014; Wang et al. 2019). A significant population of galaxies at $z > 3$ has been detected by their bright observed-frame sub-millimetre emission ('sub-millimetre galaxies'; SMGs, Blain et al. 2002). The most extreme display sub-millimetre fluxes corresponding to SFRs of thousands of solar masses per year. Sustained SFRs of this magnitude have the potential to form massive galaxies (with stellar masses of $\sim 10^{11}\,M_{\odot}$) on sub-Gyr timescales. This has led to speculation that these galaxies may be the progenitors of the rare, massive elliptical galaxies seen at $z \sim 3$ (Oteo et al., 2016).

The discovery of ultra-red DSFGs that formed within the first 2 billion years (i.e. $z > 4$) through the *Herschel* '500μm risers' selection (e.g. Dowell et al., 2014; Asboth et al., 2016), poses a significant challenge to current models of galaxy assembly. As the dust SED of typical galaxies peaks at around 100 μm in the rest frame, a galaxy SED that appears to still be rising out to 500 μm suggests that these sources likely lie at $z > 4$. Currently, neither semi-analytic models nor hydrodynamic simulations can match the number density of such massive, early-forming, dusty galaxies, under-predicting their numbers by 1-2 orders of magnitude (e.g. Somerville et al., 2012); fundamental changes to models, such as a different IMF, may be required (Lacey et al., 2016). The characterisation of these highly luminous ($L_{\mathrm{IR}} \sim 10^{13}$ L$_{\odot}$) galaxies with extraordinarily high gas and dust masses at such early cosmic times can thus place important constraints on the physical mechanisms underpinning the formation of dust and metals in the early Universe. However, confirming the nature of such galaxies (in particular, the fraction of DSFGs that comprise multiple sources, the contribution of AGN to FIR flux, and the role of mergers in driving the extreme emission) has been difficult due to the poor spatial resolution of *Herschel* and limited sample sizes with spectroscopic confirmation. The availability of deep and wide LOFAR and multi-wavelength data, including the deep

surveys from *Herschel*, allows a promising route of investigation via sub-millimetre follow-up.

For this purpose, we selected LOFAR sources without an optical-to-MIR detected counterpart (but with secure radio position; see the stacking analysis above) or with a counterpart that has $z_{\rm phot} > 4$ (and a posterior minimum $z_{\rm phot,min} > 3.5$), indicative of heavily dust obscured and high-z sources (see Fig 5.2). We then selected the subset of sources with sharply rising FIR fluxes, using the well-established '500μm risers' (i.e. $S_{250} < S_{350} < S_{500}$) criteria (Cox et al., 2011; Pavesi et al., 2018; Ma et al., 2019) and $S_{500}/S_{350} > 1.2$, requiring a $\rm SNR_{500\mu m} > 4$ and a total SNR > 10 across the three SPIRE bands to remove sources with large flux uncertainties or shallow peaks. After these selections, the final sample was refined further using visual inspection of the FIR SED and *Herschel* and LOFAR cutouts to retain the most promising candidates with sharply rising SEDs. This selection results in a sample of 25 candidate $z > 4$ DSFGs. For this sample, we have performed MCMC fits to the FIR SED using a combination of hot and cold dust modified blackbody (MBB) models. This provides predictions for sub-mm fluxes, FIR luminosities, SFRs, and broad estimates of redshifts, while folding in parameter degeneracies. The IR-based star-formation rates were estimated from the dust fits to the SEDs using the standard Kennicutt (1998) relation. Fig. 5.2 shows the fit to the FIR dust SED for one of these sources (*top panel*), with the best-fitting model shown by the solid red line, and the lighter lines showing 600 realisations from the posterior distribution.

We have obtained sub-millimetre observations of all 25 of these sources from the sub-millimetre array (SMA) at 1.3 mm (observed between March – April 2021); a subset of these (12 sources) were also observed by the James Clerk Maxwell Telescope (JCMT) SCUBA-2 instrument at 850 μm (observed between February – June 2021). The SMA observations typically have a spatial resolution of around 2.5 – 3″ and an rms noise level between 0.4 – 0.8 mJy beam^{-1}, whereas the SCUBA-2 observations have a spatial resolution of around 14″ and a typical rms noise of $\sim$ 1.75 mJy beam^{-1}. The grey triangles in the SED plot in Fig. 5.2 (not included in the fit) show the subsequently-observed flux density at 850 μm and 1.3 mm from SCUBA-2 and SMA observations, respectively, which for this source shows good agreement, within errors, with the predicted flux density based on the SED fit from *Herschel* data alone. We note that the observed sub-mm points lie towards the lower end of the posterior distribution, likely indicating the lower end of the redshift range; this is not unexpected as the highest redshift population is rare, and the candidate 500 μm riser selection will likely be contaminated by the more common lower redshift sources that get scattered into our

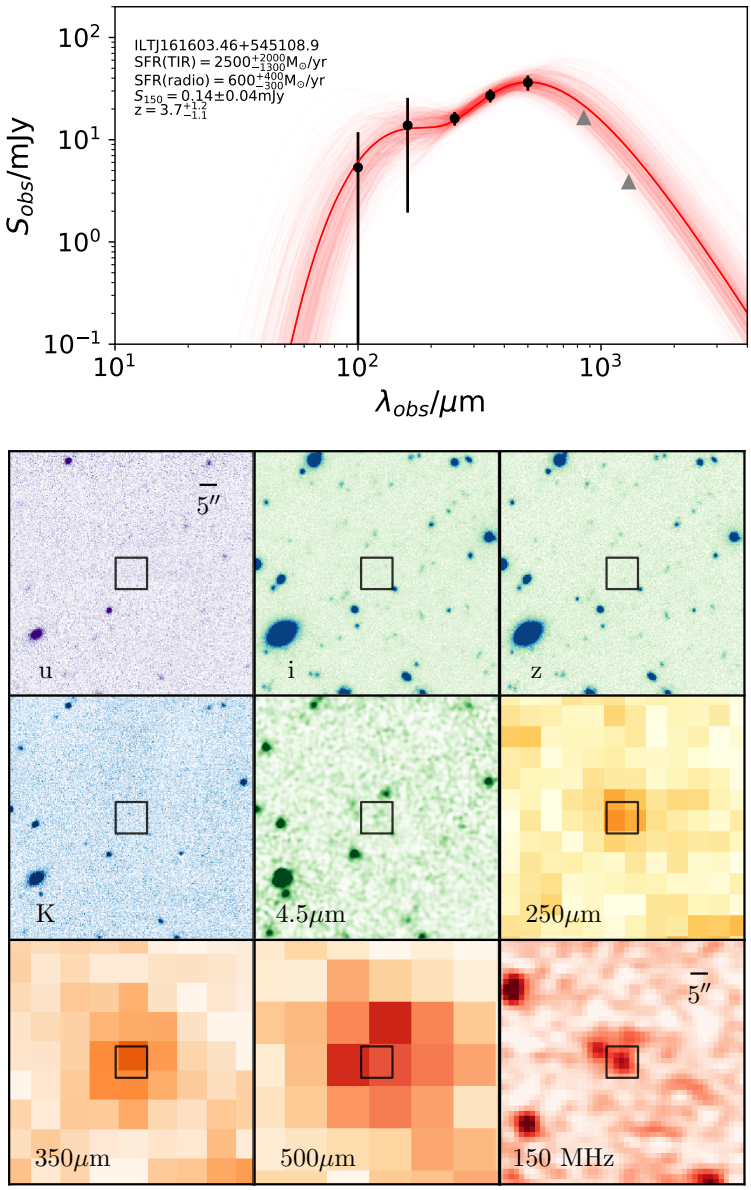

Figure 5.2 *Top: Fit to the FIR dust SED for one of the candidate high-redshift DSFGs selected for follow-up with the JCMT and the SMA. The black points show the data points, from Herschel PACS and SPIRE, used for SED fitting with the solid red line showing the best-fitting model, and the lighter red lines showing 600 realisations from the posterior. The grey triangles show the observed flux density at 850 μm and 1.1 mm from SCUBA-2 and SMA observations, respectively. Bottom: Optical, mid-IR, FIR, to 150 MHz cutouts for the same source; this source is clearly optically dark but shows bright FIR and radio detections. Cutouts are 35″ on each side.*

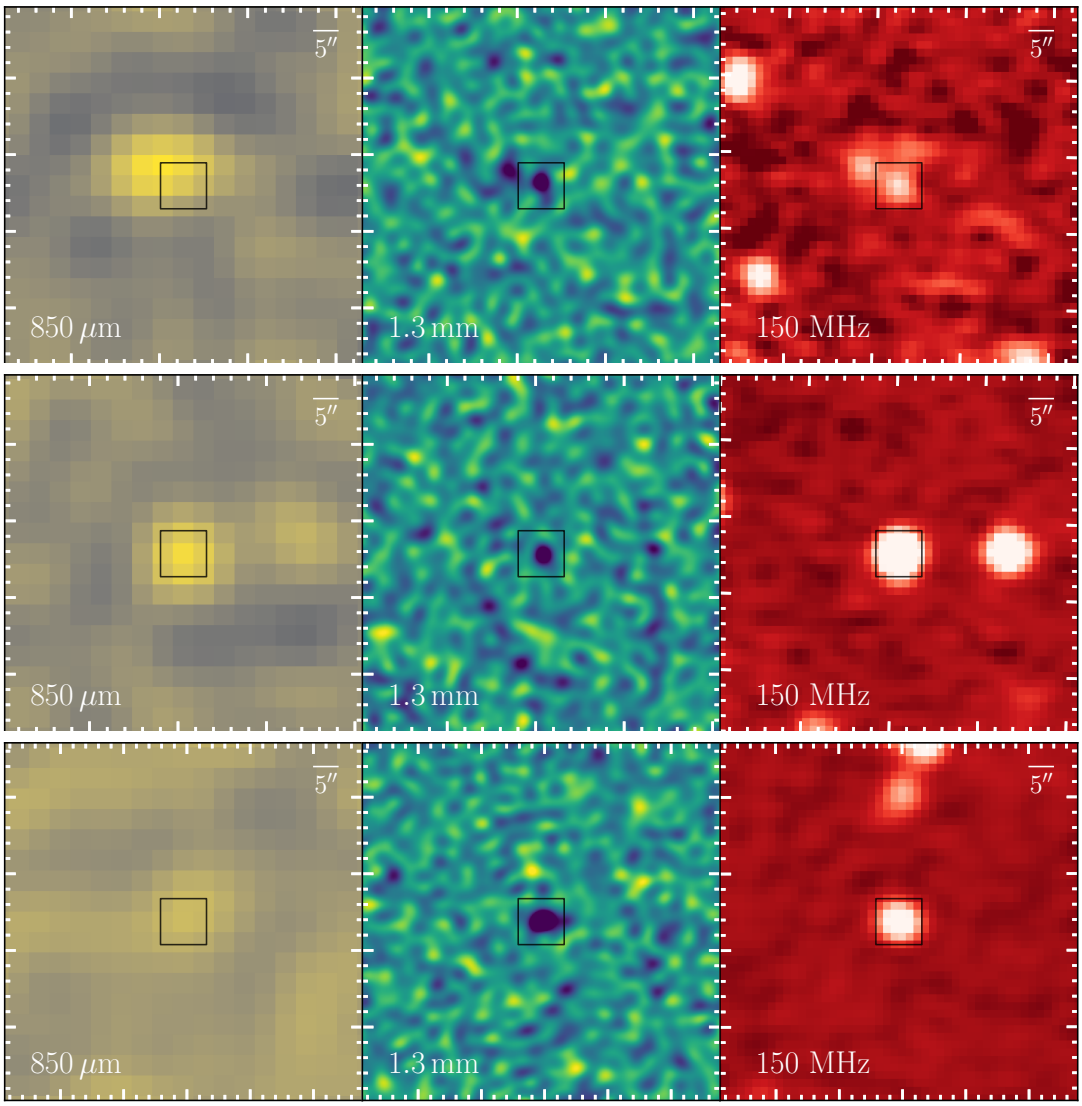

Figure 5.3 *Example cutouts for three targets (in rows) observed with SCUBA-2 (left) and the SMA (middle), and the LOFAR detections (right). The top row shows the observations for the same source whose FIR SED is shown in Fig. 5.2. Cutouts are 35″ on each side.*

selection due to flux uncertainties.

Fig. 5.3 shows cutouts from the SCUBA-2 (*left*), SMA (*middle*), and LOFAR (*right*) for three targets. The first two columns in the *top* panel shows the sub-millimetre observations for the source shown in Fig. 5.2, with strong detections in both wavelengths. Moreover, given the higher spatial resolution of the SMA dataset, the source detected at 850 μm and also in the *Herschel* bands, has been split into multiple components, in line with the radio-detections. The other two rows in the figure show cutouts for two other sources, with $> 5\sigma$ detections in all cases except at 850 μm for the third source (*bottom panel*).

More detailed analysis of these sub-millimetre observations are currently ongoing. By probing the Rayleigh-Jeans part of the dust SED with SCUBA-2 and SMA, we will be able to characterise this population and constrain their redshifts, SFRs, and total infrared luminosities significantly more tightly. In combination with our deep radio data, we will be able to distinguish galaxies where emission is driven by AGN activity rather than star formation. Given the large areas covered by the LoTSS Deep Fields, this can also set constraints on the space densities of these sources, down to fainter fluxes than previous wide-field DSFG studies. The data will also provide a basis for future high spatial resolution studies in multiple wavebands of sub-samples of these sources, for example those with the highest 1.3mm flux, or those that show signs of multiplicity; this will allow us to disentangle the thermal dust emission from the synchrotron AGN emission.

5.1.2 WEAVE-LOFAR: Dissecting LERG activity in exquisite detail at $z \lesssim 1$

The William Herschel Telescope Enhanced Area Velocity Explorer instrument (WEAVE), is a next-generation spectrograph, currently being installed on the William Herschel Telescope (WHT) in La Palma. WEAVE is a multi-object spectrograph (MOS), consisting of 1000 fibres that can be positioned across > 3 square degrees field-of-view of the instrument, and allows a contiguous wavelength coverage between 395 – 960 nm at high spectral resolution, making it an ideal survey instrument. Over its five years of operation, the WEAVE spectrograph will undertake a wide range of surveys, including the WEAVE-LOFAR survey (Smith et al., 2016). The WEAVE-LOFAR survey is designed to obtain high-resolution spectra for around 1 million sources from LoTSS, with the scientific drivers ranging from studying the dust-independent view

of star-formation history, the growth of SMBHs and AGN feedback, and cosmology, among others. In particular, the deep tier of the WEAVE-LOFAR survey will obtain spectra for *all* of the LOFAR-identified sources in the LoTSS Deep Fields and will greatly enhance the scientific analysis of radio-AGN feedback derived in Chapter 4 in numerous ways.

Spectroscopy is not only vital for providing robust redshifts, but it is also one of the most reliable methods of separating star-forming galaxies from AGN, and for differentiating between the various modes of AGN (e.g. Best & Heckman, 2012), resolving any residual source classification uncertainties in the analysis from Chapter 4. Moreover, spectra will also allow us to determine accurate physical galaxy and SMBH properties. Building on the analysis of the radio-AGN population in Chapter 4, a detailed understanding of AGN at moderate redshifts ($z \lesssim 1$) is crucial in fully interpreting the evolution at even higher redshifts. This is particularly critical for understanding the star-forming LERG population as these are found to be very rare at $z \sim 0$, but begin to reach comparable importance to the queiscent LERG population by $z \sim 1$. The LoTSS Deep Fields sample is large enough to perform statistical analysis even when split by multiple parameters simultaneously; the addition of WEAVE-LOFAR spectra, at these redshfits, will allow us to trace in detail the evolution of the sub-groups of the LERG population, and to extend the scope of the analysis in Chapter 4 by investigating the prevalence of radio-AGN activity as separate functions of not only stellar mass and radio luminosity but also environment, and to map the evolution of the host galaxy properties (e.g. stellar mass, star-formation rate, stellar age) of LERGs with redshift in unprecedented detail. Such detailed analysis has only previously been possible in the local Universe ($z \sim 0.1$; cf. Best & Heckman 2012): the cosmic evolution of these trends remains uncharted territory. As 'Survey Builder' for WEAVE-LOFAR, and a member of the WEAVE-LOFAR quality assurance group and science verification team, I plan to use the WEAVE-LOFAR spectra upon first light (due November 2021) to further characterise the LERG population in the LoTSS Deep Fields.

5.1.3 The nature of fuelling of radio-AGN

The analysis of the host galaxy properties of the LERG population in Chapter 4 has revealed that there exists a population of LERGs that are hosted in quiescent galaxies and fuelled by accretion of cooling hot gas, and a population of LERGs hosted in star-forming galaxies that are likely fuelled by a different mechanism, associated with the cold gas, as found for the HERG population. The present LoTSS Deep Fields

dataset alone, however, is not sufficient to understand how AGN activity is triggered in the star-forming LERGs and if the cold gas indeed plays a role.

WEAVE spectroscopy from the combination of the LoTSS Deep Fields dataset, suited to examine the low-luminosity LERG population, and the upcoming shallower but wider LoTSS data release 2, spanning $> 5000\,\mathrm{deg}^2$ and detecting large samples of the higher-luminosity HERG population, will be valuable in studying the incidence of radio-AGN across the galaxy population. Moreover, this combination of AGN samples from deep and wide LOFAR surveys will allow us to select matched samples (in e.g. redshift, radio luminosity, stellar mass) of HERGs, star-forming LERGs, and quiescent LERGs for further follow-up and characterisation of the cold molecular gas properties (e.g. by measuring molecular and dynamical gas masses, gas kinematics) of the AGN, and in comparison to inactive galaxies. This will help us to understand the physical processes triggering AGN activity in these different systems and how feedback from AGN affects the star-formation properties.

5.1.4 Testing feedback models and selection effects with simulations

As discussed in the introduction to the thesis, feedback from AGN is a critical ingredient in cosmological simulations, necessary to recover the observed local galaxy properties. However, due in large part to the poorly understood physical processes, methods of modelling and implementing SMBH growth and AGN feedback vary widely across simulations. The latest suite of simulations include advanced prescriptions for radio-AGN feedback (e.g. SIMBA; cf. Thomas et al. 2021); this now enables detailed comparisons with deep radio and multi-wavelength surveys, making it possible to directly constrain AGN feedback and galaxy formation models.

I plan to complement results from Chapter 4 and planned future work described in Sects. 5.1.2– 5.1.3 by extracting physical properties of radio-AGN host galaxies from the SIMBA simulations (Davé et al., 2019); moreover, I aim to also recover and fit photometry of the simulated radio-AGN, as done for the observations, applying the same source classification scheme to reconstruct the evolution of the radio LF from simulations. This comparison not only provides a direct constraint on the evolution of AGN feedback models from different simulations, but will also allow valuable tests of source classification uncertainties, and completeness and selection biases within observations.

The implementation of radiative-transfer in the latest simulations will allow me to use FIR selections to test the completeness of the radio-selected dusty high-z galaxies, informing refinement of observational selections used for the analysis described in Sect. 5.1.1.1. In parallel, I plan to trace the triggering of AGN and starburst activity, and their duty cycles from a theoretical perspective; in conjunction with measurements of the prevalence of AGN activity in the high-z radio sources, this will enable comparison to model predictions.

Bibliography

Aihara H., et al., 2018, First data release of the Hyper Suprime-Cam Subaru Strategic Program, PASJ, 70, S8

Aihara H., et al., 2019, Second Data Release of the Hyper Suprime-Cam Subaru Strategic Program, arXiv e-prints, p. arXiv:1905.12221

Aird J., et al., 2010, The evolution of the hard X-ray luminosity function of AGN, MNRAS, 401, 2531

Aird J., Coil A. L., Georgakakis A., 2019, X-rays across the galaxy population - III. The incidence of AGN as a function of star formation rate, MNRAS, 484, 4360

Albareti F. D., et al., 2017, The 13th Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey Mapping Nearby Galaxies at Apache Point Observatory, ApJS, 233, 25

Allen S. W., Dunn R. J. H., Fabian A. C., Taylor G. B., Reynolds C. S., 2006, The relation between accretion rate and jet power in X-ray luminous elliptical galaxies, MNRAS, 372, 21

Antonucci R., 1993, Unified models for active galactic nuclei and quasars, ARA&A, 31, 473

Asboth V., et al., 2016, HerMES: a search for high-redshift dusty galaxies in the HerMES Large Mode Survey - catalogue, number counts and early results, MNRAS, 462, 1989

Ashby M. L. N., et al., 2009, The Spitzer Deep, Wide-field Survey, ApJ, 701, 428

Audibert A., et al., 2019, ALMA captures feeding and feedback from the active galactic nucleus in NGC 613, A&A, 632, A33

Baldry I. K., et al., 2012, Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at z < 0.06, MNRAS, 421, 621

Barnes J. E., Hernquist L., 1996, Transformations of Galaxies. II. Gasdynamics in Merging Disk Galaxies, ApJ, 471, 115

Baugh C. M., 2006, A primer on hierarchical galaxy formation: the semi-analytical approach, Reports on Progress in Physics, 69, 3101

Becker R. H., White R. L., Helfand D. J., 1995, The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters, ApJ, 450, 559

Benson A. J., 2010, Galaxy formation theory, Phys. Rep., 495, 33

Bertin E., Arnouts S., 1996, SExtractor: Software for source extraction., A&AS, 117, 393

Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, Astronomical Society of the Pacific Conference Series Vol. 281, Astronomical Data Analysis Software and Systems XI. p. 228

Best P. N., 2004, The environmental dependence of radio-loud AGN activity and star formation in the 2dFGRS, MNRAS, 351, 70

Best P. N., Heckman T. M., 2012, On the fundamental dichotomy in the local radio-AGN population: accretion, evolution and host galaxy properties, MNRAS, 421, 1569

Best P. N., Arts J. N., Röttgering H. J. A., Rengelink R., Brookes M. H., Wall J., 2003, CENSORS: A Combined EIS-NVSS Survey Of Radio Sources - I. Sample definition, radio data and optical identifications, MNRAS, 346, 627

Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005a, The host galaxies of radio-loud active galactic nuclei: mass dependences, gas cooling and active galactic nuclei feedback, MNRAS, 362, 25

Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005b, A sample of radio-loud active galactic nuclei in the Sloan Digital Sky Survey, MNRAS, 362, 9

Best P. N., Kaiser C. R., Heckman T. M., Kauffmann G., 2006, AGN-controlled cooling in elliptical galaxies, MNRAS, 368, L67

Best P. N., von der Linden A., Kauffmann G., Heckman T. M., Kaiser C. R., 2007, On the prevalence of radio-loud active galactic nuclei in brightest cluster galaxies: implications for AGN heating of cooling flows, MNRAS, 379, 894

Best P. N., Ker L. M., Simpson C., Rigby E. E., Sabater J., 2014, The cosmic evolution of radio-AGN feedback to z = 1, MNRAS, 445, 955

Bian F., et al., 2013, The LBT Boötes Field Survey. I. The Rest-frame Ultraviolet and Near-infrared Luminosity Functions and Clustering of Bright Lyman Break Galaxies at Z ∼3, ApJ, 774, 28

Bielby R., et al., 2012, The WIRCam Deep Survey. I. Counts, colours, and mass-functions derived from near-infrared imaging in the CFHTLS deep fields, A&A, 545, A23

Birnboim Y., Dekel A., 2003, Virial shocks in galactic haloes?, MNRAS, 345, 349

Bîrzan L., Rafferty D. A., McNamara B. R., Wise M. W., Nulsen P. E. J., 2004, A Systematic Study of Radio-induced X-Ray Cavities in Clusters, Groups, and Galaxies, ApJ, 607, 800

Bîrzan L., McNamara B. R., Nulsen P. E. J., Carilli C. L., Wise M. W., 2008, Radiative Efficiency and Content of Extragalactic Radio Sources: Toward a Universal Scaling Relation between Jet Power and Radio Power, ApJ, 686, 859

Blain A. W., Smail I., Ivison R. J., Kneib J. P., Frayer D. T., 2002, Submillimeter galaxies, Phys. Rep., 369, 111

Boehringer H., Voges W., Fabian A. C., Edge A. C., Neumann D. M., 1993, A ROSAT HRI study of the interaction of the X-ray emitting gas and radio lobes of NGC 1275., MNRAS, 264, L25

Böhringer H., Matsushita K., Churazov E., Ikebe Y., Chen Y., 2002, The new emerging model for the structure of cooling cores in clusters of galaxies, A&A, 382, 804

Bondi H., 1952, On spherically symmetrical accretion, MNRAS, 112, 195

Bonzini M., et al., 2015, Star formation properties of sub-mJy radio sources, MNRAS, 453, 1079

Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, CIGALE: a python Code Investigating GALaxy Emission, A&A, 622, A103

Bouwens R. J., et al., 2012, UV-continuum Slopes at z ˜ 4-7 from the HUDF09+ERS+CANDELS Observations: Discovery of a Well-defined UV Color-Magnitude Relationship for z >= 4 Star-forming Galaxies, ApJ, 754, 83

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Breaking the hierarchy of galaxy formation, MNRAS, 370, 645

Brandt W. N., et al., 2001, The Chandra Deep Field North Survey. V. 1 Ms Source Catalogs, AJ, 122, 2810

Brown M. J. I., Dey A., Jannuzi B. T., Brand K., Benson A. J., Brodwin M., Croton D. J., Eisenhardt P. R., 2007, The Evolving Luminosity Function of Red Galaxies, ApJ, 654, 858

Brown M. J. I., et al., 2008, Red Galaxy Growth and the Halo Occupation Distribution, ApJ, 682, 937

Bruzual G., Charlot S., 2003, Stellar population synthesis at the resolution of 2003, MNRAS, 344, 1000

Burgarella D., Buat V., Iglesias-Páramo J., 2005, Star formation and dust attenuation properties in galaxies from a statistical ultraviolet-to-far-infrared analysis, MNRAS, 360, 1413

Butler A., et al., 2018, The XXL Survey. XVIII. ATCA 2.1 GHz radio source catalogue and source counts for the XXL-South field, A&A, 620, A3

Butler A., Huynh M., Kapińska A., Delvecchio I., Smolčić V., Chiappetti L., Koulouridis E., Pierre M., 2019, The XXL Survey. XXXVI. Evolution and black hole feedback of high-excitation and low-excitation radio galaxies in XXL-S, A&A, 625, A111

Calistro Rivera G., Lusso E., Hennawi J. F., Hogg D. W., 2016, AGNfitter: A Bayesian MCMC Approach to Fitting Spectral Energy Distributions of AGNs, ApJ, 833, 98

Calistro Rivera G., et al., 2017, The LOFAR window on star-forming galaxies and AGNs - curved radio SEDs and IR-radio correlation at 0<z<2.5, MNRAS, 469, 3468

Calistro Rivera G., et al., 2018, Resolving the ISM at the Peak of Cosmic Star Formation with ALMA: The Distribution of CO and Dust Continuum in z ~ 2.5 Submillimeter Galaxies, ApJ, 863, 56

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, The Dust Content and Opacity of Actively Star-forming Galaxies, ApJ, 533, 682

Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, Observational evidence of quasar feedback quenching star formation at high redshift, A&A, 537, L8

Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, Inferring the star formation histories of massive quiescent galaxies with BAGPIPES: evidence for multiple quenching mechanisms, MNRAS, 480, 4379

Carnall A. C., et al., 2020, Timing the earliest quenching events with a robust sample of massive quiescent galaxies at 2 < z < 5, MNRAS, 496, 695

Carniani S., et al., 2016, Fast outflows and star formation quenching in quasar host galaxies, A&A, 591, A28

Casali M., et al., 2007, The UKIRT wide-field camera, A&A, 467, 777

Casey C. M., Narayanan D., Cooray A., 2014, Dusty star-forming galaxies at high redshift, Phys. Rep., 541, 45

Cattaneo A., et al., 2009, The role of black holes in galaxy formation and evolution, Nature, 460, 213

Cavagnolo K. W., McNamara B. R., Nulsen P. E. J., Carilli C. L., Jones C., Bîrzan L., 2010, A Relationship Between AGN Jet Power and Radio Power, ApJ, 720, 1066

Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest, ApJ, 437, L9

Chabrier G., 2003, Galactic Stellar and Substellar Initial Mass Function, PASP, 115, 763

Chambers K. C., et al., 2016, The Pan-STARRS1 Surveys, arXiv e-prints, p. arXiv:1612.05560

Chiaberge M., Macchetto F. D., Sparks W. B., Capetti A., Allen M. G., Martel A. R., 2002, The Nuclei of Radio Galaxies in the Ultraviolet: The Signature of Different Emission Processes, ApJ, 571, 247

Ching J. H. Y., et al., 2017, The Large Area Radio Galaxy Evolution Spectroscopic Survey (LARGESS): survey design, data catalogue and GAMA/WiggleZ spectroscopy, MNRAS, 464, 1306

Ciliegi P., et al., 1999, A deep VLA survey at 20 CM of the ISO ELAIS survey regions, MNRAS, 302, 222

Cisternas M., et al., 2013, X-Ray Nuclear Activity in S^4G Barred Galaxies: No Link between Bar Strength and Co-occurrent Supermassive Black Hole Fueling, ApJ, 776, 50

Colless M., et al., 2001, The 2dF Galaxy Redshift Survey: spectra and redshifts, MNRAS, 328, 1039

Condon J. J., 1989, The 1.4 GHz Luminosity Function and Its Evolution, ApJ, 338, 13

Condon J. J., 1997, Errors in Elliptical Gaussian FITS, Publications of the Astronomical Society of the Pacific, 109, 166

Conselice C. J., 2014, The Evolution of Galaxy Structure Over Cosmic Time, ARA&A, 52, 291

Cool R. J., 2007, zBootes: z-Band Photometry in the NOAO Deep Wide-Field Survey Bootes Field, ApJS, 169, 21

Coupon J., Czakon N., Bosch J., Komiyama Y., Medezinski E., Miyazaki S., Oguri M., 2018, The bright-star masks for the HSC-SSP survey, PASJ, 70, S7

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields, AJ, 112, 839

Cox P., et al., 2011, Gas and Dust in a Submillimeter Galaxy at z = 4.24 from the Herschel Atlas, ApJ, 740, 63

Cresci G., et al., 2015, Blowin' in the Wind: Both "Negative" and "Positive" Feedback in an Obscured High-z Quasar, ApJ, 799

Croston J. H., et al., 2019, The environments of radio-loud AGN from the LOFAR Two-Metre Sky Survey (LoTSS), A&A, 622, A10

Croton D. J., et al., 2006, The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies, MNRAS, 365, 11

Croton D. J., et al., 2016, Semi-Analytic Galaxy Evolution (SAGE): Model Calibration and Basic Results, ApJS, 222, 22

Das S., et al., 2011, Detection of the Power Spectrum of Cosmic Microwave Background Lensing by the Atacama Cosmology Telescope, Physical Review Letters, 107, 021301

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, SIMBA: Cosmological simulations with black hole growth and feedback, MNRAS, 486, 2827

Dekel A., Birnboim Y., 2006, Galaxy bimodality due to cold flows and shock heating, MNRAS, 368, 2

Di Matteo T., Springel V., Hernquist L., 2005, Energy input from quasars regulates the growth and activity of black holes and their host galaxies, Nature, 433, 604

Domínguez Sánchez H., et al., 2011, The evolution of quiescent galaxies at high redshifts (z≥ 1.4), MNRAS, 417, 900

Donley J. L., et al., 2018, Evidence for Merger-driven Growth in Luminous, High-z, Obscured AGNs in the CANDELS/COSMOS Field, ApJ, 853, 63

Dowell C. D., et al., 2014, HerMES: Candidate High-redshift Galaxies Discovered with Herschel/SPIRE, ApJ, 780, 75

Driver S. P., Robotham A. S. G., 2010, Quantifying cosmic variance, MNRAS, 407, 2131

Duncan K., et al., 2014, The mass evolution of the first galaxies: stellar mass functions and star formation rates at $4 < z < 7$ in the CANDELS GOODS-South field, MNRAS, 444, 2960

Duncan K. J., et al., 2018a, Photometric redshifts for the next generation of deep radio continuum surveys - I. Template fitting, MNRAS, 473, 2655

Duncan K. J., Jarvis M. J., Brown M. J. I., Röttgering H. J. A., 2018b, Photometric redshifts for the next generation of deep radio continuum surveys - II. Gaussian processes and hybrid estimates, MNRAS, 477, 5177

Duncan K., et al., 2019, Observational Constraints on the Merger History of Galaxies since $z \approx 6$: Probabilistic Galaxy Pair Counts in the CANDELS Fields, ApJ, 876, 110

Duncan K. J., et al., 2021, The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. IV. Photometric redshifts and stellar masses, A&A, 648, A4

Dunlop J. S., Peacock J. A., 1990, The Redshift Cut-Off in the Luminosity Function of Radio Galaxies and Quasars, MNRAS, 247, 19

Eisenhardt P. R., et al., 2004, The Infrared Array Camera (IRAC) Shallow Survey, ApJS, 154, 48

Engelbracht C. W., et al., 2007, Absolute Calibration and Characterization of the Multiband Imaging Photometer for Spitzer. I. The Stellar Calibrator Sample and the 24 μm Calibration, PASP, 119, 994

Evans D. W., et al., 2018, Gaia Data Release 2. Photometric content and validation, A&A, 616, A4

Fabian A. C., 1994, Cooling Flows in Clusters of Galaxies, ARA&A, 32, 277

Fabian A. C., 2012, Observational Evidence of Active Galactic Nuclei Feedback, ARA&A, 50, 455

Fabian A. C., Sanders J. S., Taylor G. B., Allen S. W., Crawford C. S., Johnstone R. M., Iwasawa K., 2006, A very deep Chandra observation of the Perseus cluster: shocks, ripples and conduction, MNRAS, 366, 417

Fazio G. G., et al., 2004, The Infrared Array Camera (IRAC) for the Spitzer Space Telescope, ApJS, 154, 10

Ferrarese L., Merritt D., 2000, A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies, ApJ, 539, L9

Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, Quasar feedback revealed by giant molecular outflows, A&A, 518

Finkbeiner D. P., et al., 2016, Hypercalibration: A Pan-STARRS1-based Recalibration of the Sloan Digital Sky Survey Photometry, ApJ, 822, 66

Fischer J., et al., 2010, Herschel-PACS spectroscopic diagnostics of local ULIRGs: Conditions and kinematics in Markarian 231, A&A, 518

Fitzpatrick E. L., 1999, Correcting for the Effects of Interstellar Extinction, PASP, 111, 63

Fleuren S., et al., 2012, Herschel-ATLAS: VISTA VIKING near-infrared counterparts in the Phase 1 GAMA 9-h data, MNRAS, 423, 2407

Fluetsch A., et al., 2019, Cold molecular outflows in the local Universe and their feedback effect on galaxies, MNRAS, 483, 4586

Fontana A., et al., 2009, The fraction of quiescent massive galaxies in the early Universe, A&A, 501, 15

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, emcee: The MCMC Hammer, PASP, 125, 306

Fritz J., Franceschini A., Hatziminaoglou E., 2006, Revisiting the infrared spectra of active galactic nuclei with a new torus emission model, MNRAS, 366, 767

Gabor J. M., Davé R., 2015, Hot gas in massive haloes drives both mass quenching and environment quenching, MNRAS, 447, 374

Gaia Collaboration et al., 2016, The Gaia mission, A&A, 595, A1

Gaia Collaboration et al., 2018, Gaia Data Release 2. Summary of the contents and survey properties, A&A, 616, A1

Galloway M. A., et al., 2015, Galaxy Zoo: the effect of bar-driven fuelling on the presence of an active galactic nucleus in disc galaxies, MNRAS, 448, 3442

Gao F., et al., 2020, Mergers trigger active galactic nuclei out to z ∼ 0.6, A&A, 637, A94

Gao Y., Egusa F., Liu G., Kohno K., Bao M., Morokuma-Matsui K., Kong X., Chen X., 2021, The Nuclear Region of NGC 1365: Star Formation, Negative Feedback, and Outflow Structure, ApJ, 913, 139

Garn T., Green D. A., Riley J. M., Alexand er P., 2008a, A 610-MHz survey of the ELAIS-N1 field with the Giant Metrewave Radio Telescope - observations, data analysis and source catalogue, MNRAS, 383, 75

Garn T., Green D. A., Riley J. M., Alexand er P., 2008b, A 610-MHz survey of the Lockman Hole with the Giant Metrewave Radio Telescope - I. Observations, data reduction and source catalogue for the central 5 deg^2, MNRAS, 387, 1037

Gaspari M., Ruszkowski M., Oh S. P., 2013, Chaotic cold accretion on to black holes, MNRAS, 432, 3401

Gaspari M., Brighenti F., Temi P., 2015, Chaotic cold accretion on to black holes in rotating atmospheres, A&A, 579, A62

Gaspari M., Temi P., Brighenti F., 2017, Raining on black holes and massive galaxies: the top-down multiphase condensation model, MNRAS, 466, 677

Gebhardt K., et al., 2000, A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion, ApJ, 539, L13

Gehrels N., 1986, Confidence Limits for Small Numbers of Events in Astrophysical Data, ApJ, 303, 336

Gendre M. A., Best P. N., Wall J. V., Ker L. M., 2013, The relation between morphology, accretion modes and environmental factors in local radio AGN, MNRAS, 430, 3086

Gonzalez A. H., et al., 2010, in American Astronomical Society Meeting Abstracts #216. p. 415.13

Gordon Y. A., et al., 2019, The Effect of Minor and Major Mergers on the Evolution of Low-excitation Radio Galaxies, ApJ, 878, 88

Goulding A. D., et al., 2017, Galaxy-scale Bars in Late-type Sloan Digital Sky Survey Galaxies Do Not Influence the Average Accretion Rates of Supermassive Black Holes, ApJ, 843, 135

Green G. M., 2018, dustmaps: A Python interface for maps of interstellar dust, The Journal of Open Source Software, 3, 695

Griffin M. J., et al., 2010, The Herschel-SPIRE instrument and its in-flight performance, A&A, 518, L3

Gültekin K., King A. L., Cackett E. M., Nyland K., Miller J. M., Di Matteo T., Markoff S., Rupen M. P., 2019, The Fundamental Plane of Black Hole Accretion and Its Use as a Black Hole-Mass Estimator, ApJ, 871, 80

Gürkan G., et al., 2018, LOFAR/H-ATLAS: the low-frequency radio luminosity-star formation rate relation, MNRAS, 475, 3010

Hardcastle M., 2018, Interpreting radiative efficiency in radio-loud AGNs, Nature Astronomy, 2, 273

Hardcastle M. J., Croston J. H., 2020, Radio galaxies and feedback from AGN jets, New A Rev., 88, 101539

Hardcastle M. J., Evans D. A., Croston J. H., 2007, Hot and cold gas accretion and feedback in radio-loud active galaxies, MNRAS, 376, 1849

Hardcastle M. J., et al., 2019, Radio-loud AGN in the first LoTSS data release. The lifetimes and environmental impact of jet-driven sources, A&A, 622, A12

Harrison C. M., et al., 2012, Energetic galaxy-wide outflows in high-redshift ultraluminous infrared galaxies hosting AGN activity, MNRAS, 426, 1073

Hasinger G., et al., 2001, XMM-Newton observation of the Lockman Hole. I. The X-ray data, A&A, 365, L45

Heckman T. M., Best P. N., 2014, The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe, ARA&A, 52, 589

Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, Present-Day Growth of Black Holes and Bulges: The Sloan Digital Sky Survey Perspective, ApJ, 613, 109

Herrera-Camus R., et al., 2020, Molecular gas inflows and outflows in ultraluminous infrared galaxies at z ∼ 0.2 and one QSO at z = 6.1, A&A, 633, L4

Hewett P. C., Warren S. J., Leggett S. K., Hodgkin S. T., 2006, The UKIRT Infrared Deep Sky Survey ZY JHK photometric system: passbands and synthetic colours, MNRAS, 367, 454

Hildebrandt H., et al., 2016, RCSLenS: The Red Cluster Sequence Lensing Survey, MNRAS, 463, 635

Hinshaw G., et al., 2013, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results, ApJS, 208, 19

Hodapp K. W., et al., 2004, Design of the Pan-STARRS telescopes, Astronomische Nachrichten, 325, 636

Hopkins A. M., 2004, On the Evolution of Star-forming Galaxies, ApJ, 615, 209

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, A Cosmological Framework for the Co-Evolution of Quasars, Supermassive Black Holes, and Elliptical Galaxies. I. Galaxy Mergers and Quasar Activity, ApJS, 175, 356

Hoyle F., Fowler W. A., 1963, Nature of Strong Radio Sources, Nature, 197, 533

Huber M., PS1 Science Consortium Pan-STARRS IPP Team 2017, in American Astronomical Society Meeting Abstracts #229. p. 237.06

Hurley P. D., et al., 2017, HELP: XID+, the probabilistic de-blender for Herschel SPIRE maps, MNRAS, 464, 885

Ibar E., Ivison R. J., Biggs A. D., Lal D. V., Best P. N., Green D. A., 2009, Deep multi-frequency radio imaging in the Lockman Hole using the GMRT and VLA - I. The nature of the sub-mJy radio population, MNRAS, 397, 281

Ilbert O., et al., 2013, Mass assembly in quiescent and star-forming galaxies since z ≈ 4 from UltraVISTA, A&A, 556, A55

Intema H. T., van der Tol S., Cotton W. D., Cohen A. S., van Bemmel I. M., Röttgering H. J. A., 2009, Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results, A&A, 501, 1185

Intema H. T., van Weeren R. J., Röttgering H. J. A., Lal D. V., 2011, Deep low-frequency radio observations of the NOAO Boötes field. I. Data reduction and catalog construction, A&A, 535, A38

Ishwara-Chandra C. H., Taylor A. R., Green D. A., Stil J. M., Vaccari M., Ocran E. F., 2020, A wide-area GMRT 610-MHz survey of ELAIS N1 field, MNRAS, 497, 5383

Jannuzi B. T., Dey A., 1999, in Weymann R., Storrie-Lombardi L., Sawicki M., Brunner R., eds, Astronomical Society of the Pacific Conference Series Vol. 191, Photometric Redshifts and the Detection of High Redshift Galaxies. p. 111

Janssen R. M. J., Röttgering H. J. A., Best P. N., Brinchmann J., 2012, The triggering probability of radio-loud AGN. A comparison of high and low excitation radio galaxies in hosts of different colors, A&A, 541, A62

Jarvis M. J., Teimourian H., Simpson C., Smith D. J. B., Rawlings S., Bonfield D., 2009, The discovery of a typical radio galaxy at z = 4.88, MNRAS, 398, L83

Jones C., et al., 2004, in 35th COSPAR Scientific Assembly. p. 4119 (arXiv:astro-ph/0312576)

Kaiser N., et al., 2010, in Ground-based and Airborne Telescopes III. p. 77330E, doi:10.1117/12.859188

Kennicutt Robert C. J., 1998, Star Formation in Galaxies Along the Hubble Sequence, ARA&A, 36, 189

Kim M., Choi Y.-Y., 2020, The Relative Role of Bars and Galaxy Environments in AGN Triggering of SDSS Spirals, ApJ, 901, L38

King A., 2003, Black Holes, Galaxy Formation, and the M_{BH}-σ Relation, ApJ, 596, L27

Kollmeier J. A., et al., 2006, Black Hole Masses and Eddington Ratios at $0.3 < z < 4$, ApJ, 648, 128

Kondapally R., et al., 2021, The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. III. Host-galaxy identifications and value added catalogues, A&A, 648, A3

Kormendy J., Ho L. C., 2013, Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies, ARA&A, 51, 511

Lacey C. G., et al., 2016, A unified multiwavelength model of galaxy formation, MNRAS, 462, 3854

Laigle C., et al., 2016, The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies, ApJS, 224, 24

Lambrides E., et al., 2021, Lower-Luminosity Obscured AGN Host Galaxies are Not Predominantly in Major-Merging Systems at Cosmic Noon, arXiv e-prints, p. arXiv:2107.07533

Lang D., Hogg D. W., Mykytyn D., 2016, The Tractor: Probabilistic astronomical source detection and measurement (ascl:1604.008)

Lawrence A., et al., 2007, The UKIRT Infrared Deep Sky Survey (UKIDSS), MNRAS, 379, 1599

Lee G.-H., Woo J.-H., Lee M. G., Hwang H. S., Lee J. C., Sohn J., Lee J. H., 2012, Do Bars Trigger Activity in Galactic Nuclei?, ApJ, 750, 141

Lonsdale C. J., et al., 2003, SWIRE: The SIRTF Wide-Area Infrared Extragalactic Survey, PASP, 115, 897

Lutz D., et al., 2020, Molecular outflows in local galaxies: Method comparison and a role of intermittent AGN driving, A&A, 633, A134

Lynden-Bell D., 1969, Galactic Nuclei as Collapsed Old Quasars, Nature, 223, 690

Lynden-Bell D., Rees M. J., 1971, On quasars, dust and the galactic centre, MNRAS, 152, 461

Ma J., et al., 2019, Spitzer Catalog of Herschel-selected Ultrared Dusty Star-forming Galaxies, ApJS, 244, 30

Maccagni F. M., Morganti R., Oosterloo T. A., Oonk J. B. R., Emonts B. H. C., 2018, ALMA observations of AGN fuelling. The case of PKS B1718-649, A&A, 614, A42

Mainieri V., et al., 2011, Black hole accretion and host galaxies of obscured quasars in XMM-COSMOS, A&A, 535, A80

Manners J. C., et al., 2003, The ELAIS deep X-ray survey - I. Chandra source catalogue and first results, MNRAS, 343, 293

Martin D. C., et al., 2005, The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission, ApJ, 619, L1

Mauch T., Sadler E. M., 2007, Radio sources in the 6dFGS: local luminosity functions at 1.4GHz for star-forming galaxies and radio-loud AGN, MNRAS, 375, 931

Mauduit J.-C., et al., 2012, The Spitzer Extragalactic Representative Volume Survey (SERVS): Survey Definition and Goals, PASP, 124, 714

McAlpine K., Smith D. J. B., Jarvis M. J., Bonfield D. G., Fleuren S., 2012, The likelihood ratio as a tool for radio continuum surveys with Square Kilometre Array precursor telescopes, MNRAS, 423, 132

McCheyne I., et al., 2020, subm., The LoTSS Deep Fields: FIR data, A&A

McLeod D. J., McLure R. J., Dunlop J. S., Cullen F., Carnall A. C., Duncan K., 2021, The evolution of the galaxy stellar-mass function over the last 12 billion years from a combination of ground-based and HST surveys, MNRAS, 503, 4413

McNamara B. R., Nulsen P. E. J., 2007, Heating Hot Atmospheres with Active Galactic Nuclei, ARA&A, 45, 117

McNamara B. R., Rohanizadegan M., Nulsen P. E. J., 2011, Are Radio Active Galactic Nuclei Powered by Accretion or Black Hole Spin?, ApJ, 727, 39

Merlin E., et al., 2015, T-PHOT: A new code for PSF-matched, prior-based, multiwavelength extragalactic deconfusion photometry, A&A, 582, A15

Merlin E., et al., 2016, T-PHOT version 2.0: Improved algorithms for background subtraction, local convolution, kernel registration, and new options, A&A, 595, A97

Merlin E., et al., 2018, Chasing passive galaxies in the early Universe: a critical analysis in CANDELS GOODS-South, MNRAS, 473, 2098

Merloni A., Heinz S., di Matteo T., 2003, A Fundamental Plane of black hole activity, MNRAS, 345, 1057

Merritt D., Ferrarese L., 2001, Black hole demographics from the M-σ relation, MNRAS, 320, L30

Mingo B., Hardcastle M. J., Croston J. H., Dicken D., Evans D. A., Morganti R., Tadhunter C., 2014, An X-ray survey of the 2 Jy sample - I. Is there an accretion mode dichotomy in radio-loud AGN?, MNRAS, 440, 269

Mohan N., Rafferty D., 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library (ascl:1502.007)

Morganti R., Oosterloo T., Oonk J. B. R., Frieswijk W., Tadhunter C., 2015, The fast molecular outflow in the Seyfert galaxy IC 5063 as seen by ALMA, A&A, 580, A1

Morrissey P., et al., 2007, The Calibration and Data Products of GALEX, ApJS, 173, 682

Moustakas J., et al., 2013, PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from z = 0-1, ApJ, 767, 50

Murray S. S., et al., 2005, XBootes: An X-Ray Survey of the NDWFS Bootes Field. I. Overview and Initial Results, ApJS, 161, 1

Muzzin A., et al., 2009, Spectroscopic Confirmation of Two Massive Red-Sequence-Selected Galaxy Clusters at z ∼1.2 in the SpARCS-North Cluster Survey, ApJ, 698, 1934

Muzzin A., et al., 2013, The Evolution of the Stellar Mass Functions of Star-forming and Quiescent Galaxies to z = 4 from the COSMOS/UltraVISTA Survey, ApJ, 777, 18

Narayan R., Yi I., 1994, Advection-dominated accretion: A self-similar solution, ApJ, 428, L13

Narayan R., Yi I., 1995, Advection-dominated Accretion: Underfed Black Holes and Neutron Stars, ApJ, 452, 710

Nisbet D. M., 2018, PhD thesis, The University of Edinburgh

Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, Analysis of galaxy spectral energy distributions from far-UV to far-IR with CIGALE: studying a SINGS test sample, A&A, 507, 1793

North E. V., et al., 2019, WISDOM project - V. Resolving molecular gas in Keplerian rotation around the supermassive black hole in NGC 0383, MNRAS, 490, 319

Nyland K., et al., 2017, An Application of Multi-band Forced Photometry to One Square Degree of SERVS: Accurate Photometric Redshifts and Implications for Future Science, ApJS, 230, 9

Ocaña Flaquer B., Leon S., Combes F., Lim J., 2010, TANGO I: Interstellar medium in nearby radio galaxies. Molecular gas, A&A, 518, A9

Ocran E. F., Taylor A. R., Vaccari M., Ishwara-Chand ra C. H., Prandoni I., 2019, Deep GMRT 610 MHz Observations of the ELAIS N1 Field : Catalogue and Source Counts, MNRAS, p. 2558

Oke J. B., Gunn J. E., 1983, Secondary standard stars for absolute spectrophotometry., ApJ, 266, 713

Okuda T., Kohno K., Iguchi S., Nakanishi K., 2005, Rotating Molecular Gas Associated with a Silhouette Disk at the Center of the Radio Galaxy 3C 31, ApJ, 620, 673

Oliver S., et al., 2000, The European Large Area ISO Survey - I. Goals, definition and observations, MNRAS, 316, 749

Oliver S. J., et al., 2012, The Herschel Multi-tiered Extragalactic Survey: HerMES, MNRAS, 424, 1614

Oteo I., et al., 2016, Witnessing the Birth of the Red Sequence: ALMA High-resolution Imaging of [C II] and Dust in Two Interacting Ultra-red Starbursts at z = 4.425, ApJ, 827, 34

Pacifici C., et al., 2016, The Evolution of Star Formation Histories of Quiescent Galaxies, ApJ, 832, 79

Pavesi R., et al., 2018, Hidden in Plain Sight: A Massive, Dusty Starburst in a Galaxy Protocluster at z = 5.7 in the COSMOS Field, ApJ, 861, 43

Peebles P. J. E., Yu J. T., 1970, Primeval Adiabatic Perturbation in an Expanding Universe, ApJ, 162, 815

Pereira-Santaella M., et al., 2018, Spatially resolved cold molecular outflows in ULIRGs, preprint, (arXiv:1805.03667)

Pérez-Torres M., Mattila S., Alonso-Herrero A., Aalto S., Efstathiou A., 2021, Star formation and nuclear activity in luminous infrared galaxies: an infrared through radio review, A&A Rev., 29, 2

Perlmutter S., et al., 1999, Measurements of Ω and Λ from 42 High-Redshift Supernovae, ApJ, 517, 565

Peterson J. R., Fabian A. C., 2006, X-ray spectroscopy of cooling clusters, Phys. Rep., 427, 1

Peterson J. R., et al., 2001, X-ray imaging-spectroscopy of Abell 1835, A&A, 365, L104

Peterson J. R., Kahn S. M., Paerels F. B. S., Kaastra J. S., Tamura T., Bleeker J. A. M., Ferrigno C., Jernigan J. G., 2003, High-Resolution X-Ray Spectroscopic Constraints on Cooling-Flow Models for Clusters of Galaxies, ApJ, 590, 207

Pilbratt G. L., et al., 2010, Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy, A&A, 518, L1

Planck Collaboration et al., 2016, Planck 2015 results. XIII. Cosmological parameters, A&A, 594, A13

Planck Collaboration et al., 2020, Planck 2018 results. VI. Cosmological parameters, A&A, 641, A6

Poglitsch A., et al., 2010, The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory, A&A, 518, L2

Pozzetti L., et al., 2010, zCOSMOS - 10k-bright spectroscopic sample. The bimodality in the galaxy stellar mass function: exploring its evolution with redshift, A&A, 523, A13

Pracy M. B., et al., 2016, GAMA/WiggleZ: the 1.4 GHz radio luminosity functions of high- and low-excitation radio galaxies and their redshift evolution to z = 0.75, MNRAS, 460, 2

Prestage R. M., Peacock J. A., 1988, The cluster environments of powerful radio galaxies., MNRAS, 230, 131

Rafferty D. A., McNamara B. R., Nulsen P. E. J., Wise M. W., 2006, The Feedback-regulated Growth of Black Holes and Bulges through Gas Accretion and Starbursts in Cluster Central Dominant Galaxies, ApJ, 652, 216

Reichard T. A., Heckman T. M., Rudnick G., Brinchmann J., Kauffmann G., Wild V., 2009, The Lopsidedness of Present-Day Galaxies: Connections to the Formation of Stars, the Chemical Evolution of Galaxies, and the Growth of Black Holes, ApJ, 691, 1005

Remillard R. A., McClintock J. E., 2006, X-Ray Properties of Black-Hole Binaries, ARA&A, 44, 49

Retana-Montenegro E., Röttgering H. J. A., Shimwell T. W., van Weeren R. J., Prandoni I., Brunetti G., Best P. N., Brüggen M., 2018, Deep LOFAR 150 MHz imaging of the Boötes field: Unveiling the faint low-frequency sky, A&A, 620, A74

Rieke G. H., et al., 2004, The Multiband Imaging Photometer for Spitzer (MIPS), ApJS, 154, 25

Riello M., et al., 2018, Gaia Data Release 2. Processing of the photometric data, A&A, 616, A3

Riess A. G., et al., 1998, Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, AJ, 116, 1009

Rigby E. E., Best P. N., Brookes M. H., Peacock J. A., Dunlop J. S., Röttgering H. J. A., Wall J. V., Ker L., 2011, The luminosity-dependent high-redshift turnover in the steep spectrum radio luminosity function: clear evidence for downsizing in the radio-AGN population, MNRAS, 416, 1900

Rigby E. E., Argyle J., Best P. N., Rosario D., Röttgering H. J. A., 2015, Cosmic downsizing of powerful radio galaxies to low radio luminosities, A&A, 581, A96

Ruffa I., et al., 2019a, The AGN fuelling/feedback cycle in nearby radio galaxies I. ALMA observations and early results, MNRAS, 484, 4239

Ruffa I., et al., 2019b, The AGN fuelling/feedback cycle in nearby radio galaxies - II. Kinematics of the molecular gas, MNRAS, 489, 3739

Runco J. N., Malkan M. A., Fernández-Ontiveros J. A., Spinoglio L., Pereira-Santaella M., 2020, Herschel/PACS OH Spectroscopy of Seyfert, LINER, and Starburst Galaxies, ApJ, 905, 57

Sabater J., Best P. N., Argudo-Fernández M., 2013, Effect of the interactions and environment on nuclear activity, MNRAS, 430, 638

Sabater J., Best P. N., Heckman T. M., 2015, Triggering optical AGN: the need for cold gas, and the indirect roles of galaxy environment and interactions, MNRAS, 447, 110

Sabater J., et al., 2019, The LoTSS view of radio AGN in the local Universe. The most massive galaxies are always switched on, A&A, 622, A17

Sabater J., et al., 2021, The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. II. The ELAIS-N1 LOFAR deep field, A&A, 648, A2

Saxena A., et al., 2018, Discovery of a radio galaxy at z = 5.72, MNRAS, 480, 2733

Scannapieco C., et al., 2012, The Aquila comparison project: the effects of feedback and numerical methods on simulations of galaxy formation, MNRAS, 423, 1726

Schaye J., et al., 2015, The EAGLE project: simulating the evolution and assembly of galaxies and their environments, MNRAS, 446, 521

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds, ApJ, 500, 525

Schmidt M., 1963, 3C 273 : A Star-Like Object with Large Red-Shift, Nature, 197, 1040

Schmidt M., 1968, Space Distribution and Luminosity Functions of Quasi-Stellar Radio Sources, ApJ, 151, 393

Schreiber C., et al., 2015, The Herschel view of the dominant mode of galaxy growth from z = 4 to the present day, A&A, 575, A74

Shakura N. I., Sunyaev R. A., 1973, Black holes in binary systems. Observational appearance., A&A, 24, 337

Sharma P., McCourt M., Quataert E., Parrish I. J., 2012, Thermal instability and the feedback regulation of hot haloes in clusters, groups and galaxies, MNRAS, 420, 3174

Sharma R. S., et al., 2021, The Connection between Mergers and AGN Activity in Simulated and Observed Massive Galaxies, arXiv e-prints, p. arXiv:2101.01729

Shimwell T. W., et al., 2017, The LOFAR Two-metre Sky Survey. I. Survey description and preliminary data release, A&A, 598

Shimwell T. W., et al., 2019, The LOFAR Two-metre Sky Survey. II. First data release, A&A, 622, A1

Shirley R., et al., 2021, HELP: The Herschel Extragalactic Legacy Project, arXiv e-prints, p. arXiv:2105.05659

Silva L., Maiolino R., Granato G. L., 2004, Connecting the cosmic infrared background to the X-ray background, MNRAS, 355, 973

Silva A., Marchesini D., Silverman J. D., Martis N., Iono D., Espada D., Skelton R., 2021, Galaxy Mergers up to z < 2.5. II. AGN Incidence in Merging Galaxies at Separations of 3-15 kpc, ApJ, 909, 124

Simpson C., Westoby P., Arumugam V., Ivison R., Hartley W., Almaini O., 2013, The prevalence of AGN feedback in massive galaxies at z ≈ 1, MNRAS, 433, 2647

Sirothia S. K., Dennefeld M., Saikia D. J., Dole H., Ricquebourg F., Roland J., 2009, 325-MHz observations of the ELAIS-N1 field using the Giant Metrewave Radio Telescope, MNRAS, 395, 269

Smith K. M., Zahn O., Doré O., 2007, Detection of gravitational lensing in the cosmic microwave background, Phys. Rev. D, 76, 043510

Smith D. J. B., et al., 2011, Herschel-ATLAS: counterparts from the ultraviolet-near-infrared in the science demonstration phase catalogue, MNRAS, 416, 857

Smith D. J. B., et al., 2016, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 271–280 (arXiv:1611.02706)

Smith D. J. B., et al., 2021, The LOFAR Two-metre Sky Survey Deep Fields. The star-formation rate-radio luminosity relation at low frequencies, A&A, 648, A6

Smolčić V., 2009, The Radio AGN Population Dichotomy: Green Valley Seyferts Versus Red Sequence Low-Excitation Active Galactic Nuclei, ApJ, 699, L43

Smolčić V., et al., 2009, Cosmic Evolution of Radio Selected Active Galactic Nuclei in the Cosmos Field, ApJ, 696, 24

Smolčić V., et al., 2017a, The VLA-COSMOS 3 GHz Large Project: Continuum data and source catalog release, A&A, 602, A1

Smolčić V., et al., 2017b, The VLA-COSMOS 3 GHz Large Project: Cosmic evolution of radio AGN and implications for radio-mode feedback since z 5, A&A, 602, A6

Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, Galaxy properties from the ultraviolet to the far-infrared: Λ cold dark matter models confront observations, MNRAS, 423, 1992

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014a, A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from z ∽0-6, ApJS, 214, 15

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014b, A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from z ∼0-6, ApJS, 214, 15

Springel V., et al., 2005, Simulations of the formation, evolution and clustering of galaxies and quasars, Nature, 435, 629

Springel V., Frenk C. S., White S. D. M., 2006, The large-scale structure of the Universe, Nature, 440, 1137

Stalevski M., Fritz J., Baes M., Nakos T., Popović L. Č., 2012, 3D radiative transfer modelling of the dusty tori around active galactic nuclei as a clumpy two-phase medium, MNRAS, 420, 2756

Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, The dust covering factor in active galactic nuclei, MNRAS, 458, 2288

Sturm E., et al., 2011, Massive Molecular Outflows and Negative Feedback in ULIRGs Observed by Herschel-PACS, ApJ, 733, L16

Suh H., et al., 2019, Multi-wavelength Properties of Type 1 and Type 2 AGN Host Galaxies in the Chandra-COSMOS Legacy Survey, ApJ, 872, 168

Sutherland W., Saunders W., 1992, On the likelihood ratio for source identification, MNRAS, 259, 413

Szalay A. S., Connolly A. J., Szokoly G. P., 1999, Simultaneous Multicolor Detection of Faint Galaxies in the Hubble Deep Field, AJ, 117, 68

Tadhunter C., 2016, Radio AGN in the local universe: unification, triggering and evolution, A&A Rev., 24, 10

Tasca L. A. M., et al., 2015, The evolving star formation rate: $M_\star$ relation and sSFR since z ≃ 5 from the VUDS spectroscopic survey, A&A, 581, A54

Tasse C., Best P. N., Röttgering H., Le Borgne D., 2008, Radio-loud AGN in the XMM-LSS field. II. A dichotomy in environment and accretion mode?, A&A, 490, 893

Tasse C., et al., 2018, Faceting for direction-dependent spectral deconvolution, A&A, 611, A87

Tasse C., et al., 2021, The LOFAR Two-meter Sky Survey: Deep Fields Data Release 1. I. Direction-dependent calibration and imaging, A&A, 648, A1

Thomas N., Davé R., Jarvis M. J., Anglés-Alcázar D., 2021, The radio galaxy population in the SIMBA simulations, MNRAS, 503, 3492

Tonry J. L., et al., 2012, The Pan-STARRS1 Photometric System, ApJ, 750, 99

Tremblay G. R., et al., 2016, Cold, clumpy accretion onto an active supermassive black hole, Nature, 534, 218

Trump J. R., et al., 2009, Observational Limits on Type 1 Active Galactic Nucleus Accretion Rate in COSMOS, ApJ, 700, 49

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, Galactic Winds, ARA&A, 43, 769

Veilleux S., et al., 2013, Fast Molecular Outflows in Luminous Galaxy Mergers: Evidence for Quasar Feedback from Herschel, ApJ, 776, 27

Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, Cool outflows in galaxies and their implications, A&A Rev., 28, 2

Vogelsberger M., et al., 2014, Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe, MNRAS, 444, 1518

Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, The cosmic evolution of massive black holes in the Horizon-AGN simulation, MNRAS, 460, 2979

Wang T., et al., 2019, A dominant population of optically invisible massive galaxies in the early Universe, Nature, 572, 211

Werner M. W., et al., 2004, The Spitzer Space Telescope Mission, ApJS, 154, 1

Whitaker K. E., et al., 2011, The NEWFIRM Medium-band Survey: Photometric Catalogs, Redshifts, and the Bimodal Color Distribution of Galaxies out to z ∼3, ApJ, 735, 86

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, The Star Formation Mass Sequence Out to z = 2.5, ApJ, 754, L29

White S. D. M., Frenk C. S., 1991, Galaxy formation through hierarchical clustering, ApJ, 379, 52

Whittam I. H., Prescott M., McAlpine K., Jarvis M. J., Heywood I., 2018, The Stripe 82 1-2 GHz Very Large Array Snapshot Survey: host galaxy properties and accretion rates of radio galaxies, MNRAS, 480, 358

Williams W. L., et al., 2016, LOFAR 150-MHz observations of the Boötes field: catalogue and source counts, MNRAS, 460, 2385

Williams W. L., et al., 2018, LOFAR-Boötes: properties of high- and low-excitation radio galaxies at 0.5 < z < 2.0, MNRAS, 475, 3429

Williams W. L., et al., 2019, The LOFAR Two-metre Sky Survey. III. First data release: Optical/infrared identifications and value-added catalogue, A&A, 622, A2

Willott C. J., Rawlings S., Blundell K. M., Lacy M., 1999a, The emission line-radio correlation for radio sources using the 7C Redshift Survey, MNRAS, 309, 1017

Willott C. J., Rawlings S., Blundell K. M., Lacy M., 1999b, The emission line-radio correlation for radio sources using the 7C Redshift Survey, MNRAS, 309, 1017

Wilman R. J., et al., 2008, A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes, MNRAS, 388, 1335

Wilson G., et al., 2009, Spectroscopic Confirmation of a Massive Red-Sequence-Selected Galaxy Cluster at z = 1.34 in the SpARCS-South Cluster Survey, ApJ, 698, 1943

Windhorst R., Mathis D., Neuschaefer L., 1990, in Kron R. G., ed., Astronomical Society of the Pacific Conference Series Vol. 10, Evolution of the Universe of Galaxies. pp 389–403

Windhorst R. A., Fomalont E. B., Partridge R. B., Lowenthal J. D., 1993, Microjansky Source Counts and Spectral Indices at 8.44 GHz, ApJ, 405, 498

Woo J.-H., Schulze A., Park D., Kang W.-R., Kim S. C., Riechers D. A., 2013, Do Quiescent and Active Galaxies Have Different M $_{BH}$-σ_* Relations?, ApJ, 772, 49

Yang G., et al., 2020, X-CIGALE: Fitting AGN/galaxy SEDs from X-ray to infrared, MNRAS, 491, 740

York D. G., et al., 2000a, The Sloan Digital Sky Survey: Technical Summary, AJ, 120, 1579

York D. G., et al., 2000b, The Sloan Digital Sky Survey: Technical Summary, AJ, 120, 1579

Yuan F., Narayan R., 2014, Hot Accretion Flows Around Black Holes, ARA&A, 52, 529

Yuan F., Yu Z., Ho L. C., 2009, Revisiting the "Fundamental Plane" of Black Hole Activity at Extremely Low Luminosities, ApJ, 703, 1034

da Cunha E., Charlot S., Elbaz D., 2008, A simple model to interpret the ultraviolet, optical and infrared emission from galaxies, MNRAS, 388, 1595

de Gasperin F., et al., 2019, Systematic effects in LOFAR data: A unified calibration strategy, A&A, 622, A5

de Ruiter H. R., Willis A. G., Arp H. C., 1977, A Westerbork 1415 MHz survey of background radio sources. II. Optical identifications with deep IIIa-J plates., Astronomy and Astrophysics Supplement Series, 28, 211

van Haarlem M. P., et al., 2013, LOFAR: The LOw-Frequency ARray, A&A, 556

van Weeren R. J., et al., 2016, LOFAR Facet Calibration, The Astrophysical Journal Supplement Series, 223