Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **550**, 1–20 (2026) https://doi.org/10.1093/mnras/stag1299
Advance Access publication 2026 July 11

The DESI view of the faint radio source population in LoTSS DR2

M. I. Arnaudova,[1,2★] L. R. Holden,[2] D. J. B. Smith,[2] P. N. Best,[1] R. Kondapally,[3,4] K. J. Duncan,[1] A. Bushi,[1] S. Das,[2] S. R. Flury,[1] C. L. Hale,[1] M. J. Hardcastle,[2] H. J. A. Röttgering[5] and S. Shenoy[2]

[1] *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[2] *Centre for Astrophysics Research, University of Hertfordshire, Hatfield AL10 9AB, UK*
[3] *Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK*
[4] *Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[5] *Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

Accepted 2026 July 7. Received 2026 July 2; in original form 2026 February 23

ABSTRACT

The faint radio-source population includes galaxies dominated by both star formation (SF) and active galactic nuclei (AGN), which are two key processes shaping galaxy evolution. To investigate this population, we probabilistically classified 251 413 radio sources from the second data release of the LOFAR Two-metre Sky Survey (LoTSS DR2) using spectroscopic data from the first release of the Dark Energy Spectroscopic Instrument (DESI DR1). Our classification method includes: (i) the identification of radio excess relative to SF, (ii) the Baldwin, Philips & Terlevich (BPT) diagram, (iii) a modified Mass–Excitation ($\mathcal{M}$Ex) diagram, and (iv) the [O III] λ5007 equivalent width. These are combined with Monte Carlo methods to estimate the probability that each source is a star-forming galaxy (SFG), a radio-quiet AGN (RQ AGN), or a low- or high-excitation radio galaxy (LERG or HERG), allowing various thresholds to be applied depending on science goals. Considering classifications above a 90 per cent probability threshold, we identify 68 820 SFGs, 32 288 RQ AGN, 35 210 LERGs, and 3085 HERGs, representing the largest radio sample to date with high-confidence spectroscopic classifications. Using this sample, we show with higher statistical power than previous studies that LERGs typically accrete below 1 per cent of the Eddington limit and HERGs above it. We also identify a small subset of high-accreting LERGs whose stacked spectra reveal prominent star-forming features, highlighting difficulties in interpreting their accretion properties. Our results demonstrate the power of large spectroscopic samples to characterize the radio-source population, providing a foundation for studies in the SKA era.

Key words: techniques: spectroscopic – catalogues – galaxies: active – galaxies: evolution – radio continuum: galaxies.

1 INTRODUCTION

Radio observations provide a dust-unbiased view of the energetic processes that shape galaxy evolution – from supernova-driven synchrotron and thermal (free–free) emission in star-forming galaxies (SFGs; J. J. Condon 1992) to the relativistic jets and outflows associated with active galactic nuclei (AGN; M. C. Begelman, R. D. Blandford & M. J. Rees 1984) – making radio surveys a key tracer of cosmic activity. Investigating the radio source population is therefore essential for understanding how galaxies and their supermassive black holes have co-evolved over cosmic time. However, disentangling the dominant source of radio emission in each source is a non-trivial task and requires additional information (e.g. W. L. Williams et al. 2019; R. Kondapally et al. 2021; M. J. Hardcastle et al. 2023).

Based on the dominant source of emission, radio sources can be divided into four classes: SFGs, whose emission is associated with stellar processes; radio-quiet AGN (RQ AGN), whose radio emission may be powered by star formation or by weak AGN-related processes such as small-scale jets or AGN-driven winds (e.g. F. Panessa et al. 2019; C. Macfarlane et al. 2021; L. K. Morabito et al. 2022; J. W. Petley et al. 2022; B.-H. Yue et al. 2024 S. V. White et al. 2025); and radio-loud AGN, which possess powerful radio jets and can be further divided into high-excitation radio galaxies (HERGs) and low-excitation radio galaxies (LERGs), depending on their accretion properties (for a review, see e.g. T. M. Heckman & P. N. Best 2014; M. J. Hardcastle & J. H. Croston 2020). In this scheme, radio-quiet AGN and HERGs are radiatively efficient systems that are believed to accrete material at relatively high rates (typically above 1 per cent of the Eddington limit, $\lambda_{\rm Edd} \gtrsim 0.01$) through a geometrically thin, optically thick accretion disc (e.g. N. I. Shakura & R. A. Sunyaev 1973), in which matter is efficiently converted into radiation across the electromagnetic spectrum. In contrast, LERGs are radiatively inefficient AGN thought to accrete at lower accretion rates ($\lambda_{\rm Edd} \lesssim 0.01$) via a geometrically thick, optically thin flow (R. Narayan & I. Yi 1994,

★ E-mail: m.i.arnaudova@gmail.com

1995), which does not efficiently convert matter into radiation. As a result, they show little or no radiative AGN signatures at other wavelengths but appear highly efficient at producing relativistic radio jets.

In recent years, distinguishing between these populations observationally has become increasingly feasible thanks to the ability to cross-match radio surveys with multiwavelength imaging and spectroscopic surveys (e.g. P. N. Best et al. 2005; P. N. Best & T. M. Heckman 2012; G. Gürkan, M. J. Hardcastle & M. J. Jarvis 2014; M. B. Pracy et al. 2016; V. Smolčić et al. 2017; I. H. Whittam et al. 2022; P. N. Best et al. 2023; S. Das et al. 2024; A. B. Drake et al. 2024; M. J. Hardcastle et al. 2025; M. I. Arnaudova et al. 2025). New-generation radio interferometers, such as the Low-Frequency Array (LOFAR; M. P. van Haarlem et al. 2013), have also hugely increased the sample size of radio sources that can be studied. In particular, the LOFAR Two-metre Sky Survey (LoTSS) observes the faint radio sky at 144 MHz with high sensitivity and resolution, enabling the identification of large samples of powerful radio sources across a wide area (LoTSS-wide; T. W. Shimwell et al. 2017, 2019, 2022, 2026), as well as fainter populations in deeper fields (LoTSS Deep fields; C. Tasse et al. 2021; J. Sabater et al. 2021; T. W. Shimwell et al. 2025).

The LoTSS Deep fields observations allowed P. N. Best et al. (2023) to employ multiwavelength photometry and four independent spectral energy distribution (SED) fitting codes (MAGPHYS ; E. da Cunha et al. 2015, BAGPIPES ; A. C. Carnall et al. 2018, CIGALE ; S. Noll et al. 2009; M. Boquien et al. 2019 and AGNFITTER ; G. Calistro Rivera et al. 2016) to produce maximum likelihood classifications of $\approx 80\,000$ sources in the LoTSS Deep Fields Data Release 1. Building upon this, S. Das et al. (2024) showed that similar results could be achieved using a single SED-fitting code (PROSPECTOR ; J. Leja et al. 2018; B. D. Johnson et al. 2021), significantly simplifying the framework and reducing the computational resources required. However, these photometric classifications rely on model assumptions that are subject to degeneracies between stellar age, dust content, and metallicity, and often on photometric redshift estimates, which can limit the accurate identification of AGN, particularly when the galaxy's overall SED is dominated by star formation. Using spectroscopic information from the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration 2016a, b) for a sub-sample of sources photometrically classified by P. N. Best et al. (2023) and S. Das et al. (2024), M. I. Arnaudova et al. (2025, hereafter A25) demonstrated that the two photometric classifications miss a substantial fraction of radiatively efficient AGN. Specifically, they found that the photometric approach underestimates the number of RQ AGN by a factor of ~3–5 compared to spectroscopic identifications at >90 per cent confidence. Furthermore, when studying the Eddington-scaled accretion rates of LERGs and HERGs, A25 showed that photometric classifications obscure the distinction between the two accretion modes, whereas spectroscopic diagnostics recover two clearly separated distributions, as had been suggested by e.g. P. N. Best & T. M. Heckman (2012) using bright radio samples at low redshifts ($z < 0.3$). Spectroscopy therefore remains the preferred method for distinguishing between these classes and can be applied in wider fields with limited multiwavelength data, such as the LoTSS-wide area, which contains a substantially larger number of sources than the deep fields. The LoTSS second data release (LoTSS DR2) is currently the largest low-frequency radio catalogue with optical counterparts, comprising over 4.1 million sources across 5700 deg^2 (see M. J. Hardcastle et al. 2023 for details).

A. B. Drake et al. (2024, hereafter Dr24) presented a probabilistic spectroscopic classification scheme for the subset of LoTSS DR2 sources with spectra from the Sloan Digital Sky Survey (SDSS), classifying over 150 000 sources up to $z \sim 0.6$. This framework combined a radio-excess diagnostic, based on the relation between the dust-corrected Hα and radio luminosity, with the BPT-N II diagram (J. A. Baldwin, M. M. Phillips & R. Terlevich 1981), and employed Monte Carlo methods to determine the probability of each source being an SFG, RQ AGN, LERG, or HERG. The authors validated their method by comparing their results with independent emission-line methods (e.g. the Excitation Index; P. N. Best & T. M. Heckman 2012), *WISE* selection techniques (e.g. G. Gürkan et al. 2014), and physical properties found in the literature (e.g. G. Gürkan et al. 2018; D. J. B. Smith et al. 2021; S. Das et al. 2024).

With the advent of DESI, this approach can be extended to a larger and deeper spectroscopic sample, probing a radio source population approximately twice as optically faint as that reached by SDSS. Moreover, by adopting the method introduced by A25, in which a modified Mass–Excitation ($\mathcal{M}$Ex) diagram is used in place of the BPT at higher redshifts, we are able to extend this classification scheme up to $z \sim 1$. Furthermore, including the equivalent width of [O III] λ5007 diagnostic (e.g. R. A. Laing et al. 1994; C. N. Tadhunter et al. 1998; P. N. Best & T. M. Heckman 2012) allows classification of many sources left unclassified by Dr24. The significantly larger statistical sample compared to A25, together with the broader area covered, which enables the inclusion of rarer, more luminous jet populations, provides the opportunity to build upon their study of LERGs and HERGs and to further investigate their accretion-rate distributions.

This paper is structured as follows. Section 2 describes the LoTSS and DESI data sets and our sample selection. Section 3 details the spectral fitting and emission-line measurements. Section 4 presents the probabilistic classification framework. Section 5 presents a comparison with the SDSS classification scheme from Dr24, and discusses the selection effects introduced by using DESI spectroscopy. Section 6 focuses on the Eddington-scaled accretion rates of the radio-loud population, while Section 7 summarizes our results. Throughout, we adopt a flat ΛCDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm M} = 0.3$, and $\Omega_\Lambda = 0.7$.

2 DATA

2.1 LoTSS DR2-wide

The radio data used in this work are taken from LoTSS DR2 (T. W. Shimwell et al. 2022). This data set consists of observations centred at 144 MHz with a resolution of 6 arcsec (using the Dutch stations only), covering $\sim$ 5700 deg^2 of the northern sky (see the black outline of the footprint in the left panel of Fig. 1) with a median rms sensitivity of 83 μJy beam^{-1}. Using this data release, M. J. Hardcastle et al. (2023) employed the likelihood ratio method (W. Sutherland & W. Saunders 1992; D. J. B. Smith et al. 2011), complemented by visual classifications from the LOFAR Galaxy Zoo[1] to account for extended and complex radio morphologies (following the approach of W. L. Williams et al. 2019), providing reliable optical identifications for 3.5 million sources (about 85 per cent of the total sample). Approximately 58 per cent

[1] http://lofargalaxyzoo.nl/

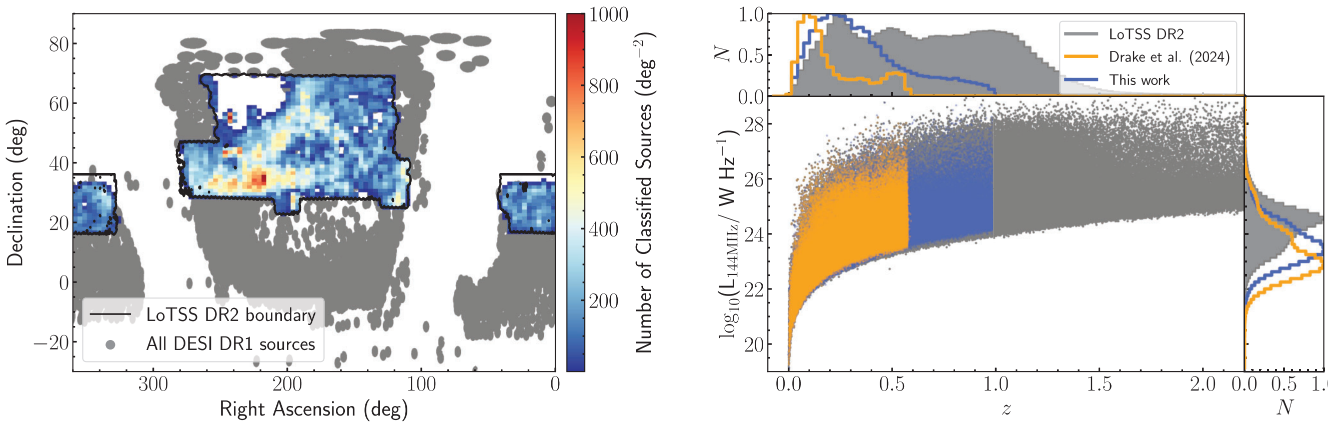

Figure 1. Left panel: sky coverage of DESI DR1 and LoTSS DR2. The grey region indicates all DESI DR1 sources, the black lines show the LoTSS DR2 footprint, while the colour scale represents the number density of sources in our classification sample. Right panel: the redshift-radio rest-frame luminosity plane for all radio sources with a multiwavelength counterpart (in grey), the sources with spectroscopic information from DESI DR1 (blue), and those that are part of the SDSS sample from Dr24 (orange), along with their respective luminosity and redshift distributions (normalized to the maximum bin count in each sample). For reference the total number of sources (N) in each sample is $N_{\rm LoTSS} = 4\,167\,359$ (of which 2 392 639 have a photometric or spectroscopic redshift), $N_{\rm DESI} = 279\,700$, and $N_{\rm SDSS} = 152\,355$.

of the LoTSS DR2 sources have robust redshift estimates, including photometric redshifts from K. J. Duncan (2022), as well as spectroscopic redshifts from SDSS DR16 (R. Ahumada et al. 2020) and DESI DR1 (DESI Collaboration 2025), and, as a result, they serve as the starting point for our classification sample.

2.2 DESI DR1

The spectroscopic data used in this work are taken from DESI DR1 (DESI Collaboration 2025). This data set is currently the largest extragalactic data set with spectroscopic redshifts, consisting of a total of 18.7 million objects across 14 000 deg^2 area (shown in grey in Fig. 1). The targeting algorithm for these objects is associated with optical and mid-infrared photometry from the Legacy Imaging Surveys (A. Dey et al. 2019) and the *Wide-field Infrared Survey Explorer* (*WISE*; E. L. Wright et al. 2010), respectively, and is broadly divided into the following categories: the Milky Way survey (MWS; A. P. Cooper et al. 2023), Bright Galaxy Survey (BGS) galaxies ($0 < z < 0.6$; O. Ruiz-Macias et al. 2020; C. Hahn et al. 2023); luminous red galaxies (LRGs, $0.4 < z < 1.1$; R. Zhou et al. 2020, 2023); emission-line galaxies (ELGs, $0.6 < z < 1.6$; A. Raichoor et al. 2020, 2023); quasi-stellar objects (QSOs; $0.9 < z < 4$; C. Yèche et al. 2020; E. Chaussidon et al. 2023), as well as a range of secondary targets (SCND; see e.g. A. D. Myers et al. 2023 for details). The spectra were taken using 10 identical spectrographs mounted on the 4 m Mayall Telescope at Kitt Peak National Observatory, each equipped with 500 fibres with a 1.5 arcsec entrance diameter, covering the wavelength range of 3600–9824 Å at a resolving power of $\Delta\lambda/\lambda \approx 2000{-}5500$ (J. Guy et al. 2023). Together with the LoTSS DR2 optical identifications, these high-quality DESI spectra allow us to construct a large sample of radio sources suitable for robust classification into SFGs, RQ AGN, LERGs, and HERGs.

2.3 Sample selection

Starting with the LoTSS DR2 optical identifications from M. J. Hardcastle et al. (2023), we perform a positional cross-match with the DESI DR1 catalogue using a maximum search radius of 1.5 arcsec and the recommended quality flags. Specifically, we select unique spectra (ZCAT_PRIMARY == True) corresponding to non-sky targets (OBJTYPE == 'TGT') with reliable redshift measurements (ZWARN == 0). This results in a total of 366 122 radio sources with available spectroscopic information, of which 363 326 are extragalactic (SPEC_TYPE ! = 'STAR'). Next, we consider the aperture corrections provided by the `FastSpecFit` Spectral Synthesis and Emission-Line Catalogue (J. Moustakas et al. 2023; Moustakas et al., in preparation), one of DESI's value-added catalogues, which we use to correct flux losses in the emission-line measurements required for our classification (see Section 4). We note that `FastSpecFit` uses updated redshifts from `QuasarNET`[2] for QSO targets, while the value added catalogue from M. Siudek et al. (2024), which we later use for the classifications (see Section 4), and the redshifts used for the spectral fitting are estimated with `redrock`.[3] Therefore, we remove 1285 sources which have $\Delta z > 0.001$. As in A25, we exclude sources requiring an aperture correction factor greater than 5, as well as sources having an aperture correction less than 1, resulting in a sample of 307 732 radio sources. While A25 applied aperture corrections based on the total-to-fibre r-band flux ratio, we find that this correction systematically overestimates the Hα-derived SFRs (by a median offset of ~ 0.3 dex) relative to those in the MPA-JHU catalogue (G. Kauffmann et al. 2003a; J. Brinchmann et al. 2004; C. A. Tremonti et al. 2004). In contrast, the `FastSpecFit` aperture corrections yield much better agreement (with a median offset of only 0.06 dex), and therefore we adopt them in this work.

Finally, since our classification scheme (see Section 4) relies on the availability of specific emission lines, we restrict our sample to $z < 0.947$ (hereafter the parent sample) and consider the following redshift ranges:

(i) $0 < z < 0.483$ (hereafter the low-z sample), which includes 207 657 sources with spectral data covering the Hα and [N II] λ 6583 emission lines, given the wavelength coverage of the DESI

[2]https://github.com/ngbusca/QuasarNET
[3]https://github.com/desihub/redrock

spectrographs, necessary for the use of the radio diagnostic and the BPT-N II diagram (J. A. Baldwin et al. 1981).

(ii) $0.483 < z < 0.947$ (hereafter the high-z sample), which includes 72 043 sources with spectral data covering the Hβ and [O III] λ5007 emission lines, necessary to use the modified Mass–Excitation diagram (hereafter $\mathcal{M}$Ex; A25).

The source density of these sources and their distribution in the redshift-radio rest-frame luminosity ($L_{144\,\mathrm{MHz}}$ [4]) plane are presented in Fig. 1. In the right panel, we can see that we are probing a parameter space which differs from that available to Dr24; due to DESI's selection criteria, which rely on the Legacy Surveys which reach approximately 1.5 mag deeper than SDSS in the g, r, and z bands, we have spectroscopic measurements for about twice as many radio sources at $z < 0.6$ (corresponding to the redshift range of Dr24). Furthermore, as a result of the updates to the classification scheme introduced in A25, we extend the redshift range studied up to $z \sim 1$. We note that the impact of selection effects, both relative to Dr24 and in general, is discussed in Section 5. While this parameter space still differs from that of the full LoTSS DR2 catalogue, the WEAVE-LOFAR survey (D. J. B. Smith et al. 2016), which will select targets solely based on their 144 MHz flux density, will provide a more uniform and complete coverage of the full radio-source population.

3 SPECTRAL FITTING

For this work and in preparation for the WEAVE-LOFAR survey, we have developed a new spectral-fitting code, `WL-SLAYER`: the WEAVE-LOFAR Spectral Line Analyser (see M. I. Arnaudova et al. 2024b, L. R. Holden & C. N. Tadhunter 2025, A25, and L. R. Holden et al. 2026 for predecessors), in order to obtain robust emission line measurements and realistic uncertainties needed for the classification scheme (see Section 4). This new code is particularly necessary for radio sources, as it is designed to model complex emission line profiles that are commonly observed in these objects (e.g. due to AGN-driven winds or outflows; S. J. Molyneux, C. M. Harrison & M. E. Jarvis 2019; E. L. Escott et al. 2025). In this section, we give details of the approach and describe how we used it to produce line-flux measurements and various quality-assurance flags that we make use of in our analysis.

3.1 Stellar continuum modelling and subtraction

To isolate the flux from the emission lines, `WL-SLAYER` models and subtracts the stellar continuum using the `pPXF` code (M. Cappellari & E. Emsellem 2004; M. Cappellari 2017, 2023). To do this, it first corrects the spectra for Galactic extinction using the $R_v = 3.1$ E. L. Fitzpatrick (1999) extinction law (as implemented in the PYTHON EXTINCTION package) with the dust maps from D. J. Schlegel, D. P. Finkbeiner & M. Davis (1998), and uses the redshift values provided in the DESI DR1 catalogue to shift the spectra to the rest frame. Next, it masks regions spanning $\pm 750\,\mathrm{km\,s^{-1}}$ around common emission lines, ensuring that even broad Balmer lines do not affect the fit. Finally, it models the full wavelength range using a linear combination of G. Bruzual & S. Charlot (2003) stellar templates (BC03), along with a fourth-order polynomial to correct low-frequency continuum variations, such as flux-calibration errors or dust extinction of the stellar continuum. The fits are performed in log-wavelength space, after which the resulting stellar-continuum model is interpolated onto a linear wavelength grid and subtracted from the spectrum, leaving only the emission-line flux for subsequent analysis. We note that this continuum fitting approach does not include AGN templates as the majority of the sample is expected to consist of star-forming galaxies and type 2 AGN whose continua are dominated by stellar emission. Type 1 AGN are identified and excluded from the final sample (see Section 3.4) and will be explored in future work (Bushi et al., in preparation).

3.2 Emission-line fitting

For the emission line modelling in this work, we focus only on the Hβ, [O III]$\lambda\lambda$4959, 5007, Hα, and [N II]$\lambda\lambda$6548, 6583 emission lines (for the high-z sample we model only Hβ and [O III]$\lambda\lambda$4959, 5007 as the rest fall outside the observing window). However, we note that `WL-SLAYER` is designed to work with any user-defined list of lines (e.g. Holden et al., in preparation). Here, we implement MCMC sampling using the Affine Invariant Ensemble provided by the EMCEE PYTHON package (D. Foreman-Mackey et al. 2013), with the differential evolution Markov chain (DE-MC) algorithm for updating the walker positions (C. ter Braak & J. Vrugt 2008). The DE-MC is used to enable efficient sampling of the high-dimensional parameter spaces that we are considering when fitting multiple emission line components simultaneously. Our emission-line model consists of a number of Gaussian components (N_g) for each emission line and is fitted in velocity space, with each Gaussian component having the same velocity shift and width for all lines but a variable peak flux (i.e. if a broad component is required for the Balmer lines, the corresponding forbidden-line component would have an amplitude of zero).

To ensure that the MCMC walkers quickly converged to regions of maximum probability, our routine first performs a least-squares fit of the emission-line model using the SCIPY PYTHON library (P. Virtanen et al. 2020), the results of which we use as initial positions for the walkers in parameter space (with a small scatter). Physically motivated lower and upper bounds were placed on the parameters of the least-squares fit: the velocity widths of the Gaussian components were required to be positive, and constraints on the relative fluxes in line doublets were set based on atomic physics ([O III](5007/4959) = 2.99; [N II](6583/6548) = 2.92: D. E. Osterbrock & G. J. Ferland 2006). The peak flux is allowed to be negative to account for the presence of absorption lines, as well as to avoid a 'positive bias' effect in the presence of noise. The same constraints were used as priors for the MCMC sampling itself, with the addition of Gaussian priors on velocity shift ($\mu = 0\,\mathrm{km\,s^{-1}}$; $\sigma = 400\,\mathrm{km\,s^{-1}}$) and width ($\mu = 200\,\mathrm{km\,s^{-1}}$; $\sigma = 50\,\mathrm{km\,s^{-1}}$) to ensure that walker positions were weighted against physically unlikely extreme values. The values of the Gaussian priors were selected as a result of testing the fitting routine on DESI spectra presenting emission-lines of different types (e.g. weak lines, strong narrow lines, Type 1 AGN), and result in walkers converging on reasonable velocity values (and hence avoiding anomalously large flux values arising from fitting very broad low-peak components) when there is little (or zero) emission line flux while still being able to converge on

[4]The $L_{144\,\mathrm{MHz}}$ is calculated from the integrated 144 MHz flux density ($S_{144\,\mathrm{MHz}}$; M. J. Hardcastle et al. 2023), assuming a spectral index of $\alpha = -0.7$ ($S_\nu \propto \nu^\alpha$). Redshifts are taken from DESI DR1 for the sample analysed in this work, from the Dr24 catalogue for the SDSS sample, and from the spectroscopic and photometric redshifts compiled in M. J. Hardcastle et al. (2023) for the full LoTSS sample.

large velocity widths/shifts in the case of genuinely kinematically extreme emission lines.

The log-likelihood function used in our MCMC routine was

$$\ln \mathcal{L} = -\frac{1}{2}\sum_{i=1}^{k}\left(\frac{F_{\lambda,\mathrm{i}} - F^{\mathrm{m}}_{\lambda,\mathrm{i}}}{F^{\mathrm{err}}_{\lambda,\mathrm{i}}}\right)^2, \quad (1)$$

where, at the velocity step i (of maximum steps k), $F_{\lambda,\mathrm{i}}$ is the flux density at wavelength λ in a given DESI spectrum, $F^{\mathrm{m}}_{\lambda,\mathrm{i}}$ is the modelled flux density, and $F^{\mathrm{err}}_{\lambda,\mathrm{i}}$ is the flux-density uncertainty for the DESI spectrum. In order to ensure that the MCMC walkers had converged on the region of maximum likelihood and therefore were effectively sampling the posterior distribution, we calculated the autocorrelation times (τ) from the latter half of the walker chain for every parameter at 1000 step intervals, and stopped sampling when all autocorrelation times were less than 5 per cent of the current number of steps.

Another feature of `WL-SLAYER` is that it uses an iteratively increasing number of Gaussian components to model emission lines. We chose the first iteration to use a single Gaussian component for each emission line ($N_g = 1$): MCMC sampling for this model is performed as described earlier in this section. Once the walkers have converged, a fit with $N_g + 1$ components is performed, and the Bayesian Information Criterion (BIC) is calculated following:

$$\mathrm{BIC} = k\ln(n) - 2\ln(\hat{L}), \quad (2)$$

where k is the number of parameters (dimensions) of the model of a given iteration, n is the number of flux (or wavelength) data points from a given spectrum that is considered in the emission-line model, and $\ln\hat{L}$ is the maximized value of the log-likelihood function (equation 1). If the decrease in the value of the BIC was greater than six (following the recommendation of A. E. Raftery 1995), we accepted the more complex model and continued with further iterations until this criterion was not met. With this approach, we are able to simultaneously fit the profiles of different emission lines with a statistically justified number of Gaussian profiles, preventing under- and overfitting.

An example output from this routine is shown in Fig. 2, where a randomly selected galaxy spectrum (solid grey line) is displayed alongside its best-fit model (solid black line). The corresponding stellar-continuum (solid orange line) and emission-line components (dash–dotted blue line) are also presented.

3.3 Emission-line parameter calculation

In order to produce measurements of key emission-line parameters, including the total line flux, equivalent width (EW), and full-width-half maximum (FWHM), we directly used the last 1000 steps of the flattened MCMC chains. For each of these steps, we generated a total line profile (consisting of the final number of Gaussians determined using the BIC) using the parameter values extracted from their respective flattened chains. The line flux for a given realization was determined by calculating the total area of the line profile, which would be positive for a well-detected emission line, close to zero for a non-detection, and negative for an absorption line. We then used all the last 1000 realizations for a given emission line to produce a line-flux distribution, from which we took the 50th percentile as the line-flux value and estimated 1σ uncertainties using the 16th and 84th percentiles. A similar procedure was done for the EW and FWHM measurements.

3.4 Identifying Type 1 AGN

In addition to the emission-line parameters for our catalogue, we have defined a quality-control flag that identifies likely Type 1 AGN. This is useful for excluding sources in which a significant fraction of the recombination-line flux (e.g. Hα, Hβ) does not originate from the same region as the forbidden lines, leading to contamination of crucial flux ratios that are used for source classification.

For each realization drawn from the MCMC chains, we measured the FWHM of the total line profiles for Hβ and [O III]λ5007. We then identified an object as a likely Type 1 AGN if the FWHM of Hβ was greater than that of [O III] in 99 per cent (or more) of the realizations and the FWHM of Hβ was greater than 1000 km s^{-1} – these objects are flagged with FLAG_TYPE1 = TRUE in our catalogue. Using this flag, we identify 2664 likely Type 1 AGN in our sample. In Fig. 3, we show the Hβ and [O III]λ5007 FWHM distributions of the likely Type 1 objects (solid blue contours) that we identified with this flag, alongside the non-Type 1 objects (dash–dotted red contours) that we use in the classification scheme.

3.5 Evaluating goodness of fit

To assess the reliability of the spectral fits, we defined a quality-control flag that evaluates the performance of the fitting routine in the regions around the emission lines.[5] Following the approach of A25 (also used in e.g. D. J. B. Smith et al. 2012), we modelled the distribution of the resulting chi-squared (χ^2) values (evaluated at $\pm$4000 km s^{-1} around the Balmer lines and $\pm$2000 km s^{-1} for the forbidden lines) with a χ2 distribution separately for the low-z and high-z samples (independent of the number of Gaussian components used). We then identified the 99th percentile of the best-fitted χ2 distributions, above which spectra are classified as having 'bad' fits and are excluded. Applying these thresholds results in a retained sample (hereafter the classification sample) of 251 413 sources (this includes $\sim$ 90 per cent of sources for the low-z and 93 per cent for the high-z sample). In our catalogue, these sources can be chosen with FLAG_SPEC = TRUE. All processing steps leading to the final classification sample are summarized in Table 1.

4 CLASSIFICATIONS

To classify sources as SFGs, RQ AGN, LERGs, and HERGs, we adopt a probabilistic spectroscopic classification method similar to A25. This method uses a combination of a radio-excess diagnostic, the BPT-N II (hereafter BPT), a modified Mass–Excitation ($\mathcal{M}$Ex) diagram and the [O III] λ5007 EW, which we discuss in more detail in the following sections, alongside Monte Carlo realizations to estimate the probability for a source to belong in each of the four categories.

4.1 Identifying radio-excess galaxies

Following A25, radio-excess galaxies (RXGs) are identified in this work using the $L^{\mathrm{corr}}_{\mathrm{H}\alpha}$–$L_{144\,\mathrm{MHz}}$ diagnostic, where $L^{\mathrm{corr}}_{\mathrm{H}\alpha}$ is the dust-corrected Hα luminosity. This relation is a proxy for the linear

[5]We note that while `WL-SLAYER` fits the continuum globally and subtracts it across the full spectrum, the quality of the fit in the emission-line regions is what is critical for this analysis.

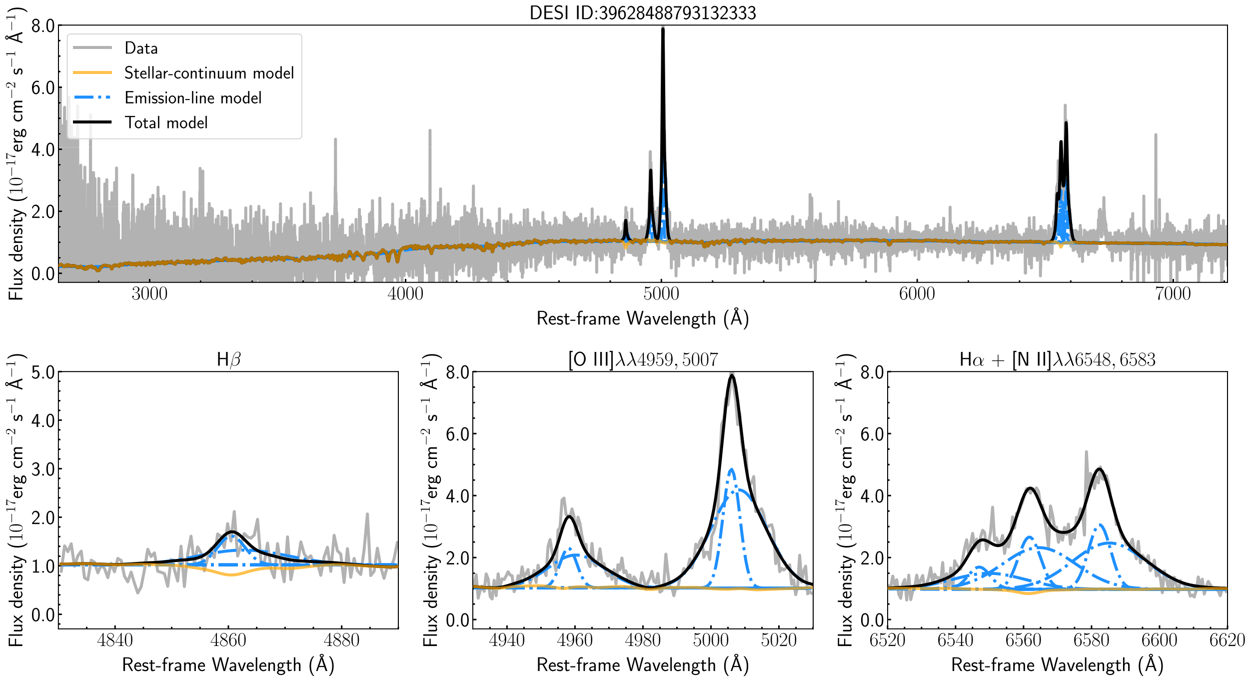

Figure 2. An example fit of a random galaxy spectrum using `WL-SLAYER`. The observed data are shown in grey, the continuum model in orange, the emission-line model in blue, and the total best-fitting model in black. The top panel displays the fit to the full spectrum, which is used for continuum subtraction, while the bottom panels focus on the Hβ, [O III] $\lambda\lambda$4959, 5007, and Hα + [N II] complexes to highlight the emission-line modelling.

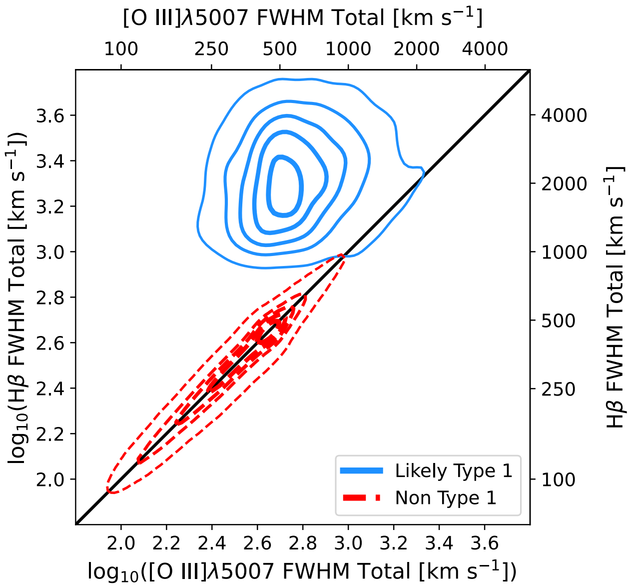

Figure 3. FWHM values for the total [O III]λ5007 and Hβ line profiles measured by `WL-SLAYER`. The probability density function for the objects in our sample that are identified as likely Type 1 AGN by our FLAG_TYPE1 flag is shown as solid blue contours, while objects not flagged as Type 1 AGN are shown as dashed red contours. Contours contain 10, 30, 50, 70, and 90 per cent of the total objects for each flag value.

correlation between star formation rate (SFR) and radio luminosity found for star-forming galaxies (albeit with indication of some stellar mass dependence as found in e.g. G. Gürkan et al. 2018; D. J. B. Smith et al. 2021; I. Delvecchio et al. 2021; S. Das et al. 2024; S. Shenoy et al. 2026) such that sources offset from this relation towards higher radio luminosities are considered to have radio emission due to AGN-related processes such as jets. However, as discussed in A25, the precise location of this demarcation line depends on factors such as aperture corrections. Since our approach to aperture corrections differs from that of A25, we need to redefine the threshold: we adopt the approach of Dr24 (also used in A25), which estimates the peak of the $\log_{10} L_{144\,\mathrm{MHz}}$–$\log_{10} L^{\mathrm{corr}}_{\mathrm{H}\alpha}$ distribution for all sources with a 5σ detection in Hβ, Hα and radio flux using a kernel density estimation (KDE). The left-hand side of this KDE peak (where it is assumed that the SFGs dominate) is then modelled with a Gaussian distribution, and the 99th percentile of the symmetrized Gaussian is taken as the demarcation threshold. Therefore, RXGs identified in this work are sources for which:

$$\log_{10}(L_{144\,\mathrm{MHz}}/\mathrm{WHz}^{-1}) > \log_{10}(L^{\mathrm{corr}}_{\mathrm{H}\alpha}/\mathrm{L}_{\odot}) + 15.00(\pm 0.02), \quad (3)$$

where 15.00 corresponds to the 99th percentile of the SFG Gaussian distribution. We note that re-deriving the RX threshold using only sources classified as SFGs according to the BPT diagram

Table 1. A summary of the processing steps applied to the cross-matched sample between the catalogue from M. J. Hardcastle et al. (2023) and DESI DR1.

Processing step	Sources removed	Sources remaining
LoTSS DR2 matches with DESI	–	398 399
DESI quality flags	35 073	363 326
Consistent zs	1285	362 041
Aperture corrections	54 309	307 732
Sources with $z \leq 0.947$	28 032	279 700
Excluding Type 1 AGN	2664	277 036
Good spectral fits	25 623	251 413

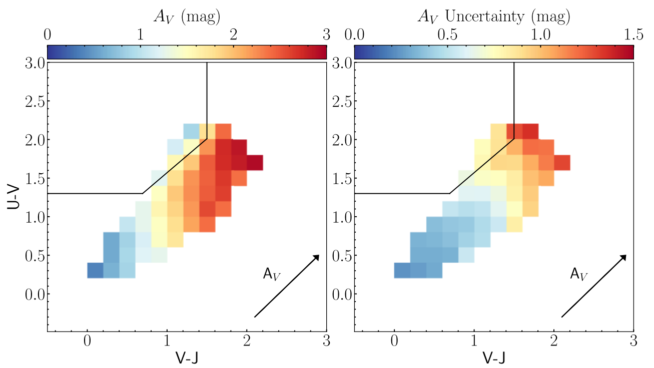

Figure 4. The rest-frame $U - V$ versus $V - J$ (the UVJ diagram) for all sources in our sample that satisfy the CIGALE SED-fitting goodness-of-fit criteria and have a 5σ detection in Hβ and Hα, colour-coded based on the 50th percentile of the Balmer-decrement-derived A_V distribution for that bin (left), as well as the spread in those values as estimated by half the difference between the 16th and 84th percentiles (right). The solid lines show the adopted divisions between star-forming and quiescent galaxies from A. Muzzin et al. (2013).

(see Section 4.2) yields a value of 14.92 ± 0.01. However, as non-radio excess sources dominate our sample (which is the case irrespective of the demarcation line used), we choose to retain equation (3) as it provides a more conservative criterion for identifying radio-excess sources and is consistent with previous works (e.g. Dr24, A25).

This approach is straightforward for the low-z sample where Hα lies within the DESI spectral coverage allowing us to directly measure $L_{\mathrm{H}\alpha}$ and calculate the extinction of the Hα line ($A_{\mathrm{H}\alpha}$) using the Balmer decrement (see A25 for details). Beyond $z > 0.483$, applying this method becomes more challenging. Although Hβ remains within the DESI spectral coverage, the lack of consistent high-quality multiwavelength photometry across the full LoTSS DR2-wide area limits our ability to derive robust extinction values from SED fitting, as was done by A25. However, as part of DESI DR1, several value-added catalogues (VACs) are available, including the AGN Host Galaxies Physical Properties VAC (M. Siudek et al. 2024), which provides rest-frame $U - V$ and $V - J$ colours for all galaxies derived with the CIGALE SED fitting code. This enables us to use the UVJ diagram (S. Wuyts et al. 2007; R. J. Williams et al. 2009; A. Muzzin et al. 2013), commonly used to distinguish between dusty star-forming and old passively evolved galaxies, as a way to infer the extinction for the high-z sample.

We achieve this by binning the Balmer-derived V-band extinction (A_V) values[6] of the low-z sample (with 5σ detections in Hβ and Hα) in the rest-frame $U - V$ and $V - J$ colour space using 0.2 dex bins. For each bin containing at least 50 objects, we compute the 16th, 50th, and 84th percentiles of the A_V distribution; these are shown in Fig. 4. We then assign to each source in the high-z sample an A_V corresponding to the median (50th percentile) of the bin, with an uncertainty estimated as half the difference between the 84th and 16th percentiles. If a source falls outside these bins (affecting only ~ 10 per cent of sources), we adopt the values from the nearest bin. This approach works well, as several works (e.g. N. S. Martis et al. 2016; J. J. Fang et al. 2018) have found that the A_V increases diagonally, except for sources located in the quiescent region (as denoted by the region marked by the solid black line in Fig. 4) which are associated with low-A_V values, and that this trend does not evolve significantly over the redshift range considered in this work. To further validate this approach, we compare the A_V values assigned to the high-z sample with Balmer-derived A_V estimates obtained from Hγ/Hβ measurements in stacked high-z spectra within each UVJ bin (see Section 6.2 for details of the stacking procedure). The stacked spectra are fitted using WL-SLAYER in a similar setup as discussed in Section 3. We qualitatively recover similar trends; however, the combination of the relatively low signal-to-noise ratio (S/N) of the Balmer lines and the increased difficulty of accurately modelling the stellar continuum in the stacked spectra limits a more detailed quantitative assessment, which we defer to future work. Nonetheless, this comparison provides empirical support for applying the low-z Balmer-derived A_V assignment method to capture population-level dust trends across UVJ space. This technique then allows us to infer dust-corrected Hα luminosities for the high-z sample as in A25 and, in turn, apply equation (3) to identify sources exhibiting a radio excess. As discussed in Section 4.4, this technique also provides a route to estimating the extinction for lower z sources when the Balmer lines are too weak to allow for the Balmer decrement approach to be used.

We note that our radio-excess diagnostic shows a moderate level of disagreement with the absolute $W3$ magnitude-$L_{144\,\mathrm{MHz}}$ diagnostic of M. J. Hardcastle et al. (2025). Specifically, $\sim$10 per cent of sources classified as radio excess using our $L^{\mathrm{corr}}_{\mathrm{H}\alpha}$–$L_{144\,\mathrm{MHz}}$ diagnostic lie below the radio-excess region defined by the $W3$–$L_{144\,\mathrm{MHz}}$ diagnostic of M. J. Hardcastle et al. (2025), which was developed using the probabilistic spectroscopic classifications of Dr24. Whilst of course some level of disagreement is expected between different radio excess diagnostics (as indeed there is not perfect agreement between samples of Seyfert galaxies obtained using the BPT and modified $\mathcal{M}$Ex methods, for example), further investigation reveals that the 'discrepant' sources are typically bright in $W3$ (observed $W3 < 17$ mag) and exhibit red observed $W2 - W3$ (Vega) colours ($W2 - W3 > 2$), which are largely absent from the sample used by Dr24. Taken together, we conclude that differences between the DESI and SDSS spectroscopic selection criteria, as well as the complexity of the mid-infrared SED (which is not fully captured by the power-law extrapolation used to derive the absolute $W3$ magnitude; see M. J. Hardcastle et al. 2025 for details) are playing a role here. We defer a fuller investigation to future work, combining the strengths of both photometric and spectroscopic classifications (e.g. Das et al., in preparation).

4.2 Identifying radiatively efficient AGN

To identify radiatively efficient AGN, we first use the well-known BPT diagnostic diagram (J. A. Baldwin et al. 1981), which uses the emission-line ratios [O III] λ5007/Hβ and [N II] λ6583/Hα to distinguish galaxies according to their dominant ionization mechanism. Following A25, this diagram identifies radiatively efficient AGN among sources without a radio excess (see Sec-

[6] This is done by calculating the colour excess as $\mathrm{E(B\text{-}V)} = 1.97 \log_{10}\left[\frac{(\mathrm{H}\alpha/\mathrm{H}\beta)_{\mathrm{obs}}}{(\mathrm{H}\alpha/\mathrm{H}\beta)_{\mathrm{int}}}\right]$, where $(\mathrm{H}\alpha/\mathrm{H}\beta)_{\mathrm{obs}}$ is the observed Hα-to-Hβ flux ratio and $(\mathrm{H}\alpha/\mathrm{H}\beta)_{\mathrm{int}} = 2.86$ is the intrinsic ratio, corresponding to Case B recombination at $T = 10^4$ K and $n_e = 10^2\,\mathrm{cm}^{-3}$. We then convert this to A_V using the attenuation curve from D. Calzetti et al. (2000), $A_V = 4.05\,E(B - V)$, and because we ultimately convert A_V back to $A_{\mathrm{H}\alpha}$ using the same attenuation curve, the assumed differential extinction only affects the absolute scaling of A_V in Fig. 4.

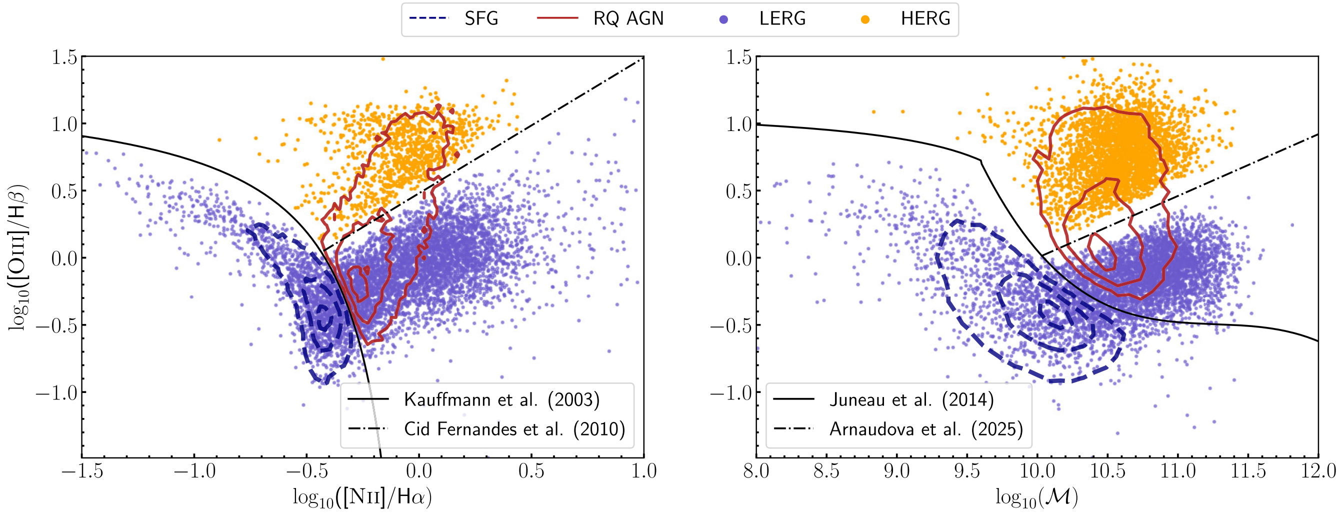

Figure 5. The left panel shows the BPT diagram for sources classified at >90 per cent confidence, where the demarcation line from G. Kauffmann et al. (2003b) separates sources into SFGs (dashed blue contours) and RQ AGN (solid red contours), while the combination of the G. Kauffmann et al. (2003b) line and the diagnostic line from R. Cid Fernandes et al. (2010) distinguishes LERGs (purple dots) from HERGs (orange dots). The contours encompass 10, 50, and 90 per cent of the relevant population. Similarly, the right panel presents the $\mathcal{M}$Ex diagram, where the separation between SFGs and RQ AGN is done via the demarcation line from S. Juneau et al. (2014), and the division between LERGs and HERGs is based on the M. I. Arnaudova et al. (2025) demarcation line. We note that the larger gap between the contours in the $\mathcal{M}$Ex diagram compared to the BPT diagram reflects the greater uncertainties in stellar mass measurements relative to those of the [N II] and Hα fluxes.

tion 4.1), namely classifying them as either SFGs or RQ AGN,[7] using the empirical demarcation line from G. Kauffmann et al. (2003b). Sources exhibiting a radio excess are instead separated into LERGs and HERGs using, in addition to the G. Kauffmann et al. (2003b) line, the division from R. Cid Fernandes et al. (2010), which traces the transition between low- and high-ionization systems in a manner consistent with the Excitation Index (EI; S. Buttiglione et al. 2010; P. N. Best & T. M. Heckman 2012), a parameter commonly used to distinguish between these two classes. The location of all four classes – SFGs, RQ AGN, LERGs, and HERGs – in this BPT classification scheme at >90 per cent confidence (see details in Section 4.4) are shown in the left panel of Fig. 5.

At higher redshifts ($z > 0.483$), where Hα and [N II] λ6583 are no longer observable in the optical range, we turn to the $\mathcal{M}$Ex diagram. This diagnostic builds upon the original Mass–Excitation diagram introduced by S. Juneau et al. (2011), S. Juneau et al. (2014), in which the x-axis of the BPT diagram is replaced by stellar mass. The modified version incorporates a redshift-dependent offset which is necessary to account for the evolution of the mass-metallicity relation, as discussed in recent studies (e.g. A. Henry et al. 2021; N. J. Cleri et al. 2023, A25). This correction shifts the dividing lines presented by S. Juneau et al. (2014), which allow us to separate SFGs and RQ AGN, to higher stellar masses at increasing redshift (see A25 for details), and is defined as

$$\log_{10} \mathcal{M} = \log_{10} M - [1 - \exp(-1.2z)]. \quad (4)$$

Here, M refers to the stellar mass, taken from DESI's AGN Host Galaxies Physical Properties VAC, while $\mathcal{M}$ denotes the modified stellar mass. To further separate the RXGs into LERGs and HERGs, we use, in addition to the S. Juneau et al. (2014) line, the EI-based demarcation line from A25, which allows us to classify sources with the $\mathcal{M}$Ex diagram into SFGs, RQ AGN, LERGs, and HERGs (hereafter the $\mathcal{M}$Ex scheme). The resulting $\mathcal{M}$Ex classification scheme for sources above the 90 per cent threshold is shown in the right panel of Fig. 5. Note that in both the BPT and $\mathcal{M}$Ex diagrams, LERGs overlap with the 'LINER' AGN region (above the G. Kauffmann et al. 2003b/S. Juneau et al. 2014 line but below that of R. Cid Fernandes et al. 2010/A25) and also with the SFG region. As weak emission-line AGN, LERGs can have their emission-line properties dominated by star formation in the host galaxy, but the radio excess reveals the presence of AGN activity.

4.3 Recovering LERGs with low-significance lines

In addition to the BPT and $\mathcal{M}$Ex diagnostics, we aim to improve the classification of LERGs, particularly those associated with low signal-to-noise emission lines that may be missed by these two schemes. To recover such sources, we use the equivalent width of [O III] λ5007 ($\mathrm{EW}_{[\mathrm{O\,III}]}$). Previous studies (e.g. R. A. Laing et al. 1994; C. N. Tadhunter et al. 1998; P. N. Best & T. M. Heckman 2012) have shown that $\mathrm{EW}_{[\mathrm{O\,III}]}$ provides an effective separation between LERGs and HERGs, and this also holds for the subset of sources already classified using the BPT/$\mathcal{M}$Ex diagnostics (see Fig. 6). The transition between LERGs and HERGs appears around 10 Å, with the majority of LERGs having $\mathrm{EW}_{[\mathrm{O\,III}]} < 3$ Å, consistent with the division proposed by R. A. Laing et al. (1994). We therefore adopt the more stringent division from R. A. Laing

[7]We note that this scheme classifies sources (without radio excess) that fall within the BPT LINER region as RQ AGN, although previous works have shown that emission in this region may not always be AGN-driven (e.g. G. Stasińska et al. 2008). We adopt this classification approach to maintain consistency with previous spectroscopic radio-source classifications (e.g. Dr24, A25), but provide the full set of classification probabilities, including the probabilistic locations within the diagnostic diagrams, allowing users to adopt alternative treatments of these sources.

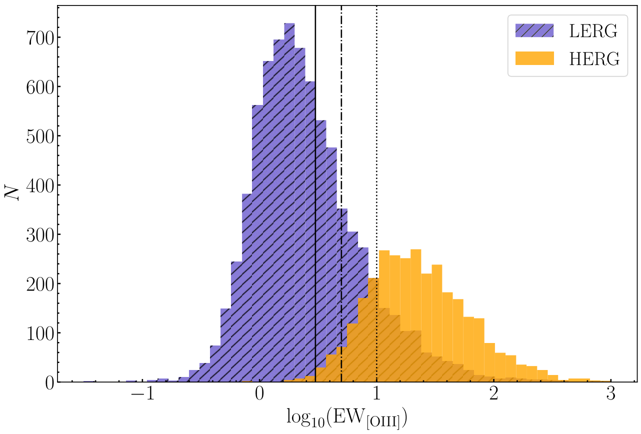

Figure 6. The distribution of $\mathrm{EW}_{[\mathrm{O\,III}]}$ for sources classified at >90 per cent confidence as LERGs and HERGs using the BPT and $\mathcal{M}$Ex classification schemes. The vertical solid, dashed, and dotted lines denote the suggested divisions from R. A. Laing et al. (1994) at $\mathrm{EW}_{[\mathrm{O\,III}]} = 3$ Å, P. N. Best & T. M. Heckman (2012) at $\mathrm{EW}_{[\mathrm{O\,III}]} = 5$ Å, and C. N. Tadhunter et al. (1998) at $\mathrm{EW}_{[\mathrm{O\,III}]} = 10$ Å, respectively.

et al. (1994) to recover additional LERGs within the radio-excess population.

4.4 Probabilistic classifications

To construct the full probabilistic classification scheme, we generate 1000 Monte Carlo realizations in which we perturb the stellar masses, emission-line fluxes, radio fluxes and the $\mathrm{EW}_{[\mathrm{O\,III}]}$ using their respective uncertainties. For each realization, we apply the four diagnostics discussed in the previous section to derive the probability of each source belonging to one of the four classes for the entire classification sample, including sources with weak or undetected emission lines.

For the RX diagnostic, we calculate the extinction values for each Monte Carlo realization. However, as noted in A25, low-significance emission lines can sometimes yield negative fluxes, producing unphysical Balmer-derived extinction values for the low-z sample. In such cases, we use the *UVJ* diagram in the same manner as for the high-z sample to assign $A_{\mathrm{H}\alpha}$ when Hβ, Hα, or both are negative. To track how often this procedure is applied and to allow users to decide whether to include these sources we flag the fraction of realizations for which the *UVJ*-based substitution is used (FLAG_A_V). Negative fluxes can also propagate into the derived quantities used in the diagnostics, such as the dust-corrected Hα luminosity ($L^{\mathrm{corr}}_{\mathrm{H}\alpha}$) and radio luminosity ($L_{144\mathrm{MHz}}$) in the RX diagram, as well as the logarithmic line ratios in the BPT and $\mathcal{M}$Ex diagrams, resulting in undefined or extreme values. To address this, we apply the following conditions: (i) when the numerator is positive but the denominator negative, we assign an unambiguously large value on the relevant axis in the RX, BPT, or $\mathcal{M}$Ex diagram; (ii) when the numerator is negative but the denominator positive, we assign an unambiguously small value; and (iii) when both are negative, the realization is considered invalid, as no classification can be assigned. We then again record the fraction of affected realizations as three warning flags (RX_WARNING_FRAC, BPT_WARNING_FRAC, and $\mathcal{M}$EX_WARNING_FRAC), allowing the end user to determine how to treat these cases.

In this work, we construct the final classification using a hierarchical approach. First, we apply the BPT scheme (i.e. the combination of the RX and BPT diagrams) to all sources where Hα and [N II] λ6583 are available, adopting a high-reliability threshold of 90 per cent for each class. Sources for which this BPT classification falls below this threshold, or which lack Hα and [N II] measurements (i.e. in the high-z sample), are then classified using the $\mathcal{M}$Ex scheme (i.e. the RX and $\mathcal{M}$Ex diagram), also with a 90 per cent reliability threshold. Following A25, we additionally require RX/BPT/$\mathcal{M}$EX_WARNING_FRAC < 0.10, limiting cases where realizations produce negative fluxes to at most 10 per cent, and retain all low-z sources regardless of the FLAG_A_v . Sources that remain unclassified but exhibit a radio excess are further examined using the $\mathrm{EW}_{[\mathrm{O\,III}]}$ criterion: those with more than 90 per cent of Monte Carlo realizations having $\mathrm{EW}_{[\mathrm{O\,III}]} < 3$ Å (EW_FRAC > 0.90) are further classified as LERGs. Here, we also allow cases where the Hα flux can be negative up to 50 per cent (i.e. RX_WARNING_FRAC < 0.5) as this behaviour is expected for weak-lined LERGs whose Hα emission may be consistent with zero within the uncertainties while still remaining securely in the radio-excess regime. We note that this procedure increases the number of LERGs by a factor of ~ 5. Considering also the quality flags discussed in Sections 3.4 and 3.5 (i.e. sources with FLAG_SPEC set, and not flagged as Type 1 AGN by FLAG_TYPE1), we find 68 820 SFGs, 32 288 RQ AGN, 35 210 LERGs (7765 classified via the BPT/$\mathcal{M}$Ex schemes, 15 520 via $\mathrm{EW}_{[\mathrm{O\,III}]}$ with RX_WARNING_FRAC < 0.1, and 11 771 via $\mathrm{EW}_{[\mathrm{O\,III}]}$ with 0.1 < RX_WARNING_FRAC < 0.5), 3085 HERGs, and 112 010 unclassified sources.

We emphasize that the adopted 90 per cent reliability threshold is intentionally conservative and is designed to provide fiducial classification. Users of this catalogue may construct alternative samples tailored to specific science goals, for example, by lowering the reliability threshold (e.g. to 70 per cent) or by assigning each source to the class with the maximum probability, rather than applying a hard cut. Such approaches increase completeness at the expense of reliability and can significantly reduce the fraction of unclassified sources. The large number of unclassified sources in this work is therefore an expected consequence of this choice. We can, however, further characterize these unclassified sources by considering whether a radio excess is present. We identify approximately 7959 sources that exhibit a radio excess at >90 per cent (i.e. $P(\mathrm{RX}) > 0.90$ and RX_WARNING_FRAC < 0.50) but cannot be robustly separated into LERGs or HERGs (uncertain radio-excess objects or RX_Unc), even after applying the $\mathrm{EW}_{[\mathrm{O\,III}]}$ criterion. Similarly, there are 24 737 sources that securely do not have a radio excess ($P(\mathrm{RX}) < 0.10$ and RX_WARNING_FRAC < 0.10[8] for which the available diagnostics do not allow a confident separation between SFGs and RQ AGN (uncertain non-radio-excess objects or $\overline{\mathrm{RX}}$_Unc) at this threshold (see also Table 2 for a detailed breakdown of the classes across the DESI surveys). The remaining 79 314 unclassified sources (or Unc.) are generally located near the boundaries of the classification criteria and/or are at higher redshifts where the spectral S/N is lower (see Section

[8]We apply the relaxed RX_WARNING_FRAC threshold only when selecting RX candidates. For these sources, realizations with negative Hα fluxes (i.e. low-S/N Hα emission) are consistent with the expected properties of weak-lined LERGs. In contrast, for non-RX sources we adopt the stricter requirement of RX_WARNING_FRAC < 0.10 to ensure that the absence of a radio excess is not driven by poorly constrained Hα measurements.

Table 2. A breakdown of source classifications at > 90 per cent confidence across the DESI surveys described in Section 2.2, including both radio-excess sources that cannot be classified as LERGs or HERGs (RX_Unc) and non-radio-excess sources that cannot be classified as SFGs or RQ AGN ($\overline{\mathrm{RX}}$_Unc).

Survey	SFG	RQ AGN	LERG	HERG	RX_Unc	$\overline{\mathrm{RX}}$_Unc	Unc	Total
All	68 820	32 288	35 210	3085	7959	24 737	79 314	251 413
BGS	63 893	28 533	21 663	1246	1986	21 782	41 881	180 984
LRG	1075	5445	18 496	1843	5738	3528	41 341	77 466
ELG	30	19	105	29	177	161	402	923
QSO	45	46	140	36	176	186	341	970
MWS	41	93	94	4	7	31	82	352
SCND	2312	1395	639	56	193	615	1494	6704

5), such that they cannot be robustly classified using the available spectroscopic and radio diagnostics at the adopted 90 per cent reliability threshold. Relaxing the reliability requirement from 90 to 70 per cent would reduce the number of unclassified sources from 79 314 to 43 289, while increasing the numbers of classified SFGs, RQ AGN, LERGs, and HERGs to 77 010, 46 498, 49 412, and 5501, respectively. Future work (Das et al., in preparation) will investigate the inclusion of additional information from broad-band photometry and joint photometric-spectroscopic SED fitting within a Bayesian framework, moving beyond the frequentist Monte Carlo approach adopted here. Such methods offer a natural way to incorporate informative priors that may help to further reduce the unclassified fraction of sources.

We note that the hierarchical classification adopted in this work results in 1242 (less than 1 per cent of the classification sample) sources changing class between the 70 and 90 per cent reliability thresholds. This occurs when a source has a BPT classification probability above 70 per cent but below 90 per cent, in which case the classification is instead determined by the $\mathcal{M}$Ex or $\mathrm{EW}_{[\mathrm{O\,III}]}$ diagnostic at the higher reliability threshold. In particular, 72 sources classified as LERGs at 70 per cent threshold are instead classified as HERGs at 90 per cent, and 28 vice versa. Similarly, 1010 RQ AGN at the 70 per cent threshold are instead classified as SFGs at 90 per cent, while 132 sources change from SFGs to RQ AGN. Given the small number of affected sources, this does not significantly impact the results presented in the following sections. Nevertheless, we note that these sources can be identified in the released catalogue and removed if a fully self-consistent hierarchical classification is preferred.

5 RESULTS

In this section, we compare the results of the classification scheme discussed in the previous section with those obtained in Dr24 using SDSS spectroscopy, and examine the selection effects introduced by DESI's targeting strategy.

5.1 Comparison with SDSS

5.1.1 Direct comparison

To make a direct comparison with the probabilistic classification scheme from Dr24, we identify the sources in common with the Dr24 catalogue[9] and our parent sample using the LoTSS Object

[9]Available from https://lofar-surveys.org/dr2_release.html

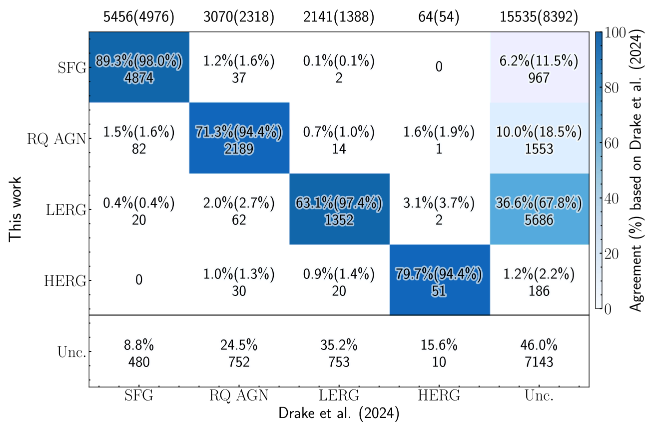

Figure 7. Confusion matrix comparing the spectroscopic classifications with >90 per cent confidence between this work (rows) and Dr24 (columns) for the common DESI and SDSS LOFAR sources. Each cell represents the level of agreement based on Dr24 in percentages, where the number of sources in each cell is also included. The percentages in brackets are calculated for the subset of sources with classifications (i.e. not including the unclassified sources) in both works and are represented by the colour bar. The total number of sources in each of the Dr24 classes is shown at the top, with the bracketed numbers indicating the number of Dr24 sources with classifications in this paper.

identifier. In doing so, we identify 29 996 common sources between DESI DR1 and SDSS, of which 26 266 are included in our classification sample once Type 1 AGN and 'bad' spectral fits are removed (see Sections 3.4 and 3.5). To construct the 90 per cent confidence sample from Dr24, we select sources that have > 90 per cent probability in each class in the catalogue – SFGs, RQ AGN, emission-line LERGs (sources with measured BPT emission lines), and HERGs – while also satisfying the recommended quality conditions (ZSCORE < 2.5 and CLASS_z_WARNING == 0). This results in 5456 SFGs, 3070 RQ AGN, 2141 LERGs, and 64 HERGs, according to Dr24.

A detailed comparison between the two schemes is shown in Fig. 7 in the form of a confusion matrix, where the results are presented as percentages, showing the level of agreement when considering Dr24 classifications as the true values. It can be seen that most of the cases where the classifications disagree are because one classification scheme is able to produce a >90 per cent classification confidence and the other is not, leading to the source being deemed as unclassified. To demonstrate this, we present the subset that meets the 90 per cent threshold for both works in brackets. When restricting the comparison to this subset, the agreement increases to ∼ 95 per cent in each class. There are a number of reasons for discrepancies in whether a >90 per cent confidence classification could be obtained. First of all, we adopt aperture corrections from `FastSpecFit` , while Dr24 only considers sources with small optical sizes (having a half-light radius of $R_{50} < 15$ arcsec) and applies no aperture correction, which results in a different RX demarcation line. Their selection of only small optical sources can also bias the sample towards lower mass systems, particularly at low redshift. Secondly, we apply a different demarcation line than Dr24 to separate LERGs and HERGs in the BPT diagram, where Dr24 adopt the G. Kauffmann et al. (2003b) line, while we use the one from R. Cid Fernandes et al. (2010) which A25 found to better follow the EI diagnostic. In addition, we use the warning fractions to account for negative fluxes in individual MCMC realizations, which are propagated into the classification results (see Section 4.4). Finally, we use additional diagnostics such as the $\mathcal{M}$Ex diagram and the $\mathrm{EW}_{[\mathrm{O\,III}]}$, which allow us to classify significantly more sources, as shown in the fifth column. This is particularly evident for LERGs, where we recover nearly three times as many sources compared to Dr24 that would otherwise be missed. These, combined with the inherent randomness introduced by noise of the different spectra and the Monte Carlo realizations used to estimate classification probabilities, which may lead to sources being just over the 90 per cent threshold in one data set and just below the other, explain the inconsistencies between the two works.

5.1.2 *Overall comparison*

To make an overall comparison between the two probabilistic classifications, we examine the demographics of the classes based on the redshift, 144 MHz flux density, radio luminosity and stellar mass, as shown in Fig. 8. In each panel, our classifications are shown by the solid lines, while those from Dr24 are shown by the dashed lines. Here, for improved statistics, we use the entire sample from Dr24 (not just the sources in common with DESI), where applying the 90 per cent threshold and recommended flags as in the previous section gives us 38 704 SFGs, 18 654 RQ AGN, 12 110 LERGs, 359 HERGs, and 82 528 unclassified sources, classified according to Dr24.

The top-left panel shows the redshift distributions, where we can see that our unclassified fraction is significantly smaller than that of Dr24 across the overlapping redshift range. This improvement probably reflects the combined effect of using multiple diagnostic diagrams, robust spectral fitting techniques, and an improved treatment of the Monte Carlo realizations. At higher z, our unclassified fraction increases as our sample becomes both fainter and increasingly dominated by DESI's LRGs (see Section 5.2). These systems are typically passive and often lack the strong emission lines necessary for confident classification. We also observe a higher fraction of SFGs and RQ AGN, particularly in the $0.2 < z < 0.4$ range. This partly results from the same improvements that allow more sources to be classified, but may also reflect selection effects: in this redshift regime, our sample includes many sources from the DESI Bright Galaxy Survey (BGS), which targets fainter optical and thus less massive galaxies compared to SDSS (as demonstrated by the stellar mass distribution of BGS sources in Fig. 9) that, when detected by LOFAR, are more likely to be associated with star formation activity (see further details in Section 5.2).

In the top-right panel, the flux density distribution shows that both our method and that of Dr24 maintain a relatively flat unclassified fraction across the full range of 144 MHz fluxes, indicating that the ability to classify sources is mostly governed by the spectroscopic data quality. We do, however, observe a modest decrease in our unclassified fraction at the highest flux densities ($S_{144,\mathrm{MHz}} \gtrsim 2.2$ mJy), accompanied by a corresponding increase in the fraction of sources classified as LERGs in this work, suggesting that our classification scheme may be more effective at recovering radio-excess AGN at higher radio flux densities. This LERG fraction is significantly larger than that in Dr24, likely due to the inclusion of the $\mathrm{EW}_{[\mathrm{O\,III}]}$ diagnostic. We also identify a larger fraction of SFGs at the lowest flux densities, whereas above $S_{144,\mathrm{MHz}} \simeq 1.3$ mJy the SFG fraction is larger in Dr24. The distributions for RQ AGN and HERGs remain broadly consistent between the two works. These differences likely reflect a combination of classification methodology and survey selection effects (DESI versus SDSS), rather than intrinsic physical differences.

By contrast, the radio luminosity (bottom left) and stellar mass (bottom right) distributions indicate that the underlying physical properties of the classified populations remain broadly consistent between the two classification schemes, despite the differences seen in redshift and flux density. The SFGs dominate at $L_{144\mathrm{MHz}} \lesssim 10^{24}\,\mathrm{W\,Hz^{-1}}$, while LERGs and HERGs become increasingly prevalent at higher radio luminosities. Similarly, SFGs are predominantly associated with lower stellar mass systems, whereas the AGN populations are mostly hosted by more massive galaxies ($M_* \gtrsim 10^{10}\,\mathrm{M_\odot}$). We do, however, identify a larger fraction of LERGs at $L_{144\mathrm{MHz}} \lesssim 10^{24-27}\,\mathrm{W\,Hz^{-1}}$ and $M_* \gtrsim 10^{11}\,\mathrm{M_\odot}$ compared to Dr24 which again is due to the additional weak-lined LERGs identified through the $\mathrm{EW}_{[\mathrm{O\,III}]}$ diagnostic.

5.2 DESI selection

As discussed in Section 2.2, DESI consists of multiple surveys with distinct selection functions, which influence the populations we classify. As shown in Table 2, the majority of our classified sources come from the BGS and LRG samples. The BGS targets low-redshift ($0 < z < 0.6$), relatively bright galaxies, including a mix of star-forming and passive systems, whereas the LRG sample selects high-mass, red, passively evolving galaxies at higher redshift ($0.4 < z < 1.1$). Therefore, to explore whether these selection strategies impact our results, we examine the demographics of classifications split by the two surveys, again considering redshift, 144 MHz flux density, radio luminosity, and stellar mass, as shown in Fig. 9. In each panel, BGS classifications are represented by solid lines, and LRG classifications by dashed lines.

The top-left panel shows the redshift distributions of the BGS and LRG subsets. The LRG sample exhibits a persistently high fraction of unclassified sources (> 50 per cent) across all redshifts, reflecting the passive nature of these galaxies and the lack of strong emission lines required for classification. In contrast, the BGS sample is largely classified at low redshift, with the unclassified fraction increasing only beyond $z \gtrsim 0.5$, where the number of BGS targets declines sharply. As expected, BGS sources are dominated by SFGs at all redshifts, whereas LRGs contain very few SFGs. RQ AGN and LERGs are similarly represented in both samples.

The top-right panel, showing the 144 MHz flux density distribution, reveals a more pronounced decrease in the unclassified fraction towards higher flux densities, especially for the

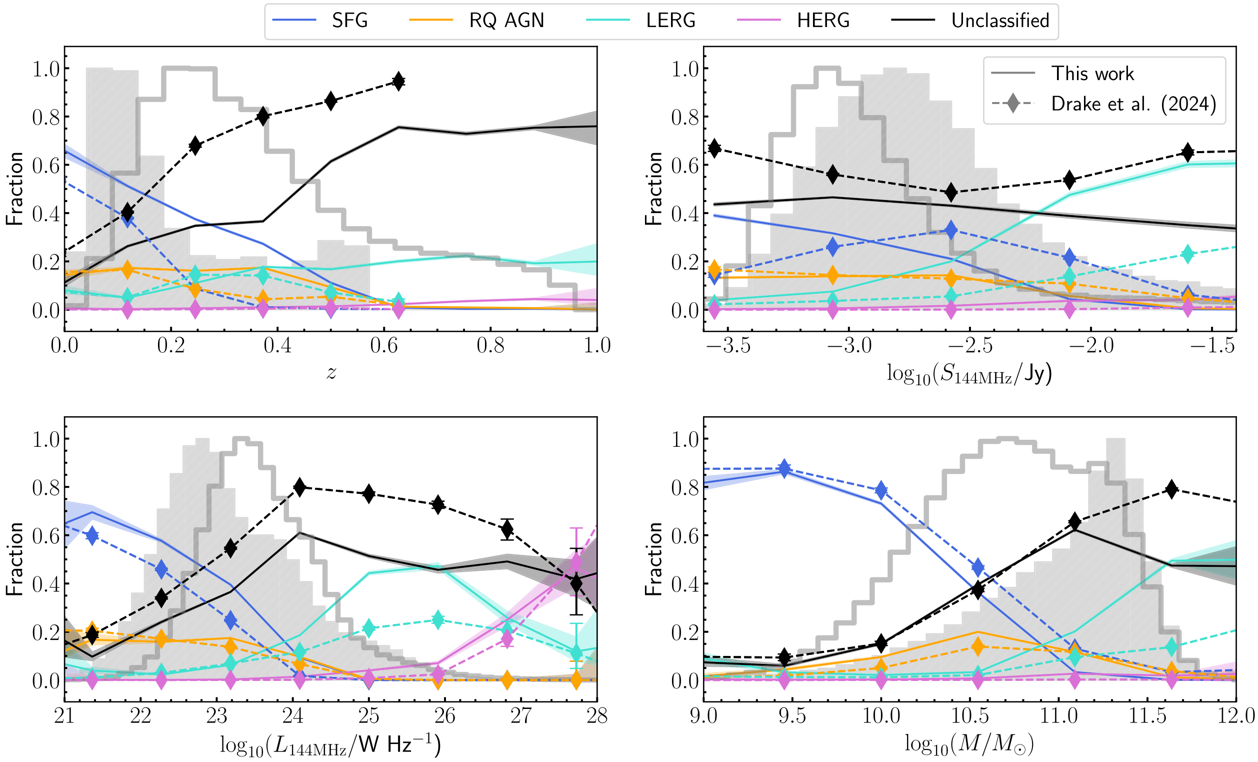

Figure 8. The distributions of sources classified using our > 90 per cent reliability threshold (solid lines and step histograms) and those from Dr24 (dashed lines and filled histograms) as a function of redshift (top left), 144 MHz flux density (top right), radio luminosity (bottom left), and stellar mass (bottom right). Each panel shows the distributions for the different classes: SFGs (blue), radio-quiet AGN (orange), LERGs (turquoise), HERGs (purple), and unclassified sources (black), as indicated in the legend at the top. The error bars and shaded regions accompanying each panel represent binomial uncertainties.

LRG sample, accompanied by a corresponding increase in the fraction of sources classified as LERGs. This trend further supports the idea that our classification scheme is more effective at recovering radio-excess AGN at higher radio flux densities, particularly within the LRG population. The BGS sample seems to be dominated by SFGs and RQ AGN at low fluxes, with LERGs becoming more common above $S_{144\mathrm{MHz}} \gtrsim -3.2$ mJy. The LRG sample, on the other hand, is dominated by LERGs across the entire radio flux range, with some contribution of RQ AGN at lower radio fluxes and HERGs at $S_{144\mathrm{MHz}} \gtrsim -2.2$ mJy. The radio luminosity distributions in the bottom-left panel highlight the expected trends observed in the previous section: radio-excess sources dominate above $L_{144\mathrm{MHz}} = 10^{24}$ W Hz^{-1}, while SFGs and RQ AGN lie primarily below this threshold. Notably, the LERG distribution in the LRG sample does not decrease strongly towards low radio powers, revealing a population of weaker radio AGN. These objects also appear to be associated with lower stellar masses, as can be seen in the bottom-right panel. A similar behaviour is observed for the RQ AGN within the LRG sample, which tends to reside in lower mass galaxies than those in the BGS population.

Overall, these trends illustrate that DESI's targeting strategy influences the observed class distributions. These selection effects must therefore be taken into account when interpreting the physical properties and relative prevalence of radio source populations in the DESI-LOFAR sample.

6 THE PROPERTIES OF LERGS AND HERGS

6.1 Eddington-scaled accretion rates

Using a similar classification method, A25 showed that the Eddington-scaled accretion rates of LERGs and HERGs remain distinct up to $z = 1$, contrary to some recent findings in the literature (e.g. I. H. Whittam et al. 2022). A25 attributed this result to the use of photometric methods in those literature studies, showing that spectroscopic classifications break the degeneracy. However, the sample of A25 was relatively small and did not allow a detailed investigation of the spectral properties of LERGs and HERGs as a function of accretion rate. This is of particular interest as recent works (e.g. R. Kondapally et al. 2022, 2025) found evidence for some LERGs residing in lower mass, star-forming systems that may have a different fuelling mechanism. Given our significantly larger spectroscopic sample ($\sim 70\times$ larger than A25), we are able to examine the distribution of accretion rates and related properties in much greater detail.

To do this, we estimate the log Eddington-scaled accretion rate (λ_{Edd}) following A25 by assuming that the total energetic output of the black hole includes both the bolometric radiative luminosity (L_{bol}) and the jet mechanical luminosity (L_{mech}), such that $\lambda_{\mathrm{Edd}} = \log_{10}[(L_{\mathrm{bol}} + L_{\mathrm{mech}})/L_{\mathrm{Edd}}]$. Here $L_{\mathrm{Edd}} = 1.3 \times 10^{31}(M_{\mathrm{BH}}/\mathrm{M_\odot})$ W is the Eddington luminosity, where M_{BH} is the black hole mass, which we estimate using the stellar mass-black hole mass relation from N. Häring & H.-W. Rix (2004). The bolometric luminosity is obtained from the observed [O III] lumi-

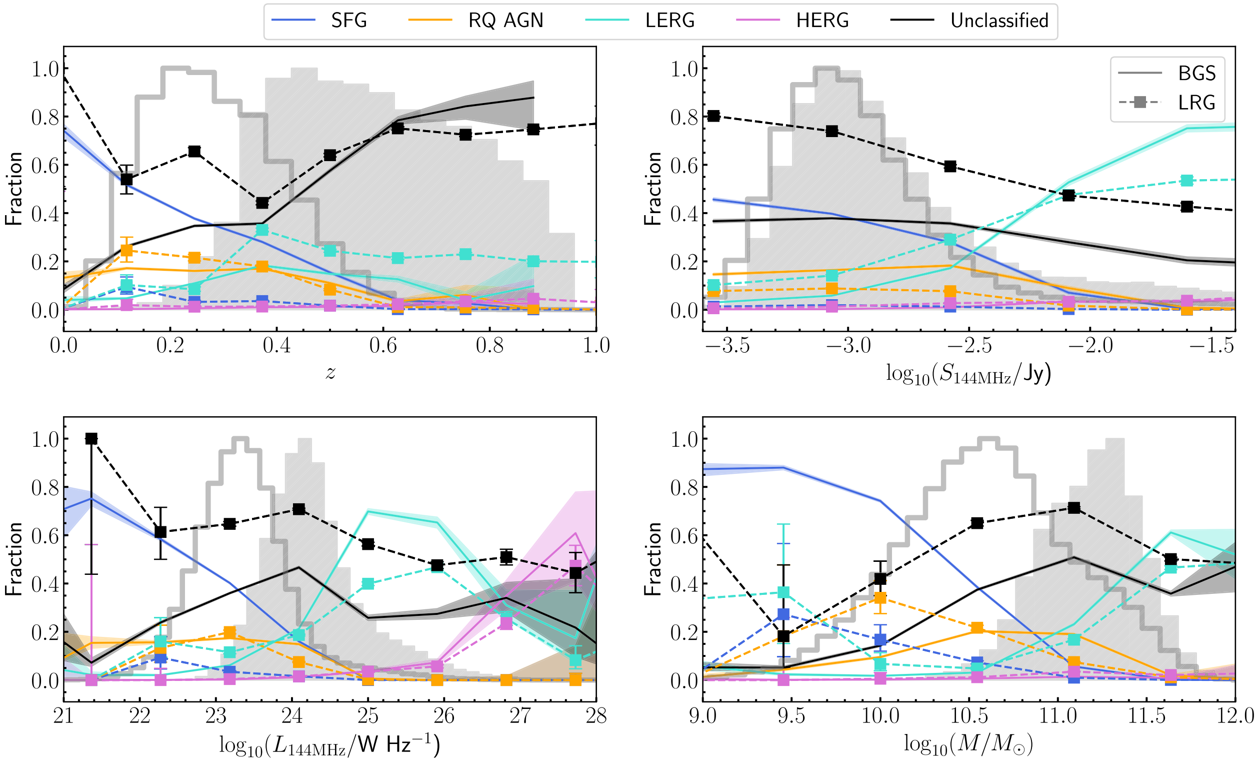

Figure 9. The influence of optical selection criteria on the sampling of the faint radio source population, illustrated by the distributions of sources classified above our > 90 per cent reliability threshold in the Bright Galaxy Survey (BGS; solid lines and step histograms) and the luminous red galaxy sample (LRG; dashed lines and filled histograms). Panel layout, line styles, and shading are the same as in Fig. 8.

nosity using $L_{\rm bol} = 3500\,L_{\rm [O\,III]}$ (T. M. Heckman et al. 2004), where a 3σ upper limit is used in cases where [O III] λ5007 is not significantly detected. The jet mechanical power is estimated as $L_{\rm mech} = 7.3 \times 10^{36}(L_{\rm 1.4GHz}/10^{24}\,{\rm W\,Hz^{-1}})^{0.7}$W (K. W. Cavagnolo et al. 2010), where the 144 MHz luminosity is converted to 1.4 GHz assuming a spectral index of $\alpha = -0.7$. We note that each of these scaling relations is characterized by substantial intrinsic scatter, and that the use of [O III] as a proxy for bolometric luminosity may be less reliable for LERGs (e.g. T. M. Heckman et al. 2004) and may differ in low-mass systems (e.g. E. C. Moran et al. 2014). Nevertheless, this framework is widely adopted in the literature (e.g. P. N. Best & T. M. Heckman 2012; B. Mingo et al. 2014; J. Chilufya et al. 2024, A25) to compare LERGs and HERGs, and our sources predominantly reside in massive host galaxies ($M > 10^{10}$ M$_\odot$).

The resulting $\lambda_{\rm Edd}$ distributions for LERGs and HERGs classified with > 90 per cent confidence are presented in Fig. 10. We additionally show the distributions for the separate LERG classification methods, namely the BPT/$\mathcal{M}$Ex diagnostics and the EW$_{\rm [O\,III]}$-based classifications with RX_WARNING_FRAC <0.10 and 0.10< RX_WARNING_FRAC <0.50. We note that the BPT/$\mathcal{M}$Ex LERGs approximately represent true measurements of the accretion-rate distributions based on significant [O III] emission, whereas the EW$_{\rm [O\,III]}$-selected LERGs should be treated as upper limits. In addition, the distribution of the full sample of sources identified as having a radio excess at >90 per cent confidence is presented. This sample is primarily composed of securely classified HERGs and LERGs, but also includes 7959 additional objects (denoted RX_Unc) for which the available diagnostics do not allow a confident distinction between the two classes. We find that LERGs display a strong peak at low accretion rates (centred at ~1 per cent of the Eddington limit), while HERGs are predominantly associated with higher $\lambda_{\rm Edd}$ values, as expected (e.g. P. N. Best & T. M. Heckman 2012; B. Mingo et al. 2014, A25). The radio-excess population follows a similar distribution to LERGs at low accretion rates, reflecting their greater numbers, but exhibits a broad tail towards higher $\lambda_{\rm Edd}$ where HERGs become more common, forming a continuous distribution between the two canonical accretion modes. Interestingly, the LERG distribution also shows a non-negligible high-$\lambda_{\rm Edd}$ tail that extends into the HERG regime, which is associated with the BPT/$\mathcal{M}$Ex LERGs and therefore reflects direct accretion-rate measurements rather than upper limits. Although these high-accreting LERGs ($\lambda_{\rm Edd} > -1.5$, corresponding to ~ 3 per cent of the Eddington luminosity) represent only a small fraction of the overall LERG population (~ 9 per cent), their relative abundance is roughly comparable to that of HERGs at similar accretion rates, as reflected in the fraction of the total radio-excess sources in the upper panel.[10] We also note that this behaviour is largely driven by the [O III] or bolometric luminosity, as the same trends are observed when computing the Eddington ratio without including $L_{\rm mech}$.

[10]To verify that the high-accretion LERGs are not misclassified, we examined the sample at a 99 per cent confidence threshold: 9 per cent remain high-accreting. Lowering the threshold (e.g. to 70 or 50 per cent) increases the overlap with HERGs (as observed in A25) and inflates the apparent high-$\lambda_{\rm Edd}$ LERG population, mainly by including otherwise uncertain sources. Adopting a > 90 per cent threshold therefore minimizes contamination and ensures the high-accretion tail represents a robust subpopulation.

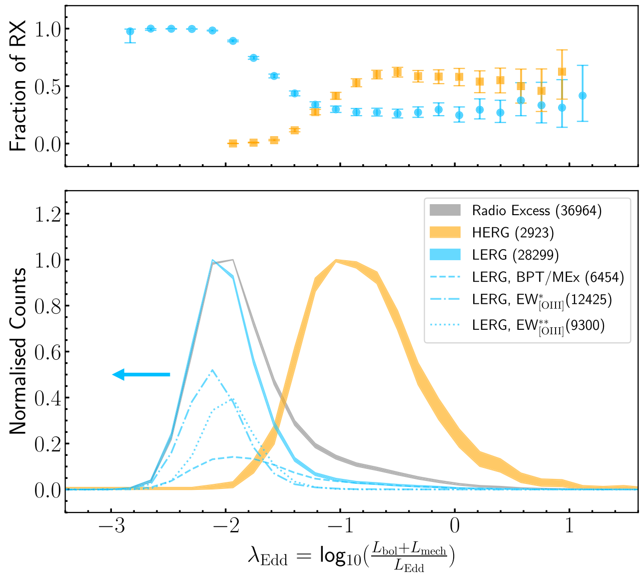

Figure 10. The log Eddington-scaled accretion rate distributions for LERGs (blue), HERGs (orange), and radio-excess sources (grey; including objects that cannot be confidently assigned to either class) that satisfy our $>$ 90 per cent threshold. For the LERG population, the distributions are additionally separated by classification method: the dashed line indicates sources classified via the BPT/$\mathcal{M}$Ex diagnostics, while the dot–dashed ($*$) and dotted ($**$) lines indicate sources classified using the $\mathrm{EW}_{[\mathrm{O\,III}]}$ method with RX_WARNING_FRAC < 0.10 and $0.10 <$ RX_WARNING_FRAC < 0.50, respectively. The arrow indicates that the LERG accretion rates should be considered upper limits due to the 3σ upper limit on the [O III] emission, while the shaded regions indicate binomial uncertainties. The number of sources in each class is given in the legend at the upper right; these numbers differ from Table 2 as not all sources have reliable stellar mass estimates. The upper panel shows the fraction of radio-excess sources classified as LERGs (blue) and HERGs (orange), with error bars again representing binomial uncertainties. Only bins containing at least five sources are included.

As noted in Section 5.2, it is important to account for DESI's targeting strategy when interpreting these distributions. To assess this, in Appendix A we separate the Eddington-scaled distributions into BGS and LRG subsamples. We find that the high-$\lambda_{\rm Edd}$ LERG tail is more prominent in the BGS population, suggesting that these high-accreting LERGs may preferentially reside in systems with more ongoing star formation. We also consider the impact of populations excluded from the main analysis. Type 1 AGN are excluded because our diagnostics are primarily designed for narrow-line systems. The broad-line emission complicates the radio-excess classification, while the strong AGN continuum may bias the SED-derived stellar masses used in the Eddington-scaled accretion-rate estimates. Nevertheless, if all Type 1 AGN were assumed to be radio excess and their stellar masses treated as reliable, they would only strengthen the high-accretion tail of the HERG distribution, as all Type 1 AGN in our sample have $\lambda_{\rm Edd} > -1.5$ (assuming the same calibrations apply). Similarly, as discussed in Appendix B, including the large Unc population would primarily increase the sample size of the existing populations without qualitatively changing the main conclusions of this work.

While the aforementioned DESI selection effects may influence the observed $\lambda_{\rm Edd}$ distributions, the large DESI spectroscopic sample nevertheless provides an important opportunity to investigate the properties of the classified radio-source populations, which we present in the following sections.

6.2 Composite spectra as a function of $\lambda_{\rm Edd}$

To investigate how the spectral properties of the radio-loud population vary with accretion rate, we use `SpecStacker` [11] to construct median-stacked spectra in bins of $\lambda_{\rm Edd}$. To summarize, this stacking code shifts the spectra to the rest frame and resamples them onto a common wavelength grid. Next, it normalizes the spectra at the reddest wavelength range where all spectra contribute, and takes the median value at each wavelength to create the stacked spectrum. To estimate uncertainties, `SpecStacker` uses bootstrapping, along with an additional set of simulations that propagate errors introduced by the redshifting and normalization steps (see M. I. Arnaudova et al. 2024a for further details). The results of this procedure are shown in Fig. 11 for LERGs (top panel) and HERGs (second panel) classified with >90 per cent confidence, radio-excess objects that we cannot confidently classify into a LERG or a HERG (or radio-excess unclassified; third panel), and the total radio excess population (bottom panel) binned in increments of $\Delta\lambda_{\rm Edd} = 0.5$ dex.[12] To further investigate our results, we applied `WL-SLAYER` in a setup similar to that described in Section 3 to measure emission lines and place them on the BPT diagram (see Fig. 12). We consider only fits with an acceptable continuum model ($\chi^2_\nu < 3$), in order to prevent the emission-line modelling from fitting continuum artefacts,[13] and require a 3σ detection in all BPT emission lines.

In the top panel of Fig. 11, we can see that the LERG stacked spectra at low $\lambda_{\rm Edd}$ show the expected behaviour for this population: the spectra are dominated by an older stellar population, indicated by a strong 4000 Å break and prominent Ca H & K absorption features, with weak or absent emission lines. As $\lambda_{\rm Edd}$ increases, however, emission features begin to emerge, in particular the Balmer lines (Hδ, Hγ, Hβ, and Hα), as well as [O II] $\lambda\lambda$3726, 3728, [O III] $\lambda\lambda$4959, 5007, [N II] $\lambda\lambda$6548 6583, and the [SII] $\lambda\lambda$6717, 6731 doublet. The relative strengths of the Balmer and forbidden lines indicate that these galaxies remain in a low-ionization state, although their emission line ratios gradually evolve from the LINER region towards the SFG region in the BPT diagram,[14] as seen in Fig. 12. The continuum shape also changes with increasing accretion rate, becoming progressively bluer, accompanied by a weakening of the Ca H & K absorption, consistent with a growing contribution from a younger stellar population (cf. R. Kondapally et al. 2025).

While a similar trend is seen in the HERG spectra, contrary to the LERGs, the forbidden lines appear much stronger than the Balmer lines (even in the lowest accretion rate stacks, the [O III] λ5007 and [N II] λ6583 emission line can be seen, while Hβ and Hα are absent), and are located as expected into the

[11] https://github.com/m-arnaudova/SpecStacker

[12] We note that due to the broad redshift range considered ($0 < z < 1$), not all spectra contribute to every wavelength bin, resulting in inhomogeneous uncertainties across the stacked spectrum, particularly for the radio-excess unclassified sources.

[13] Stacked spectra, which have higher SNR and contain a mixture of stellar populations, are challenging to model accurately using the BC03 stellar population library.

[14] We note that the same features persist when restricting the analysis to the $0.2 < z < 0.4$ range, suggesting that cosmic evolution does not significantly affect these results.

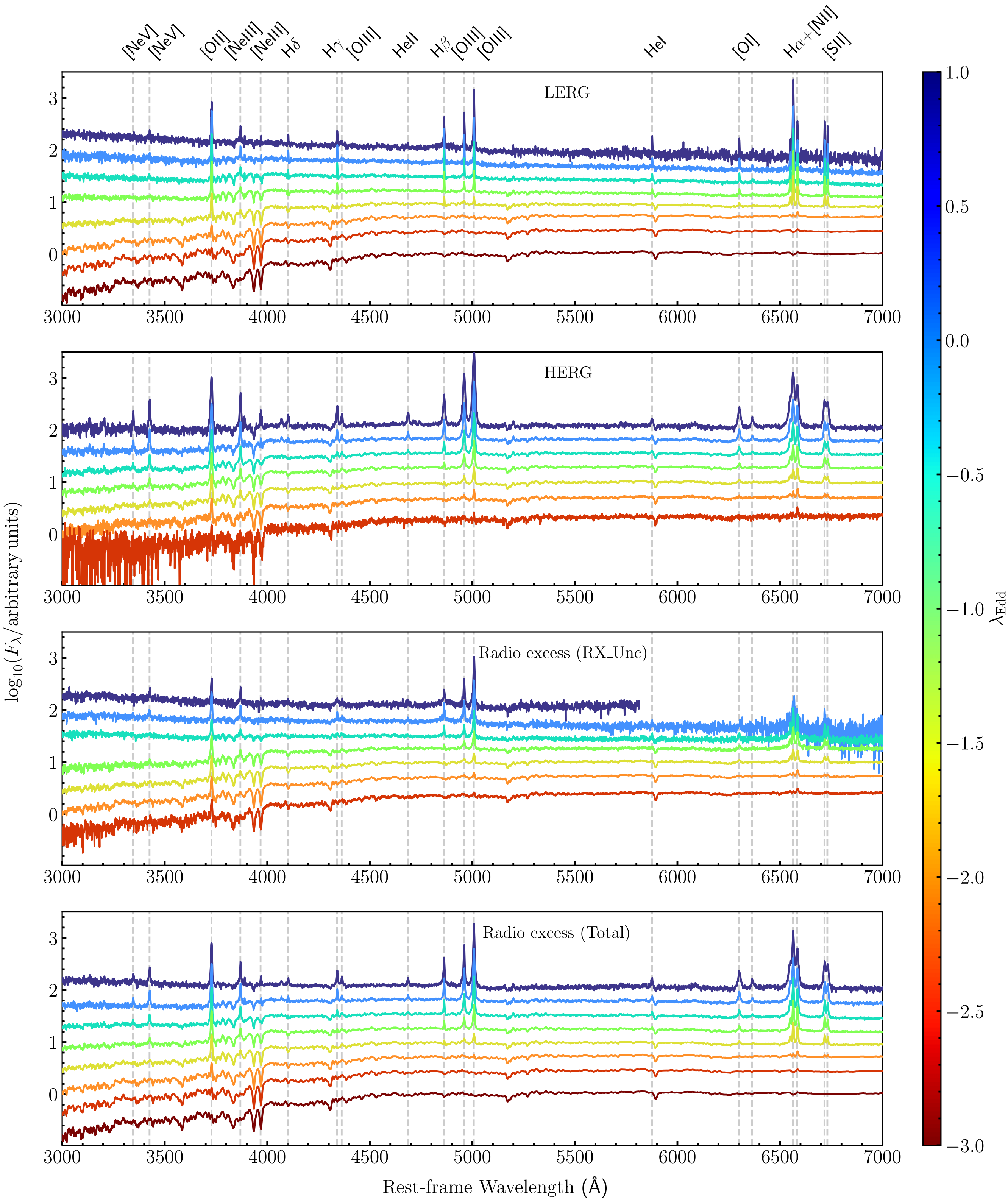

Figure 11. The stacked spectra of LERGs (top panel), HERGs (second panel), radio-excess sources that cannot be classified at > 90 per cent confidence into the two classes (or RX_Unc, third panel) and the whole radio-excess population (bottom panel) shown as a function of log Eddington-scaled accretion rate, as indicated by the colour bar. Each stack represents the median of all spectra within a given accretion-rate bin in the rest frame, as indicated by the colour bar. Prominent emission lines are marked with dashed grey lines and labelled in the top panel.

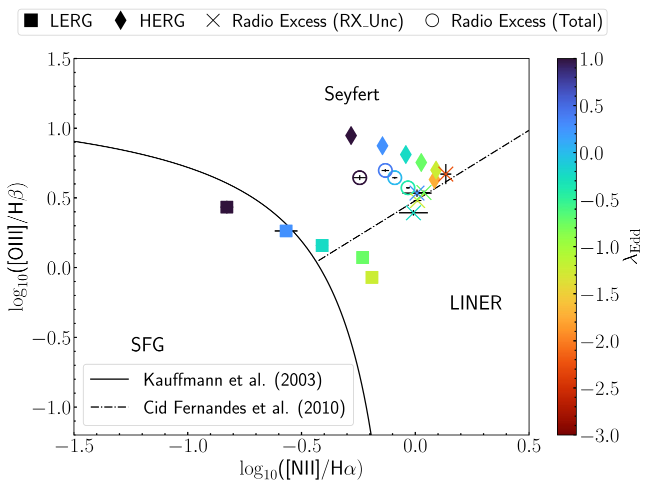

Figure 12. The BPT diagram for the stacked spectra of LERGs (squares), HERGs (diamonds), radio-excess sources that remain unclassified (or RX_Unc; cross), and the total radio-excess sources (circles), colour coded by the log Eddington-scaled accretion rate. The diagnostic lines delineating the SFG, LINER, and Seyfert regions are from G. Kauffmann et al. (2003b) and R. Cid Fernandes et al. (2010) and are denoted in the legend.

Seyfert region in the BPT (see Fig. 12). There are also a number of higher ionization lines such as [Ne V] and [Ne III] that are very pronounced in the HERGs, while absent or weak for the LERG stacks for the same accretion rate bin. Similar results were observed when comparing the stacked spectra of HERGs and star-forming LERGs in R. Kondapally et al. (2025).

The stacked spectra of the radio-excess unclassified objects (third panel) resemble those of HERGs, showing higher ionization lines such as [Ne V], albeit at lower strength. Their emission line ratios place them near the boundary between LERGs and HERGs on the BPT diagram, consistent with their intermediate spectral properties. Finally, the total radio-excess population (bottom panel) exhibits a clear transition between the LERG- and HERG-dominated regimes, reflecting the accretion rates at which each mode becomes more prevalent. At low accretion rates ($\lambda_{\rm Edd} \lesssim -1.5$), the stacked spectra closely resemble those of the LERGs, where the spectra are continuum dominated. As $\lambda_{\rm Edd}$ increases, the spectra gradually shift towards HERG-like characteristics, where the forbidden lines and in particular higher ionization lines become more prominent, which is clearly seen in the BPT diagram.

Overall, these results highlight that, within a given $\lambda_{\rm Edd}$ bin, LERGs and HERGs remain distinct, while the emergence of Balmer lines and bluer continua in high-accreting LERGs points to ongoing star formation, helping to explain their increased presence in the BGS sample.

6.3 High-accreting LERGs as a function of diagnostics

Given the results of the previous section, we examined the location of high-accreting LERGs ($\lambda_{\rm Edd} > -1.5$) in the diagnostics introduced in Section 4, in addition to the BPT results from the stacked spectra shown in Fig. 12, to characterize their properties in more detail. For the radio-excess diagnostic, we consider the full sample of high-accreting LERGs, whereas for the BPT and $\mathcal{M}$Ex diagrams we include only sources classified at > 90 per cent confidence in the respective diagnostics. The [O III] EW method does not yield a significant population of high-accreting LERGs and is therefore excluded from this analysis.

The results are presented in Fig. 13, where we can see in the left panel that the majority of high-accreting LERGs, shown as the colour scale, are concentrated near the radio-excess demarcation line, suggesting that their radio emission may be modestly enhanced relative to their Hα luminosity, potentially due to low-power AGN activity or contributions from intense star-formation-driven shocks. By contrast, low-accreting LERGs ($\lambda_{\rm Edd} < -1.5$), represented by grey contours, populate a broader region of the radio-excess plane, with sources both near the demarcation line and at higher radio powers for a given Hα luminosity. In the BPT diagram (middle panel), high-accreting LERGs predominantly occupy the BPT_SFG region, while in the $\mathcal{M}$Ex diagram (right panel) they are distributed between the $\mathcal{M}$Ex_SFG and $\mathcal{M}$Ex_LINER regions. This mixed behaviour may be influenced by redshift-dependent effects, since the $\mathcal{M}$Ex diagram is used primarily at $z > 0.5$, or by optical selection biases (e.g. LRGs and BGS subsamples). Low-accreting LERGs, in contrast, are mostly found in the BPT_LINER and $\mathcal{M}$Ex_LINER regions, suggesting that they are associated with different ionization properties.

Taken together, these results indicate that high-accreting LERGs may either be star-forming galaxies with a modest radio excess due to shocks or recent star formation, or genuine LERGs at lower radio powers residing in more actively star-forming hosts. In either scenario, star-forming processes could contribute to the [O III] emission, potentially leading to an overestimation of $L_{\rm [O\,III]}$ associated with the AGN activity and thus $\lambda_{\rm Edd}$. Furthermore, as previously discussed, the conversion from $L_{\rm [O\,III]}$ to $L_{\rm bol}$ is subject to significant intrinsic scatter and may be less reliable for LERGs. If this calibration does not hold in this regime, the inferred bolometric luminosities and consequently the Eddington ratios may not be representative of the true accretion power. Future work using independent constraints on the accretion power, for example from X-ray measurements, would therefore provide a valuable cross-check. We cannot also rule out the possibility that, if these systems are indeed LERGs, their fuelling mechanisms may differ from the standard hot-gas accretion via advection-dominated flows (ADAFs; R. Narayan & I. Yi 1994, 1995) observed in typical low-accreting LERGs. These results will be explored further by Kondapally et al. (in preparation).

7 CONCLUSIONS

In this work, we have used the spectroscopic information from the first data release of the DESI survey to probabilistically classify the radio source population observed by the second data release of the LOFAR LoTSS. Building upon the classification methods by Dr24 and A25, we have used a combination of radio excess, the BPT diagnostic, and a modified MEx ($\mathcal{M}$Ex) diagnostic, alongside Monte Carlo methods to estimate the probability of a source being an SFG, a RQ AGN, or an LERG or HERG. In addition, we used the [O III] λ5007 equivalent width to recover LERGs associated with low-significance emission lines, significantly increasing the sample size and enabling a more complete recovery of the LERG population. Applying a 90 per cent probability threshold to this classification framework, we identify 68 820 SFGs, 32 288 RQ AGN, 35 210 LERGs, and 3085 HERGs.

Using this high-confidence DESI-LOFAR sample:

(i) We compared our classifications with Dr24, which are based on SDSS spectroscopy, and found comparable results when con-

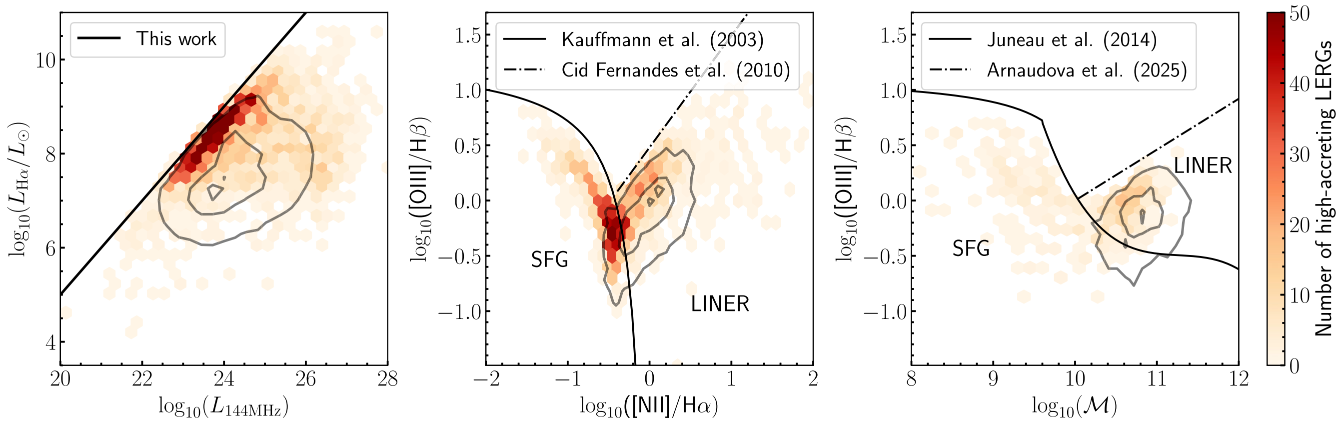

Figure 13. The radio excess (left), BPT (middle) and $\mathcal{M}$Ex diagnostic diagrams (right) for the high-accreting LERGs ($\lambda_{\mathrm{Edd}} > -1.5$), as denoted by the colour bar. The grey contours encompass 10, 50, and 90 per cent of the low-accreting LERGs. The demarcation lines are the ones used in the classification scheme discussed in Section 4, and are denoted by the legend in each panel. For reference, the total number of high-accreting LERGs is 2611, while the total number of low-accreting LERGs is 25 688. The remaining 6911 LERGs do not have reliable stellar mass estimates and therefore lack an accretion rate measurement.

sidering the sources classified by both works. However, some discrepancies arise when including the unclassified fraction, resulting from our use of aperture corrections, additional diagnostics and the more rigorous treatment of the Monte Carlo realizations.

(ii) We assessed the impact of DESI selection effects by comparing the properties of the numerically dominant samples from DESI's BGS and LRG surveys, showing that selection biases affect individual class distributions, and thus must be taken into account.

(iii) We examined the Eddington-scaled accretion rate (λ_{Edd}) distributions up to $z = 1$, confirming very distinct accretion rate distributions for the two classes of objects where LERGs typically accrete below 1 per cent of Eddington and HERGs above it with higher statistical significance than before.

(iv) We uncovered a high-accreting population of LERGs that appear to be hosted by more strongly star-forming systems, as shown by the stacked spectra and their location on the diagnostic diagrams. These high-accreting LERGs appear similar to the star-forming LERG population identified by R. Kondapally et al. (2022), R. Kondapally et al. (2025), despite differences in selection and classification methods. We outline several possible explanations for their nature, including contamination of the [O III] emission by star-forming processes, limitations in the $L_{[\mathrm{O\,III}]}$–L_{bol} conversion in this regime, or a distinct fuelling mechanism compared to typical low-accreting LERGs. Future multiwavelength observations, particularly in the X-ray, will be essential to disentangle these scenarios.

In addition to the findings reported in this paper, our catalogue will be invaluable for statistical investigations of other key aspects of galaxy evolution. For example, we are currently undertaking a study of the properties of AGN-driven outflows for the different classifications presented here (Holden et al., in preparation), which is now possible thanks to the large sample size provided by DESI. Building on this work, the upcoming WEAVE-LOFAR survey (D. J. B. Smith et al. 2016) will directly target LoTSS sources, selecting solely on radio flux density and thus offering a more complete and unbiased view of the radio source population. By reducing the optical selection biases inherent in the current DESI-based sample, WEAVE-LOFAR will allow robust characterization of radio sources' physical properties, accretion modes, and host galaxies across a wide range of environments. This will open the door to more detailed studies of high-accreting LERGs, and more broadly to a deeper understanding of the interplay between star formation and AGN activity and their impact on galaxy evolution.

ACKNOWLEDGEMENTS

The authors sincerely thank the anonymous reviewer for a report which significantly improved the quality of this paper.

MIA and PNB acknowledge support from the UK Science and Technology Facilities Council (STFC) under grant ST/Y000951/1. DJBS and LRH acknowledge support from STFC under grant ST/Y001028/1. DJBS acknowledges support from STFC grant ST/V000624/1. DJBS and SD acknowledge support from the Leverhulme Trust via Research Project Grant RPG-2025-078. PNB is grateful for support from the UK STFC via grant ST/V000594/1. RK acknowledges support from the Leverhulme Trust through a Leverhulme Trust Early Career Fellowship. KJD acknowledges support from STFC through an Ernest Rutherford Fellowship (grant number ST/W003120/1). MJH thanks the UK STFC for support via grant ST/V000624/1 and ST/Y001249/1. SS and DJBS acknowledge support from the UK STFC via the grant ST/X508408/1. CLH acknowledges support from STFC through grant ST/Y000951/1.

LOFAR is the Low Frequency Array designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with its own funding sources), and which are collectively operated by the LOFAR ERIC under a joint scientific policy. The LOFAR resources have benefited from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland; The Istituto Nazionale di Astrofisica (INAF), Italy. This research made use of the Dutch national e-infrastructure with support from the SURF Cooperative (e-infra 180169) and the LOFAR e-infra group. The Jülich LOFAR Long Term Archive and the German LOFAR network are both coordinated and operated by the Jülich Supercomputing Centre (JSC), and computing resources on the supercomputer JUWELS at JSC were provided by the Gauss Centre for Supercomputing e.V. (grant CHTB00) through the John von Neumann Institute for Computing (NIC). This research made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC (ST/P000096/1), and of the Italian LOFAR IT computing infrastructure supported and operated by INAF, and by the Physics Department of Turin University (under an agreement with Consorzio Interuniversitario per la Fisica Spaziale) at the C3S Supercomputing Centre, Italy. This research is part of the project LOFAR Data Valorization (LDV; project numbers 2020.031, 2022.033, and 2024.047) of the research programme Computing Time on National Computer Facilities using SPIDER that is (co-)funded by the Dutch Research Council (NWO), hosted by SURF through the call for proposals of Computing Time on National Computer Facilities.

DATA AVAILABILITY

The radio and spectroscopic data used in this work are publicly available from the LOFAR Surveys (https://www.lofar-surveys.org) and the Dark Energy Spectroscopic Instrument (https://data.desi.lbl.gov), respectively. The output catalogue, including the emission-line measurements and the classification probabilities, as well as the stacked spectra of the radio-excess population, is made available as part of this paper at https://www.lofar-surveys.org/dr2_release.html.

REFERENCES

Ahumada R. et al., 2020, ApJS, 249, 3
Arnaudova M. I. et al., 2024a, MNRAS, 528, 4547
Arnaudova M. I. et al., 2024b, MNRAS, 535, 2269
Arnaudova M. I. et al., 2025, MNRAS, 542, 2245
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Begelman M. C., Blandford R. D., Rees M. J., 1984, Rev. Mod. Phys., 56, 255
Best P. N., Heckman T. M., 2012, MNRAS, 421, 1569
Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9
Best P. N. et al., 2023, MNRAS, 523, 1729
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Buttiglione S., Capetti A., Celotti A., Axon D. J., Chiaberge M., Macchetto F. D., Sparks W. B., 2010, A&A, 509, A6
Calistro Rivera G., Lusso E., Hennawi J. F., Hogg D. W., 2016, ApJ, 833, 98
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cappellari M., 2017, MNRAS, 466, 798
Cappellari M., 2023, MNRAS, 526, 3273
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379
Cavagnolo K. W., McNamara B. R., Nulsen P. E. J., Carilli C. L., Jones C., Bîrzan L., 2010, ApJ, 720, 1066
Chaussidon E. et al., 2023, ApJ, 944, 107
Chilufya J., Hardcastle M. J., Pierce J. C. S., Croston J. H., Mingo B., Zheng X., Baldi R. D., Röttgering H. J. A., 2024, MNRAS, 529, 1472
Cid Fernandes R., Stasińska G., Schlickmann M. S., Mateus A., Vale Asari N., Schoenell W., Sodré L., 2010, MNRAS, 403, 1036
Cleri N. J. et al., 2023, ApJ, 948, 112
Condon J. J., 1992, ARA&A, 30, 575
Cooper A. P. et al., 2023, ApJ, 947, 37
da Cunha E. et al., 2015, ApJ, 806, 110
Das S. et al., 2024, MNRAS, 531, 977
Delvecchio I. et al., 2021, A&A, 647, A123
DESI Collaboration, 2016a, preprint (arXiv:1611.00036)
DESI Collaboration, 2016b, preprint (arXiv:1611.00037)
DESI Collaboration, 2025, preprint (arXiv:2503.14745)
Dey A. et al., 2019, AJ, 157, 168
Drake A. B. et al., 2024, MNRAS, 534, 1107
Duncan K. J., 2022, MNRAS, 512, 3662
Escott E. L. et al., 2025, MNRAS, 536, 1166
Fang J. J. et al., 2018, ApJ, 858, 100
Fitzpatrick E. L., 1999, PASP, 111, 63
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gürkan G., Hardcastle M. J., Jarvis M. J., 2014, MNRAS, 438, 1149
Gürkan G. et al., 2018, MNRAS, 475, 3010
Guy J. et al., 2023, AJ, 165, 144
Hahn C. et al., 2023, AJ, 165, 253
Hardcastle M. J., Croston J. H., 2020, New A Rev., 88, 101539
Hardcastle M. J. et al., 2023, A&A, 678, A151
Hardcastle M. J. et al., 2025, MNRAS, 539, 1856
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Heckman T. M., Best P. N., 2014, ARA&A, 52, 589
Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109
Henry A. et al., 2021, ApJ, 919, 143
Holden L. R., Tadhunter C. N., 2025, MNRAS, 536, 1857
Holden L. R. et al., 2026, MNRAS, 545, staf2075
Johnson B. D., Leja J., Conroy C., Speagle J. S., 2021, ApJS, 254, 22
Juneau S., Dickinson M., Alexander D. M., Salim S., 2011, ApJ, 736, 104
Juneau S. et al., 2014, ApJ, 788, 88
Kauffmann G. et al., 2003a, MNRAS, 341, 33
Kauffmann G. et al., 2003b, MNRAS, 346, 1055
Kondapally R. et al., 2021, A&A, 648, A3
Kondapally R. et al., 2022, MNRAS, 513, 3742
Kondapally R. et al., 2025, MNRAS, 536, 554
Laing R. A., Jenkins C. R., Wall J. V., Unger S. W., 1994, in Bicknell G. V., Dopita M. A., Quinn P. J., eds, ASP Conf. Ser. Vol. 54, The Physics of Active Galaxies. Astron. Soc. Pac., San Francisco, p. 201
Leja J., Johnson B. D., Conroy C., van Dokkum P., 2018, ApJ, 854, 62
Macfarlane C. et al., 2021, MNRAS, 506, 5888
Martis N. S. et al., 2016, ApJ, 827, L25
Mingo B., Hardcastle M. J., Croston J. H., Dicken D., Evans D. A., Morganti R., Tadhunter C., 2014, MNRAS, 440, 269
Molyneux S. J., Harrison C. M., Jarvis M. E., 2019, A&A, 631, A132
Morabito L. K. et al., 2022, MNRAS, 515, 5758
Moran E. C., Shahinyan K., Sugarman H. R., Vélez D. O., Eracleous M., 2014, AJ, 148, 136
Moustakas J., Buhler J., Scholte D., Dey B., Khederlarian A., 2023, FastSpecFit: Fast Spectral Synthesis and Emission-line Fitting of DESI Spectra, Astrophysics Source Code Library, record ascl:2308.005
Muzzin A. et al., 2013, ApJ, 777, 18
Myers A. D. et al., 2023, AJ, 165, 50
Narayan R., Yi I., 1994, ApJ, 428, L13
Narayan R., Yi I., 1995, ApJ, 452, 710
Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, California, US
Panessa F., Baldi R. D., Laor A., Padovani P., Behar E., McHardy I., 2019, Nat. Astron., 3, 387

Petley J. W. et al., 2022, MNRAS, 515, 5159
Pracy M. B. et al., 2016, MNRAS, 460, 2
Raftery A. E., 1995, Sociol. Methodol., 10, 3, 227
Raichoor A. et al., 2020, Res. Notes Am. Astron. Soc., 4, 180
Raichoor A. et al., 2023, AJ, 165, 126
Ruiz-Macias O. et al., 2020, Res. Notes Am. Astron. Soc., 4, 187
Sabater J. et al., 2019, A&A, 622, A17
Sabater J. et al., 2021, A&A, 648, A2
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shenoy S. et al., 2026, MNRAS, 546, 3, stag137
Shimwell T. W. et al., 2017, A&A, 598, A104
Shimwell T. W. et al., 2019, A&A, 622, A1
Shimwell T. W. et al., 2022, A&A, 659, A1
Shimwell T. W. et al., 2025, A&A, 695, A80
Shimwell T. W. et al., 2026, A&A, 707, A198
Siudek M. et al., 2024, A&A, 691, A308
Smith D. J. B. et al., 2011, MNRAS, 416, 857
Smith D. J. B. et al., 2012, MNRAS, 427, 703
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 Lyon, France, p. 271
Smith D. J. B. et al., 2021, A&A, 648, A6
Smolčić V. et al., 2017, A&A, 602, A2
Stasińska G. et al., 2008, MNRAS, 391, L29
Sutherland W., Saunders W., 1992, MNRAS, 259, 413
Tadhunter C. N., Morganti R., Robinson A., Dickson R., Villar-Martin M., Fosbury R. A. E., 1998, MNRAS, 298, 1035
Tasse C. et al., 2021, A&A, 648, A1
ter Braak C., Vrugt J., 2008, Stat. Comput., 18, 435
Tremonti C. A. et al., 2004, ApJ, 613, 898
van Haarlem M. P. et al., 2013, A&A, 556, A2
Virtanen P. et al., 2020, Nat. Methods, 17, 261
White S. V. et al., 2025, MNRAS, 541, 3636
Whittam I. H. et al., 2022, MNRAS, 516, 245
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Williams W. L. et al., 2019, A&A, 622, A2
Wright E. L. et al., 2010, AJ, 140, 1868
Wuyts S. et al., 2007, ApJ, 655, 51
Yèche C. et al., 2020, Res. Notes Am. Astron. Soc., 4, 179
Yue B.-H. et al., 2024, MNRAS, 529, 3939
Zhou R. et al., 2020, Res. Notes Am. Astron. Soc., 4, 181
Zhou R. et al., 2023, AJ, 165, 58

APPENDIX A: IMPACT OF DESI SELECTION

To evaluate the impact of DESI's targeting strategy on the Eddington-scaled accretion rates, we separate our radio-loud sample into the two numerically dominant surveys-BGS and LRGs-and compute the accretion-rate distributions for LERGs, HERGs, and radio-excess (RX) sources following the procedure outlined in Section 6.1. The results are shown in Fig. A1, where we show the Eddington-scaled accretion-rate distributions (with source counts listed in the legends), alongside the fraction of RX sources classified as LERGs or HERGs in each subsample.

In both the BGS and LRG samples, we recover the characteristic dual peaks associated with LERGs and HERGs. However, in the BGS sample the high-$\lambda_{\rm Edd}$ tail of the LERG distribution is much more pronounced. This suggests that the high-accreting LERGs may reside in more star-forming systems, which are galaxy types that the LRG selection is designed to exclude. We explore this further in Sections 6.2 and 6.3.

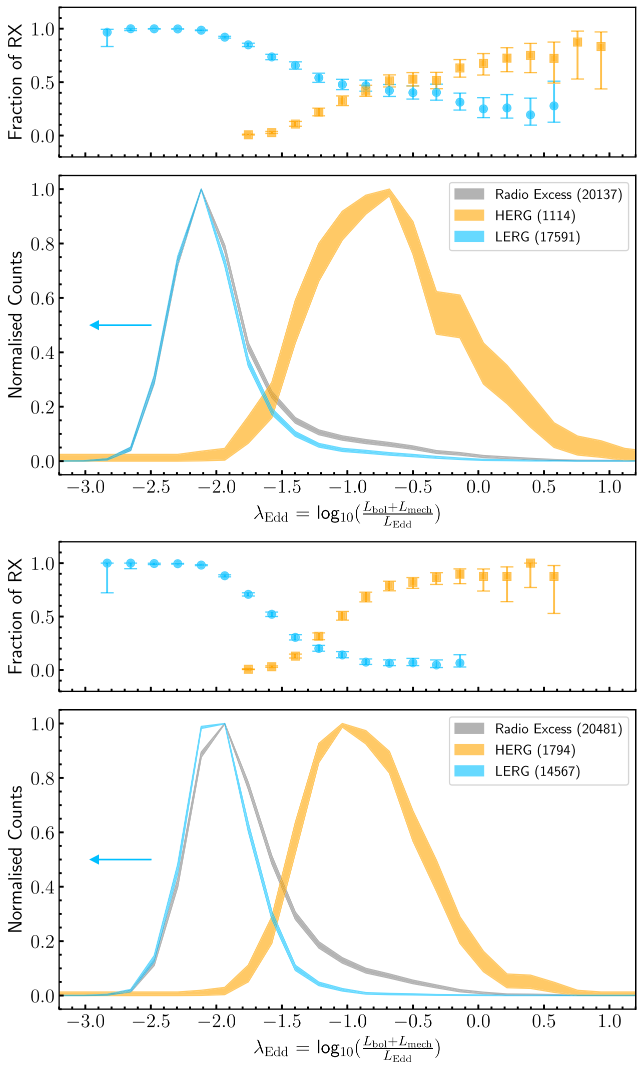

Figure A1. The Eddington-scaled accretion rate distributions for LERGs (blue), HERGs (orange), and all radio-excess sources (grey), separated into the subsamples belonging to the DESI BGS (top) and LRG (bottom) survey selections. The arrows indicate that the LERG accretion rates should be considered upper limits due to the 3σ upper limit on the [O III] emission, while the shaded regions indicate binomial uncertainties. The upper panel shows the fraction of radio-excess sources classified as LERGs (blue) and HERGs (orange), with error bars representing binomial uncertainties. Only bins containing at least five sources are included.

APPENDIX B: IMPACT OF UNCLASSIFIED SOURCES

As discussed in Section 4, we have a large fraction of unclassified sources which lie close to the boundaries of our diagnostics and therefore cannot be confidently assigned to a given class at the adopted > 90 per cent reliability threshold without additional information. Nevertheless, it is informative to investigate the impact of excluding these systems from the final analysis. In Fig. B1, we present the accretion-rate distributions of: (i) radio-excess sources that cannot be robustly separated into LERGs and HERGs (RX_Unc), (ii) non-radio-excess sources that cannot be confidently classified as either SFGs or RQ AGN ($\overline{\rm RX}$_Unc), and (iii) the remaining unclassified sources (Unc), for which even the presence of a radio excess cannot be determined at the 90

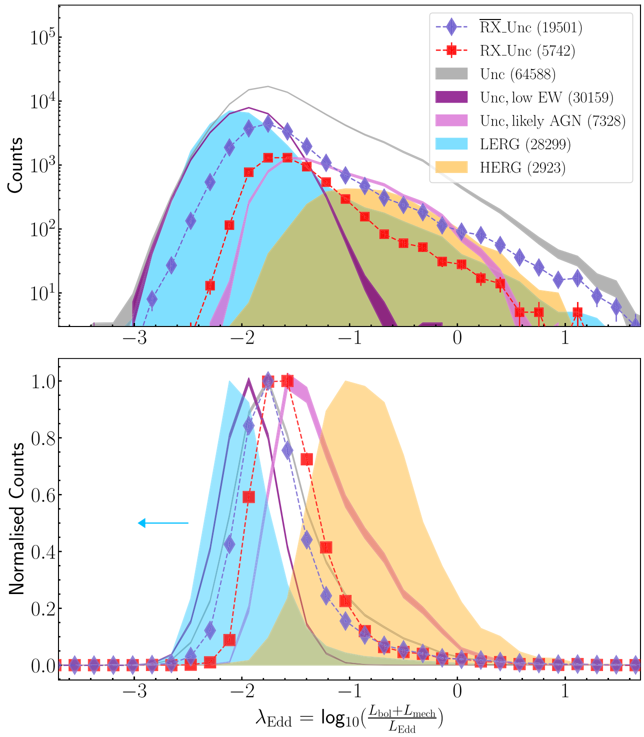

Figure B1. The Eddington-scaled accretion-rate total-count (upper panel) and normalized (lower panel) distributions for radio-excess sources that cannot be divided into LERGs and HERGs (uncertain radio-excess sources, RX_Unc; red) at 90 per cent confidence, non-radio-excess sources that cannot be confidently classified as either SFGs or RQ AGN (uncertain non-radio-excess sources, $\overline{\text{RX}}$_Unc; dark blue), and the remaining unclassified sources (Unc; grey), further split into low-EW (purple) and 'likely' AGN (magenta) subsets. The distributions of LERGs (light blue) and HERGs (orange) are also shown for reference. The corresponding Poisson (upper panel) and binomial (lower panel) uncertainties are shown as error bars and shaded regions, respectively. The number of sources in each class is given in the legend in the upper panel; these numbers differ from Table 2 as not all sources have reliable stellar mass estimates.

per cent confidence level. The figure shows both the total-count and normalized distributions, allowing comparison of both the absolute numbers and relative distribution shapes of the different populations. We note that 42 338 of the 251 413 sources in our sample (corresponding to 19, 28, 21, 20, and 5 per cent of the Unc, RX_Unc, $\overline{\text{RX}}$_Unc, LERG, and HERG populations, respectively) do not have reliable stellar-mass estimates from CIGALE and are therefore excluded from the accretion-rate analysis presented in this work. These sources do not appear to occupy a distinct region of parameter space in redshift, $L_{144\,\text{MHz}}$, or [O III] λ5007 luminosity.

We find that the overall Unc distribution is more similar to that of the $\overline{\text{RX}}$_Unc population, suggesting that the fully unclassified sources are likely dominated by systems without a strong radio excess and therefore predominantly by SFG-like galaxies. However, if this population were associated with radio-excess systems, we could investigate possible subsets within the Unc population. One such subset consists of systems with weak emission lines, analogous to the weak-lined LERGs discussed previously, but without requiring a radio excess. Specifically, we select sources with $\text{EW}_{[\text{O\,III}]} < 3$ Å in more than 90 per cent of the Monte Carlo realizations (EW_FRAC>0.90). This criterion is satisfied by approximately 50 per cent of the Unc population and the associated Eddington-scaled accretion-rate distribution of this population appears more LERG-like (see Fig. B1). In addition, Fig. 8 shows that essentially no SFGs are identified at $z > 0.5$, $\log_{10}(L_{144\text{MHz}}/\text{W Hz}^{-1}) \gtrsim 24$, and $\log_{10}(M_*/\text{M}_\odot) \gtrsim 11$, a parameter space typically associated with radio AGN (e.g. J. Sabater et al. 2019). This 'likely' AGN subset represents approximately ∼ 10 per cent of the Unc population and, as can be seen in Fig. B1, its λ_{Edd} distribution peaks near the transition between the LERG and HERG populations ($\lambda_{\text{Edd}} \sim -1.5$), similar to the RX_Unc population, but also extends to higher Eddington-scaled accretion rates, suggesting that it comprises a mixture of sources with intermediate properties and more HERG-like systems. The sources near the transition likely represent objects with intermediate properties, as is also apparent from the RX_Unc stacked spectra presented in Section 6.2.

Overall, this analysis suggests that excluding the fully unclassified population is unlikely to qualitatively change the main conclusions of this work. Instead, including these sources would primarily increase the sample size of the existing populations, with most Unc sources appearing consistent with LERG-like systems and a smaller subset occupying the transition region between the LERG and HERG populations or extending into the HERG regime.

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