




































Little Red Dots host Black Hole Stars: A unified family of gas-reddened AGN revealed by JWST/NIRSpec spectroscopy

Anna de Graaff ^{1,2*}†, Raphael E. Hviding ¹, Rohan P. Naidu ^{‡3}, Jenny E. Greene ⁴, Tim B. Miller ⁵, Joel Leja ^{6,7,8}, Jorryt Matthee ⁹, Gabriel Brammer ¹⁰, Harley Katz ^{11,12}, Rachel Bezanson ¹³, Leindert A. Boogaard ¹⁴, Sownak Bose ¹⁵, John Chisholm ^{16,17}, Nikko J. Cleri ^{6,7,8}, Pratika Dayal ^{18,19,20}, Robert Feldmann ²¹, Yoshinobu Fudamoto ²², Seiji Fujimoto ^{19,23}, Lukas J. Furtak ^{17,16}, Karl Glazebrook ²⁴, Rashmi Gottumukkala ¹⁰, Kasper E. Heintz ^{10,25}, Vasily Kokorev ^{16,17}, Ivo Labbe ²⁴, Michael V. Maseda ²⁶, Ian McConachie ²⁶, Themiya Nanayakkara ²⁴, Erica Nelson ²⁷, Przemysław Nowaczyk ²⁸, Pascal A. Oesch ^{25,10}, Hans-Walter Rix ¹, David J. Setton ⁴, § Alberto Torralba ⁹, Fabian Walter ^{1,29}, Bingjie Wang ⁴, ¶ Andrea Weibel ²⁵, Arjen van der Wel ³⁰

(Affiliations can be found after the references)

1 December 2025

ABSTRACT

We use the DAWN JWST Archive to construct and characterise a sample of 116 little red dots (LRDs) across $2.3 < z < 9.3$, selecting all sources with v-shaped UV-optical continua from NIRSpec/PRISM spectra and compact morphologies in NIRCам/F444W imaging. We show that LRD continuum spectra are ubiquitously well described by modified blackbodies across $\sim 0.4 - 1.0 \mu\text{m}$, with typical $T \sim 5000 \text{ K}$ or $\lambda_{\text{peak}} \sim 0.65 \mu\text{m}$ across 2 dex in luminosity, and a tail toward $T \sim 2000 \text{ K}$. LRDs therefore trace a locus in the Hertzsprung-Russell diagram that is directly analogous to stars on the Hayashi track, strongly supporting the picture that LRDs are AGN embedded in thermalised dense gas envelopes in approximate hydrostatic equilibrium. Hotter LRDs with $\lambda_{\text{peak}} < 0.65 \mu\text{m}$ typically have strong Balmer breaks, redder UV slopes and high optical luminosities (L_{5100}); other LRDs show weak or no Balmer breaks at all, and wide variety in β_{UV} and L_{5100} . Crucially, we demonstrate that the UV-optical continuum shapes and luminosities are strongly linked to the $H\alpha$, $H\beta$, $[\text{O III}]$ and O I line properties. There is a tight linear relation between the $H\alpha$ and optical continuum luminosities, as well as $H\alpha$ and O I_{18446} , indicating that Balmer, O I and optical emission must primarily be powered by the same source. The Balmer decrement increases strongly toward higher $L_{H\alpha}$, L_{5100} and Balmer break strength, providing key evidence for luminosity-dependent effects of collisional (de-)excitation and resonant scattering in the gaseous envelopes. In contrast, we show that $[\text{O III}]$ emission likely originates from star-forming host galaxies, and that its strong correlation with Balmer break strength arises naturally from variation in the AGN-to-host ratio. Our work presents an empirical description of the nature and structure of LRDs, defining a new benchmark for ongoing LRD model developments.

Key words: Galaxies: evolution – Galaxies: active – quasars: general – Hertzsprung-Russell and colour-magnitude diagrams

1 INTRODUCTION

Understanding the physical nature of Little Red Dots (LRDs), compact high-redshift sources with v-shaped UV-optical continua found by the James Webb Space Telescope (JWST), remains a major open challenge. LRDs were originally proposed to be either dust-reddened active galactic nuclei (AGN; e.g. Matthee et al. 2024b; Labbe et al. 2025; Kocovski et al. 2024), motivated by their red rest-optical con-

tinua and the discovery of broad Balmer lines in a high fraction of LRDs ($> 80\%$; Greene et al. 2024; Hviding et al. 2025), or massive post-starburst or dusty star-forming galaxies (e.g. Labbé et al. 2023; Pérez-González et al. 2024; Wang et al. 2024), due to the frequent detection of Balmer breaks (Setton et al. 2024), weak X-ray emission (Ananna et al. 2024; Yue et al. 2024), and lack of hot dust emission in the rest near-infrared (e.g. Williams et al. 2024b; Wang et al. 2025b; Setton et al. 2025). However, recent work has shown that both these models are likely incorrect or at least incomplete, as neither can explain the spectral energy distributions (SEDs) of the most extreme LRDs discovered to date (de Graaff et al. 2025a; Naidu et al. 2025a) nor the lack of emission from cold and hot dust in the mid- to far-

* anna.de_graaff@cfa.harvard.edu

† Clay Fellow

‡ NASA Hubble Fellow, Pappalardo Fellow

§ Brinson Prize Fellow

¶ NHFP Hubble Fellow

infrared in a broad range of LRDs (Casey et al. 2024, 2025; Akins et al. 2024; Setton et al. 2025; Xiao et al. 2025).

A new picture has emerged instead, first proposed by Inayoshi & Maiolino (2025), in which an accreting massive black hole is not reddened by surrounding dust, but by absorption in dense neutral hydrogen gas ($n_{\text{H}} \sim 10^{9-11} \text{ cm}^{-3}$, $N_{\text{H}} \sim 10^{24-26} \text{ cm}^{-2}$) with high, near-unity covering fraction. Attempts at 1D models of such a “black hole star” (BH*) scenario using the spectral synthesis code Cloudy (Ferland et al. 2017) have shown that the resulting model spectra can reproduce key features of LRDs, such as very strong Balmer breaks and high equivalent-width Balmer emission lines (e.g. Ji et al. 2025a; Naidu et al. 2025a; de Graaff et al. 2025a; Taylor et al. 2025). The presence of dense gas is also naturally able to explain absorption features observed in the Balmer and He I emission lines (Matthee et al. 2024b; Wang et al. 2025b; Juodžbalis et al. 2024; D’Eugenio et al. 2025a; Naidu et al. 2025a). Moreover, radiative transfer effects such as electron scattering (Rusakov et al. 2025; Chang et al. 2025) or resonant scattering (Naidu et al. 2025a; Chang et al. 2025) in a dense medium could contribute to the broadening of the Balmer lines. Finally, in this picture the dense gas, provided the column density $\gtrsim 10^{25} \text{ cm}^{-2}$, may be Compton thick and absorb emitted X-rays, reconciling the observed X-ray weakness. Alternatively, the low X-ray luminosities could be linked to high black hole accretion rates, as proposed in some super-Eddington models (e.g. Pacucci & Narayan 2024; Inayoshi et al. 2024; Lambrides et al. 2024), possibly facilitated by the geometry of the gaseous envelopes.

Importantly, the interpretation of these various observations has been largely phenomenological, that is, none of the aforementioned studies self-consistently model the AGN properties, the structure of the gaseous envelope, and its sustained accretion onto the black hole. From a theoretical perspective, the existence of black holes embedded in dense gas has been discussed in the context of the direct collapse of the first massive stars and the subsequent rapid growth of black hole seeds (e.g. Begelman et al. 2006; Alexander & Natarajan 2014; Coughlin & Begelman 2024). The quasistar model proposed by Begelman et al. (2008) is qualitatively most similar to the scenario discussed for LRDs (Inayoshi & Maiolino 2025; Ji et al. 2025a; Naidu et al. 2025a; de Graaff et al. 2025a), but with the key difference that the envelope is massive compared to the black hole mass ($\gtrsim 10 \times M_{\text{BH}}$; as opposed to the low-mass envelopes suggested by, e.g., Inayoshi & Maiolino 2025; Naidu et al. 2025a) and consists of a radiative layer and outer photosphere.

Begelman & Dexter (2025) recently used analytical arguments to propose that LRDs may be the stable (tens of Myr) final stage in the evolution of quasistars, as they suggest that this late phase naturally produces red rest-optical continua and broad emission lines from electron scattering in the photosphere. This is similar to the work of Kido et al. (2025), who argued that LRDs are black holes residing in massive, optically thick envelopes and demonstrate that this results in photospheres near the Hayashi limit, with temperatures $T \sim 5000 - 7000 \text{ K}$ that likely resemble the red rest-optical spectra of observed LRDs. To date, only Liu et al. (2025) have made detailed predictions for the expected continuum SEDs (from radiation transport calculations) for super-Eddington accreting black holes embedded in massive envelopes ($M_{\text{env}} \sim M_{\text{BH}}$). The spectra predicted from these models match observed LRD spectra well to first order, as they show red rest-optical continua and strong Balmer breaks.

Although these new observational insights and rapidly developing models are highly promising, key open questions on the structure of the gas and the properties of the central engines remain. A major focus has been to explain the continuum properties of individual LRDs, especially those with strong Balmer breaks (e.g. Ji et al.

2025a; de Graaff et al. 2025a; Naidu et al. 2025a), as well as a recently discovered local analogue (Lin et al. 2025; Ji et al. 2025b). Whether the proposed gaseous envelope models can also describe the rest-optical continua of a broader population of LRDs (Hviding et al. 2025) is still unclear, as is the physical connection between the UV and optical continuum properties. Furthermore, the Balmer emission lines have sparked a separate debate. The kinematics of the broad wings may reflect either a broad line region (e.g. Greene et al. 2024; Kocevski et al. 2024; Brazzini et al. 2025) or arise from scattering processes (Rusakov et al. 2025; Chang et al. 2025; Naidu et al. 2025a; Torralba et al. 2025b), with major implications for the inferred black hole masses and growth rates. In addition, the $H\alpha$ equivalent widths exceed that of lower-redshift AGNs by a factor $\gtrsim 2$ (e.g. Lin et al. 2024), raising the question how this can be reconciled with the relatively red accretion disk spectra that have been invoked to explain the observed red rest-optical continua (e.g. Liu et al. 2025; Naidu et al. 2025a; de Graaff et al. 2025a; Wang et al. 2025a).

To begin to tackle these major questions it is critical to map the large parameter space spanned by LRDs. For example, although all LRDs are selected to have v-shaped UV-optical continua, only $\sim 50\%$ have an inflection point at the Balmer limit (Setton et al. 2024). The emission line properties also vary strongly, as some sources show strong nebular emission from a range of metals including ionised Fe (Labbe et al. 2024; Lambrides et al. 2025; D’Eugenio et al. 2025b; Torralba et al. 2025b), while in others only hydrogen and helium lines are strongly detected (de Graaff et al. 2025a; Furtak et al. 2024). The correlation (or lack thereof) between different emission line and continuum properties may thus shed important new light on the physical processes underlying the population of LRDs.

In this paper, we leverage the complete public archive of observations obtained with the JWST/NIRSpec microshutter array (MSA; Ferruit et al. 2022) in order to compile a large sample of 116 LRDs and analyse the population statistics of their key spectral features. The data used and our selection procedure are described in Section 2. We then explore the spectral shapes and continuum properties of the LRD sample in Section 3, revealing that the rest-optical to near-IR continua of LRDs are ubiquitously well described by single-temperature blackbody models. Next, we investigate the properties of strong emission lines ($H\alpha$, $H\beta$, [O III], O I) in Section 4 and demonstrate, for the first time, a tight link between these lines and the continuum properties. Finally, we discuss our findings in the context of the black hole star models and explore the relative importance of host galaxies to the SEDs in Section 5, and conclude in Section 6. We measure magnitudes using the AB system (Oke & Gunn 1983), and, where relevant, adopt a flat Λ CDM cosmology (Planck Collaboration et al. 2020) with $\Omega_{\text{m}} = 0.3$ and $h = 0.7$.

2 DATA & LRD SAMPLE

2.1 NIRSpec spectra

We use all public NIRSpec MSA spectra available in version 4.4 of the DAWN JWST Archive (DJA; Brammer & Valentino 2025)¹, further supplemented with data from the Mirage or Miracle program (GO-5224, PIs Oesch & Naidu; see e.g. Naidu et al. 2025b). Spectra were reduced with msaexp (Brammer 2023a) largely following the methodology described in Heintz et al. (2025) and de Graaff et al. (2025b), using a local background subtraction from nodded exposures where possible. Briefly, version 4 of the DJA improves upon

¹ <https://dawn-cph.github.io/dja/index.html>

version 3 by employing an empirical correction to the wavelength calibration based on the source centroid in the shutter, resolving some of the wavelength offsets between the PRISM and higher-resolution spectra reported in [de Graaff et al. \(2025b\)](#). It also implements new reference files for the flux calibration, which enable spectral extraction beyond the nominal wavelength range (see also [Valentino et al. 2025](#); [Pollock et al. 2025](#)). The 1D spectra were extracted using an optimal extraction ([Horne 1986](#)), and the extraction kernel was subsequently used together with the source position in the shutter to correct for wavelength-dependent slit losses. For point sources this slit loss correction typically performs well (see, e.g., [de Graaff et al. 2025a](#); [Hviding et al. 2025](#)), and we therefore use the spectra without further flux corrections.

All spectra were visually inspected (G. Brammer) to vet the quality of the spectra and automated redshift fitting. We only use spectra with grade=3, which correspond to high-quality data with robustly determined redshifts. This results in a sample of $\sim 17,000$ low-resolution PRISM spectra with grade=3, originating from large surveys including, in alphabetical order, CANUCS ([Sarrouh et al. 2025](#)), CAPERS (GO-6368; PI: Dickinson), CEERS ([Finkelstein et al. 2025](#)), JADES ([Eisenstein et al. 2023](#); [Curtis-Lake et al. 2025](#); [Scholtz et al. 2025](#)), NEXUS ([Shen et al. 2024](#)), NIRSpec GTO-Wide ([Maseda et al. 2024](#)), RUBIES ([de Graaff et al. 2025a](#)), and UNCOVER ([Bezanson et al. 2024](#); [Price et al. 2025](#)), as well as a wide range of smaller programs². For sources from the UNCOVER survey we also obtain magnifications from the lens model of the Abell-2744 cluster of [Furtak et al. \(2023a\)](#) and [Price et al. \(2025\)](#), and use the best-fit values to correct the luminosities measured in Sections 3 and 4. For sources in CANUCS clusters, we obtain best-fit lensing magnification factors from the CANUCS first data release ([Sarrouh et al. 2025](#)).

Many of the sources in the DJA were observed not only with the low-resolution ($R \sim 100$) PRISM, but also at higher resolution ($R > 1000$). Where available, we use these higher resolution NIRSpec spectra in conjunction with the PRISM data for improved emission line fitting in Section 4. Most relevant to this work are the G395M data (and in some cases also G235M, G235H, or G395H) from the RUBIES, JADES, and NIRSpec GTO-Wide programs.

2.2 NIRCам imaging

We obtain JWST/NIRCам image mosaics from the DJA (version 7; at a pixel scale of $0.04''$), which were reduced with `grizli` ([Brammer 2023b](#)) as described in detail in [Valentino et al. \(2023\)](#). We focus only on the F444W filter, which is the most commonly used long-wavelength filter across all JWST fields. Indeed, a very large fraction (85%) of all grade=3 PRISM spectra have available NIRCам/F444W imaging, from a mixture of imaging programs including CANUCS ([Sarrouh et al. 2025](#)), CEERS ([Finkelstein et al. 2025](#)), JADES ([Eisenstein et al. 2023](#)), NEXUS ([Shen et al. 2024](#)), PRIMER (GO-1837; PI: Dunlop), and UNCOVER ([Bezanson et al. 2024](#)), as well as pure parallel programs such as PANORAMIC ([Williams et al. 2024a](#)).

Because some of the spectra were observed prior to NIRCам imaging becoming available (e.g. large parts of the NIRSpec GTO programs were selected based on HST imaging), there may be small mismatches in the astrometric centre used for targeting and the source centroid measured in the F444W image. We therefore cross-match our spectroscopic sample to the photometric catalogues on the DJA

to obtain the photometric source centroids. We then extract image stamps and measure fluxes in circular apertures of radii $0.1''$ and $0.2''$ centred on these photometric centres, which we use in Section 2.3.

2.3 LRD selection

To select our sample of LRDs we follow a strategy inspired by the work of [Hviding et al. \(2025\)](#), who showed that the three defining properties of LRDs – a v-shaped UV-optical continuum, dominant point source morphology in the rest-optical, and broad Balmer lines – are inextricably linked. Crucially, this implies that selecting sources with v-shaped continua and compact rest-optical morphologies will virtually always yield detection of broad Balmer emission lines, at least for sources brighter than $F444W < 26.5$ (100% of sources where data quality permits detection of broad lines; [Hviding et al. 2025](#)). Because the spectroscopic archive of the DJA primarily consists of PRISM data, for which the low resolution makes it difficult to robustly determine line widths, we only impose v-shape and compactness criteria to select LRDs. In what follows, we assert that these sources also have some broad kinematic component in their Balmer lines. We test this assumption in section 2.4 using sources for which medium- or high-resolution spectra are available.

We perform broken power-law fits, joint at the Balmer limit (H_∞), of the form $f_\lambda = \alpha(\lambda/H_\infty)^\beta$ to the PRISM spectra across rest-frame $0.12 - 0.70 \mu\text{m}$ in order to determine the UV slope β_{UV} and optical slope β_{opt} . A region of width 100 \AA is masked around strong emission lines (H I, He I, [Ne III], [O III], and [O II]) in the fitting. In addition, where photometry at rest-UV wavelengths are available we also fit the photometric UV slopes by integrating the model through the NIRCам bandpasses. We implement the broken power-law model in NumPyro ([Phan et al. 2019](#)) and employ Markov chain Monte Carlo (MCMC) sampling using the No-U-Turn Sampler (NUTS; [Hoffman et al. 2014](#)) comprised of one chain with 250 warmup steps and 2500 posterior samples. We select sources with v-shaped continua by requiring all the following criteria:

- (i) $> 95\%$ of all samples satisfy $\beta_{\text{opt}} > 0.0$
- (ii) $> 95\%$ of all samples satisfy $\beta_{\text{UV}} < -0.2$
- (iii) $> 95\%$ of all samples satisfy $\beta_{\text{UV}} - \beta_{\text{opt}} > 0.5$
- (iv) Criteria 2 & 3 must be simultaneously satisfied by β_{UV} measured from either photometry or spectroscopy.

This yields a sample of 247 v-shaped PRISM spectra, the redshift distribution of which is shown in the top panel of Figure 1.

Requiring a v-shape alone selects a large number of dusty star-forming galaxies at cosmic noon. We next impose compactness criteria to select sources for which the rest-optical morphologies are dominated by a point source component. Of the 247 sources, 243 have available NIRCам/F444W imaging. Because the sources are spread across a large number of fields, homogeneous empirical point spread function (PSF) models are not immediately available for all objects, which are necessary to perform detailed morphological fitting. We therefore use two different criteria, the first being a circular aperture flux ratio that was calibrated using PSF models for the selection of point-like objects ([Labbe et al. 2025](#)). Second, where possible (the five CANDELS fields and Abell-2744), we also perform two-component point source and Sérsic profile fitting with `pysersic` ([Pasha & Miller 2023](#)) using the empirical PSFs of [Weibel et al. \(2024\)](#), in the exact same manner as described in [Hviding et al. \(2025\)](#). We hence define compactness as either of

² for a detailed breakdown, see <https://dawn-cph.github.io/dja/blog/2025/05/01/nirspec-merged-table-v4/> and Table A1

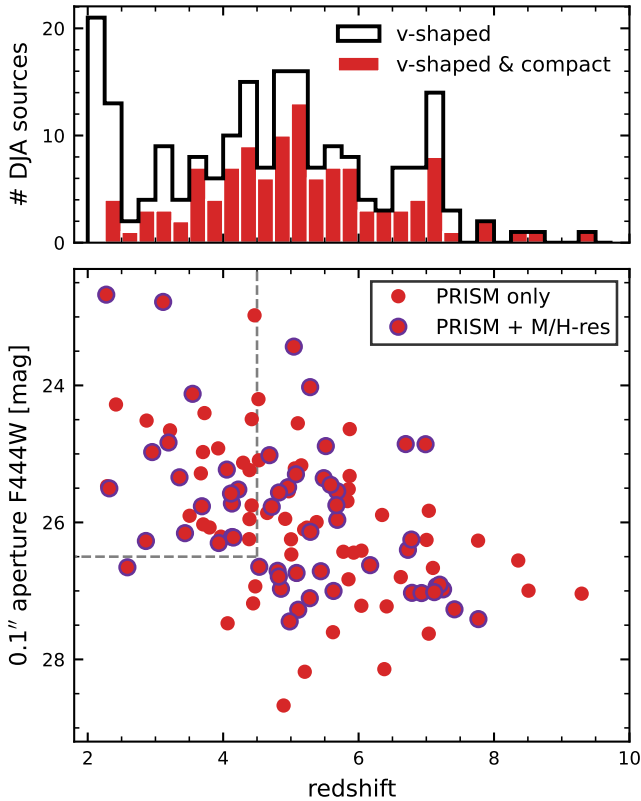


Figure 1. Top: Redshift distribution of all PRISM spectra selected from the public DJA that satisfy the spectroscopic v-shape criteria of [Hviding et al. \(2025\)](#) (black). The sample of v-shaped objects contains both LRDs and a large number of dust-reddened star-forming galaxies at cosmic noon. We define LRDs (116 unique sources; red) as the subset of sources satisfying strict compactness criteria in NIRCcam/F444W imaging. Bottom: NIRCcam/F444W magnitudes measured in circular apertures of radius $0.1''$ (half the MSA slit width) of all LRDs as a function of redshift. Of these sources, 47% (55/116) also have medium or high resolution spectra from NIRSpec. Dashed lines indicate the subsample used for the modified blackbody fitting of Section 3.1.

- $f_{F444W}(0.2'')/f_{F444W}(0.1'') < 1.7$
- the point source component dominates ($> 50\%$) the F444W model flux in $> 95\%$ of all samples.

Of sources that are both v-shaped and compact, 128 (114) satisfy the first (second) criterion, with large overlap between the two selections. In total, our sample consists of 137 v-shaped PRISM spectra with a compact rest-optical morphology, although this still contains some duplicate observations.

We visually inspect the PRISM spectra of the entire sample, removing 3 spectra (6585-59553; a duplicate observation of RUBIES-UDS-40579; RUBIES-UDS-149298) where the location of the chip gap or background subtraction issues prohibit analysis of the key features explored in this paper (i.e. the Balmer break and strong rest-optical lines). Finally, we flag 18 duplicate observations of the same source or of multiple images of the same strongly-lensed source (i.e. the source of [Furtak et al. 2023b](#)). In the following sections, we use the PRISM spectrum with the highest signal-to-noise ratio (S/N).

The redshift and F444W magnitude distribution of the sample of 116 unique LRDs is shown in Figure 1. Approximately half of the sample (55/116) benefits from one or more higher resolution observations in the rest-optical, primarily from RUBIES (33) and the NIRSpec GTO programs (JADES and GTO-WIDE, 20), as well

as GO-4106 (PI: Nelson, 4) and CEERS (1). The full LRD sample spans a broad redshift range of $2.3 < z < 9.3$ (the extremes being the sources of [Juodžbalis et al. 2024](#) and [Taylor et al. 2025](#)), and contains many well-studied LRDs such as A2744-45924 ([Labbe et al. 2024](#)), A2744-QSO1 ([Furtak et al. 2024](#)), RUBIES-BLAGN-1 ([Wang et al. 2025b](#)), *The Cliff* ([de Graaff et al. 2025a](#)), and MoM-BH*-1 ([Naidu et al. 2025a](#)). An overview of the selected sources is provided in Table A1, and we publicly release a complete table of all spectra and measurements made in this work together with this paper³, the column entries of which are described in Table A2.

We stress that both our v-shape and compactness selection criteria are strict, and therefore yield a large, but certainly not complete, sample of LRDs. For the purpose of this work we prioritise purity over completeness in order to gain an understanding of the physical properties of and diversity among LRDs. The incompleteness can be seen in Figure 1, as the sample is skewed toward relatively bright F444W magnitudes. At higher redshifts this is primarily the result of our stringent v-shape selection, which, as discussed in detail in [Hviding et al. \(2025\)](#), requires sufficient S/N to robustly determine whether $\beta_{\text{opt}} > 0$. At lower redshifts the cause of the bias is less clear. Our selection is sufficiently sensitive to detect LRDs to fainter magnitudes, and the apparent cutoff at F444W ~ 26.5 therefore could be physical, translating to a lower luminosity limit of $L \sim 10^{43} \text{ erg s}^{-1}$. However, the large number of spectroscopic programs among the sample have heterogeneous selection functions, making it difficult to assess whether selection bias could also play a role.

Finally, we define a subsample of 36 sources that have coverage out to rest-frame $1 \mu\text{m}$, i.e. $z_{\text{spec}} < 4.5$, and are well-detected, $f_{F444W}(0.1'') < 26.5$ (where the aperture radius corresponds to half the MSA slit width). This sample is central to our discussion in the following sections, because the broad wavelength coverage allows for the blackbody fitting described in Section 3.1.

2.4 Emission line measurements

We use the `unite` ([Hviding 2025](#)) software package to robustly constrain emission line properties from the NIRSpec MSA spectra, described in detail in [Hviding et al. \(2025\)](#). Briefly, `unite` was designed to fit spectra from multiple NIRSpec dispersers simultaneously, whilst rigorously accounting for the undersampling of the NIRSpec line spread function (LSF) as well as wavelength and flux calibration uncertainties ([de Graaff et al. 2025b](#)). For the LSF we assume that of an idealised point source, based on the methods described in Appendix A of [de Graaff et al. \(2024\)](#), rescaled by a nuisance parameter to capture calibration uncertainties in the LSF ([de Graaff et al. in prep.](#)).

The emission model consists of (i) linear continua, (ii) two-component Gaussians (broad + narrow) for the Balmer lines, (iii) single-component Gaussians (narrow) for the [O III] $\lambda\lambda 4959, 5007$ doublet, and (iv) single-component Gaussians for the O I $\lambda 6300$ and O I $\lambda 8446$ lines. We do not include the [N II] doublet, as this to date has not been strongly detected in LRDs. The line widths of the broad Balmer components and narrow Balmer and [O III] lines are tied together respectively, and the flux ratio of the narrow [O III] doublet is fixed to 1:2.98. The O I lines are allowed to vary in width independently of the Balmer and [O III] lines. In detail, the emission line kinematics may differ for each line, and the broad lines may not be Gaussian (e.g. [Rusakov et al. 2025](#)), or there may be absorption features in the

³ <https://doi.org/10.5281/zenodo.17665942>

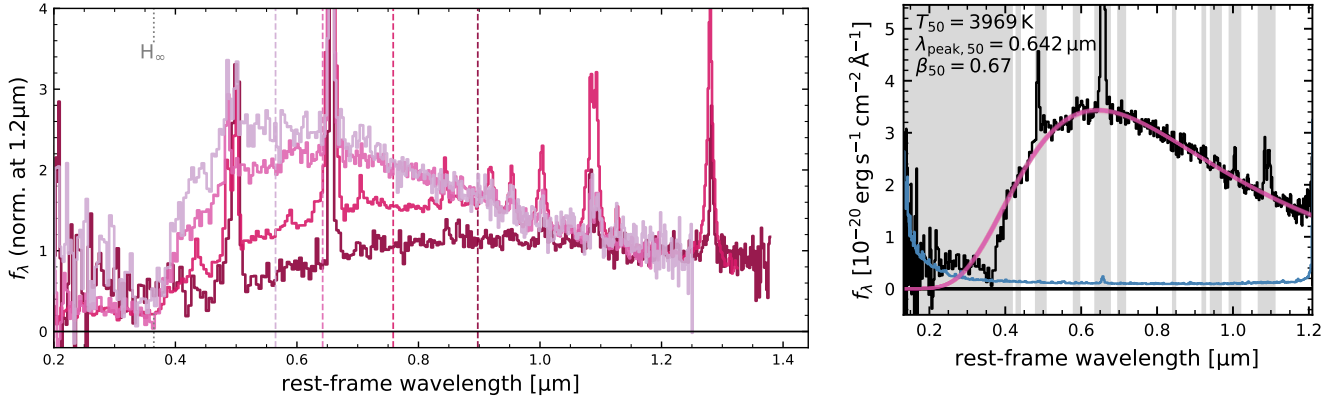


Figure 2. Left: NIRSpec/PRISM spectra (normalised at $1.2 \mu\text{m}$) of four RUBIES LRDs showcasing their diverse spectral shapes. From light to dark: RUBIES-UDS-144195 (Hviding et al. 2025), *The Cliff* (de Graaff et al. 2025a), RUBIES-BLAGN-1 (Wang et al. 2025b), and RUBIES-UDS-175975. The colour coded vertical dashed lines indicate the peak wavelengths of the modified blackbody fits to the spectra. Right: Example of the 3-parameter modified blackbody fitting to the PRISM spectrum of *The Cliff*, with its error spectrum shown in blue. Regions around strong emission lines and the region blueward of λ_{ν} , marked in grey, were masked in the fit. Model lines (pink) represent 100 random draws from the parameter posteriors.

Balmer lines (e.g. Matthee et al. 2024b). For the majority of the sample we lack either the spectral resolution or S/N to robustly constrain individual profiles of Balmer and metal lines in such detail, but we find that the described setup yields good fits overall and therefore accurately recovers total line fluxes and equivalent widths (see examples in Appendix B of Hviding et al. 2025). We defer a detailed analysis of the line kinematics and presence of absorbers for high S/N spectra to a future paper.

We incorporate all available spectra where possible, provided that the observations were taken in the same slit. That is, we fit simultaneously to PRISM and higher-resolution spectra where available, but do not combine spectra obtained in different masks where the position of the source in the shutter differs. The priors are as described in Hviding et al. (2025), with the exception of the broad line full width at half maximum ($\text{FWHM}_{\text{broad}}$), which is set to a uniform distribution in the range $[\text{FWHM}_{\text{narrow}} + 100, 5000] \text{ km s}^{-1}$. The prior for the O I lines is set to $[0, 2000] \text{ km s}^{-1}$.

In addition, we fit all objects without the broad component for the Balmer lines in order to test our previous assumption that our selected LRDs have broad lines. We perform model comparison using the Widely Applicable Information Criterion (WAIC Watanabe & Opper 2010) and the detection threshold described in detail in Hviding et al. (2025). For 49 of the 55 sources with available medium- or high-resolution spectroscopy (in any grating) the spectra cover the $H\alpha$ line. We significantly detect broad Balmer lines in 48/49 (98%), highlighting the high accuracy of our spectrophotometric LRD selection approach. Crucially, these sources with available higher-resolution spectra are broadly distributed in redshift-magnitude space (Figure 1), and we can therefore expect similarly high purity for the sources where only PRISM spectra are available.

Finally, for comparison to the LRD population we also fit the emission lines of the broad line sample of Hviding et al. (2025) that are not LRDs, consisting of 37 sources from RUBIES in the redshift range $3.3 < z < 7.2$. We use the same model as described above, but include a narrow [N II] doublet with fixed flux ratio 1:2.95.

3 UV TO NEAR-IR CONTINUUM SHAPES

The PRISM spectra provide a detailed view of the SED shapes across the rest-UV to optical, as well as the rest near-IR for sources at intermediate redshifts ($z \sim 2 - 4$). Visual inspection of the sample reveals great variety in the spectral shapes of LRDs, demonstrated in Figure 2 (left panel). The sample contains a large number of sources with strong Balmer breaks, as discussed extensively in previous literature (see references in Section 1). Intriguingly, however, their spectral shapes differ significantly, as the rest-optical continua peak at different wavelengths and the rest-UV ranges from flat to blue UV slopes. Other sources (darkest line in Figure 2) do not show Balmer breaks; their rest-optical continua peak in the near-IR, and the location of the ‘v’ is redward of the Balmer limit. For the reddest source among our sample, RUBIES-EGS-38434, the v is located at $\sim 5700 \text{ \AA}$ (shown in Appendix B).

We also find that, provided the spectra cover rest $> 1 \mu\text{m}$, the LRDs consistently show a flattening or turnover (in f_{λ}) at optical to near-IR wavelengths. In other words, the LRDs appear ubiquitously blue in the near-IR, as was also shown for small samples and individual objects (Williams et al. 2024b; Wang et al. 2025b; Setton et al. 2025; de Graaff et al. 2025a). We quantify these observations in the following subsection by fitting modified blackbody models to the spectra.

3.1 Modified blackbody fitting

Motivated by the shape of the optical to near-IR continua, we fit 3-parameter modified blackbody models of the form

$$f_{\nu} = A_{\text{tot}} B_{\nu}(T) \left(\frac{\nu}{\nu_0} \right)^{\beta_{\text{MBB}}}, \quad (1)$$

where A_{tot} is a scale parameter, and B_{ν} is the Planck function that depends on temperature T . The blackbody function is modified by a power law of slope β_{MBB} and pivot frequency $\nu_0 = c/(5500 \text{ \AA}) \approx 5.45 \times 10^{14} \text{ Hz}$ (where c is the speed of light). The purpose of the pivot frequency is only to reduce the dynamic range in A_{tot} , and the specific value chosen does not affect the posterior distribution of β_{MBB} .

The power-law modification was inspired by other astrophysical contexts where it has been used to account for frequency dependence

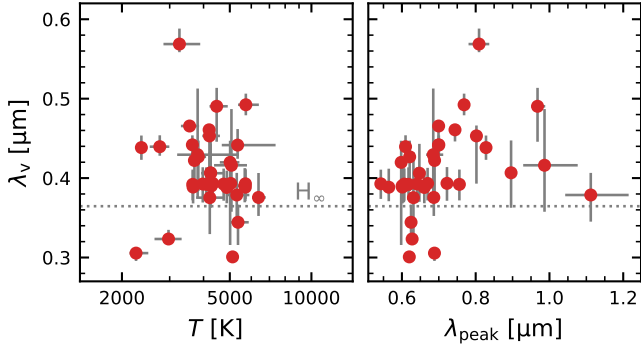


Figure 3. Location of the v-shape turnover wavelength λ_v versus modified blackbody temperature and λ_{peak} . For most LRDs λ_v is near the Balmer limit (there is a small systematic bias, see main text), but cooler LRDs with long peak wavelengths can have substantially larger λ_v than this.

of the dust absorption cross section (Hildebrand 1983). The mathematical effect of this modification is a broadening or narrowing of the shape of the SED with respect to a true blackbody. Importantly, the deviation from a blackbody can be caused by several effects including, but not limited to, dust attenuation. For example, a distribution in temperatures (i.e. a typical stellar population) would result in a broader SED shape. In the context of this work, we therefore use the power-law modification to the blackbody primarily as a nuisance parameter in order to establish how close the SED shapes of LRDs are to single-temperature blackbodies, and discuss the physical interpretation of β_{MBB} in further detail in Section 5.1.

A perfect single-temperature blackbody has $\beta_{\text{MBB}} = 0$. Deviations from $\beta_{\text{MBB}} = 0$ primarily affect the Rayleigh-Jeans limit, where $B_\nu \propto \nu^2$ (or $B_\lambda \propto \lambda^{-4}$). That is, SEDs broader than perfect blackbodies will have $\beta_{\text{MBB}} < 0$, and similarly $\beta_{\text{MBB}} > 0$ for SEDs that are narrower in shape. There is a well-known strong degeneracy between β_{MBB} and T (Shetty et al. 2009b), and it is often difficult to precisely constrain the two, especially in the case of low S/N (Shetty et al. 2009a). Instead of temperature, we therefore additionally compute the peak wavelength of the model (λ_{peak}) which we find is more stable against variations in β_{MBB} and T .

Finally, we caution that in order to robustly constrain the three parameters of the model, spectral coverage both blueward and redward of the peak wavelength are crucial. For practical purposes we restrict the analysis of our fits to the $z < 4.5$ sample highlighted in Figure 1, the spectra of which (for all but three sources) reach redward of λ_{peak} . Strong emission lines are masked in the fit (the Balmer and Paschen series, [O III] doublet, O I λ_{8446} , and optical to near-IR He I lines).

We also mask the blue side of the v-shaped continuum. To do so, we first reperform our v-shape fitting to the PRISM spectra of Section 2.3, but instead of fixing the transition wavelength to H_∞ , we fit for λ_v as a free parameter with a uniform prior in the range [2500–6000] Å. The distribution in λ_v is shown in Figure 3. We find some sources with $\lambda_v < 3500$ Å as well as an overall slight offset with respect to the Balmer limit. As also discussed by Setton et al. (2024), this is a result of the fact that SEDs with (strong) Balmer breaks and associated damping wings are not well described by the simplistic broken power-law model. However, for $\sim 30\%$ of the sample $\lambda_v \gtrsim 4000$ Å and we find that this is not an artefact, but an accurate description of the SED (a gallery of SEDs and median λ_v values can be found in appendix Figure B3). We therefore mask all wavelengths blueward of rest 4200 Å or the 84th percentile of λ_v , whichever is greater, for the rest-optical blackbody fitting. We perform MCMC fitting with emcee (Foreman-Mackey et al. 2013), using a uniform

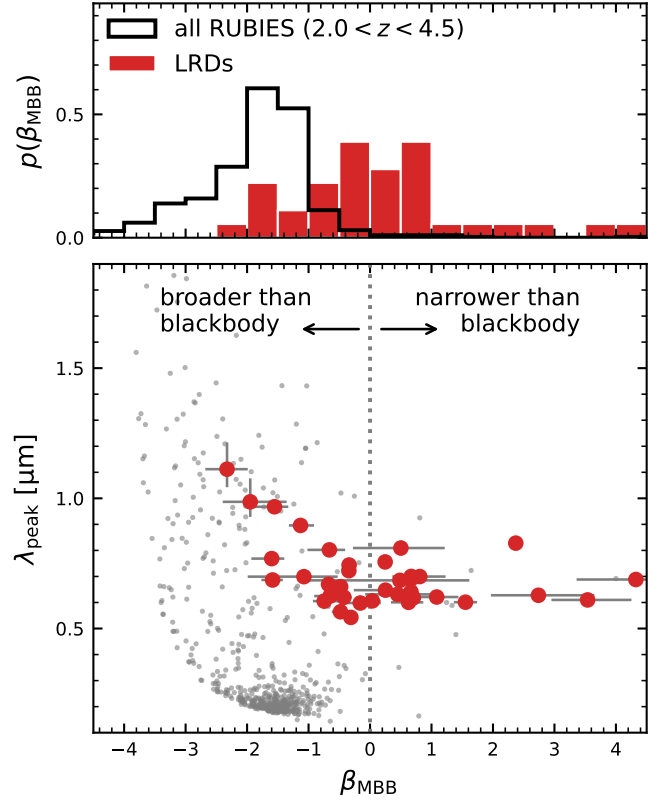


Figure 4. Distribution of the power law slopes and peak wavelengths from modified blackbody fits to the LRD sample restricted to $z < 4.5$ (red). For reference, the results of fits to the complete RUBIES galaxy sample at $2 < z < 4.5$ and $F444W < 26$ are shown in grey. A slope of $\beta_{\text{MBB}} = 0$ corresponds to a perfect single-temperature blackbody. Whereas normal galaxies typically have broad SED shapes, as they consist of a mixture of stellar types (and hence temperatures), many LRDs are well approximated by single-temperature blackbodies. The peak wavelengths of galaxies span a broad range due to variation in recent star formation activity and dust attenuation, but LRDs have a characteristic continuum $\lambda_{\text{peak}} \sim 0.6 - 0.7 \mu\text{m}$.

prior for β_{MBB} ($[-5, 5]$) and loguniform priors for T ($[10^0, 10^7]$ K) and A_{tot} . The 3-parameter modified blackbody models fit the LRD sample very well, an example of which is shown in Figure 2 (figures for the complete sample can be found in Appendix B).

3.2 HR diagram of LRDs

Figure 4 shows the distribution in β_{MBB} and peak wavelength (a corresponding figure with temperature can be found in Appendix B). For comparison, we have also run modified blackbody fits for ~ 600 bright ($F444W < 26$) galaxies at $2.0 < z < 4.5$ from RUBIES that are not LRDs over the same rest wavelength range $\sim 0.4 - 1.0 \mu\text{m}$. For typical galaxies $\beta_{\text{MBB}} \lesssim -1$, because normal stellar populations are comprised of a distribution of stellar temperatures, resulting in a broadening of the SED compared to a single-temperature star. The peak wavelengths span a broad range, but are short ($\sim 0.2 \mu\text{m}$) for star-forming galaxies as well as post-starburst galaxies that are dominated by O, B and A stars, and longer for galaxies with increasingly stronger dust attenuation. A small number of grey points lie at $\beta_{\text{MBB}} > 0$, but we find these sources tend to be large galaxies, resulting in spectral artefacts due to self-subtraction in the background removal of the reduction pipeline.

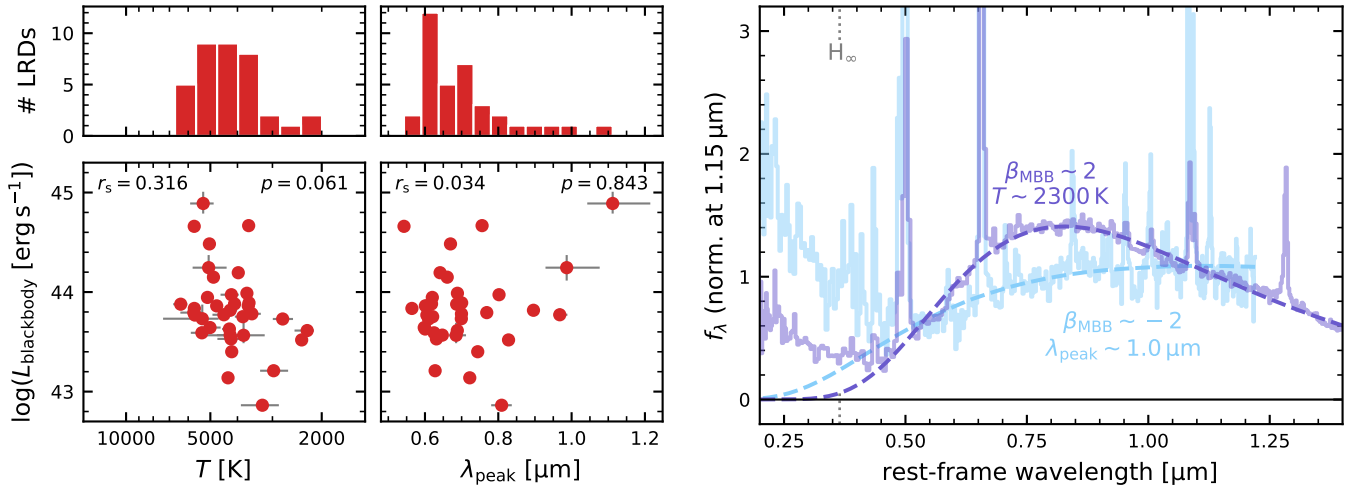


Figure 5. HR diagram of LRDs at $z < 4.5$ for the modified blackbody temperature (left) and peak wavelength (middle). LRDs span a wide range in luminosity, with a lower luminosity limit that may be part physical and part driven by the imposed magnitude limit and depth of the PRISM spectra. The rest-optical continua of LRDs typically peak at a wavelength of $\sim 0.65 \mu\text{m}$, corresponding to a temperature of $\sim 5000 \text{ K}$ for a perfect blackbody ($\beta = 0$), although some systems are as cold as $\sim 2000 \text{ K}$. The right panel shows examples of extremes in the HR diagram: a remarkably cold source ($T \sim 2300 \text{ K}$, $\beta_{\text{MBB}} \sim 2$, $\lambda_{\text{peak}} \sim 0.83 \mu\text{m}$; UNCOVER-20698), and a source that peaks in the near-IR ($\lambda_{\text{peak}} \sim 1.05 \mu\text{m}$, $\beta_{\text{MBB}} \sim -2$; CAPERS-COSMOS-30440).

The LRDs clearly stand out from the normal galaxy population, as the average $\beta_{\text{MBB}} \sim 0$ and the distribution in peak wavelength is strongly clustered around a value of $0.6 - 0.7 \mu\text{m}$. In other words, many LRDs are close to true single-temperature blackbodies, with near-IR continua that are near the Rayleigh-Jeans limit. The scatter about β_{MBB} is partially physical, but we also note that in our modelling we have not accounted for wavelength-dependent slit losses beyond the nominal correction made by the reduction pipeline, and any imperfections in this correction are therefore absorbed by β_{MBB} . By comparing the inferred values of β_{MBB} for sources with duplicate observations, as well as fits to version 3 and 4 of the DJA, we find variations $\Delta\beta_{\text{MBB}} < 1.0$ for individual sources due to systematics in the data reduction and spectral extraction. Crucially, however, the sample distribution of median β_{MBB} values does not vary strongly.

We identify two types of outliers in Figure 4. First, there are three sources with $\lambda_{\text{peak}} \approx 1.0 \mu\text{m}$ and $\beta_{\text{MBB}} \lesssim -2$, which would imply that their SEDs are red in the rest near-IR. However, these sources (one of which is shown in blue in the right panel of Figure 5) lack coverage redward of λ_{peak} in the near-IR. As a result, small changes to, for instance, the mask used have a significant effect on the inferred β_{MBB} . Further constraints on the IR SED are therefore needed to robustly constrain their spectral shapes and temperatures.

The second group of outliers at $\beta_{\text{MBB}} > 1$ is intriguing, as these sources have remarkably narrow SED shapes with v-shapes that lie redward of the Balmer limit. An example is shown in purple in Figure 5. These objects also correspond to some of the coldest temperatures in our sample, reaching $\sim 2000 \text{ K}$ or $\lambda_{\text{peak}} \sim 0.8 \mu\text{m}$. Such a subpopulation of very cool LRDs has, to our knowledge, not been reported thus far, and provides insight into the physical properties of the gaseous structures of LRDs (Section 5.1).

Given the close analogy to stars, we show an LRD version of the Hertzsprung-Russell (HR) diagram in Figure 5. We use the integrated luminosity of the modified blackbody fits, which is close to the bolometric luminosity of LRDs (Greene et al. 2025). Wien’s law approximately holds, as $\beta_{\text{MBB}} \sim 0$, such that the characteristic peak wavelengths of $\sim 0.6 - 0.7 \mu\text{m}$ correspond to typical temperatures $T \sim 4000 - 6000 \text{ K}$, which agrees well with values estimated for the photosphere models of Liu et al. (2025), Kido et al. (2025), and

Begelman & Dexter (2025) (see Section 5.1 for further discussion). Interestingly, despite the wide range in luminosity spanned by the sample (2 dex), the temperature and λ_{peak} do not appear to depend strongly on the luminosity, although the very coldest sources ($T < 3000 \text{ K}$) are all less luminous than the average (i.e., $L < 10^{44} \text{ erg s}^{-1}$).

3.3 Linking the UV and optical

We have demonstrated that the SEDs of LRDs are well approximated by modified black bodies across $\sim 0.4 - 1.0 \mu\text{m}$. Next, we explore whether the optical properties are correlated with the spectral shapes at $< 0.4 \mu\text{m}$. The UV emission from LRDs has been ascribed to star-forming host galaxies in some studies, as many LRDs show spatially extended UV emission (e.g. Rinaldi et al. 2024; Torralba et al. 2025a), while others have found strong UV emission lines in LRDs that can only reasonably be attributed to AGN emission (Labbe et al. 2024; Akins et al. 2025). To mitigate the emission from host galaxies, we here consider only the UV emission in the aperture of the slit, extracted with the same Gaussian kernel used for the rest-optical and near-IR.

Following de Graaff et al. (2025a), we calculate the Balmer break strength across $[0.362, 0.372] \mu\text{m}$ and $[0.40, 0.41] \mu\text{m}$ from the PRISM spectra, estimating uncertainties using random draws of the error spectra. Figure 6 shows that the Balmer break correlates positively with the optical continuum luminosity (L_{5100} , measured from the blackbody models). Moreover, there is a weak positive correlation between the UV slope (as measured in Section 3.1) and Balmer break, such that sources with redder UV slopes have stronger Balmer breaks. In contrast, the peak wavelength obtained from the modified blackbody fits correlates negatively with Balmer break strength, the only exception with a strong Balmer break yet long λ_{peak} being RUBIES-BLAGN-1 (Wang et al. 2025b). There is great scatter in λ_{peak} for sources with weak or no Balmer breaks. For a large fraction of our LRD sample $\lambda_{\text{v}} > H_{\infty}$ and hence these sources do not show a Balmer break, resulting in break strengths $\lesssim 1$.

Crucially, the reddest sources in the UV do not correspond to the reddest sources at longer wavelengths, as would be expected in

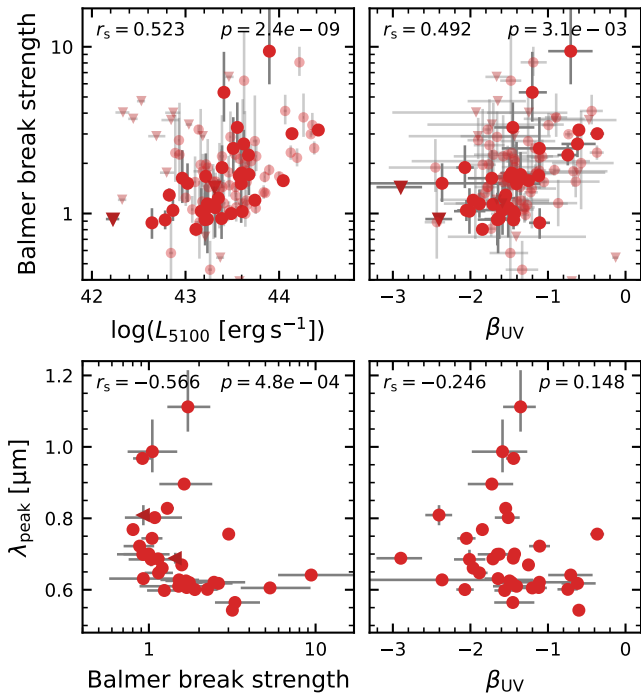


Figure 6. Distribution of parameters describing the continuum SEDs of LRDs: Balmer break strengths, optical luminosities, peak wavelengths, and UV slopes. Results from modified blackbody fitting (bottom) are shown only for the subset of LRDs with $z_{\text{spec}} < 4.5$; the upper panels also include the higher redshift subsample. Spearman correlation coefficients and p-values are listed in each panel. Upper limits (triangles) on the Balmer break strength, corresponding to the 84th percentile, are shown for spectra with low S/N at rest $\sim 0.36 \mu\text{m}$. The Balmer break strength correlates significantly with the luminosity and shape of the UV-optical SED, reflecting a systematic variation in the SEDs across the LRD population.

case of dust reddening. Sources with the reddest UV slopes have comparatively short λ_{peak} (bottom right panel of Figure 6), whereas sources with bluer UV slopes have a broad range in λ_{peak} . Coupled with the correlation between λ_{peak} and the Balmer break strength, this lends further support for the interpretation that the SEDs of LRDs are not reddened by dust, but by the effects of dense, optically-thick gas. First, a large population of $n = 2$ hydrogen in such dense gas absorbs UV light; second, at high optical depth high-energy photons are scattered to redder wavelengths. Finally, we mention, but do not show here, that we do not find any correlation between the other parameters of the modified blackbody (β_{MBB} , T) and UV slope, although the lack of correlation with temperature may simply be due to the limited dynamic range.

Altogether these correlations point to systematic variations in the spectral shapes of LRDs across the broad population, and as a function of luminosity. Notably, although the scatter in some regimes is large, one subset stands out among the different panels of Figures 6 and B2. The hottest sources (i.e. $\lambda_{\text{peak}} < 0.65 \mu\text{m}$) all tend to have high L_{5100} , strong Balmer breaks, and comparatively red UV continua. The most extreme examples of this are the sources reported by Labbe et al. (2024), Naidu et al. (2025a), and de Graaff et al. (2025a). We discuss the luminosity dependence as well as the origin of the UV emission and possible host galaxy contribution in further detail in Section 5.

4 RELATION BETWEEN LINE AND CONTINUUM PROPERTIES

The diversity in continuum properties found in Section 3 also extends to the emission lines. In this section, we focus on characterising a set of key strong optical emission lines ($\text{H}\beta$, $[\text{O III}]$, $\text{H}\alpha$, O I), and show that their properties can be linked to the shapes of the UV to near-IR continuum shapes.

4.1 $\text{H}\alpha$ emission

We first explore the properties of the $\text{H}\alpha$ line, focusing on the total luminosity and equivalent width (EW). We do not consider the broad and narrow components separately, due to the limited spectral resolution for a large fraction of the sample, and the fact that individual components of the double Gaussian model might not be physically meaningful if the true line profile is Lorentzian or exponential. However, observations with the medium-resolution ($R \sim 1500$) NIRCam grism have shown that the $\text{H}\alpha$ lines of LRDs are consistently dominated by the broad component (forming $\sim 70\%$ of the total luminosity; Lin et al. 2024). Emission from potential star-forming host galaxies is thus subdominant to the emission from the central engine and unlikely to affect population trends – our primary interest – but may introduce scatter.

Figure 7 shows there is a tight relation between the $\text{H}\alpha$ luminosity and optical continuum luminosity, L_{5100} . Such a relation had long been established for lower redshift AGN (e.g. Greene & Ho 2005), but we here demonstrate this correlation holds also for LRDs (both the full sample and $z < 4.5$ subsample), as well as the sample of high-redshift broad-line sources that are not LRDs of Hviding et al. (2025). We note that outliers, i.e. non-LRD broad-line sources at high L_{5100} , are AGN residing in massive host galaxies and therefore explained by a significant contribution from the host to the continuum. The slope of the relation (from linear regression to the $z < 4.5$ sample⁴, with uncertainties estimated from bootstrapping) is close to unity,

$$L_{\text{H}\alpha} = 10^{43.24 \pm 0.08} \text{erg s}^{-1} \left(\frac{L_{5100}}{10^{44} \text{erg s}^{-1}} \right)^{0.91 \pm 0.09}, \quad (2)$$

and the offset is approximately equal to that of the non-LRD sample. The scatter is modest at 0.23 dex and, as expected given the proximity in wavelength, correlates strongly with $\text{H}\alpha$ EW.

Interestingly, the relation remains equally tight (0.23 dex scatter about a unit slope) when considering the integrated luminosity of the blackbody fits instead of L_{5100} . The sample spans a wide range in λ_{peak} and hence also $L_{5100}/L_{\text{blackbody}}$,

$$\frac{L_{5100}}{L_{\text{blackbody}}} = 10^{-0.645 \pm 0.010} \text{erg s}^{-1} \left(\frac{\lambda_{\text{peak}}}{0.8 \mu\text{m}} \right)^{-1.93 \pm 0.07}. \quad (3)$$

The fact that the scatter is unchanged in spite of this wide variation may be due to a weak correlation between λ_{peak} and the $\text{H}\alpha$ EW (right panel of Figure 7). The slope of the relation between $\log L_{\text{H}\alpha}$ and $\log L_{\text{blackbody}}$ is similarly close to unity and only the offset differs,

$$L_{\text{H}\alpha} = 10^{42.83 \pm 0.04} \text{erg s}^{-1} \left(\frac{L_{\text{blackbody}}}{10^{44} \text{erg s}^{-1}} \right)^{0.94 \pm 0.10}. \quad (4)$$

The implication of these trends is that the $\text{H}\alpha$ and continuum emission must primarily originate from the same source. Combined with the fact that the non-LRD broad-line sources follow the same

⁴ When fitting the full LRD sample these coefficients are 43.30 ± 0.04 and 1.04 ± 0.06 , respectively.

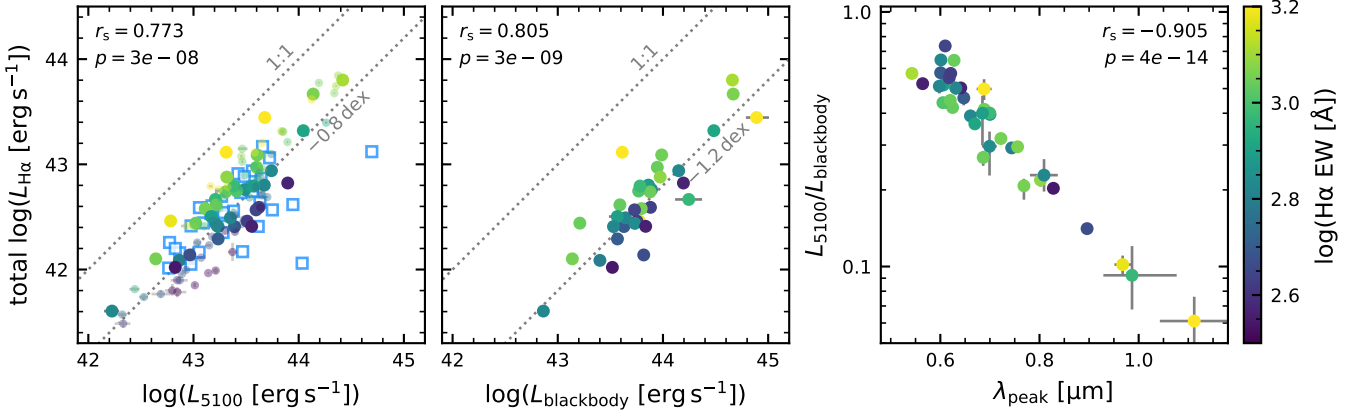


Figure 7. Total (broad + narrow) H α line luminosity vs. the luminosity at rest-frame 5100 Å (left) and integrated blackbody luminosity (middle), colour-coded by total H α EW. Large (small) circles show the $z < 4.5$ (full) LRD sample, and blue squares the RUBIES broad line sources of [Hviding et al. \(2025\)](#) that are not LRDs. Grey dotted lines indicate linear correlations, offset by a constant factor. The scatter about the unit slope is equally small (0.2 dex) for both L_{5100} and $L_{\text{blackbody}}$, despite the broad range in $L_{5100}/L_{\text{blackbody}}$ spanned by the sample (right) as the two continuum luminosities are related via the peak wavelength. These tight relations point to a common origin for the production of the H α lines and rest-optical continua of LRDs.

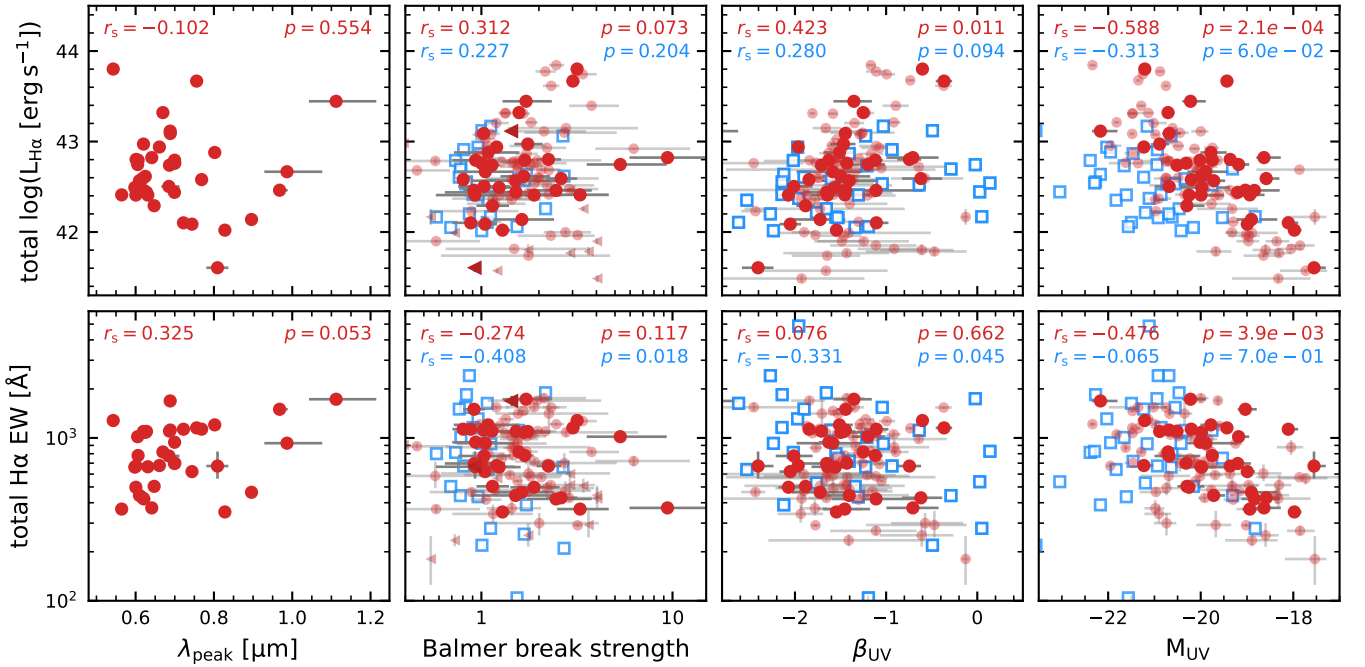


Figure 8. H α line properties as a function of the optical continuum shape parameters (λ_{peak} , Balmer break) and UV continuum properties (β_{UV} , M_{UV}). Large (small) circles show the $z < 4.5$ (full) LRD sample, and blue squares the RUBIES broad line sources of [Hviding et al. \(2025\)](#) that are not LRDs. Spearman correlation coefficients and p-values are shown in each panel for the LRD sample. The H α emission is independent of the rest-optical continuum shape, but is correlated with the rest UV, as LRDs with the reddest UV slopes are most luminous in H α . The UV luminosities of LRDs correlate with both the H α luminosities and EWs, and suggests that a substantial fraction of the UV can be attributed to emission from AGN.

scaling relation, this at face value points to photoionisation by AGN, although we will discuss the possibility of collisional excitation and the effects of high optical depth in and near the Balmer transitions in Section 5.2. Moreover, it implies that $L_{\text{H}\alpha}$, L_{5100} are good tracers of the bolometric luminosity of LRDs, in addition to $L_{\text{blackbody}}$ that was already presented in [Greene et al. \(2025\)](#).

Next, we consider correlations between H α and the shape of the rest-optical and UV continuum. We do not find any significant trends between the H α luminosity or EW and parameters describing the shape of the rest-optical, such as λ_{peak} and the Balmer break strength

shown in Figure 8, or T and β_{MBB} (not shown). The H α emission does depend weakly on the shape of the UV, as sources with redder UV slopes tend to have greater H α luminosities, but not necessarily higher EWs. This is in contrast with the broad-line sample that are not LRDs for which the H α luminosity is uncorrelated with β_{UV} , but sources with bluer UV slopes have higher EWs.

We also obtain UV magnitudes, M_{UV} , by evaluating the power-law slopes of Section 3 at 1500 Å (noting this is an extrapolation for sources at $z < 3$, and that we consider only UV emission within the NIRSspec shutter). Figure 8 shows that the H α luminosities of

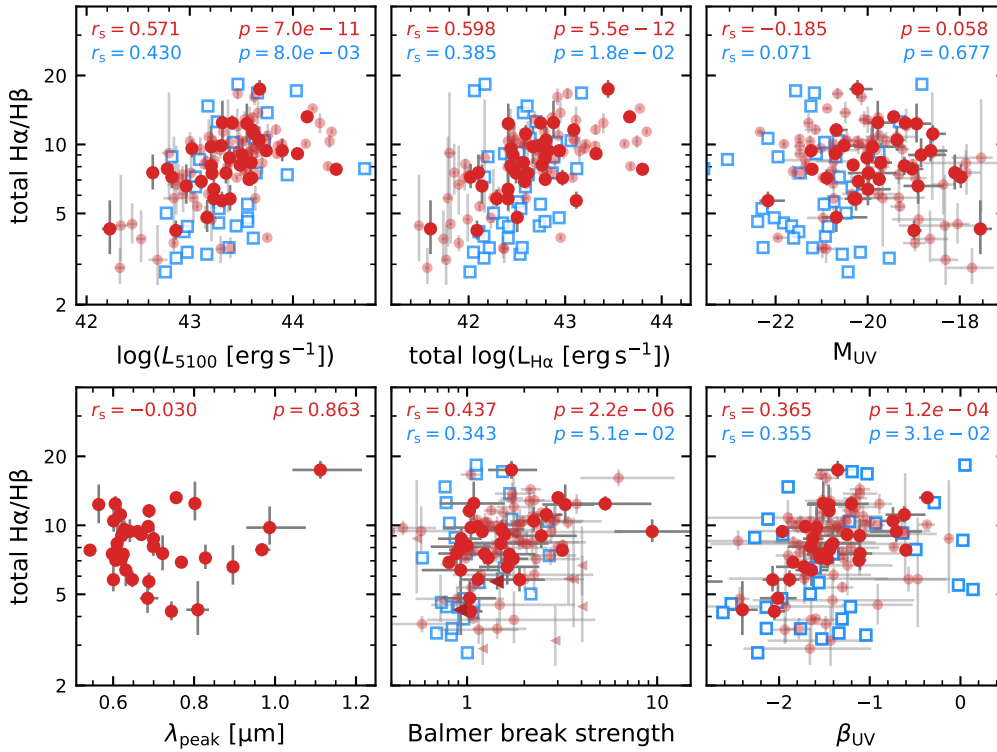


Figure 9. Balmer decrement of the total (broad+narrow) emission lines vs. different luminosities (top) and spectral shape parameters (bottom). Symbols indicate the same as in Figure 8. LRDs typically have high Balmer decrements ($H\alpha/H\beta \sim 10$), which increase for sources with redder UV slopes. However, a lack of correlation with λ_{peak} argues against dust reddening as the cause. The Balmer decrement correlates strongly with both the rest-optical continuum and $H\alpha$ luminosity, and weakly with Balmer break strength. This likely points to a luminosity-dependent population of hydrogen in the $n=2$ state in the gaseous envelopes of LRDs, which in turns affects the balance between the collisional (de-)excitation and optical depths of the $H\alpha$ and $H\beta$ lines.

LRDs are strongly correlated with M_{UV} , but systematically offset from the non-LRD population (as was already discussed in [Hviding et al. 2025](#)),

$$\log(L_{H\alpha}/\text{erg s}^{-1}) = (-0.31 \pm 0.04)M_{\text{UV}} + (42.69 \pm 0.04). \quad (5)$$

Under the assumption that the $H\alpha$ luminosity primarily reflects emission from AGN, the moderate scatter of 0.34 dex about the best-fit relation suggests a causal connection between the $H\alpha$ and UV emission, therefore implying a substantial fraction of the UV light in LRDs may originate from AGN (in addition to likely emission from host galaxies).

In summary, despite following the same tight relation between $L_{H\alpha}$ and L_{5100} , the $H\alpha$ lines of LRDs clearly differ in their properties from other broad-line AGN in several ways. This may reflect a difference in the physical mechanisms producing Balmer line emission, which we discuss in Section 5.2.

4.2 Balmer decrement

The correlations between $H\alpha$ and the UV-optical continuum properties do not hold in the same way for higher order Balmer lines. We demonstrate this through correlations between the Balmer decrement ($H\alpha/H\beta$) and continuum SED properties. Here, the Balmer decrement is defined as the ratio between the total (i.e. broad+narrow) $H\alpha$ and $H\beta$ line flux, again assuming that emission from an underlying host galaxy is subdominant.

The Balmer decrements of LRDs are typically very high, as the sample average and median of $H\alpha/H\beta \approx 8.7$ are three times greater

than expected from case B recombination, with some sources as extreme as $H\alpha/H\beta \approx 17$ (NEXUS-2314). Surprisingly, we find that the Balmer decrement depends strongly on both the continuum luminosity L_{5100} and $H\alpha$ line luminosity (Figure 9). A similar trend is substantially weaker among the broad-line sources that are not LRDs, which span a broader range in $H\alpha/H\beta$ at fixed luminosity.

In contrast, there is no dependence on UV luminosity, which may be related to the fact that sources with high Balmer decrements also tend to have redder UV slopes. The reddening of both UV continua and Balmer lines is often interpreted as arising from dust extinction. However, as also discussed in Section 3.3, LRDs with redder UV slopes typically do not have the reddest optical continua. Figure 9 shows that the Balmer decrement indeed does not depend on λ_{peak} obtained from the modified blackbody fitting, and suggests different mechanisms than dust extinction are at play.

In this context, the weak correlation between Balmer decrement and strength of the Balmer break is noteworthy. A strong Balmer break points to a large population of dense hydrogen gas in the $n=2$ state, which has formed a key piece of evidence for the presence of gaseous envelopes in LRDs (e.g. [Ji et al. 2025a](#); [Naidu et al. 2025a](#); [de Graaff et al. 2025a](#)). The correlation with Balmer decrement therefore suggests it is driven by the physical conditions of the gas, that is, the effects of scattering and collisional (de-)excitation in dense gas. We further discuss this scenario in Section 5.2.

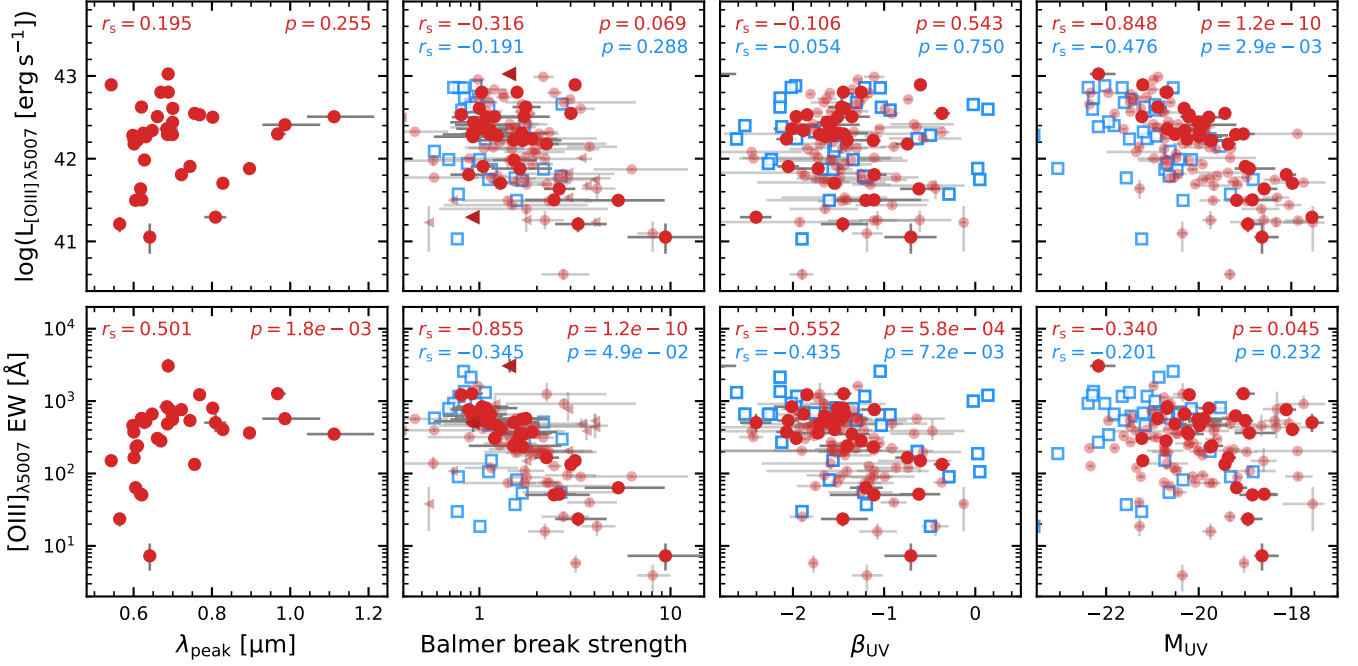


Figure 10. $[\text{O III}]_{\lambda 5007}$ line properties as a function of the optical continuum shape parameters (λ_{peak} , Balmer break) and UV continuum properties (β_{UV} , M_{UV}). Symbols are as described in Figure 8. The $[\text{O III}]$ luminosity depends only on M_{UV} , but the EW is correlated strongly with the UV and optical SED shape, especially the Balmer break strength. For LRDs these trends are opposite to those found for $\text{H}\alpha$ (Figure 8), and indicate the $[\text{O III}]$ emission has a different origin, likely from host galaxies.

4.3 $[\text{O III}]$ emission

Next, we explore the properties of the rest-optical $[\text{O III}]$ emission. The kinematic properties of this forbidden transition have already been shown to differ greatly from the Balmer lines, as the $[\text{O III}]$ lines do not have a substantial broad component (e.g. Wang et al. 2024; Juodžbalis et al. 2024; Hviding et al. 2025). This suggests that the $[\text{O III}]$ emission has a different origin, and has been attributed to star-forming regions in host galaxies.

Figure 10 shows the distribution of the $[\text{O III}]_{\lambda 5007}$ luminosity and EW as a function of UV and optical SED properties (these are the same panels as in Figure 8). The $[\text{O III}]$ luminosity depends strongly on M_{UV} but not the spectral shape itself, different from $\text{H}\alpha$ in Section 4.1. Moreover, in this case there is no systematic offset in the line luminosity at fixed M_{UV} between the LRD and non-LRD broad-line populations.

The $[\text{O III}]$ EWs of LRDs do depend on the shape of the UV-optical SED, as there is a modest negative trend with β_{UV} such that sources with bluer UV slopes show stronger $[\text{O III}]$ emission. Most remarkable is the fact that the wide spread in EW, spanning nearly three orders of magnitude, is strongly anti-correlated with the strength of the Balmer break. A similar, but positive, correlation is seen with λ_{peak} , although this may simply reflect our previous finding that sources with strong Balmer breaks have short λ_{peak} (see Section 3).

These trends between the line EW and continuum shape are opposite to the results for $\text{H}\alpha$ of Figure 8, and point to a different physical origin or geometry. For the $[\text{O III}]$ emission, the increase in line strength and luminosity for sources with bluer UV slopes and brighter M_{UV} is consistent with expectations from photoionisation in star-forming regions. The fact that the weakest $[\text{O III}]$ lines are found in LRDs with the strongest Balmer breaks provides further insight into the physical origin (i.e. photoionisation by hot young stars or

AGN). Because LRDs with strong Balmer breaks are also more luminous (Section 3.3), the AGN in these sources may outshine the host galaxies in which they reside. We investigate whether variations in the host galaxy-to-AGN ratio can reproduce the observed trend between $[\text{O III}]$ EW and Balmer break strength in Section 5.3.

4.4 O I emission

Finally, we also examine the properties of the $\text{O I}_{\lambda 8446}$ emission line. Different from the Balmer and $[\text{O III}]$ lines, O I traces neutral rather than ionised gas, although, as we will discuss further below, the ionisation state of O I is coupled to H II via charge-exchange reactions and O I emission is therefore expected to correlate with the hydrogen recombination lines. The O I emission has been of interest in the context of LRDs as it independently probes the physical conditions of the gas near the black hole and interstellar medium (Tripodi et al. 2025; Juodžbalis et al. 2024; Labbe et al. 2024; Kokorev et al. 2025). The 8446\AA transition in particular stands out as one of the brightest near-IR lines after hydrogen, and is accessible with JWST/NIRSpec up to $z \approx 5.5$. Among our LRD sample, spanning a range of surveys with different depths, we detect $\text{O I}_{\lambda 8446}$ in approximately 60% of sources (where detection is defined as the integrated line flux > 0 for $> 95\%$ of MCMC samples).

As in the previous sections, we test whether the line luminosity and EW are correlated with UV or optical continuum properties. We find that neither depend significantly on the spectral shape (e.g. β_{UV} , λ_{peak} , Balmer break strength), but the luminosity of $\text{O I}_{\lambda 8446}$ does depend strongly on the luminosity of both the UV and optical, shown in Figure 11. Above all, however, the $\text{O I}_{\lambda 8446}$ luminosity is strongly correlated with the $\text{H}\alpha$ luminosity. We fit a power-law relation, accounting for non-detections, and obtain uncertainties from

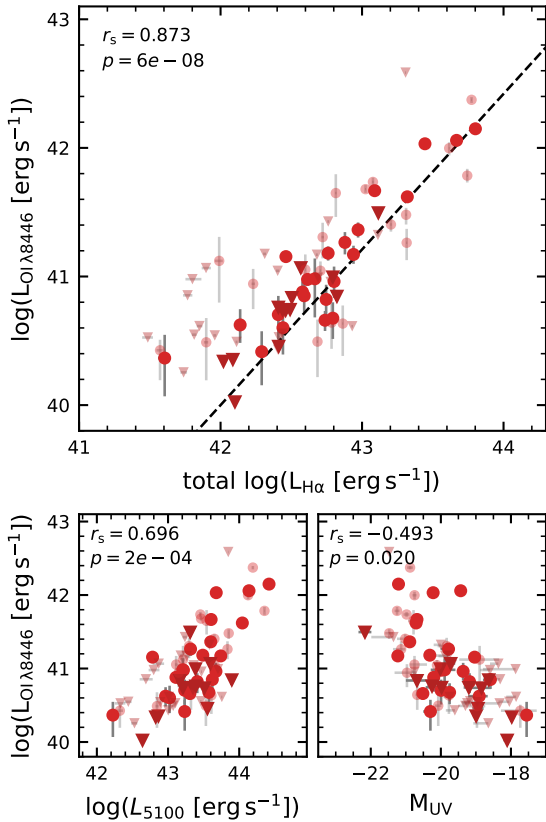


Figure 11. Correlation between the $O\text{I}_{8446}$ line luminosity and the total $H\alpha$ luminosity (top), and the continuum luminosity L_{5100} and M_{UV} (bottom). Large (small) symbols show the $z < 4.5$ (full) LRD sample, where triangles represent upper limits (95th percentile) in case of non-detections. The correlation with $L_{H\alpha}$ is particularly strong; the best-fit relation to the full sample (dashed line) is nearly linear, as the power-law slope $\gamma = 1.17 \pm 0.07$.

bootstrapping:

$$L_{O\text{I}_{8446}} = 10^{41.21 \pm 0.04} \text{ erg s}^{-1} \left(\frac{L_{H\alpha}}{10^{43} \text{ erg s}^{-1}} \right)^{1.21 \pm 0.09}. \quad (6)$$

The fact that the relation between $L_{O\text{I}_{8446}}$ and $L_{H\alpha}$ is close to linear (this even holds true for the line fluxes rather than luminosities, with power-law slope 1.15 ± 0.05) provides direct physical insight. There are various excitation mechanisms for $O\text{I}_{8446}$, including $\text{Ly}\beta$ fluorescence, collisional excitation, recombination or continuum fluorescence. The first two mechanisms, and in particular $\text{Ly}\beta$ pumping, have been shown to dominate in nearby AGN (e.g. [Rodríguez-Ardila et al. 2002](#)). Because the upper level of the cascade in which $O\text{I}_{8446}$ is produced has the same energy as that of $H\alpha$ (i.e. $n = 3$ state of hydrogen), the near-linear correlation between $O\text{I}_{8446}$ and $H\alpha$ argues strongly in favour of $\text{Ly}\beta$ fluorescence as the primary excitation mechanism for LRDs. Notably, such a tight linear relation has been found before for Herbig Ae/Be stars (i.e. pre-main sequence stars with accretion disks), with identical normalisation of $L_{H\alpha}/L_{O\text{I}_{8446}} \approx 70$ ([Mathew et al. 2018](#)). As discussed in [Mathew et al. \(2018\)](#), by comparing this normalisation to the theoretical flux ratio under optically thin conditions (~ 7500 ; [Strittmatter et al. 1977](#)), this implies $H\alpha$ is optically thick ($\tau_{H\alpha} \sim 100$) and therefore independently supports the conclusion that $\text{Ly}\beta$ fluorescence dominates.

Importantly, this also suggests that $H\alpha$ and $O\text{I}_{8446}$ are produced in approximately the same region, which may be surprising given

the very different emission line kinematics. Whereas the Balmer lines have widths of $\gtrsim 1000 - 4000 \text{ km s}^{-1}$, the $O\text{I}$ lines are substantially narrower ($\text{FWHM} < 1000 \text{ km s}^{-1}$; [Triptodi et al. 2025](#); [Juodžbalis et al. 2024](#); [Labbe et al. 2024](#)). Although beyond the scope of this paper, as at the resolution of the PRISM the width of $O\text{I}$ is difficult to robustly constrain (further exacerbated by blending from possible Ca II emission), we find that $O\text{I}_{8446}$ is systematically narrower ($\text{FWHM} \lesssim 1000 \text{ km s}^{-1}$) also in our large sample of LRDs. This may be explained by the fact that the $O\text{I}$ emission is weaker, and broad components are therefore difficult to detect unless deep medium-resolution spectroscopy is available (e.g. [Kokorev et al. 2025](#)). However, if the absence of a broad component or low broad-to-narrow ratios are confirmed in a larger number of high S/N spectra, this may set important constraints on the gaseous structures in LRDs (see Section 5.2).

5 DISCUSSION

Using a large spectroscopically-selected sample of LRDs, we have shown that the population of LRDs are unified by the fact that the rest-optical continua are well described by blackbodies of typical $T \sim 5000 \text{ K}$, although there is great variation in other continuum SED properties, such as the UV slope or Balmer break strength. Strikingly, we have found that the properties of the UV-optical continua are strongly correlated with the properties of strong rest-optical emission lines. In this section we interpret these different correlations in the context of the black hole star, quasistar and supermassive star models recently proposed in the literature.

5.1 LRD photospheres

The fact that LRDs are generally well described by modified blackbodies with $|\beta_{\text{MBB}}| < 1$ in the rest-optical to near-IR, in contrast with normal galaxies that have broader SED shapes (Section 3.2), indicates the presence of a single dominant thermal component. In turn, this strongly supports the recently proposed scenario where emission from a powerful source – plausibly an AGN due to the high luminosity – is reprocessed in dense, optically thick hydrogen gas, and fully dominates the rest-optical continuum. Such a physical picture had so far only been considered for individual LRDs, primarily sources with strong Balmer break features (e.g. [Ji et al. 2025a](#); [Naidu et al. 2025a](#); [de Graaff et al. 2025a](#)). Our work now demonstrates that this holds for the LRD population at large, and that LRDs with strong Balmer breaks form a subset of typically more luminous ($L_{\text{blackbody}} > 10^{44} \text{ erg s}^{-1}$) sources.

Most remarkably, our HR diagram of LRDs (Figure 5) reveals that the majority of LRDs occupy a narrow range in temperature of $T \sim 5000 \text{ K}$ for a broad range (2 dex) in luminosity, although the full population spans temperatures $T \sim 2000 - 7000 \text{ K}$ or $\lambda_{\text{peak}} \sim 0.5 - 1.1 \mu\text{m}$. This near-vertical relation between temperature and luminosity is akin to the Hayashi track ([Hayashi 1961](#)) of pre-main sequence low-mass stars and red giants that are undergoing core contraction and expansion of their outer layers. Being nearly completely convective stars, their luminosity is largely decoupled from the interior temperature-pressure structure, and the vertical temperature-luminosity relation results from the fact that opacity in the cooler outer radiative layer is dominated by H^- which has a steep temperature dependence (e.g. [Kippenhahn & Weigert 1994](#)). The LRDs are at higher temperatures than the Hayashi tracks of low-mass stars ($\sim 3000 \text{ K}$), which could simply reflect the fact that the location

of the Hayashi line depends in detail on the specific temperature-density structure of the envelope (for detailed explanation, see also Nandal & Loeb 2025, Kido et al. 2025); for instance, for pre-main sequence stars structural variations gives rise to a mass-dependence of the Hayashi limit. In this scenario, we also obtain a direct estimate of the physical scale of LRDs through the Stefan-Boltzmann law, resulting in a typical radius of $R \sim 1000$ AU ($\sim 10^{16}$ cm; or $R \sim 500 - 3000$ AU for the distribution of L and T found in Figure 5). These envelopes must likely also be in approximate hydrostatic equilibrium if they are to maintain their large size over a lifetime of several Myr, although the degree of support provided by rapid rotation is yet to be explored.

Recent theoretical studies (Kido et al. 2025; Liu et al. 2025; Coughlin & Begelman 2024; Begelman & Dexter 2025; Nandal & Loeb 2025; Zwick et al. 2025; Zhang et al. 2025) have proposed a range of different models to explain the observed properties of LRDs, most of which invoke the presence of a photosphere to produce rest-optical to near-IR continua that closely resemble blackbodies. Crucially, several of these models also predict that the photospheres must be close to the Hayashi limit (Kido et al. 2025; Begelman & Dexter 2025; Liu et al. 2025) or truly following Hayashi tracks (Nandal & Loeb 2025), albeit with slightly higher characteristic temperatures of $T \sim 5000 - 7000$ K than found for our observed sample. The instability to convection in these models may, for example, be explained by outward pressure induced by high surface accretion rates (Nandal & Loeb 2025). Only the proposed accretion disk models of Zwick et al. (2025) and Zhang et al. (2025) do not invoke an optically-thick envelope or arguments related to the Hayashi limit. These models do accurately predict the temperature distribution, although it is at present unclear whether the observed steep luminosity-temperature dependence would also be expected in this scenario. Interpreting LRDs as having large, convective envelopes therefore may offer a natural explanation for their wide luminosity range as well as their characteristic temperature.

Crucially, our work goes one step further by showing that the overall UV-optical spectral shape can be linked to the temperature, or rather the peak wavelength, of the rest-optical SED. These different correlations may help to distinguish between the range of theoretical models proposed. Specifically, LRDs with strong Balmer breaks and red UV slopes have shorter $\lambda_{\text{peak}} \lesssim 0.65$. The fact that these features are also correlated with the optical continuum luminosity provides physical insight, as it suggests that more luminous sources are able to populate a higher fraction of hydrogen gas in the $n = 2$ state, and is therefore possibly related to the overall mass scale or accretion rate.

Moreover, a subset of cool LRDs ($T \sim 2000 - 3000$ K) seems to fall outside many model predictions. Their SEDs do not show Balmer breaks, and the position of the characteristic v-shape for some is substantially redward of the Balmer limit (Figure 3). Only Zwick et al. (2025) explicitly model the observed SED of such a cooler LRD (the source of Killi et al. 2024) of $T \sim 3000$ K, using a model that differs fundamentally from the picture that LRDs are AGN within optically-thick envelopes. Instead, Zwick et al. (2025) propose that LRDs are supermassive stars surrounded by accretion disks, which produce thermal emission at $\sim 2000 - 5000$ K. In this context, the great similarity between the $O\text{I}$ and $H\alpha$ correlation of LRDs and Herbig stars found in Section 4.4 is intriguing.

Among the AGN models, the late-stage quasistar model of Begelman & Dexter (2025) and envelope model of Kido et al. (2025) explicitly predict $T > 5000$ K. However, Liu et al. (2025) show that the observed temperature or peak wavelength of the photosphere depend on gas density, and densities $\rho > 10^{-11.5}$ g cm $^{-3}$ can lead to redder SEDs without Balmer breaks. The existence of cool LRDs therefore

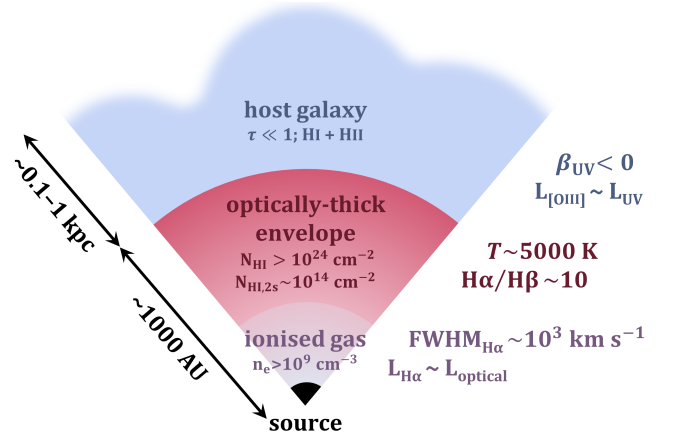


Figure 12. Schematic of the LRD’s gaseous structures discussed in Sections 5.1, 5.2 and 5.3.

does not rule out any of these theoretical models, but indicates that further refinement may be needed, as for instance these objects may have different dominant opacity (H^- rather than bound-free). These sources could possibly even point to evolutionary trends among the LRD population. Although we found only a weak dependence of temperature on the total rest-optical luminosity (Figure 5), which may be due to the small number of such sources in our sample, the coolest sources are all less luminous ($L_{\text{blackbody}} < 10^{44}$ erg s $^{-1}$) than LRDs with $T > 5000$ K. This luminosity range is where Begelman & Dexter (2025) predict the evolutionary stage of normal quasistars, rather than late-stage quasistars (which are $L > 10^{44}$ erg s $^{-1}$), for which the observed temperatures could be different.

Finally, in this discussion we have largely omitted one key parameter of our modified blackbody fits: the power-law slope β_{MBB} modifying the Planck function. The sample is distributed around $\beta_{\text{MBB}} = 0$, but with several strong outliers toward positive and negative values, indicating that some of the LRD SEDs are significantly different in shape from a single-temperature blackbody. SEDs for which $\beta_{\text{MBB}} < 0$ are broader in shape, similar to more typical galaxies, and therefore reflects the presence of a broader temperature distribution. We find that the sources for which $\beta_{\text{MBB}} < -1$ have longer λ_{peak} and higher UV luminosities, which may imply that the hot (i.e. UV) component of the SED contributes substantially also in the rest-optical. As a result, the sum of the hot and cooler component therefore may be broader, resulting in SEDs similar to, for example, those shown in Zwick et al. (2025).

In contrast, the small number of sources that are significantly narrower in shape (i.e. $\beta_{\text{MBB}} > 1$) still point to a single thermal component, but with substantial absorption in the Rayleigh-Jeans tail. These steep power-law slopes required for LRDs may be the result of broad molecular absorption features in the gaseous envelopes, analogous to those found in the atmospheres of cool stars. Indeed, we find that several of the sources with high β_{MBB} also have some of the lowest temperatures, the strongest outlier of which being UNCOVER-20698 (shown in right-hand panel of Figure 5). Detailed follow-up observations of the rest near-IR properties will be crucial to constrain the physical properties of LRDs with narrow SED shapes and cool temperatures, and to establish their connection to the broader LRD population, i.e. whether they represent an evolutionary stage or simply the tail of the distribution.

5.2 Production and scattering of Balmer lines

The Balmer lines of LRDs have formed a key topic of debate. As shown in Section 4, LRDs have $H\alpha$ luminosities that are higher than typical type-1 AGN of the same M_{UV} and redshift (see also [Hviding et al. 2025](#)), and very strong Balmer decrements ($H\alpha/H\beta \sim 10$). Previously these features were primarily interpreted as strong dust extinction (e.g. [Kocevski et al. 2023](#); [Greene et al. 2024](#)), but recent follow-up observations of mid- and far-IR wavelengths have ruled out that LRDs are dust-reddened AGN (e.g. [Casey et al. 2025](#); [Setton et al. 2025](#)). The high Balmer decrements have also been associated with super-Eddington accretion rates ([Lambrides et al. 2024](#)), or could arise from resonant scattering of the Balmer lines in dense gas clouds ([Chang et al. 2025](#)).

Moreover, the Balmer line profiles of some LRDs are exponential in shape and show significant absorption features ([Rusakov et al. 2025](#); [Labbe et al. 2024](#); [de Graaff et al. 2025a](#)), likely pointing to the effects of electron scattering in dense, and at least partially, ionised gas ([Rusakov et al. 2025](#); [Chang et al. 2025](#)) and possible effects of resonance scattering ([Naidu et al. 2025a](#)). In this dense gas picture the gas is optically thick to the Balmer transitions, and the $H\alpha$ emission itself may be partially produced in outer layers of the envelope ([Begelman & Dexter 2025](#); [Wang et al. 2025a](#)), while $H\beta$ emission may be reduced through radiative decay of $H\beta$ photons into $Pa-\alpha$ and $H\alpha$ photons in resonance scattering ([Naidu et al. 2025a](#); [Nikopoulos et al. 2025](#)). However, [Brazzini et al. \(2025\)](#) argue that scattering processes cannot dominate, due to dissimilarities in the profiles of different hydrogen recombination lines, and that the line widths and luminosities are instead consistent with more typical AGN broad-line regions.

The detailed geometry and physical conditions of LRDs therefore remain major open questions. The measured trends between various emission line and continuum properties among our large LRD sample provide a first, critical step forward to empirically constraining the nature of LRDs. We explore these trends largely agnostic to specific models, but under the assumption that an envelope of dense gas is present. Although we cannot conclusively rule out fundamentally different models (e.g. the accretion disk model of [Zwick et al. 2025](#)), such models have not yet been shown to produce Balmer breaks or made predictions for the HR diagram.

First, identical to other high-redshift broad-line AGN, we find tight, near-linear relations between the $H\alpha$ and optical continuum luminosities of LRDs (Section 4.1), strongly indicating that $H\alpha$ and optical continua are powered by the same source. A positive correlation between the $H\alpha$ and UV luminosities with only moderate scatter suggests that this central engine also produces UV continuum emission that can at least partially escape the gaseous envelopes (see e.g. [Torralba et al. 2025b](#)). Altogether, this points to the presence of a highly ionised region in the vicinity of a central engine (be it accretion onto a black hole or otherwise) where Balmer lines are predominantly produced by hydrogen recombination.

However, the Balmer transitions in this dense gas are optically thick, corroborated by the consistently low $H\alpha$ to $O\text{I}$ ratio (Section 4.4), raising the question how these recombination lines can be observed. If the Balmer lines are sufficiently broad, or broadened via electron scattering, the broad wings may have substantially lower optical depth, leading to Balmer line profiles with strong absorption features near the line centres (as shown by [Chang et al. 2025](#) using radiative transfer simulations). Such absorption features are observed in a subset of LRDs (e.g. [Matthee et al. 2024b](#); [Naidu et al. 2025a](#); [Juodžbalis et al. 2024](#)), but are not a ubiquitous feature (although this may be an observational bias due to the lack of high-resolution

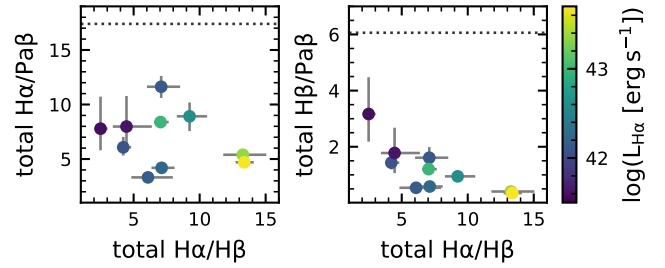


Figure 13. Paschen line ratios (left: $H\alpha/Pa\beta$, right: $H\beta/Pa\beta$) versus Balmer decrement, colour-coded by $L_{H\alpha}$, for LRDs at $z < 3.3$. Dotted lines indicate the expected value under the assumption of case B recombination. The very low $H\beta/Pa\beta$ ratios support a scenario in which $H\beta$ is resonantly scattered via $H\alpha$ and $Pa-\alpha$.

spectroscopy for many LRDs), and are not fully saturated. The emission at the Balmer line centres then likely must be produced, at least partially, in an outer region of lower optical depth. The tight relation with the optical continuum luminosity suggests that this emission is still predominantly physically associated with the central source, rather than for instance star formation in a host galaxy. The observed red optical continua and Balmer break features independently support the presence of such a region or layer of cooler, possibly less dense, gas. We provide a simple schematic overview of this structure in Figure 12.

Second, we find that the LRD sample shows very high Balmer decrements, which do not appear to be the effect of strong dust attenuation (see Section 4.2). Instead, these findings can be reconciled in the presence of very dense gas, and also suggest that the Balmer emission is not produced solely through photoionisation. Collisional excitation of the $n = 2$ to $n = 3$ state preferentially produces $H\alpha$ emission over that of $H\beta$, thereby boosting the Balmer decrement. The dependence of the decrement on optical and $H\alpha$ luminosity in this picture then may simply reflect the fact that more luminous, hotter sources have a larger population of $n = 2$ hydrogen. Moreover, at the high electron densities expected in LRDs ($\gtrsim 10^9 \text{ cm}^{-3}$; e.g. [Naidu et al. 2025a](#); [Ji et al. 2025a](#)) collisional de-excitation will also play a role, resulting in reduced $H\beta$ emission and hence also greater Balmer decrements. The third process that can substantially boost $H\alpha/H\beta$ is resonant scattering of the Balmer lines, as $H\beta$ is converted into $H\alpha$ and $Pa-\alpha$. As discussed above, the gas surrounding the central engine is optically thick to the Balmer transitions and therefore resonant scattering likely contributes substantially.

It is difficult to assess which, if any, of these three processes dominates the observed Balmer decrements. The importance of resonant scattering can be quantified by the measurement of $Pa-\alpha$, under the assumption that the gas is optically thin to $Pa-\alpha$, but at $z > 2$ this line is redshifted out of the observable wavelength range of JWST/NIRSpec. For a small number (10) of sources we are able to measure $Pa-\beta$ instead, and measure the $H\alpha/Pa\beta$ and $H\beta/Pa\beta$ ratios. In Figure 13 we show that these two ratios are substantially lower than expected from case B recombination, and depend on both $H\alpha$ luminosity and the Balmer decrement itself. The ratio of $H\beta/Pa\beta$ is especially low, and therefore clearly suggests that resonant scattering contributes to the high Balmer decrements, bolstering previous claims of resonant scattering in LRDs in the literature (e.g. [Naidu et al. 2025a](#); [Nikopoulos et al. 2025](#); [Torralba et al. 2025b](#)), but in disagreement with the findings of [Juodžbalis et al. \(2025\)](#) and [Brazzini et al. \(2025\)](#). If the observed Balmer decrement arises predominantly through resonance scattering, the implied hydrogen gas density in the $n = 2$ state from recent models of [Chang et al. \(2025\)](#) is $N_{H\text{I},2s} \sim 10^{14} \text{ cm}^{-2}$.

Scattering and collisional processes may also naturally explain the weaker correlations found between the properties of $H\alpha$ and UV-optical continua. In particular, in Section 4.1 we found that sources with higher $L_{H\alpha}$ have stronger Balmer breaks and redder UV slopes. This is in contrast with, for example, the $[O\text{ III}]$ emission which does not show such trends. Moreover, whereas there is no correlation between $L_{H\alpha}$ and λ_{peak} in Figure 8 for the entire LRD population, we do find a weak anti-correlation when considering these properties at fixed optical luminosity. In other words, the scatter in the relation between $L_{H\alpha}$ and $L_{\text{blackbody}}$ (Figure 7) correlates weakly with λ_{peak} , such that hotter LRDs (i.e. shorter λ_{peak}) may be brighter in $H\alpha$. The combination of these weaker correlations then points to the production of $H\alpha$ emission from collisional excitation in the outer region of the gaseous envelopes, especially for more luminous sources. This emission should be kinematically distinct from both the broad emission and possible $H\alpha$ emission from the host galaxy, and therefore may be confirmed by deep high-resolution spectra.

In summary, in this section we have argued from an empirical standpoint that the Balmer emission in LRDs originates primarily from a central photoionised region and partially from collisional excitation, likely facilitated by the presence of a large population of $n = 2$ hydrogen. The correlations between the emission line and continuum properties indicate that scattering processes must be important in altering the emission line ratios. These empirical findings agree well with the recently proposed 2-layer picture of Chang et al. (2025), of an H II and H I region in which respectively electron scattering and resonant scattering dominate, and is also qualitatively in line with the quasistar model of Begelman & Dexter (2025). To quantify the relative importance of the different physical processes inside LRDs will require further investigation of the Balmer and Paschen lines, by determining their flux ratios and kinematic components in detail.

5.3 Little Red Dots as BH*s in faint host galaxies

The remaining component not yet discussed is the host galaxy in which the BH* or quasistar must reside. That such hosts are present is clear from the fact that many LRDs have extended UV emission (e.g. Rinaldi et al. 2024; Torralba et al. 2025a). Clustering analyses have shown that these galaxies are likely to be low in mass (Matthee et al. 2024a), independently supported by the detection of narrow $[O\text{ III}]$ emission (implying low dynamical masses; e.g. Juodžbalis et al. 2024; Wang et al. 2024; Ji et al. 2025a), although the brightest LRDs may reside in more massive haloes (Labbe et al. 2024; Schindler et al. 2025).

We find spectroscopic evidence of star-forming host galaxies from the observed UV and $[O\text{ III}]$ emission. Our LRD sample shows great diversity in UV luminosity and slope, and for sources with bright UV emission (relative to the rest-optical) this emission contributes to an inferred broadening of the rest-optical SED shape (Section 5.1). Perhaps most convincing is the $[O\text{ III}]$ emission, which shows different correlations than the $H\alpha$ emission and, due to its relatively low critical density, likely originates from an interstellar medium. The fact that there is a strong correlation between the $[O\text{ III}]$ and UV emission supports this picture. Interestingly, this correlation (when measured within the aperture of the NIRSpec shutter) is approximately equally tight as that between $H\alpha$ and M_{UV} , and either suggests that the central UV emission originates from a mixture of AGN and host galaxy, or that the galaxy-scale properties correlate strongly with the properties of the AGN.

Recent work has argued in favour of a physical connection between the properties of the host and AGN. Maiolino et al. (2025)

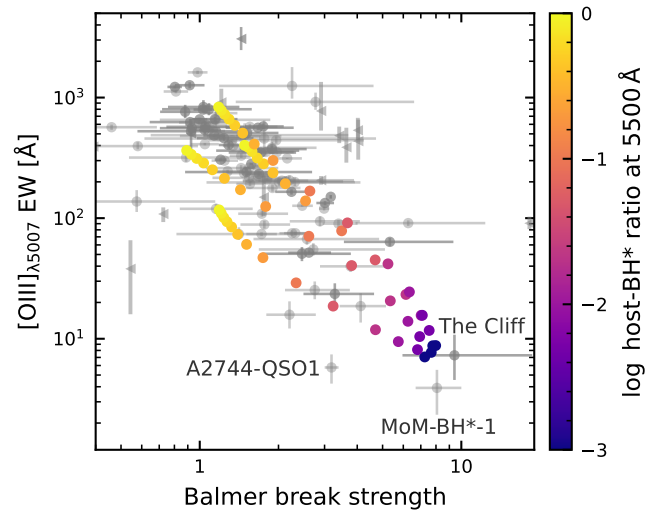


Figure 14. We investigate the strong correlation found between the $[O\text{ III}]$ λ_{5007} EW and Balmer break strength (Figure 10) empirically, by co-adding the spectra of The Cliff and low-mass ($M_* \sim 10^8 M_\odot$) star-forming galaxies for a wide range in host galaxy-to-BH* ratios and measuring the resulting $[O\text{ III}]$ EW and Balmer break strength. Our toy model, evaluated for four star-forming galaxy spectra, traces the slope and scatter of the observed distribution remarkably well. This suggests that LRDs can be described as the simple sum of a host galaxy and BH* component, and that variation in the host contribution alone can explain the weak $[O\text{ III}]$ lines observed in LRDs with strong Balmer breaks (Furtak et al. 2024; de Graaff et al. 2025a; Naidu et al. 2025a).

interpret the weak $[O\text{ III}]$ emission relative to narrow $H\beta$ emission in A2744-QS01 (an LRD with a strong Balmer break; Furtak et al. 2024) as evidence for a galaxy of extremely low metallicity hosting a primordial black hole. In this context, the strong correlation found between the $[O\text{ III}]$ EW and Balmer break strength in Figure 10 is striking. However, we may also expect there to be natural diversity in the $[O\text{ III}]$ EWs among the LRD population as a result of variation in the relative contributions from host and BH* to the total spectrum.

To test this scenario we construct a toy model, adding the spectra of The Cliff (de Graaff et al. 2025a, representing a “pure” AGN or BH*) and low-mass star-forming galaxies at different flux ratios. We then evaluate the $[O\text{ III}]$ EW and Balmer break strength as a function of the host: BH* ratio at 5500 \AA . For this exercise we select four PRISM spectra from the UNCOVER survey (Bezanson et al. 2024; Price et al. 2025), chosen to have approximately the same redshift as The Cliff ($3.3 < z < 3.7$), a stellar mass of $\sim 10^{8.0-8.2} M_\odot$ (obtained from the DJA catalogues, measured from NIRCcam photometry with eazy Brammer et al. 2008) and high continuum S/N, as well as diversity in emission line and UV properties.

Crucially, we find that LRDs can be described as the simple sum of a host galaxy and AGN or BH* component. Figure 14 shows the resulting tracks of the toy models in the $[O\text{ III}]$ EW vs Balmer break strength plane, which trace the observed population of LRDs remarkably well, and therefore indicate that variation in the host contribution dominates. Moreover, although A2744-QS01 remains an outlier in this plane, our simple test also cautions against a broader interpretation of the $[O\text{ III}]$ emission in LRDs as a metallicity indicator.

6 CONCLUSIONS

We have used the complete public DAWN JWST Archive to select a sample of 116 LRDs with v-shaped rest UV-optical NIR-Spec/MSA PRISM spectra and compact rest-optical morphologies in NIRCcam/F444W imaging. Higher-resolution NIRSpec spectroscopy (available for 47% of the sample) confirms the high purity of our LRD selection, as we detect broad Balmer lines in 98% of sources.

The LRD sample spans a broad range in redshift ($2.3 < z < 9.3$), luminosity, and SED properties. We perform a systematic characterisation of their SED properties, quantifying the shapes of the UV-optical continua as well as strong optical emission lines, summarised as follows:

(i) Most strikingly, we find that the rest-optical to near-IR continua are well described by modified blackbody models with $|\beta_{\text{MBB}}| \lesssim 1$ and are therefore dominated by a single thermal component. The LRDs have temperatures in the range $T \sim 2000 - 7000$ K ($\lambda_{\text{peak}} \sim 0.5 - 1.2 \mu\text{m}$), but the distribution is strongly clustered around $T \sim 5000$ K ($\lambda_{\text{peak}} \sim 0.6 \mu\text{m}$) with an asymmetric tail to colder temperatures. The blackbody luminosities span a broad range $\sim 10^{43-45} \text{ erg s}^{-1}$ and correlate only weakly with temperature in the HR diagram.

(ii) The UV continuum properties vary systematically with the rest-optical continuum properties. Sources with strong Balmer breaks typically have redder UV slopes, are hotter (i.e. shorter λ_{peak}) and more luminous. Crucially, sources with the reddest UV slopes are less red in the rest-optical and near-IR than other LRDs, adding further evidence against dust reddening in LRDs and instead favouring reddening by absorption in dense gas.

(iii) Balmer line properties are strongly linked to the continuum properties. We find a very tight correlation between the $H\alpha$ and rest-optical continuum luminosities, demonstrating that the two must be powered by the same source. We further find a strong correlation between the $H\alpha$ and UV emission, indicating that at least some UV emission can be attributed to the AGN. The Balmer decrements are typically very high ($H\alpha/H\beta \sim 9$) and correlate with luminosity and Balmer break strength.

(iv) The $[\text{O III}]_{\lambda 5007}$ line properties are also connected to the continuum SED, as the $[\text{O III}]$ and UV luminosities are strongly correlated. However, the properties of $[\text{O III}]$ differ in important ways from the Balmer lines, pointing to a different physical origin. Notably, we find a strong correlation between the Balmer break strength and $[\text{O III}]_{\lambda 5007}$ EW.

(v) The $\text{O I}_{\lambda 8446}$ line luminosity is strongly correlated with the $H\alpha$, UV and optical continuum luminosities. The relation between the O I and $H\alpha$ lines is near-linear (power-law slope of 1.21 ± 0.09) and indicates that the O I emission is primarily produced by $\text{Ly}\beta$ pumping. The normalisation of the relation shows that the gas from which these lines originate must have very high optical depth ($\tau_{H\alpha} \sim 100$).

Crucially, our findings show that the LRDs differ systematically from galaxies of similar brightness and redshift, as the SEDs of typical galaxies are substantially broader than single-temperature blackbodies. The LRD properties and correlations thereof also indicate that dust attenuation plays little role in LRDs. Instead, many of the properties found can be naturally reconciled by the effects of dense, optically-thick gas. We therefore interpret our findings in the context of recently proposed black hole star and quasistar models. We hence identify four key components of LRDs:

(i) The central source produces ionising photons that give rise to Balmer emission. Given the high luminosities and ubiquity of LRDs, we consider accreting massive black holes to be the most likely

source, although we cannot rule out categorically different models such as supermassive stars.

(ii) A (partially) ionised region that is optically thick to Balmer transitions. Collisional excitation and de-excitation may boost the Balmer decrements. Scattering processes that cause line broadening may explain why Balmer lines are observed in spite of the high optical depth.

(iii) An optically-thick envelope and radiative layer produce the observed rest-optical continuum emission. This envelope is likely in approximate hydrostatic equilibrium to maintain its large radius (~ 1000 AU) and photosphere over several Myr. The photosphere follows a temperature-luminosity track analogous to the Hayashi track observed in pre-main sequence stars and red giants (albeit at slightly higher characteristic $T \sim 5000$ K), with pressure support from the (potentially convective) envelope balancing gravitational collapse and ram pressure from accretion. A high fraction of $n = 2$ hydrogen in the outer layer of the envelope causes the Balmer break and may produce further narrow $H\alpha$ emission through collisional excitation. Moreover, resonant scattering likely contributes substantially to the high Balmer decrements.

(iv) The host galaxy is subdominant at rest-optical wavelengths and is primarily observed by its UV and nebular emission, such as $[\text{O III}]$. There is large variation in the relative contribution of the host to the UV and nebular emission of the overall spectrum. A simple toy model describing LRDs as the sum of a low-mass star-forming galaxy and BH^* with a range in host: BH^* ratios can explain both the average trend and scatter of observed correlations, such as that between $[\text{O III}]_{\lambda 5007}$ EW and Balmer break strength.

In summary, by compiling a large set of spectral diagnostics for a broad sample of LRDs with high-quality NIRSpec spectra, we have sketched an empirical picture of the structures of LRDs. A wide variety of models attempting to explain the observed properties of LRDs have recently been proposed in the literature. Our empirical constraints therefore offer an important means to test whether these models can explain not just individual outliers, but also sample trends. This serves as a critical step toward understanding the true physical nature of the population of LRDs and the detailed properties of their gaseous structures and black holes.

ACKNOWLEDGEMENTS

AdG thanks Devesh Nandal for discussions on the properties of convective envelopes and Hayashi limit, and the students of the I-HOW JWST Workshop (Barbora Adamcova, Biswaraj Palit, Olena Pastoven, Weronika Puchalska) for their help in producing an early version of figures 8 and 10.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programs 1180, 1181, 1208, 1212, 1213, 1215, 1286, 1345, 1433, 2198, 2561, 2750, 2767, 4106, 4233, 5105, 5224, 6368, and 6585. Support for program #4233 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127.

AdG acknowledges support from a Clay Fellowship awarded by the Smithsonian Astrophysical Observatory. REH acknowledges support by the German Aerospace Center (DLR) and the Federal Min-

istry for Economic Affairs and Energy (BMW) through program 500R2403 ‘RUBIES’. The Cosmic Dawn Center is funded by the Danish National Research Foundation (DNRF) under grant #140. This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00072, as well as from the Swiss National Science Foundation (SNSF) through project grant 200020_207349. TBM was supported by a CIERA fellowship. PD warmly acknowledges support from an NSERC discovery grant (RGPIN-2025-06182). KG acknowledges support from Australian Research Council Laureate Fellowship FL180100060. LAB acknowledges support from the Dutch Research Council (NWO) under grant VI.Veni.242.055 (<https://doi.org/10.61686/LAJVP77714>) and the ERC Consolidator grant 101088676 (“VOYAJ”). YF is supported by JSPS KAKENHI Grant Numbers JP22K21349 and JP23K13149.

DATA AVAILABILITY

All data used in this work are publicly available through the DAWN JWST Archive (<https://dawn-cph.github.io/dja/index.html>). The spectra used in this work are part of version 4.4 of the DJA (Brammer & Valentino 2025). We also release our complete LRD catalogue, together with all key measurements made in the modified blackbody fitting and emission line fitting. The table can be found at <https://doi.org/10.5281/zenodo.17665942>, and its contents are described in Appendix A.

REFERENCES

- Akins H. B., et al., 2024, *arXiv e-prints*, p. arXiv:2406.10341
- Akins H. B., et al., 2025, *ApJ*, 980, L29
- Alexander T., Natarajan P., 2014, *Science*, 345, 1330
- Ananna T. T., Bogdán Á., Kovács O. E., Natarajan P., Hickox R. C., 2024, *ApJ*, 969, L18
- Begelman M. C., Dexter J., 2025, *arXiv e-prints*, p. arXiv:2507.09085
- Begelman M. C., Volonteri M., Rees M. J., 2006, *MNRAS*, 370, 289
- Begelman M. C., Rossi E. M., Armitage P. J., 2008, *MNRAS*, 387, 1649
- Bezanson R., et al., 2024, *ApJ*, 974, 92
- Brammer G., 2023b, grizli, doi:10.5281/zenodo.1146904
- Brammer G., 2023a, msaexp: NIRSpc analysis tools, doi:10.5281/zenodo.8319596
- Brammer G., Valentino F., 2025, The DAWN JWST Archive: Compilation of Public NIRSpc Spectra, doi:10.5281/zenodo.15472354, <https://doi.org/10.5281/zenodo.15472354>
- Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503
- Brazzini M., D’Eugenio F., Maiolino R., Juodžbalis I., Ji X., Scholtz J., 2025, *arXiv e-prints*, p. arXiv:2507.08929
- Casey C. M., Akins H. B., Kokorev V., McKinney J., Cooper O. R., Long A. S., Franco M., Manning S. M., 2024, *ApJ*, 975, L4
- Casey C. M., et al., 2025, *arXiv e-prints*, p. arXiv:2505.18873
- Chang S.-J., Gronke M., Matthee J., Mason C., 2025, *arXiv e-prints*, p. arXiv:2508.08768
- Coughlin E. R., Begelman M. C., 2024, *ApJ*, 970, 158
- Curtis-Lake E., et al., 2025, *arXiv e-prints*, p. arXiv:2510.01033
- D’Eugenio F., et al., 2025a, *arXiv e-prints*, p. arXiv:2503.11752
- D’Eugenio F., et al., 2025b, *arXiv e-prints*, p. arXiv:2510.00101
- Eisenstein D. J., et al., 2023, *arXiv e-prints*, p. arXiv:2306.02465
- Ferland G. J., et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Ferruit P., et al., 2022, *A&A*, 661, A81
- Finkelstein S. L., et al., 2025, *ApJ*, 983, L4
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Furtak L. J., et al., 2023a, *MNRAS*, 523, 4568
- Furtak L. J., et al., 2023b, *ApJ*, 952, 142
- Furtak L. J., et al., 2024, *Nature*, 628, 57
- Greene J. E., Ho L. C., 2005, *ApJ*, 630, 122
- Greene J. E., et al., 2024, *ApJ*, 964, 39
- Greene J. E., et al., 2025, *arXiv e-prints*, p. arXiv:2509.05434
- Hayashi C., 1961, *PASJ*, 13, 450
- Heintz K. E., et al., 2025, *A&A*, 693, A60
- Hildebrand R. H., 1983, *QJRAS*, 24, 267
- Hoffman M. D., Gelman A., et al., 2014, *J. Mach. Learn. Res.*, 15, 1593
- Horne K., 1986, *PASP*, 98, 609
- Hviding R. E., 2025, TheSkyentist/Unite: Version 0, Zenodo, doi:10.5281/zenodo.15585035
- Hviding R. E., et al., 2025, *arXiv e-prints*, p. arXiv:2506.05459
- Inayoshi K., Maiolino R., 2025, *ApJ*, 980, L27
- Inayoshi K., Kimura S., Noda H., 2024, *arXiv e-prints*, p. arXiv:2412.03653
- Ji X., et al., 2025a, *arXiv e-prints*, p. arXiv:2501.13082
- Ji X., et al., 2025b, *arXiv e-prints*, p. arXiv:2507.23774
- Juodžbalis I., et al., 2024, *MNRAS*, 535, 853
- Juodžbalis I., et al., 2025, *arXiv e-prints*, p. arXiv:2504.03551
- Kido D., Ioka K., Hotokezaka K., Inayoshi K., Irwin C. M., 2025, *arXiv e-prints*, p. arXiv:2505.06965
- Killi M., et al., 2024, *A&A*, 691, A52
- Kippenhahn R., Weigert A., 1994, *Stellar Structure and Evolution*
- Kocevski D. D., et al., 2023, *ApJ*, 954, L4
- Kocevski D. D., et al., 2024, *arXiv e-prints*, p. arXiv:2404.03576
- Kokorev V., et al., 2025, *arXiv e-prints*, p. arXiv:2511.07515
- Labbé I., et al., 2023, *Nature*, 616, 266
- Labbe I., et al., 2024, *arXiv e-prints*, p. arXiv:2412.04557
- Labbe I., et al., 2025, *ApJ*, 978, 92
- Lambrides E., et al., 2024, *arXiv e-prints*, p. arXiv:2409.13047
- Lambrides E., et al., 2025, *arXiv e-prints*, p. arXiv:2509.09607
- Lin X., et al., 2024, *ApJ*, 974, 147
- Lin X., et al., 2025, *arXiv e-prints*, p. arXiv:2507.10659
- Liu H., Jiang Y.-F., Quataert E., Greene J. E., Ma Y., 2025, *arXiv e-prints*, p. arXiv:2507.07190
- Maiolino R., et al., 2025, *arXiv e-prints*, p. arXiv:2505.22567
- Maseda M. V., et al., 2024, *arXiv e-prints*, p. arXiv:2403.05506
- Mathew B., et al., 2018, *ApJ*, 857, 30
- Matthee J., et al., 2024a, *arXiv e-prints*, p. arXiv:2412.02846
- Matthee J., et al., 2024b, *ApJ*, 963, 129
- Naidu R. P., et al., 2025a, *arXiv e-prints*, p. arXiv:2503.16596
- Naidu R. P., et al., 2025b, *arXiv e-prints*, p. arXiv:2505.11263
- Nandal D., Loeb A., 2025, *arXiv e-prints*, p. arXiv:2507.12618
- Nikopoulos G. P., Watson D., Sneppen A., Rusakov V., Heintz K. E., Witstok J., Brammer G., 2025, *arXiv e-prints*, p. arXiv:2510.06362
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Pacucci F., Narayan R., 2024, *ApJ*, 976, 96
- Pasha I., Miller T. B., 2023, *The Journal of Open Source Software*, 8, 5703
- Pérez-González P. G., et al., 2024, *ApJ*, 968, 4
- Phan D., Pradhan N., Jankowiak M., 2019, *arXiv preprint arXiv:1912.11554*
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Pollock C. L., et al., 2025, *arXiv e-prints*, p. arXiv:2506.15779
- Price S. H., et al., 2025, *ApJ*, 982, 51
- Rinaldi P., et al., 2024, *arXiv e-prints*, p. arXiv:2406.13554
- Rodríguez-Ardila A., Viegas S. M., Pastoriza M. G., Prato L., Donzelli C. J., 2002, *ApJ*, 572, 94
- Rusakov V., et al., 2025, *arXiv e-prints*, p. arXiv:2503.16595
- Sarrrouh G. T. E., et al., 2025, *arXiv e-prints*, p. arXiv:2506.21685
- Schindler J.-T., et al., 2025, *Nature Astronomy*,
- Scholtz J., et al., 2025, *arXiv e-prints*, p. arXiv:2510.01034
- Setton D. J., et al., 2024, *arXiv e-prints*, p. arXiv:2411.03424
- Setton D. J., et al., 2025, *arXiv e-prints*, p. arXiv:2503.02059
- Shen Y., et al., 2024, *arXiv e-prints*, p. arXiv:2408.12713
- Shetty R., Kauffmann J., Schnee S., Goodman A. A., 2009a, *ApJ*, 696, 676
- Shetty R., Kauffmann J., Schnee S., Goodman A. A., Ercolano B., 2009b, *ApJ*, 696, 2234
- Strittmatter P. A., et al., 1977, *ApJ*, 216, 23
- Taylor A. J., et al., 2025, *ApJ*, 989, L7

Torralba A., et al., 2025a, [arXiv e-prints](#), p. [arXiv:2505.09542](#)
 Torralba A., et al., 2025b, [arXiv e-prints](#), p. [arXiv:2510.00103](#)
 Tripodi R., et al., 2025, [arXiv e-prints](#), p. [arXiv:2507.20684](#)
 Valentino F., et al., 2023, *ApJ*, 947, 20
 Valentino F., et al., 2025, *A&A*, 699, A358
 Wang B., et al., 2024, *ApJ*, 969, L13
 Wang B., et al., 2025a, [arXiv e-prints](#), p. [arXiv:2508.18358](#)
 Wang B., et al., 2025b, *ApJ*, 984, 121
 Watanabe S., Opper M., 2010, *Journal of machine learning research*, 11
 Weibel A., et al., 2024, *MNRAS*, 533, 1808
 Williams C. C., et al., 2024a, [arXiv e-prints](#), p. [arXiv:2410.01875](#)
 Williams C. C., et al., 2024b, *ApJ*, 968, 34
 Xiao M., et al., 2025, [arXiv e-prints](#), p. [arXiv:2503.01945](#)
 Yue M., Eilers A.-C., Ananna T. T., Panagioutou C., Kara E., Miyaji T., 2024, *ApJ*, 974, L26
 Zhang C., et al., 2025, [arXiv e-prints](#), p. [arXiv:2505.12719](#)
 Zwick L., Tiede C., Mayer L., 2025, [arXiv e-prints](#), p. [arXiv:2507.22014](#)
 de Graaff A., et al., 2024, *A&A*, 684, A87
 de Graaff A., et al., 2025a, [arXiv e-prints](#), p. [arXiv:2503.16600](#)
 de Graaff A., et al., 2025b, *A&A*, 697, A189

AFFILIATIONS

¹*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany*

²*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge MA 02138 USA*

³*MIT Kavli Institute for Astrophysics and Space Research, 70 Vassar Street, Cambridge, MA 02139, USA*

⁴*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA*

⁵*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA*

⁶*Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*

⁷*Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA*

⁸*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*

⁹*Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria*

¹⁰*Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark*

¹¹*Department of Astronomy & Astrophysics, University of Chicago, 5640 S. Ellis Avenue, Chicago, IL 60637, USA*

¹²*Kavli Institute for Cosmological Physics, University of Chicago, Chicago IL 60637, USA*

¹³*Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA*

¹⁴*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, The Netherlands*

¹⁵*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

¹⁶*Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA*

¹⁷*Cosmic Frontier Center, The University of Texas at Austin, Austin, TX 78712, USA*

¹⁸*Canadian Institute for Theoretical Astrophysics, 60 St George St, University of Toronto, Toronto, ON M5S 3H8, Canada*

¹⁹*David A. Dunlap Department of Astronomy and Astrophysics, University of Toronto, 50 St George St, Toronto ON M5S 3H4, Canada*

²⁰*Department of Physics, 60 St George St, University of Toronto, Toronto, ON M5S 3H8, Canada*

²¹*Department of Astrophysics, University of Zurich, Zurich CH-8057, Switzerland*

²²*Center for Frontier Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan*

²³*Dunlap Institute for Astronomy and Astrophysics, 50 St. George Street, Toronto, Ontario, M5S 3H4, Canada*

²⁴*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, VIC 3122, Australia*

²⁵*Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland*

²⁶*Department of Astronomy, University of Wisconsin-Madison, 475 N. Charter St., Madison, WI 53706 USA*

²⁷*Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA*

²⁸*Astronomical Observatory Institute, Faculty of Physics and Astronomy, Adam Mickiewicz University, ul. Słoneczna 36, 60-286, Poznań, Poland*

²⁹*California Institute of Technology, Pasadena, CA 91125, USA*

³⁰*Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, 9000 Gent, Belgium*

APPENDIX A: DATA TABLES

Table A1. Program information for the LRD spectra used in this work, including duplicate spectra and multiple images of strongly lensed LRDs (134 spectra of 116 unique objects).

Program	# LRD spectra
CANUCS (GTO-1208; Sarrouh et al. 2025)	5
CAPERS (GO-6368; PI: Dickinson)	18
CEERS (GO-1345; Finkelstein et al. 2025)	2
DDT-2750 (PI: Arrabal Haro)	2
DDT-2767 (PI: Kelly)	1
DDT-6585 (PI: Coulter)	3
GO-1433 (PI: Coe)	1
GO-2198 (PI: Barrufet)	3
GO-4106 (PIs: Nelson & Labbe)	5
GTO-WIDE (GTO-1212, GTO-1213, GTO-1215; Maseda et al. 2024)	5
JADES (GTO-1180, GTO-1181, GTO-1286; Curtis-Lake et al. 2025; Scholtz et al. 2025)	15
MoM (GO-5224; PIs: Oesch & Naidu)	13
NEXUS (GO-5105; Shen et al. 2024)	9
RUBIES (GO-4233; de Graaff et al. 2025b)	37
UNCOVER (GO-2561 Bezanson et al. 2024; Price et al. 2025)	15

Table A2. Description of data table released together with this work.

Column name	Unit	Description
pid		JWST program ID
srcid		MSA ID number
root		DJA root name
file		DJA filename
ra	deg	Right ascension
dec	deg	Declination
zspec		Spectroscopic redshift
f444w_aper0.1	nJy	NIRCam/F444W circular aperture flux (radius 0.1'')
mu		Lensing magnification factor (obtained from Furtak et al. 2023a and Sarrouh et al. 2025)
Ndup		Number of duplicate PRISM spectra in LRD table
dup_filenames		DJA filenames of duplicate spectra
use_dG25		Flag to filter the 116 unique sources (with their highest S/N PRISM spectra) used in this work
lambda_v	μm	Inflection wavelength of v-shape (λ_v , [5, 16, 50, 84, 95] percentiles)
beta_UV		UV slope (measured from PRISM spectrum; [5, 16, 50, 84, 95] percentiles)
M_UV	mag	Absolute magnitude at 1500 Å (measured from PRISM spectrum, [5, 16, 50, 84, 95] percentiles)
break_strength		Balmer break strength ([5, 16, 50, 84, 95] percentiles)
beta_MBB		Power-law slope β_{MBB} of modified blackbody ([5, 16, 50, 84, 95] percentiles)
Teff	K	Effective temperature of modified blackbody ([5, 16, 50, 84, 95] percentiles)
peak_wave	μm	Peak wavelength λ_{peak} of modified blackbody ([5, 16, 50, 84, 95] percentiles)
logL_MBB	log erg s ⁻¹	Integrated luminosity of modified blackbody ($L_{\text{blackbody}}$; [5, 16, 50, 84, 95] percentiles)
logL_5100	log erg s ⁻¹	Optical luminosity λL_λ at $\lambda = 5100$ Å ([5, 16, 50, 84, 95] percentiles)
L5100_LMBB		Ratio of L_{5100} to $L_{\text{blackbody}}$ ([5, 16, 50, 84, 95] percentiles)
Ha_total_ew	Å (rest)	Total (broad+narrow) H α EW ([5, 16, 50, 84, 95] percentiles)
Hb_total_ew	Å (rest)	Total (broad+narrow) H β EW ([5, 16, 50, 84, 95] percentiles)
Balmer_dec_total		Total (broad+narrow) Balmer decrement H α /H β ([5, 16, 50, 84, 95] percentiles)
LHa_total	erg s ⁻¹	Total (broad+narrow) H α luminosity ([5, 16, 50, 84, 95] percentiles)
logLHa_total	log erg s ⁻¹	Total (broad+narrow) H α luminosity ([5, 16, 50, 84, 95] percentiles)
OIII_5007_ew	Å (rest)	[O III] λ_{5007} EW ([5, 16, 50, 84, 95] percentiles)
LOIII_5007	erg s ⁻¹	[O III] λ_{5007} luminosity ([5, 16, 50, 84, 95] percentiles)
logLOIII_5007	log erg s ⁻¹	[O III] λ_{5007} luminosity ([5, 16, 50, 84, 95] percentiles)
LOI_8446	erg s ⁻¹	O I λ_{8446} luminosity ([5, 16, 50, 84, 95] percentiles)
logLOI_8446	log erg s ⁻¹	O I λ_{8446} luminosity ([5, 16, 50, 84, 95] percentiles)

APPENDIX B: SUPPLEMENTARY FIGURES OF THE MODIFIED BLACKBODY MODELS

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

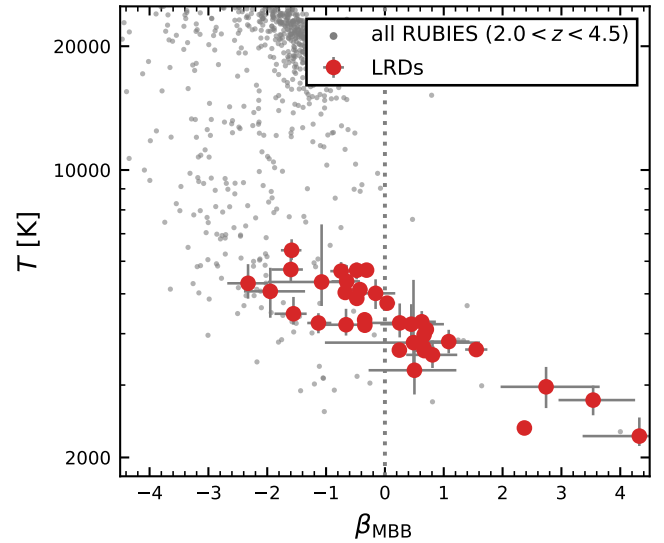


Figure B1. Distribution of the power law slopes and temperatures from modified blackbody fits to the LRD sample restricted to $z < 4.5$ (red). For reference, the results of fits to the complete RUBIES sample at $2 < z < 5$ and $F444W < 26$ are shown in grey. A slope of $\beta_{\text{MBB}} = 0$ corresponds to a perfect single-temperature blackbody. LRDs are well approximated by single-temperature blackbodies of characteristic temperature $T \sim 5000$ K, although the precise inferred temperature is strongly anti-correlated with β_{MBB} , and therefore more sensitive to the details of the data reduction and fitting than the measured λ_{peak} of Figure 4.

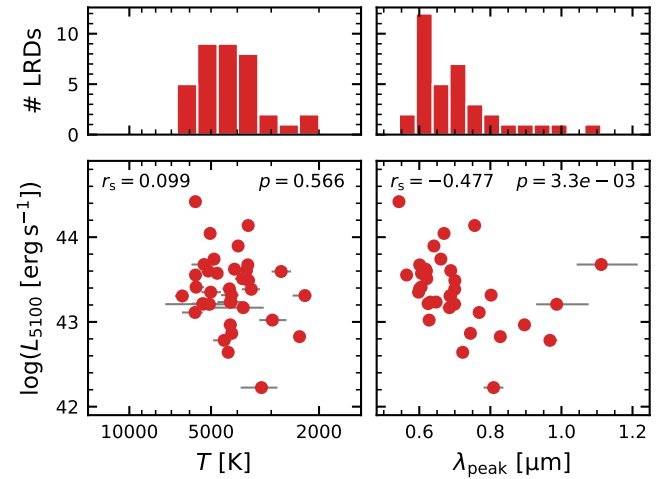


Figure B2. HR diagram, but with L_{5100} rather than the integrated luminosity ($L_{\text{blackbody}}$, Fig. 5). Because the shape of the SED differs systematically for sources that peak at longer wavelengths, a significant correlation arises between L_{5100} and λ_{peak} that is not apparent for $L_{\text{blackbody}}$ in Fig. 5.

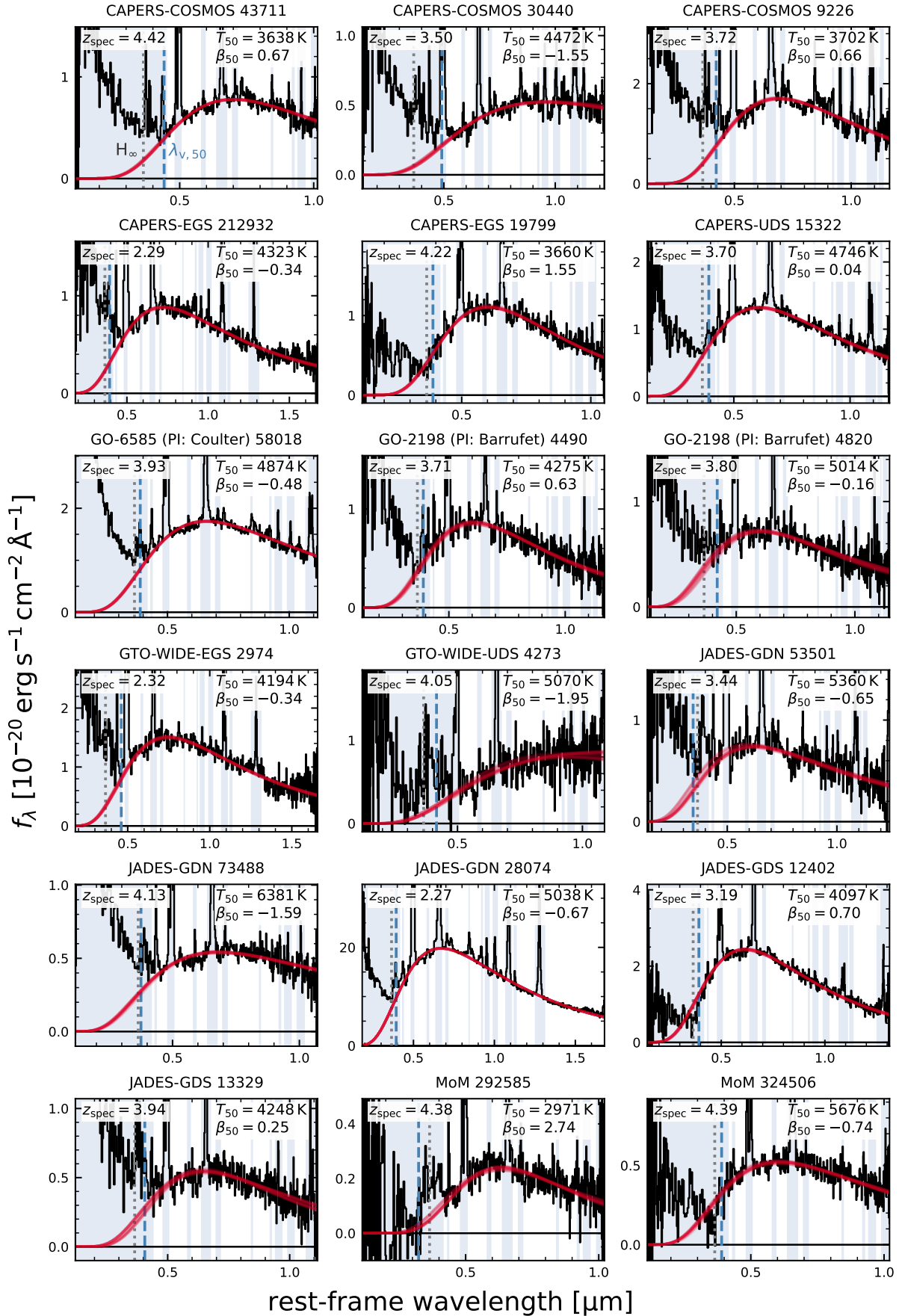
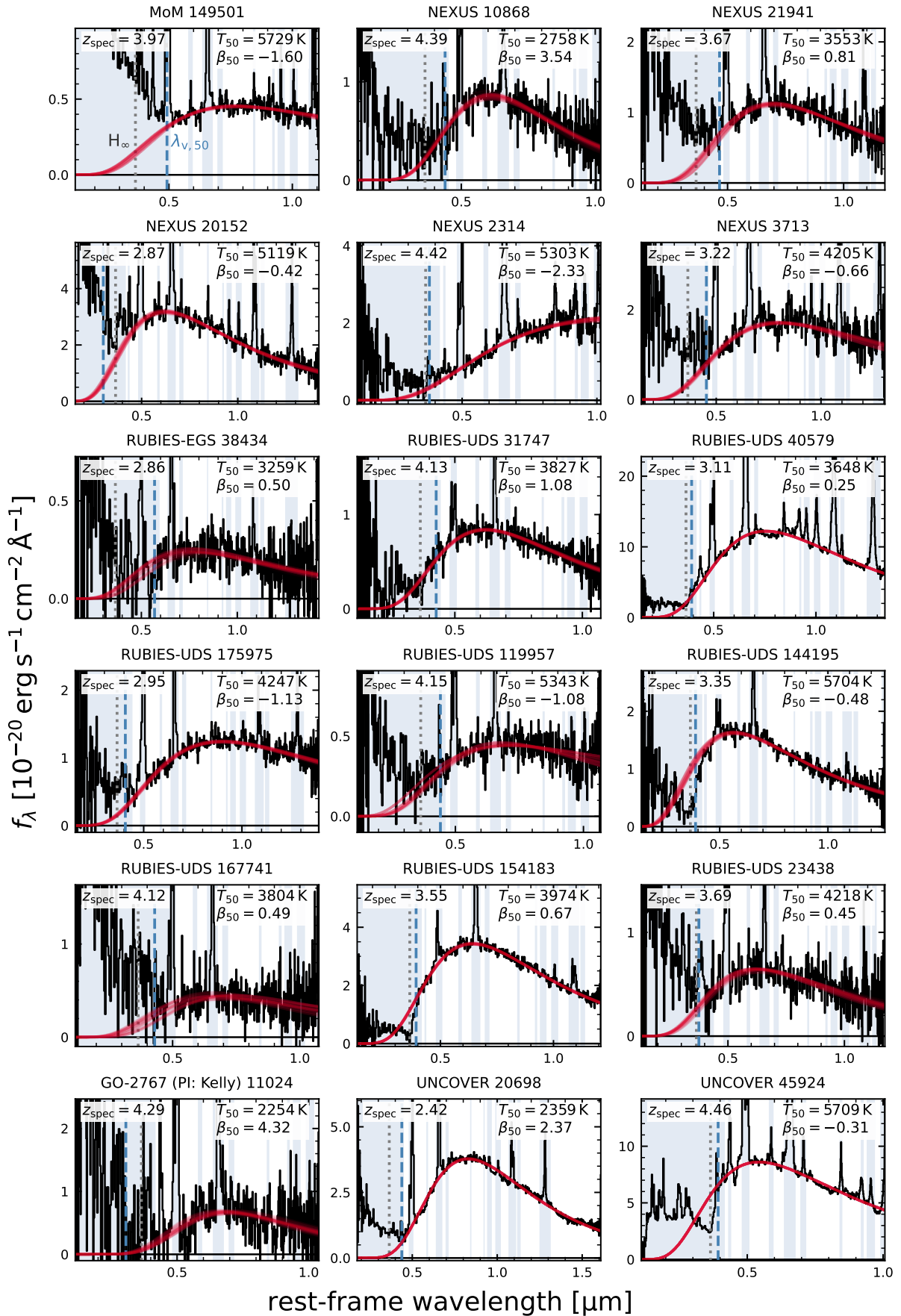


Figure B3. Modified blackbody fits (red lines, 100 random draws from posterior) to the $z < 4.5$ LRD spectra (see Section 2.3). Shaded regions indicate masked areas, and blue dashed lines the posterior median of the v -shape wavelength, λ_v .

Figure B4. As Figure B3, showing the second half of the $z < 4.5$ LRD sample.