

Beyond orientation: Evidence for distinct physical regimes among Little Red Dots and Little Blue Dots

L. Barrufet^{1*}, J. S. Dunlop¹, M. L. Hamadouche², M. Mezcu³, D. Herrero-Carrion^{4,5}, M. Santos-Lleo⁶, R. Gottumukkala^{7,8}, D. Spinoso⁹, C. T. Donnan¹⁰

¹ *Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK*

² *Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA*

³ *Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*

⁴ *Donostia International Physics Centre (DIPC), Paseo Manuel de Lardizabal 4, 20018 Donostia-San Sebastian, Spain*

⁵ *University of the Basque Country UPV/EHU, Department of Physics, Barrio Sarriena s/n. 48940 Leioa, Bizkaia, Spain*

⁶ *Formerly, European Space Astronomy Centre (ESAC), ESA, Camino Bajo del Castillo s/n, 28692 Villanueva de la Canada, Madrid, Spain*

⁷ *Cosmic Dawn Center (DAWN)*

⁸ *Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen North, Denmark*

⁹ *Dipartimento di Fisica “G. Occhialini”, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

¹⁰ *NSF’s National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Ave., Tucson, AZ 85719, US*

August 2026

ABSTRACT

Little Red Dots (LRDs) and Little Blue Dots (LBDs) may represent the same population of compact active galactic nuclei (AGN) observed along different lines of sight. We test this scenario using a spectroscopic sample from the DAWN *JWST* Archive, selected based on common criteria for $H\alpha$ equivalent width, UV continuum slope, and compactness. We use the optical continuum slope to distinguish between samples of 89 LRDs and 191 LBDs spanning $1 \lesssim z \lesssim 7.5$, and compare their continuum properties, $H\alpha$ emission, and Balmer decrements. Consistent with previous studies, we find that LRDs exhibit broader $H\alpha$ profiles than LBDs systematically (with median FWHM values of 2319^{+71}_{-64} and 1424^{+56}_{-43} km s^{−1}, respectively). We also confirm that LRDs show larger Balmer decrements, with median $\log_{10}(F_{H\alpha}/F_{H\beta}) = 1.01 \pm 0.03$, compared with 0.47 ± 0.01 for LBDs. Line-of-sight effects could explain both results. However, most significantly, we find that the LRD fraction increases strongly with $H\alpha$ line luminosity: LRDs are approximately six times more luminous in $H\alpha$ than LBDs. This robust finding is much harder to explain through unification with LBDs by orientation. Despite their higher line luminosities, LRDs have a moderately lower median $H\alpha$ equivalent width than LBDs, consistent with more of the raw emission being reprocessed into the rest-frame optical continuum in LRDs. These results, coupled with the systematically lower [O III]/ $H\beta$ ratios found in LRDs and the decline of $H\alpha$ equivalent width towards the reddest optical slopes, disfavour a simple orientation-based link between LRDs and LBDs, and are instead more consistent with the predictions of gas-cocoon models in which increasing gas column density explains the apparent transition from LBDs to LRDs.

Key words: galaxies: high-redshift – galaxies: active – galaxies: emission lines – quasars: supermassive black holes

1 INTRODUCTION

Little Red Dots (LRDs) have emerged as one of the most unexpected discoveries enabled by the *James Webb Space Telescope* (*JWST*). Initially identified in photometric samples of high-redshift sources, their bright rest-frame optical emission was interpreted as evidence for unexpectedly massive galaxies at early times (Labbé et al. 2023). Subsequent observations showed that strong emission lines and active galactic nucleus (AGN) emission contribute significantly to their rest-frame optical emission. At the same time, spectroscopy revealed compact sources with a blue UV continuum slope and a red rest-frame optical continuum, characterised by a distinctive V-shaped spectral

energy distribution (Matthee et al. 2024). Larger photometric and spectroscopic samples have since shown that LRDs are common at $z \gtrsim 3$, with relatively little evolution in their number density at $z \gtrsim 4$ (Pérez-González et al. 2024). However, recent studies suggest more complex evolution at lower redshifts (Kapoor et al. 2026; Inayoshi 2025; Loiacono et al. 2026), while LRDs appear to span a broad, possibly continuous range of continuum and emission-line properties (de Graaff et al. 2025a; Barro et al. 2026a; Rinaldi et al. 2026; Pérez-González et al. 2026; Matthee et al. 2026).

Despite this rapid observational progress, the physical nature of LRDs remains unclear. Their compact morphologies, broad Balmer emission, and high-ionisation lines favour the presence of accreting black holes (Matthee et al. 2024; Greene et al. 2024; de Graaff et al. 2025a; Maiolino et al. 2025; Papovich et al. 2026). However, most

* E-mail: lbarrufe@roe.ac.uk

remain individually undetected in X-rays and appear substantially X-ray weaker than expected from standard AGN spectral energy distributions (Ananna et al. 2024; Maiolino et al. 2025). Their weak or absent thermal dust emission challenges conventional dust-obscured AGN models (Casey et al. 2024; Akins et al. 2025), while their broad emission lines, combined with weak X-ray and infrared emission, are difficult to reconcile with a scaled-down interpretation of classical quasars (QSOs) as a whole. Nevertheless, growing spectroscopic evidence favours a compact central engine in many LRDs, and some show signatures of evolving towards more typical AGN and QSOs (Fu et al. 2026; Ji et al. 2026; Korber et al. 2026). The origin of their unusual continua and broad emission lines, and the relative contributions of stellar and AGN emission, therefore remain under debate (Pérez-González et al. 2024; Herrero-Carrión et al. 2026; Pérez-González et al. 2026; Baggen et al. 2024; Scholtz et al. 2026).

Several recent models instead invoke a compact accreting source surrounded by a dense gaseous environment to explain the unusual combination of LRD properties. In these scenarios, surrounding material can absorb and reprocess the intrinsic emission, potentially suppressing the escaping X-rays while contributing to the optical continuum and Balmer-line emission. Proposed interpretations include rapidly accreting black holes enclosed by dense ionised cocoons (Rusakov et al. 2026; Sneppen et al. 2026a), self-obscured super-Eddington systems (Inayoshi & Ho 2025; Madau & Maiolino 2026), and “black-hole stars”, in which a compact central source is embedded within a gas-dominated photosphere (Naidu et al. 2025; Sun et al. 2026). Despite their differences, these models share the idea that the observed properties of LRDs are strongly influenced by the physical conditions of the gas surrounding the central engine.

The recent identification of compact sources with similar emission-line properties but blue rest-frame optical continua has further broadened the diversity of this population. Commonly referred to as Little Blue Dots (LBDs), these sources were first defined by Brazzini et al. (2026). Like LRDs, LBDs exhibit X-ray weakness and broad Balmer emission, but their spectra lack the characteristic V-shaped continuum. The relationship between the two populations remains unclear: some observational studies suggest a continuous transition between LBDs and LRDs (Asada et al. 2026), whereas others interpret them as physically distinct classes (Brazzini et al. 2026).

Two broad scenarios have been proposed for the connection between LBDs and LRDs. In the orientation-based model, LRDs and LBDs are SMBHs accreting at super-Eddington rates and represent the same underlying population viewed along different lines of sight, analogous to the classical AGN unification model (Madau & Maiolino 2026; Madau et al. 2026). Alternatively, models proposed by Sneppen et al. (2026a,b) place a gas cocoon surrounding the accreting source, with the different spectral properties arising from changes in the physical conditions of the gas around the central engine. In this scenario, variations in the gas column density are the primary driver of the observed spectral differences, linking LBDs and LRDs as different manifestations of the same underlying population. These scenarios make different predictions for the observable properties of LBDs and LRDs, but have not yet been tested systematically using a homogeneous observational sample.

In this work, we construct and compare homogeneous samples of LRDs and LBDs to investigate whether their observed properties are linked primarily by orientation or instead reflect intrinsic physical differences. We develop a new selection based on the properties known to characterise these compact sources, combining strong H α emission, a blue UV continuum, and compact morphology, and then distinguish LRDs from LBDs according to their rest-frame optical

continua. Section 2 describes the observations, while Section 3 details the sample selection and analysis methodology. We present our results in Section 4, discuss their implications in Section 5, and summarise our conclusions in Section 6.

2 OBSERVATIONS

In this section, we describe the *JWST*/NIRSpec spectroscopy and ancillary *JWST*/NIRCam imaging used to construct and characterise the LRD and LBD samples, together with the archival X-ray data used to search for counterparts.

2.1 NIRSpec spectroscopic data

We use the public *JWST*/NIRSpec spectroscopy compiled in version 4.4 of the DAWN *JWST* Archive (DJA; Valentino et al. 2025), which contains approximately 86,000 spectra from a wide range of observing programmes and extragalactic fields. The archive primarily comprises low-resolution PRISM observations obtained with the NIRSpec micro-shutter assembly, complemented by medium- and high-resolution spectroscopy where available. The spectra were reduced uniformly with *msaexp*, including background subtraction, optimal one-dimensional extraction, wavelength and flux calibration, and corrections for wavelength-dependent slit losses (Heintz et al. 2024; Pollock et al. 2026).

We apply the same initial selection to LRDs and LBDs. To ensure reliable source identification and redshift measurements, we retain only spectra with `grade=3`, corresponding to robust spectroscopic redshifts based on the DJA quality assessment. This reduces the parent sample to 37,528 spectra. We adopt the DJA spectroscopic redshift, `zline`, and use the reduced one-dimensional spectra to measure the H α equivalent width and UV continuum slope for each source. These measurements are combined with the compactness derived from the NIRCam imaging described below to construct the final sample (Section 3).

2.2 Photometric data

Although the LRD and LBD classifications are primarily based on spectroscopic properties, ancillary imaging is required to quantify the rest-optical compactness of the sources. We retrieve the corresponding *JWST*/NIRCam F444W mosaics from the DJA. This filter provides the most widely available long-wavelength NIRCam coverage across the heterogeneous fields included in the spectroscopic archive.

We determine the source centroid in the F444W image and perform aperture photometry within circular apertures of 0.1'' and 0.2'' radii, including the associated photometric uncertainties. From these measurements, we calculate the compactness parameter

$$C_{\text{F444W}} = \frac{F_{\text{F444W}}(0.2'')}{F_{\text{F444W}}(0.1'')},$$

following the approach of de Graaff et al. (2025b). Lower values indicate that a larger fraction of the light is concentrated near the source centre. The compactness criterion applied to the sample is described in Section 3.

We additionally measure F444W fluxes within 0.4''-radius apertures and correct them to total fluxes using the NIRCam curves of growth, from which we additionally derive total AB magnitudes. At redshift $z = 6$ ($z = 2$), this corresponds to ~ 2.4 (~ 3.4) kpc, which is substantially larger than the expected radii of LRDs at these redshifts

($\lesssim 1$ kpc, see e.g., Cloonan et al. 2026). These measurements are used to characterise the photometric and structural properties of the final LRD and LBD samples.

To characterise the X-ray properties of the sample, we use archival data from the two X-ray catalogues providing the deepest per-pointing sensitivity over the fields covered by our sample: 5XMM-DR15 (Webb et al. 2026) and the Chandra Source Catalogue (CSC 2.1; Evans et al. 2024). We queried the XMM-Newton Science Archive and the CSC Quick Search interfaces for X-ray sources and available coverage at the positions of the sample galaxies. The cross-matching procedure and the identification of reliable X-ray counterparts are described in Section 3.

3 LITTLE (RED AND BLUE) DOTS METHODOLOGY AND SELECTION

To construct a large and homogeneous sample of compact AGN-like sources, we base our selection on three observables that capture the main phenomenological properties of LRDs and LBDs: strong broad $H\alpha$ emission, a blue UV continuum, and compact morphology. Broad-line selections are commonly used to identify AGN (Maiolino et al. 2025) and have also been applied in the context of LRDs. More recently, Rusakov et al. (2026) used broad $H\alpha$ emission, defined as $\text{FWHM} > 1000 \text{ km s}^{-1}$, to select LRDs. In their sample of twelve sources, however, only three are defined as LRDs, while the remaining objects are classified as LRD-like AGN. These results illustrate that broad-line selection alone does not uniquely identify LRDs. Similarly, Hviding et al. (2025) distinguish a wider population of broad-Balmer AGN from spectroscopic LRDs, defining the latter as sources that additionally exhibit a V-shaped continuum and a dominant rest-optical point-source component.

However, FWHM measurements are particularly sensitive to spectral resolution and line-profile modelling, both of which are important limitations for the predominantly low-resolution JWST/NIRSpec spectra available for large samples.

For these reasons, we use the $H\alpha$ equivalent width as a more robust indicator of strong line emission. Large $H\alpha$ equivalent widths are a characteristic property of LRDs (e.g., see de Graaff et al. 2025a), consistent with their high $H\alpha$ luminosities and faint underlying continua. This allows us to identify sources with prominent $H\alpha$ emission without requiring a resolved broad-line component.

A second defining property of LRDs is their blue UV slope. This criterion was already present in the first works defining LRDs, such as Matthee et al. (2024), which required a blue UV continuum. However, the exact threshold varies between studies, from approximately $\beta_{\text{UV}} < -0.5$ to $\beta_{\text{UV}} < -0.2$. Finally, compactness is one of the most distinctive properties of LRDs and is also expected for AGN-dominated sources. Combining these three observables allows us to identify a broad parent population of compact sources with strong and broad $H\alpha$ emission and blue UV continua, without imposing the red or blue rest-frame optical continuum that distinguishes LRDs from LBDs. The LRD and LBD populations are then distinguished according to their rest-frame optical continuum properties, as described below.

3.1 Measuring β_{UV} , β_{opt} and $H\alpha$ emission line properties

The UV slope was measured over the rest-frame wavelength interval $1300 \leq \lambda_{\text{rest}} < 3646 \text{ \AA}$, where $3646 \text{ \AA}$ corresponds to the Balmer limit. Before fitting, regions of $\pm 30 \text{ \AA}$ around prominent UV spectral features were excluded. These comprise O I $\lambda 1304$, C II $\lambda 1335$, Si IV

$\lambda 1397$, C IV $\lambda 1549$, He II $\lambda 1640$, O III $\lambda 1663$, N III $\lambda 1750$, Si III $\lambda 1883$, C III $\lambda 1908$, C II $\lambda 2326$, Mg II $\lambda 2799$, and Mg I $\lambda 2853$. We retained pixels with valid spectral flags and positive rest-frame wavelengths and flux densities. When an error spectrum was available, we additionally required finite positive flux-density uncertainties and $\text{S/N} \geq 1$. Measurements were accepted only when at least 10 valid continuum pixels remained after the emission-line masking. The UV continuum was modelled as a power law, $f_{\lambda} \propto \lambda^{\beta_{\text{UV}}}$, fitted directly in linear f_{λ} space.

We then measured the $H\alpha$ fluxes and equivalent widths for each of the 37,528 NIRSpec spectra. For each source, the expected observed-frame $H\alpha$ wavelength was computed from the spectroscopic redshift as $\lambda_{H\alpha, \text{obs}} = 0.65628(1+z) \mu\text{m}$. We fitted the spectrum with a linear continuum plus a single Gaussian emission-line profile. The initial line window was set to $\pm 0.01 \mu\text{m}$ around the expected $H\alpha$ wavelength, while the continuum was estimated from two side bands of width $0.02 \mu\text{m}$, separated from the line window by $0.01 \mu\text{m}$. When fewer than 10 line pixels or 15 continuum pixels were available, the line and continuum windows, together with their separation, were expanded iteratively by a factor of 1.5, up to six times. A preliminary linear continuum fit was sigma-clipped once (at 3σ), using a median-absolute-deviation estimate of the residual scatter, before deriving the final continuum model. The $H\alpha$ flux was then obtained from the analytic integral of the best-fitting Gaussian profile. The equivalent width was computed by integrating the continuum-subtracted line profile divided by the fitted continuum over the line window. Finally, the observed-frame EW was converted to rest-frame ($EW_{\text{rest}} = EW_{\text{obs}}/(1+z)$).

Finally, for the optical continuum, we modelled the rest-frame continuum over $3646\text{--}7000 \text{ \AA}$ directly in linear f_{λ} space using an inverse-variance-weighted power law, $f_{\lambda} = A(\lambda/5000)^{\beta_{\text{opt}}}$. We masked the regions around [O II] $\lambda 3727$, [Ne III] $\lambda 3869$, H δ , H γ , H β , [O III] $\lambda \lambda 4959, 5007$, He I $\lambda 5876$, [O I] $\lambda 6300$, [N II] $\lambda \lambda 6548, 6583$, H α , and [S II] $\lambda \lambda 6716, 6731$. In addition to the individual line masks, we excluded the broad H β + [O III] and H α + [N II]+ [S II] complexes. We retained valid spectral pixels with finite positive wavelengths, finite fluxes of either sign, and finite positive flux uncertainties. No per-pixel S/N threshold or positive-flux requirement was imposed. Fits were retained when at least 10 continuum pixels remained after masking. The resulting β_{opt} measurements therefore trace the rest-frame optical continuum while minimising the contribution from strong nebular emission lines.

3.2 Sample selection criteria

With this in mind, we adopt a two-stage procedure to construct the LRD and the LBD samples. We first identify a parent population of compact, UV blue, strong- $H\alpha$ emitters using criteria that are independent of the rest-frame optical continuum. We then classify these sources as LRDs or LBDs according to their rest-frame optical continuum slope, β_{opt} .

Based on the photometric and spectroscopic measurements described in Section 2.2 and 2.1, respectively, the parent sample is defined using the following criteria:

- (i) A rest-frame $H\alpha$ equivalent width in the range $200 \text{ \AA} < EW_{H\alpha} < 2000 \text{ \AA}$.
- (ii) A conservative blue UV continuum slope $\beta_{\text{UV}} < -0.5$.
- (iii) A compact morphology, defined via the flux ratio $F444W(0.2'')/F444W(0.1'') < 1.7$, consistent with the compactness criteria adopted in previous LRD studies (see, e.g. de Graaff et al. 2025b).

These three criteria yield an initial sample of 555 sources. This parent sample is designed to include both red and blue compact AGN-like sources with prominent $H\alpha$ emission, without imposing any requirement on the shape of the rest-frame optical continuum. The final sample is dominated by low-resolution NIRSpect/PRISM spectroscopy ($R \sim 100 - 500$), with only three sources having medium-resolution spectra ($R \sim 1000$). This selection is inherently biased towards sources for which both strong $H\alpha$ emission and a measurable UV continuum slope can be identified, and therefore does not represent an unbiased sample of the broader AGN population.

We subsequently use β_{opt} to split the sources into LRDs and LBDs (see Section 3.1 for the measurement methodology). Specifically, we compare their locations in the $\beta_{\text{UV}}-\beta_{\text{opt}}$ plane with the LRD/LBD regions defined by [Brazzini et al. \(2026\)](#): sources with $\beta_{\text{opt}} > 0$ are classified as LRD candidates, while those with $\beta_{\text{opt}} \leq 0$ are classified as LBD candidates.

We then apply additional quality cuts to ensure robust photometric and optical-slope measurements. We required $S/N > 5$ for the F444W fluxes measured within both the $0.1''$ and $0.2''$ apertures (168 sources removed). We further required $F444W \text{ mag} < 27$ for LRD candidates (18 additional sources removed). We also required a reliable β_{opt} classification, defined by the expected sign being recovered in at least 95% of 1000 MC refits: $P(\beta_{\text{opt}} > 0) \geq 0.95$ for all LRDs and $P(\beta_{\text{opt}} < 0) \geq 0.95$ for LBDs (76 sources removed). Finally, we inspected the repeated spectral observations and retained a single measurement per unique source (12 duplicate entries removed), keeping the observation with the higher signal-to-noise ratio.

After applying all quality cuts, the final robust sample comprises 280 sources: 89 LRDs in total, including 72 previously reported LRDs from [de Graaff et al. \(2025a\)](#) (68% recovery fraction), 17 newly identified LRDs, and 191 LBDs. The final selection is drawn from the surveys described in the Appendix Table A1. Hence, our sample LRDs comprise $\approx 32\%$ of the final sample, consistent with the 10 – 30% reported by [Brazzini et al. \(2026\)](#).

Fig. 1 shows the resulting classification in the $\beta_{\text{UV}}-\beta_{\text{opt}}$ plane. For all sources, including those previously identified as LRDs in the literature, both continuum slopes were remeasured from the spectra using the homogeneous procedure described in Section 3.1. Hence, the distributions are directly comparable. The underpopulated region around $\beta_{\text{opt}} = 0$ results from our requirement that the slope sign be recovered in at least 95% of the Monte Carlo realisations. Sources that do not satisfy this criterion, including 20 LRDs from [de Graaff et al. \(2025a\)](#), are excluded from the final sample and may represent a transitional population (see Section 5).

We quantify the continuum-slope distributions of the two subsamples by computing their median values and estimating the associated uncertainties from 10,000 bootstrap resamples. The LRDs have a median optical slope of $\beta_{\text{opt}} = 0.83^{+0.03}_{-0.05}$, while the LBDs show a markedly bluer distribution, with $\beta_{\text{opt}} = -2.63^{+0.15}_{-0.12}$. The median separation between the two populations is therefore $\Delta\beta_{\text{opt}} = 3.46^{+0.13}_{-0.17}$. As β_{opt} enters the classification itself, this offset should not be interpreted as an independent significance test; rather, it quantifies the extent to which the two distributions remain separated from the adopted boundary. A smaller but still significant difference is also present in the UV slopes, with median values of $\beta_{\text{UV}} = -1.60^{+0.07}_{-0.06}$ for LRDs and $\beta_{\text{UV}} = -1.97^{+0.02}_{-0.05}$ for LBDs.

3.3 Newly identified LRDs

Here, we present the 17 LRDs newly identified through our homogeneous, self-consistent selection, and examine their continuum

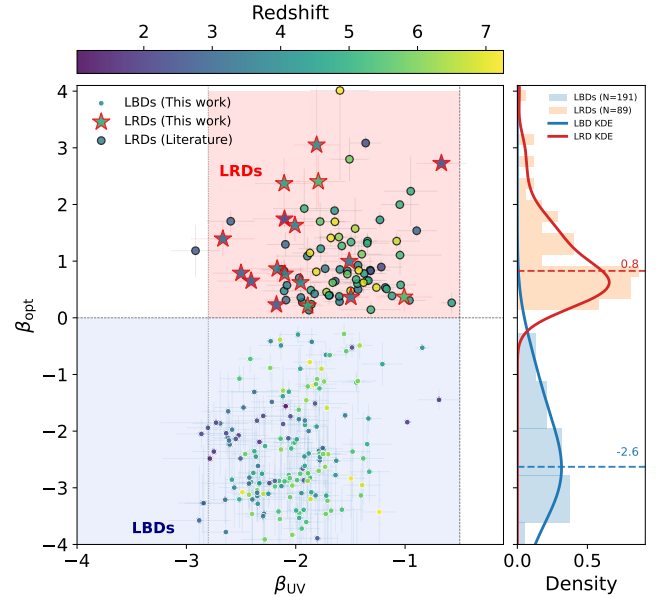


Figure 1. Rest-frame optical continuum slope, β_{opt} , as a function of the UV continuum slope, β_{UV} , for the selected spectroscopic sample. Points are colour-coded by spectroscopic redshift, as indicated by the colour bar at the top. The red- and blue-shaded regions mark the LRD and LBD classification regions, respectively, following [Brazzini et al. \(2026\)](#). Red stars indicate the 17 newly identified LRDs from this work, while red-edged circles show the previously reported LRDs ([de Graaff et al. 2025a](#)); the remaining blue-continuum sources are classified as LBDs. Error bars indicate the uncertainties on the fitted continuum slopes. The right-hand panel shows the corresponding β_{opt} distributions and kernel density estimates for the 89 LRDs and 191 LBDs. Dashed horizontal lines mark the median values, $\beta_{\text{opt}} = 0.8$ for LRDs and $\beta_{\text{opt}} = -2.6$ for LBDs.

properties relative to the previously known population. All candidates were visually inspected and lie within the LRD region of the $\beta_{\text{UV}}-\beta_{\text{opt}}$ plane shown in Fig. 1.

Fig. 2 compares a newly identified LRD with a representative LBD selected from the same parent sample. While both sources exhibit blue UV continua, the LRD shows a redder optical continuum and more prominent $H\alpha$ emission. The newly identified LRDs generally exhibit less extreme UV-optical slopes than previously reported systems. Their more moderate continuum slopes may help explain their absence from previous LRD selections and support a heterogeneous population spanning a sequence of physical properties ([Matthee et al. 2026](#); [Pérez-González et al. 2026](#)). In Section 4, we investigate whether the differences in optical continuum and $H\alpha$ properties between LRDs and LBDs reflect distinct physical regimes or can be attributed to different phases or orientations of the broader AGN population.

3.4 X-ray counterparts

We cross-matched the 280 sources in our final sample with the 5XMM-DR15 ([Webb et al. 2026](#)) and Chandra Source Catalogue (CSC 2.1; [Evans et al. 2024](#)), using matching radii of 6 arcsec and 2 arcsec, respectively. We find only two secure X-ray counterparts, both LBDs: a spectroscopically confirmed Seyfert 1 AGN detected in multiple XMM-Newton observations, and a source coincident with a known QSO at $z = 3.44$. Other candidate associations are either low-significance detections, unrelated foreground sources, or too faint

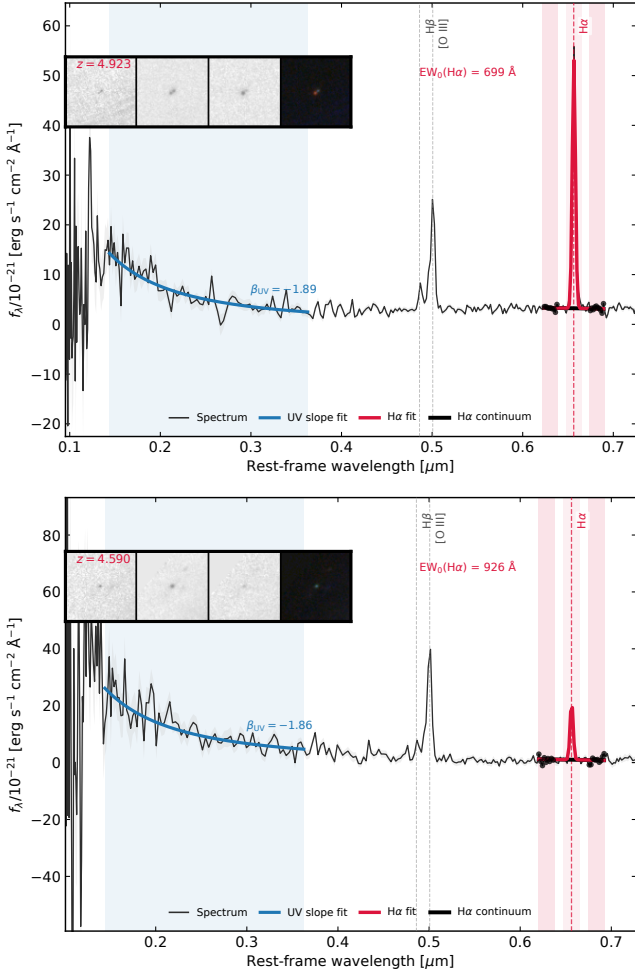


Figure 2. Example spectra of one of the 17 newly-identified LRDs uncovered in this study, shown in the upper panel, and a representative LBD, shown in the lower panel. In each case, the spectrum is shown in black, while the fitted $H\alpha$ emission line and underlying local continuum used to measure the rest-frame $H\alpha$ equivalent width are shown in red. The red-shaded regions mark the wavelength intervals used in the $H\alpha$ EW measurement. The UV continuum fit is shown in blue over the short-wavelength region of the spectrum, with the blue shaded region marking the wavelength range used to derive β_{UV} . The inset panels show the NIRC2 F090W, F200W, and F444W imaging (used to assess the compactness criterion and to perform the size measurements), together with the combined RGB colour image. The newly identified LRD displays a blue UV slope together with a strong $H\alpha$ emission line. At the same time, the representative LBD shows a more prominent blue UV slope, a substantially weaker $H\alpha$ line, and a flatter optical continuum.

and/or offset to be considered secure. We therefore find no robust X-ray detections among the LRDs and two among the LBDs. Given the inhomogeneous X-ray coverage and sensitivity, these results do not allow us to conclude that there are intrinsic differences between the two populations, but are broadly consistent with previous reports of X-ray weakness among *JWST*-selected AGN (Maiolino et al. 2025) and LRDs (Ananna et al. 2024).

4 UNVEILING THE NATURE OF LRDs AND LBDs

The contrasting spectral properties of LRDs and LBDs identified in Section 3.2 raise the question of whether they represent intrinsically distinct populations, different evolutionary phases, or a single population observed along different lines of sight. In the super-Eddington unification model proposed by Madau & Maiolino (2026), LRDs are interpreted as the dust-reddened, high-inclination counterparts of the less-obscured LBDs, with both classes powered by the same underlying AGN engine. By contrast, other scenarios associate the two populations with different phases of galaxy-SMBH co-evolution, potentially driven by differences in obscuration, accretion state, or host-galaxy properties (Asada et al. 2026; Brazzini et al. 2026). We test these alternatives through a homogeneous comparison of their continuum, emission-line, and structural properties.

4.1 $H\alpha$ constraints on a purely orientation-driven LRD-LBD scenario

We first compare the $H\alpha$ emission properties of LRDs and LBDs, focusing on the line FWHM, luminosity, and rest-frame equivalent width. These quantities provide complementary constraints on the line-emitting region: the FWHM traces the characteristic gas kinematics (albeit potentially with a contribution from scattering), while the luminosity and equivalent width quantify the strength of the line emission, both intrinsically and relative to the underlying continuum. In the simplest orientation-only scenario, in which LRDs and LBDs are drawn from the same intrinsic population and differ primarily in viewing angle, their $H\alpha$ luminosity distributions should be broadly comparable, aside from attenuation or anisotropy effects. Together, these observations allow us to test whether the differences between the two populations can be explained primary by a viewing geometry or instead require differences in their intrinsic emission-line properties. All $H\alpha$ measurements were obtained using the homogeneous fitting procedure described in Section 3.1.

Fig. 3 compares the $H\alpha$ FWHM, luminosity, and rest-frame equivalent-width distributions of the 89 LRDs and 191 LBDs in the final sample. Unless stated otherwise, uncertainties on the median values were estimated by resampling each population with replacement 10,000 times, with the 16th and 84th percentiles of the bootstrap-median distribution defining the lower and upper uncertainties, respectively.

It can be seen that the LRDs are systematically shifted towards broader $H\alpha$ line widths, with median values of 2319^{+71}_{-64} km s⁻¹ and 1424^{+56}_{-43} km s⁻¹ for LRDs and LBDs, respectively. The corresponding median offset is $\Delta\text{FWHM} = 895^{+73}_{-95}$ km s⁻¹, and the two distributions are statistically distinct ($p_{KS} = 3.42 \times 10^{-14}$). The $H\alpha$ luminosity distributions show an even stronger separation; LRDs are systematically more luminous in $H\alpha$ than LBDs, with median values of $\log(L_{H\alpha}/\text{erg s}^{-1}) = 42.634^{+0.091}_{-0.092}$ and $41.877^{+0.041}_{-0.010}$, respectively. The corresponding offset, $\Delta \log L_{H\alpha} = 0.757^{+0.090}_{-0.099}$ dex, implies that LRDs are, on average, a factor of ≈ 6 times more luminous in their observed $H\alpha$ line emission. The two distributions are statistically distinct ($p_{KS} = 2.01 \times 10^{-20}$), showing that LRDs effectively occupy the high-luminosity tail of the compact sample. Although the FWHM distributions also differ significantly, the luminosity contrast is the more pronounced distinction between the two populations.

The $H\alpha$ equivalent-width distributions show a much more substantial overlap within the common selection interval, $200 < \text{EW}_{\text{rest}}(H\alpha) < 2000$ Å. LRDs have a moderately lower median equivalent width ($\text{EW}(H\alpha) = 699^{+36}_{-42}$ Å) than LBDs, which have

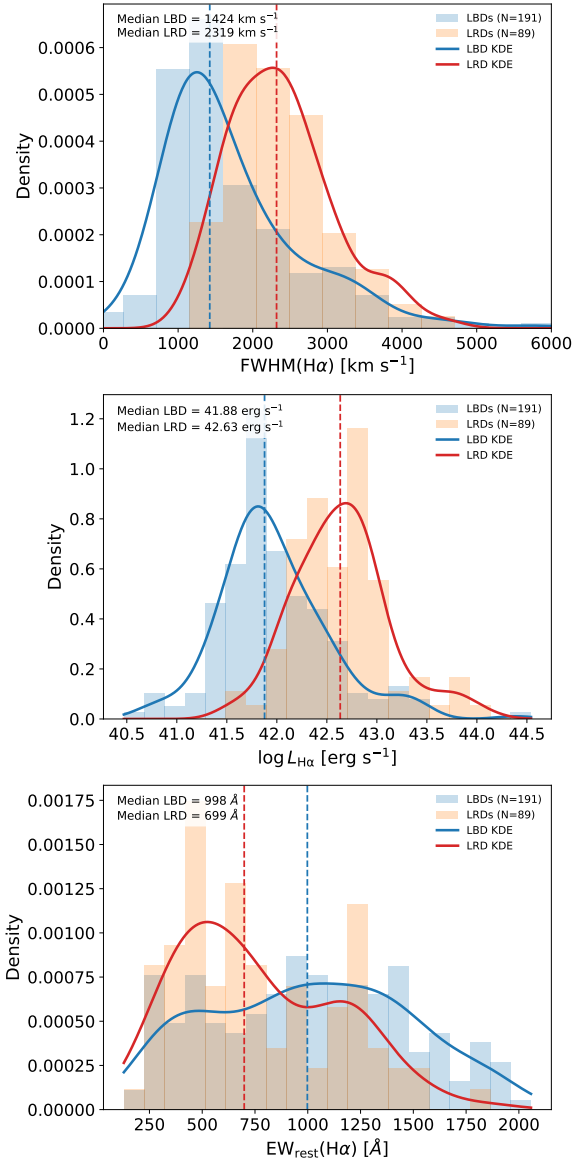


Figure 3. Normalised histograms and kernel density estimates (KDEs) of the $H\alpha$ FWHM, luminosity, and rest-frame equivalent width for the 89 LRDs and 191 LBDs in our sample. LRDs and LBDs are shown in red and blue, respectively, and dashed vertical lines mark the median values. The FWHM distributions overlap, but LRDs are systematically shifted towards broader lines, with median values of 2319^{+71}_{-64} km s $^{-1}$ for LRDs and LBDs, respectively. An even stronger separation is observed in $H\alpha$ luminosity, with median values of $\log(L_{H\alpha}/\text{erg s}^{-1}) = 42.634^{+0.091}_{-0.092}$ for LRDs and $41.877^{+0.041}_{-0.010}$ for LBDs. The equivalent-width distributions show substantial overlap within the common selection interval, $200 < \text{EW}_{\text{rest}}(H\alpha) < 2000$ Å, although LRDs have a lower median EW, 699^{+36}_{-42} Å, than LBDs, 998^{+51}_{-29} Å. Thus, LRDs are characterised by broader and substantially more luminous $H\alpha$ emission, while occupying a broadly comparable EW range imposed by the parent-sample selection.

$\text{EW}(H\alpha) = 998^{+51}_{-29}$ Å. However, because the $H\alpha$ EW is part of the parent sample selection, this difference should not be interpreted as an independent physical distinction between the two populations. The key result is instead that LRDs are substantially more luminous in $H\alpha$.

Taken together, these results show that LRDs exhibit broader $H\alpha$ profiles and, more notably, substantially higher $H\alpha$ luminosities than

LBDs. This luminosity offset is unlikely to be driven by differences in redshift coverage. The two samples span largely overlapping redshift ranges, with median values of $z_{\text{med}} = 4.53^{+0.30}_{-0.30}$ for LRDs and $z_{\text{med}} = 4.80^{+0.27}_{-0.21}$ for LBDs. Although the redshift distributions differ modestly, all sources have reliable spectroscopic redshifts, and the LRDs in our sample do not preferentially lie at higher redshift. The enhanced $H\alpha$ luminosity of LRDs therefore cannot be attributed solely to a systematic redshift difference between the two samples.

Given the heterogeneous definitions of LBDs adopted in the literature, we repeated the comparison after imposing the same broad-line criterion, $\text{FWHM}(H\alpha) > 2000$ km s $^{-1}$, on both populations. The $H\alpha$ luminosity offset persists and increases, reaching $\Delta \log L_{H\alpha} \sim 1$ dex corresponding to a factor of ~ 10 . Thus, the enhanced $H\alpha$ luminosity of LRDs is not driven by differences in redshift coverage or by the adoption of different broad-line thresholds. We discuss the implications of this result, and the impact of the heterogeneous LRD and LBD definitions used in the literature, in Section 5.

In the simplest orientation-only scenario, the probability of observing a source as an LRD should be approximately independent of the intrinsic luminosity. Since broad $H\alpha$ luminosity provides an observational proxy for bolometric luminosity in broad-line AGN, we therefore test explicitly whether the LRD fraction varies systematically with $L_{H\alpha}$.

Fig. 4 shows the LRD fraction in five independent luminosity bins centred at $\log L_{H\alpha} = 41.00, 41.75, 42.51, 43.26$, and 44.01 . The uncertainties are computed using binomial confidence intervals. The LRD fraction increases strongly with $H\alpha$ luminosity. No LRDs are found in the lowest-luminosity bin ($0/19$), while LRDs account for $14/124$ sources ($f_{\text{LRD}} = 0.11$) at $\log L_{H\alpha} = 41.75$. The fraction rises sharply to $52/102$ ($f_{\text{LRD}} = 0.51$) at $\log L_{H\alpha} = 42.51$ and to $18/29$ ($f_{\text{LRD}} = 0.62$) at $\log L_{H\alpha} = 43.26$. In the highest-luminosity bin, five of the six sources are LRDs, corresponding to $f_{\text{LRD}} = 0.83$, although this value is subject to substantial small-number uncertainty. Thus, LRDs transition from being rare at low $H\alpha$ luminosities to constituting the majority of the population above $\log L_{H\alpha} \approx 42.5$, indicating that they preferentially occupy the high $H\alpha$ -luminosity tail of the compact-source population.

This trend is hard to explain in the simplest luminosity-independent orientation scenario. If LRDs and LBDs are drawn from the same intrinsic population and differ only in viewing angle, the fraction observed as LRDs would not be expected to increase strongly or even reduce somewhat with increasing observed $H\alpha$ luminosity (due to increasing attenuation). Using $L_{H\alpha}$ as an empirical proxy, the observed increase in f_{LRD} therefore suggests that the LRD phenomenon is preferentially associated with systems which produce intrinsically more luminous $H\alpha$ emission. This may reflect intrinsic differences in accretion state, obscuration geometry, central engine properties, or host galaxy conditions, although attenuation, anisotropy, and selection effects may also influence the observed $H\alpha$ luminosity.

To investigate whether LRDs and LBDs occupy different evolutionary phases, we compare their $H\alpha$ luminosities as a function of redshift in Fig. 5. The sample used in this analysis is again our final sample, which contains 89 LRDs and 191 LBDs with reliable $H\alpha$ luminosities. At fixed redshift, LRDs are systematically more $H\alpha$ luminous than LBDs, with the median offset remaining visible across the full redshift range of $2 < z < 8$. The LRD fraction, however, shows no clear monotonic evolution with redshift. It increases from $11/38$ ($f_{\text{LRD}} = 0.29$) at $2 < z < 3$ to $22/57$ ($f_{\text{LRD}} = 0.39$) at $3 < z < 4$ and reaches $23/55$ ($f_{\text{LRD}} = 0.42$) at $4 < z < 5$. It then decreases to $20/81$ ($f_{\text{LRD}} = 0.25$) at $5 < z < 6$, before rising again to $13/42$ ($f_{\text{LRD}} = 0.31$) over $6 < z < 8$. Thus, the persistent luminosity offset between LRDs and LBDs does not appear to be

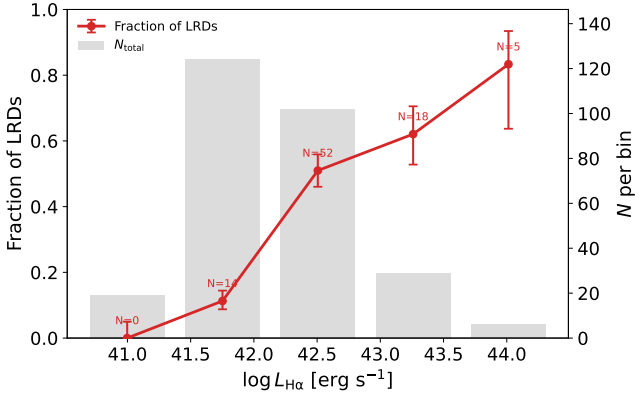


Figure 4. LRD fraction as a function of $H\alpha$ luminosity for the final sample of 89 LRDs and 191 LBDs. Red points show the fraction $f_{\text{LRD}} = N_{\text{LRD}}/(N_{\text{LRD}} + N_{\text{LBD}})$ in independent bins of $\log L_{H\alpha}$, with error bars indicating binomial confidence intervals. Grey bars show the total number of sources in each bin, referenced to the right-hand axis, while the labels indicate the corresponding number of LRDs. The LRD fraction increases strongly with $H\alpha$ luminosity, from $f_{\text{LRD}} = 0.11$ at $\log L_{H\alpha} = 41.75$ to 0.51 at $\log L_{H\alpha} = 42.51$ and 0.83 in the highest-luminosity bin. The overall trend shows that LRDs become increasingly prevalent towards the high- $L_{H\alpha}$ tail of the compact-source population.

driven simply by a monotonic change in their relative abundance with redshift. In the same figure, we also examine how the incidence of LRDs changes along the high-luminosity tail by computing the cumulative LRD fraction above successive $H\alpha$ luminosity thresholds within the redshift range $4 < z < 7.5$. At $\log L_{H\alpha} \geq 41.8$, LRDs account for 54/139 sources, corresponding to $f_{\text{LRD}} = 0.39$. This fraction rises as progressively fainter sources are excluded, reaching 50/99 ($f_{\text{LRD}} = 0.51$) and 46/81 ($f_{\text{LRD}} = 0.57$) at higher luminosity thresholds. The lower bound of the 68% Wilson confidence interval first exceeds 0.5 at $\log L_{H\alpha} = 42.30$, indicating that LRDs constitute a statistically robust majority above this luminosity. The cumulative fraction continues to increase towards the most luminous end, reaching 20/28 ($f_{\text{LRD}} = 0.71$), although the highest-luminosity measurements are increasingly affected by small-number statistics, with only 7/11 and ultimately 4/4 sources remaining. This complementary cumulative analysis therefore shows that, within the same redshift interval, LRDs become progressively more prevalent towards the high- $L_{H\alpha}$ tail of the compact-source population.

Taken together, the $H\alpha$ results show that the differences between LRDs and LBDs are not limited to their continuum colours. LRDs exhibit broader line profiles and, most notably, substantially higher $H\alpha$ luminosities, and consequently their incidence increases strongly towards the high-luminosity tail of the line luminosity distribution. These trends persist after accounting for the overlapping redshift distributions and after imposing the same broad-line criterion on both populations. They therefore place tension on the simplest luminosity-independent orientation-only scenario, in which the probability of observing a source as an LRD would not be expected to increase with observed $H\alpha$ luminosity. However, attenuation, anisotropic emission, and selection effects may still contribute to the observed differences. The $H\alpha$ results alone therefore do not establish whether LRDs represent a distinct population or a different evolutionary phase, but they indicate that viewing angle alone is unlikely to provide a complete explanation.

4.2 Enhanced Balmer decrements in LRDs

The previous $H\alpha$ analysis indicates that one of the main differences between LRDs and LBDs lies in the strength of $H\alpha$ emission rather than in the characteristic velocity scale of the broad-line gas. This may result from continuum suppression, intrinsic differences in the line-emitting gas, or a combination of these effects. To distinguish between these possibilities, we examine the Balmer decrement of both LRDs and LBDs and compare their strengths with those displayed by low-redshift broad-line QSOs.

Fig. 6 shows the distribution of LRDs and LBDs in the plane of Balmer decrement, $\log_{10}(F_{H\alpha}/F_{H\beta})$, and rest-frame $H\alpha$ equivalent width. The analysis includes a reduced sample of 174 sources with valid positive $H\alpha$ and $H\beta$ flux measurements, comprising 52 LRDs and 122 LBDs. LRDs have a substantially larger median Balmer decrement than LBDs with $\log_{10}(F_{H\alpha}/F_{H\beta}) = 1.01 \pm 0.03$ and 0.47 ± 0.01 , respectively. In contrast, their median $H\alpha$ EW is lower (699^{+36}_{-42} versus 998^{+51}_{-29} Å). Thus, LRDs preferentially occupy the high-Balmer-decrement region, but not the high-EW tail of the distribution. Within the range covered by the bins, containing all 52 LRDs and 119 of the 122 LBDs, the LRD fraction increases strongly with Balmer decrement. It rises from 0/10 and 2/83 ($f_{\text{LRD}} = 0.02$) in the two lowest populated bins to 12/37 (0.32), 27/30 (0.90), and 11/11 (1.00) in the three highest bins. Although the final bin is sensitive to small-number statistics, the progressive increase demonstrates that LRDs dominate the high Balmer decrement tail of the compact broad-line population.

Both populations are displaced from the central locus of the low-redshift broad-line QSOs of Wu & Shen (2022), indicating that their Balmer-line and continuum properties are not typical of local QSOs. LRDs nevertheless remain within the outer QSO contours, overlapping the high-decrement tail of the local population, whereas LBDs lie close to the contour boundary because of their comparatively large $H\alpha$ EWs. The LRDs may therefore represent more strongly attenuated analogues of the extreme local-QSO population, while the LBD locus may reflect enhanced line emission, a weaker optical continuum, or different line-emitting conditions. The strong- $H\alpha$ selection should, however, be considered when interpreting the large LBD EWs.

The systematically different Balmer decrements also argue against LBDs being simply LRDs with a larger host-galaxy continuum contribution. Such dilution can alter the measured EWs but does not naturally produce the observed difference in $F_{H\alpha}/F_{H\beta}$. The systematically different Balmer decrements therefore suggest differences in the attenuation or radiative-transfer conditions affecting their Balmer-emitting gas.

The median LRD location is consistent with that obtained by the super-Eddington unification model of Sun et al. (2026); Madau & Maiolino (2026). The observed LRD ratio, $F_{H\alpha}/F_{H\beta} \approx 10.2$, is close to the model value of 10.6, while its median EW is broadly comparable to the predicted ≈ 830 Å. However, the relative EW trend is opposite to that expected from inclination-driven continuum suppression: LRDs have lower, rather than higher, $H\alpha$ EWs than LBDs. Thus, although orientation and dust attenuation may explain the large Balmer decrements of LRDs, they cannot alone reproduce the relative locations of the two populations. Additional differences in their continua, line-emitting regions, or accretion states are therefore required.

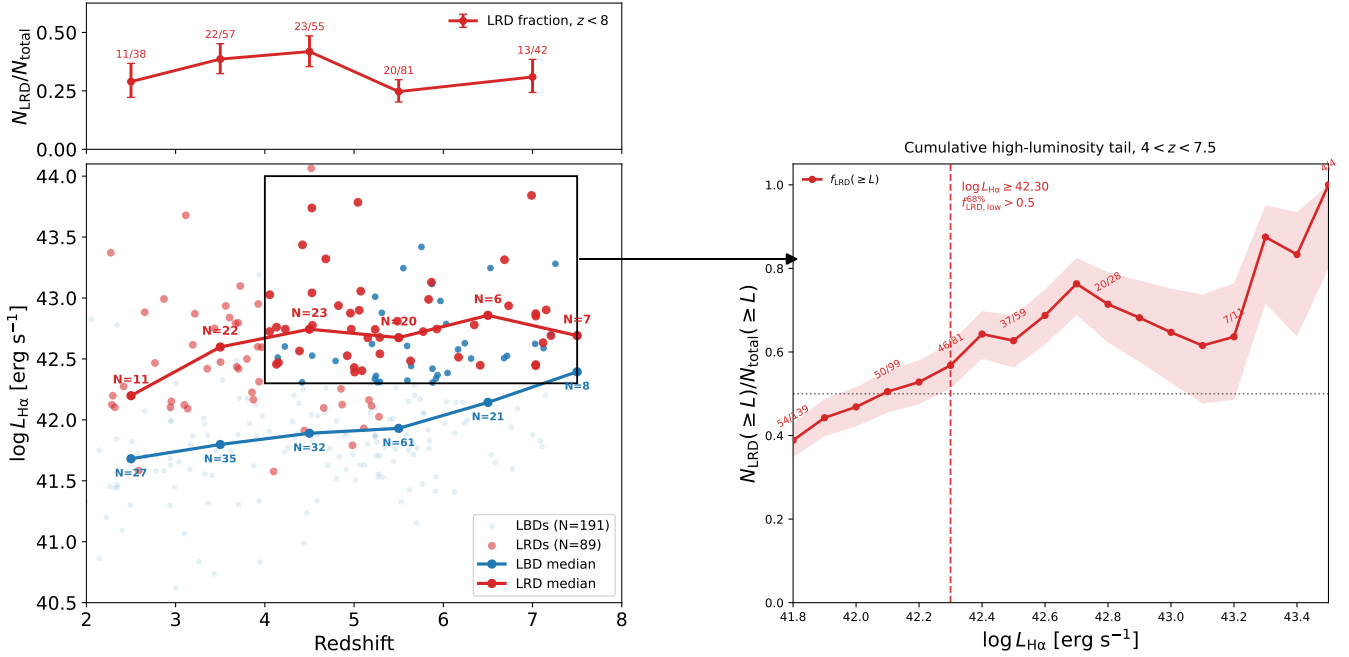


Figure 5. $H\alpha$ luminosity as a function of redshift for LRDs and LBDs. Individual sources are shown as red and blue points, respectively, while the red and blue lines indicate the median $\log L_{H\alpha}$ in each redshift bin. The number of sources contributing to each median bin is labelled in the lower-left panel. The upper-left panel shows the fraction of LRDs relative to the total LRD+LBD population as a function of redshift, computed for $z < 8$. The LRD fraction shows no clear monotonic evolution with redshift, whereas LRDs remain systematically more $H\alpha$ luminous than LBDs throughout the redshift range considered. The rectangle highlights the high-redshift, high-luminosity region connected to the cumulative analysis shown in the right-hand panel. The right-hand panel shows the cumulative LRD fraction above a given luminosity threshold, $f_{\text{LRD}}(\geq L) = N_{\text{LRD}}(\geq L)/N_{\text{total}}(\geq L)$, computed within $4 < z < 7.5$. The shaded region indicates the 68% Wilson binomial confidence interval. The cumulative LRD fraction generally increases towards higher luminosity thresholds, showing that LRDs become progressively over-represented in the high- $L_{H\alpha}$ tail. The vertical dashed line at $\log L_{H\alpha} = 42.30$ marks the first threshold at which the lower bound of the 68% Wilson confidence interval exceeds 0.5, indicating that LRDs constitute a statistically robust majority above this luminosity. The highest-luminosity measurements are based on only a few sources and therefore have comparatively large uncertainties.

5 DISCUSSION.

Our analysis reveals both continuity and systematic differences between LRDs and LBDs. Their continuum and emission-line properties overlap, making individual sources near the classification boundary difficult to assign robustly to either population. Nevertheless, LRDs exhibit broader and substantially more luminous $H\alpha$ emission, larger Balmer decrements, and lower median $H\alpha$ and [O III] equivalent widths than LBDs. In this section, we first examine the continuity between the two populations and the limitations of their spectroscopic classification, before considering the implications of their emission-line properties for orientation-based and gas-cocoon models.

The transition between LRDs and LBDs forms a continuous sequence in the $\beta_{\text{UV}}-\beta_{\text{opt}}$ plane, similar to the results of Asada et al. (2026). This extends the continuous distribution of optical colours and SED properties reported by Rinaldi et al. (2026) and is consistent with the gradual variation in continuum and emission-line properties with optical colour found across LRD subtypes (Barro et al. 2026b; Pérez-González et al. 2026). We stress that the apparent gap around $\beta_{\text{opt}} = 0$ in Fig. 1 is produced by requiring the sign of β_{opt} to be recovered in at least 95 per cent of the Monte Carlo realisations. This criterion preferentially removes sources near the boundary, making transitional objects difficult to classify robustly as either LRDs or LBDs.

The gap itself is therefore artificial, but the distribution before applying this criterion remains bimodal (see Fig. A3 in the Appendix).

A two-component Gaussian-mixture model is strongly preferred over a single Gaussian, with $\Delta\text{BIC} = 28.65$ and components centred at $\beta_{\text{opt}} = -2.53$ and 0.44 . The underlying distribution is better described by two overlapping components rather than either a single unimodal population or two classes separated by a sharp physical boundary. The overlap in their emission-line properties similarly suggests that sources with intermediate optical slopes may represent a genuine transition between the characteristic properties of LBDs and LRDs, although the present data cannot establish whether this corresponds to an evolutionary sequence.

We also investigated whether LRDs can be identified using the emission-line diagnostics developed for faint, narrow-line AGN by Treiber et al. (2025). As some of the published DJA measurements, particularly the $H\alpha$ fluxes and equivalent widths of LRDs, proved unreliable, we independently refitted the relevant lines and recalculated these quantities for consistency. Even with the revised measurements, our LRDs do not occupy the same line-ratio region of the strongest AGN candidates identified by Treiber et al. (2025). Weak He II emission may instead indicate a softer ionising spectrum rather than the absence of an AGN. Similarly, the two LRDs analysed by Papovich et al. (2026) occupy UV diagnostic regions that are also compatible with star formation or composite ionisation, although their extreme UV equivalent widths and [O III] $\lambda 4363/\text{H}\gamma$ ratios indicate an accreting contribution. Low-resolution line ratios therefore cannot robustly identify individual LRDs as narrow-line AGN. A resolved broad $H\alpha$ component without corresponding broad forbidden-line

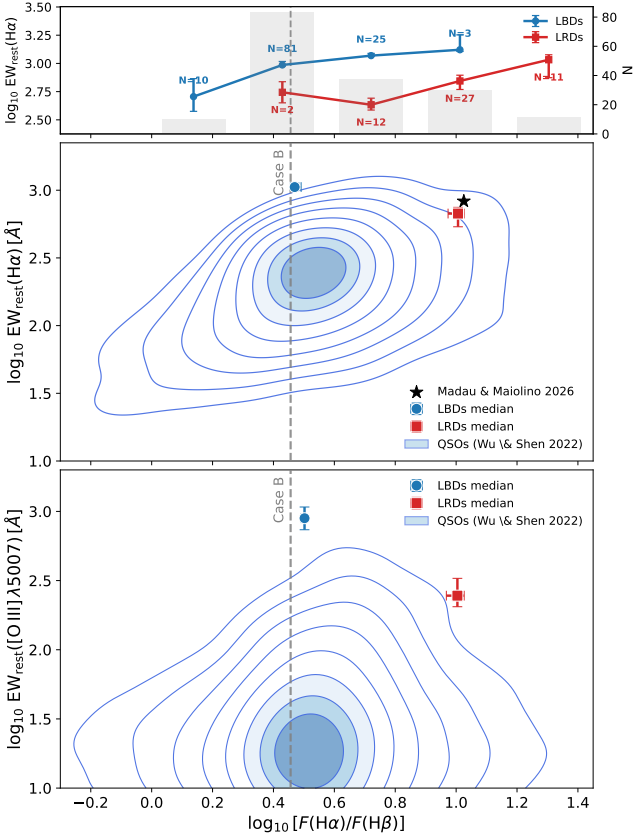


Figure 6. Balmer decrement and emission-line equivalent widths for the LBD and LRD populations. In all panels, blue circles and red squares show LBDs and LRDs, respectively, with error bars indicating the 16th–84th percentile uncertainties from 1000 bootstrap resamplings of the median. The vertical dashed line marks the Case B recombination value, $F_{H\alpha}/F_{H\beta} = 2.86$ ($\log_{10} = 0.46$), while the contours and shaded regions show the distribution of low-redshift broad-line QSOs from Wu & Shen (2022). (Top) Median rest-frame $H\alpha$ equivalent width in bins of $\log_{10}(F_{H\alpha}/F_{H\beta})$ for the 122 LBDs and 52 LRDs with positive $H\beta$ flux measurements. The labels indicate the number of sources in each bin, and the grey bars show the total number of sources. (Middle) Median rest-frame $H\alpha$ equivalent width as a function of the Balmer decrement. LBDs have median $F_{H\alpha}/F_{H\beta} = 2.95$ and $EW_{\text{rest}}(H\alpha) = 1055 \text{ \AA}$, compared with 10.14 and 673 $\text{\AA}$ for LRDs, respectively. The black star marks the fiducial model of Madau & Maiolino (2026), with $F_{H\alpha}/F_{H\beta} = 10.6$ and $EW_{\text{rest}}(H\alpha) = 830 \text{ \AA}$. (Bottom) Median rest-frame $[O III] \lambda 5007$ equivalent width as a function of the Balmer lines, comprising 77 LBDs and 41 LRDs. LBDs have a median $EW_{\text{rest}}([O III]) = 893 \text{ \AA}$, while LRDs have a substantially lower median of 246 $\text{\AA}$.

emission provides stronger evidence for accretion, although the interpretation may remain ambiguous for the faintest broad components (Maiolino et al. 2025).

To make a direct comparison with the recent study of LBDs and LRDs presented by Geris et al. (2026), it is therefore important to carefully account for differences in sample construction and line measurements. Their study analyses 19 LRDs and 17 LBDs at $2.26 < z < 7.89$, drawn primarily from the GOODS-N and GOODS-S fields, whereas our sample comprises 89 LRDs and 191 LBDs distributed across multiple extragalactic fields. Geris et al. (2026) also adopt different compactness criteria and spectroscopically classify LBDs as broad-line AGN only when the broad permitted emission has no corresponding broad $[O III]$ component, thereby reducing the

possibility of an outflow origin (see also Brazzini et al. 2026). We do not impose this requirement because type-I AGN can host ionised outflows, and our analysis is not restricted to sources with an unambiguous type-I classification. The two selections may therefore probe somewhat different spectroscopic populations.

Cross-matching the two catalogues, we recover 14 of the 36 Geris et al. (2026) sources. The remaining objects are excluded mainly because $H\alpha$ is not covered by the available spectra or because they do not satisfy our common selection criterion, $200 \text{ \AA} < EW(H\alpha) < 2000 \text{ \AA}$. We find no evidence that the differences between the two samples are driven by field-to-field variations, at least from the comparison of our GOODS-N and GOODS-S subsamples.

An additional methodological difference concerns the definition of $EW(H\alpha)$. Geris et al. (2026) report the equivalent width of the decomposed broad component, whereas our single-component fits measure the total $H\alpha$ emission. Our cross-matched comparison indicates that this difference has a substantially larger effect for LBDs than for LRDs: the LRD measurements are closer, although still not fully consistent, whereas the LBD equivalent widths are strongly separated. This likely explains much of the discrepancy between the LBD loci obtained in the two studies. Although separating the broad and narrow Balmer components would provide a more physically consistent comparison, most of our spectra were obtained with NIR-Spec/PRISM. At this resolution, such a decomposition is strongly model-dependent and cannot be performed robustly and homogeneously across the full sample. We therefore favour a consistent measurement of the total line properties.

Despite these differences, the narrow-line diagnostics of Geris et al. (2026) place LRDs and LBDs in largely overlapping regions consistent with AGN excitation. For $EW([O III])$, they find similar distributions for the two populations, whereas the right-hand panels of Fig. 6 show a lower median $EW([O III])$ for our LRDs. Neither result supports the simplest prediction that LRDs should exhibit enhanced $EW([O III])$ if approximately isotropic narrow-line emission is measured against an orientation-dependent continuum as proposed by Madau & Maiolino (2026).

To determine whether our $EW([O III])$ result is driven by differences in continuum strength, Fig. 7 compares $[O III] \lambda 5007$ directly with $H\beta$. Since the two lines are close in wavelength, their flux ratio is largely insensitive to continuum differences and differential attenuation. LRDs have a lower median $[O III] \lambda 5007/H\beta$ ratio than LBDs, demonstrating that their weaker $EW([O III])$ cannot be attributed solely to their continua. Instead, it indicates differences in the ionisation conditions, gas density, metallicity, or visibility of the narrow-line-emitting region. Nevertheless, the overlap between the two populations mirrors the continuity observed in the $\beta_{UV}-\beta_{\text{opt}}$ plane, rather than revealing two sharply separated emission-line populations.

LRDs have a median Balmer decrement of $F_{H\alpha}/F_{H\beta} = 10.1 \pm 0.7$, substantially larger than the value measured for LBDs. This is consistent with the value of ≈ 8.7 reported by de Graaff et al. (2025a) and the fiducial value of 10.6 predicted by Madau & Maiolino (2026), although it is lower than the values of ~ 15 and 16.2 ± 2.2 measured from the stacked spectra of Geris et al. (2026) and Sun et al. (2026), respectively. These measurements are not strictly equivalent: de Graaff et al. (2025a) measure the unresolved broad-plus-narrow emission; Geris et al. (2026) use the decomposed broad components; and Sun et al. (2026) obtain their value from a host-subtracted stack. Nevertheless, all consistently indicate that large Balmer decrements are a characteristic property of LRDs.

Interpreting these decrements solely as foreground dust attenuation remains challenging given the weak thermal dust emission reported

for LRDs (Akims et al. 2025; Casey et al. 2024) and the possible importance of geometry and radiative-transfer effects. Moreover, because [O III] and $H\beta$ experience nearly identical dust attenuation, foreground reddening cannot explain the lower [O III]/ $H\beta$ ratio of LRDs. Their large Balmer decrements must therefore be considered together with changes in the physical conditions or visibility of the line-emitting gas rather than interpreted simply as a direct measure of a uniform dust screen.

The definition of $EW(H\alpha)$ must also be considered when comparing our results with the broad-component predictions of Madau & Maiolino (2026). Nevertheless, the absolute location of our LRDs in the $EW(H\alpha)$ -Balmer-decrement plane is broadly consistent with their fiducial model, as shown in Fig. 6. Our median $EW(H\alpha) = 699^{+36}_{-42}$ Å is 131 Å (16%) lower than their predicted value of 830 Å, while our median Balmer decrement agrees with their fiducial value of 10.6. Our equivalent widths also broadly agree with the unresolved measurements of de Graaff et al. (2025a) and Sun et al. (2026), who report 762^{+43}_{-60} Å and $817.2^{+80.9}_{-82.7}$ Å, respectively. These values are only 63 Å (8%) and 118 Å (14%) above our median, and provide more direct observational comparisons because both studies likewise include the unresolved broad-plus-narrow $H\alpha$ emission.

However, reproducing the absolute position of the LRDs does not imply that the model reproduces the relative differences between LRDs and LBDs. In particular, our LRDs have a lower median $H\alpha$ equivalent width than LBDs, 699^{+36}_{-42} Å compared with 998^{+51}_{-29} Å, despite exhibiting substantially broader and more luminous $H\alpha$ emission. This combination is difficult to explain through continuum attenuation and orientation alone.

Several limitations should be considered when interpreting these results. First, the sample is dominated by low-resolution NIR-Spec/PRISM spectroscopy, in which $H\alpha$ is blended with [N II]. The measured flux may therefore include an unresolved [N II] contribution, systematically overestimating both $F_{H\alpha}$ and $L_{H\alpha}$. Although this affects both populations, any differential bias between LRDs and LBDs remains uncertain. More detailed multi-component modelling could alter the inferred line fluxes and widths, but cannot be applied robustly and homogeneously to most of the present spectra. Nevertheless, the observed separation is large; the median $H\alpha$ luminosities differ by 0.75 dex, with $\log_{10}(L_{H\alpha}/\text{erg s}^{-1}) = 42.63$ for LRDs and 41.88 for LBDs.

Second, the $H\alpha$ equivalent-width criterion used to construct the common parent sample may affect the comparison between the two populations, particularly near the imposed limits of $200 \text{ Å} < EW(H\alpha) < 2000 \text{ Å}$. The difference in their median equivalent widths should therefore be interpreted cautiously. However, the same criterion is applied to both populations, and it cannot by itself explain the much larger separation in $H\alpha$ luminosity or line width.

Finally, the DJA combines surveys with heterogeneous selection functions and is not representative of a uniformly selected galaxy population. Its $H\alpha$ parent sample has a median rest-frame equivalent width of 207^{+2}_{-2} Å, broadly consistent with typical high-redshift star-forming galaxies, although biases towards bright, high-redshift, or strong-line sources may remain. Given the broad range of reported LRD properties (Brazzini et al. 2026; Geris et al. 2026), the heterogeneous DJA selection may preferentially recover their brighter or more readily detectable members. Nevertheless, our selection identifies 17 LRDs beyond those reported by de Graaff et al. (2025a) within the same parent archive, demonstrating that it is not restricted to previously known sources.

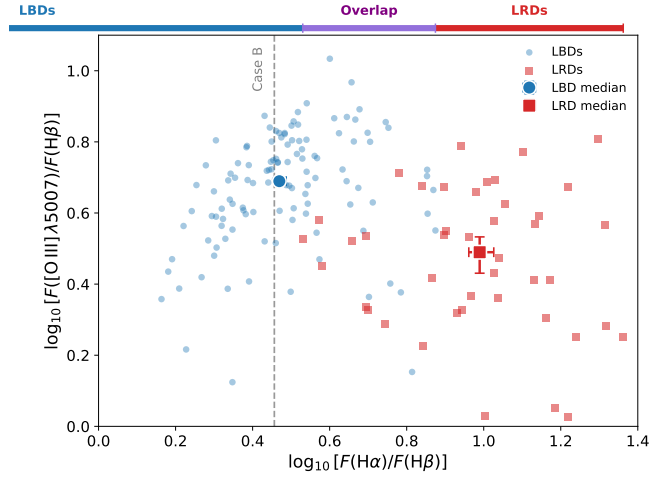


Figure 7. [O III] $\lambda 5007/H\beta$ ratio as a function of the Balmer decrement for LBDs (blue circles) and LRDs (red squares). Large symbols indicate the population medians, with uncertainties estimated through bootstrap resampling, while the vertical dashed line marks the Case B value, $F_{H\alpha}/F_{H\beta} = 2.86$. The coloured bars above the panel show the full Balmer-decrement ranges spanned by each population; the region labelled “Overlap” denotes their observed overlap in the Balmer decrement. LRDs exhibit systematically larger Balmer decrements but a lower median [O III] $\lambda 5007/H\beta$ ratio than LBDs. LBDs show a moderate positive Spearman correlation between the two ratios ($\rho = 0.38$, $p = 1.4 \times 10^{-5}$), whereas no significant correlation is detected for LRDs ($\rho = -0.06$, $p = 0.68$). The LBD correlation should be interpreted cautiously because $H\beta$ enters both axes. Because [O III] $\lambda 5007$ and $H\beta$ are closely spaced in wavelength, their ratio is only weakly affected by dust attenuation. Thus, the separation between the population medians is unlikely to be produced by foreground reddening alone, and may instead reflect differences in ionisation conditions, gas density, metallicity, or the visibility of the narrow-line region. This disfavors a simple foreground-screen interpretation of the LRD–LBD distinction and suggests that differences in ionisation conditions or narrow-line-region properties may also play a role.

5.1 Luminosity dependence and physical interpretation

An independent constraint comes from the strong increase in the LRD fraction with $H\alpha$ luminosity. LRDs are rare at the low $H\alpha$ -luminosity end but become increasingly dominant towards larger $L_{H\alpha}$. Within $4 < z < 7.5$, the cumulative analysis shows that LRDs constitute a statistically robust majority above $\log_{10}(L_{H\alpha}/\text{erg s}^{-1}) = 42.30$, where the lower bound of the 68% Wilson interval first exceeds 0.5. By contrast, the LRD fraction shows no clear monotonic evolution with redshift over the range probed, while LRDs remain systematically more $H\alpha$ luminous than LBDs.

This $H\alpha$ luminosity dependence is difficult to reconcile with a simple orientation-based scenario in which the probability of observing an intrinsically similar source as an LRD is independent of its line luminosity. The offset also persists after imposing the same broad-line criterion, $\text{FWHM}(H\alpha) > 2000 \text{ km s}^{-1}$, on both populations and therefore cannot be attributed solely to their different line-width distributions. Instead, the increasing incidence of LRDs towards high $L_{H\alpha}$ suggests that the LRD phenomenon depends on an intrinsic property such as luminosity, accretion state, gas column density, or luminosity-dependent obscuration.

Interestingly, our $H\alpha$ luminosity and line-width measurements agree rather well with the latest predictions of the gas-cocoon scenario developed by Sneppen et al. (2026a,b), in which LBDs and LRDs form a sequence driven by increasing gas column density. The

model predicts LRDs to be approximately 7.5 times more luminous in $H\alpha$ than LBDs and to exhibit broader lines, with mean FWHM values of 2080 km s^{-1} and 1530 km s^{-1} , respectively. These predictions are close to our observed average $H\alpha$ luminosity ratio of ≈ 5.6 and median FWHM values of 2319^{+71}_{-64} and $1424^{+56}_{-43} \text{ km s}^{-1}$ for LRDs and LBDs respectively.

The behaviour of $EW(H\alpha)$ across optical slope provides further support for this interpretation. As shown in Fig. 8, the median equivalent width remains broadly high and relatively stable across the LBD regime ($\beta_{\text{opt}} < 0$), with values ranging from $949^{+50}_{-87} \text{ \AA}$ to $1180^{+46}_{-140} \text{ \AA}$, before declining towards redder optical slopes in the LRD regime, from $735^{+51}_{-36} \text{ \AA}$ at $0 < \beta_{\text{opt}} < 1$ to $585^{+163}_{-82} \text{ \AA}$ at $1 < \beta_{\text{opt}} < 2$ and $410^{+224}_{-81} \text{ \AA}$ at $2 < \beta_{\text{opt}} < 3$. This decline is qualitatively consistent with the behaviour predicted by the gas-cocoon models of [Sneppen et al. \(2026a,b\)](#), in which $EW(H\alpha)$ reaches a maximum at intermediate column densities, approaching the transition between the LBD and LRD regimes, before decreasing again at larger gas column densities as an increasing fraction of the line emission is reprocessed into continuum emission. The lower equivalent widths of the reddest LRDs, despite their substantially larger $H\alpha$ luminosities, are therefore consistent with the expectation that these most extreme/red LRDs occupy the high-column-density regime of the model. In this picture, the systematically higher FWHM of the $H\alpha$ lines in the LRDs as compared to LBDs is also a consequence of increasing column density (rather than projected gas velocities) with an increasing fraction of the $H\alpha$ emission scattered into the observed exponential wings.

Viewed together, our results are inconsistent with a purely orientation-based interpretation of the LRD–LBD connection such as that proposed by [Madau & Maiolino \(2026\)](#), at least for some fraction of the LRD population. The model unifying LBDs and LRDs purely by orientation approximately reproduces the absolute Balmer decrement and $H\alpha$ equivalent width of the LRDs in our sample, and can, to some extent, explain the increased $H\alpha$ FWHM displayed by LRDs relative to LBDs. However, it cannot readily explain the substantially more luminous typical $H\alpha$ line emission produced by LRDs, the trends in EW with β_{opt} across the LBD+LRD population, nor the systematically lower $[O III]/H\beta$ ratios displayed by the LRDs. Orientation and dust attenuation may contribute to the red optical continua and large Balmer decrements of LRDs, but the observations require at least one additional luminosity- or accretion-dependent physical parameter. Taken at face value, our results favour a scenario analogous to the model developed by [Sneppen et al. \(2026a\)](#), in which LBDs are linked to LRDs via an increasing hydrogen column density in the cocoon surrounding the central engine (see also [Inayoshi & Ho 2025](#); [de Graaff et al. 2025a](#); [Sneppen et al. 2026b](#); [Rusakov et al. 2026](#)).

6 SUMMARY AND CONCLUSIONS

In this work, we have investigated whether Little Red Dots (LRDs) and Little Blue Dots (LBDs) can be understood as a single compact AGN population observed at different orientations, or whether their properties require additional physical differences. We constructed a homogeneous spectroscopic sample from the DJA archive, selecting on high $H\alpha$ equivalent width, UV continuum slope, and compactness, independently of the optical continuum shape. We subsequently classified the sources according to their position in the $\beta_{\text{UV}}-\beta_{\text{opt}}$ plane and compared their continuum properties, $H\alpha$ emission, and Balmer decrements. Our final sample spans the redshift range $1 \lesssim z \lesssim 7.5$, with the LRD and LBD sub-samples displaying very similar redshift

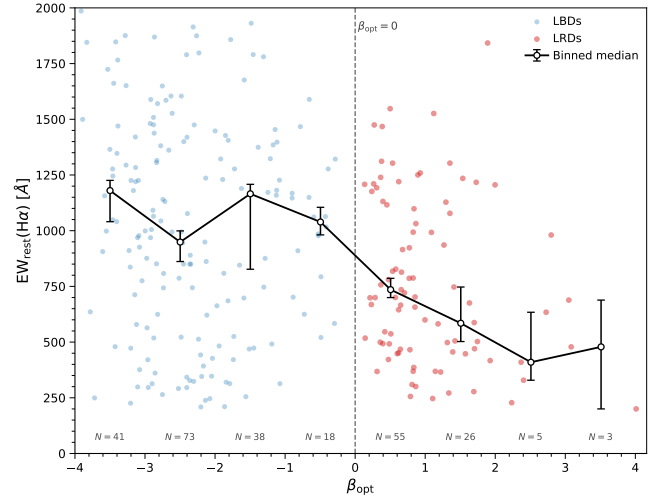


Figure 8. Rest-frame $H\alpha$ equivalent width as a function of the rest-frame optical continuum slope, β_{opt} , for LBDs (blue circles) and LRDs (red circles). The vertical dashed line marks the adopted $\beta_{\text{opt}} = 0$ boundary between the two populations. Black points and error bars show the median $EW_{\text{rest}}(H\alpha)$ and the 16th–84th percentile bootstrap uncertainties in bins of β_{opt} . The median $H\alpha$ equivalent width is broadly high and relatively stable over the LBD regime, with values of $\sim 950\text{--}1180 \text{ \AA}$ while $\beta_{\text{opt}} < 0$. However, it systematically decreases towards redder optical slopes in the LRD regime, from $735^{+51}_{-36} \text{ \AA}$ at $0 < \beta_{\text{opt}} < 1$ to $410^{+224}_{-81} \text{ \AA}$ at $2 < \beta_{\text{opt}} < 3$. This trend suggests that the $H\alpha$ equivalent width decreases as the optical continuum becomes redder, particularly among LRDs, qualitatively consistent with gas-cocoon models in which an increasing fraction of the $H\alpha$ emission is reprocessed into continuum emission at high gas column densities. The most extreme red bins contain only a small number of sources and should therefore be interpreted with caution.

distributions (with median redshifts $z_{\text{med}} = 4.53^{+0.30}_{-0.30}$ for LRDs and $z_{\text{med}} = 4.80^{+0.27}_{-0.21}$ for LBDs). Our main conclusions can be summarised as follows:

(i) Our final sample comprises 280 compact, UV-blue, strong- $H\alpha$ sources: 89 LRDs and 191 LBDs. Of the LRDs, 72 were previously reported by [de Graaff et al. \(2025b\)](#), while 17 are newly identified in this work. The two populations are strongly separated in optical continuum slope, with median values of $\beta_{\text{opt}} = 0.83^{+0.03}_{-0.05}$ for LRDs and $\beta_{\text{opt}} = -2.63^{+0.15}_{-0.12}$ for LBDs. The newly identified sources extend the known LRD population towards less extreme UV–optical continuum breaks.

(ii) LRDs have systematically broader $H\alpha$ profiles than LBDs, with median FWHM values of $2319^{+71}_{-64} \text{ km s}^{-1}$ and $1424^{+56}_{-43} \text{ km s}^{-1}$, respectively. The distributions are statistically distinct, indicating that the two populations likely differ in gas kinematics.

(iii) The strongest distinction is found in $H\alpha$ luminosity. LRDs have a median $\log_{10}(L_{H\alpha}/\text{erg s}^{-1}) = 42.63^{+0.09}_{-0.09}$, compared with $41.88^{+0.04}_{-0.01}$ for LBDs. The offset corresponds to a factor ~ 6 in luminosity and persists after imposing the same broad-line criterion, $\text{FWHM}(H\alpha) > 2000 \text{ km s}^{-1}$, on both populations. It therefore cannot be explained solely by their different FWHM distributions or by their redshift coverage.

(iv) LRDs have a moderately lower median EW than LBDs, $699^{+36}_{-42} \text{ \AA}$ compared with $998^{+51}_{-29} \text{ \AA}$. Because $H\alpha$ EW enters the parent-sample selection, this difference should be interpreted cautiously. Nevertheless, the combination of lower EW and substantially higher $H\alpha$ luminosity shows that the enhanced line luminosi-

ties of LRDs are not accompanied by correspondingly larger line-to-continuum ratios. Moreover, the $H\alpha$ EW decreases systematically towards the reddest optical slopes, reaching 544^{+94}_{-55} Å in the reddest bin, qualitatively consistent with gas-cocoon models in which increasing gas column densities cause an increasing fraction of the line emission to be reprocessed into the optical continuum.

(v) The fraction of LRDs increases strongly with $H\alpha$ luminosity, with LRDs becoming increasingly dominant towards the high-luminosity end. This trend disfavors a purely orientation-driven scenario in which the probability of observing a source as an LRD is independent of its intrinsic line luminosity. Instead, LRDs preferentially occupy the high- $L_{H\alpha}$ regime of the compact broad-line population.

(vi) LRDs are strongly concentrated at large Balmer decrements, with a median $\log_{10}(F_{H\alpha}/F_{H\beta}) = 1.01 \pm 0.03$, compared with 0.47 ± 0.01 for LBDs. The LRD fraction increases sharply with Balmer decrement, and $H\alpha$ EW of LRDs are clearly offset from the low-redshift broad-line QSO population. This indicates that the Balmer-emitting gas in LRDs is subject to substantially different attenuation and radiative-transfer conditions than in LBDs.

(vii) LRDs also show systematically lower $[O\ III] \lambda 5007/H\beta$ ratios than LBDs. Because these two lines are closely spaced in wavelength, their ratio is only weakly affected by dust attenuation, suggesting that the difference is unlikely to arise from foreground reddening alone and may instead reflect differences in ionisation conditions, gas density, metallicity, or the visibility of the narrow-line region.

Overall, our results are inconsistent with a simple orientation-only interpretation of the LRD-LBD connection. Although viewing geometry and dust attenuation likely contribute to the red optical continua and large Balmer decrements of LRDs, and can potentially explain their broader $H\alpha$ lines, they do not readily explain the substantially more luminous $H\alpha$ emission produced by LRDs, their lower $[O\ III]/H\beta$ ratios, or the decline of $H\alpha$ EW towards redder optical slopes. Our results therefore favour a scenario in which orientation may contribute to the observed diversity, but in which increasing gas column density around the central engine provides a plausible mechanism for linking LBDs and LRDs.

Rather than forming two sharply separated populations, LBDs and LRDs may represent different regimes within a broader population of compact, high-redshift accreting sources, with increasing gas column density providing a possible link between their observed properties.

DATA AVAILABILITY

All catalogues and software products generated for this work will be made publicly available in a GitHub repository upon publication of this paper. The repository will include the final catalogues, the Jupyter notebooks used for the analysis, and the Python scripts required to reproduce the results and figures presented in this work.

Software: This work made use of the Python packages *Astropy* (Astropy Collaboration et al. 2022), *NumPy* (Harris et al. 2020), *SciPy* (Virtanen et al. 2020), *Pandas* (McKinney 2010), *Matplotlib* (Hunter 2007), and *Photutils* (Bradley et al. 2026), as well as Jupyter notebooks for data analysis and visualisation.

REFERENCES

Akins H. B., et al., 2025, *The Astrophysical Journal*, 991, 37
 Ananna T. T., Bogdán Á., Kovács O. E., Natarajan P., Hickox R. C., 2024, *The Astrophysical Journal Letters*, 969, L18

Asada Y., Inayoshi K., Fei Q., Fujimoto S., Willott C., 2026, *arXiv e-prints*, p. [arXiv:2601.10573](https://arxiv.org/abs/2601.10573)
 Astropy Collaboration et al., 2022, *The Astrophysical Journal*, 935, 167
 Baggen J. F. W., et al., 2024, *The Astrophysical Journal Letters*, 977, L13
 Barro G., et al., 2026a, *The Astrophysical Journal*, 997, 48
 Barro G., et al., 2026b, *The Astrophysical Journal*, 1003, 96
 Barrufet L., et al., 2025, *Monthly Notices of the Royal Astronomical Society*, 537, 3453
 Bezanson R., et al., 2024, *ApJ*, 974, 92
 Bradley L., et al., 2026, *Photutils*, doi:10.5281/zenodo.19636730, <https://doi.org/10.5281/zenodo.19636730>
 Brazzini M., et al., 2026, *arXiv e-prints*, p. [arXiv:2601.22214](https://arxiv.org/abs/2601.22214)
 Casey C. M., Akins H. B., Kokorev V., McKinney J., Cooper O. R., Long A. S., Franco M., Manning S. M., 2024, *The Astrophysical Journal Letters*, 975, L4
 Cloonan A. P., et al., 2026, *arXiv e-prints*, p. [arXiv:2603.24700](https://arxiv.org/abs/2603.24700)
 Dickinson M., et al., 2024, The CANDELS-Area Prism Epoch of Reionization Survey (CAPERS), JWST Cycle 3 Proposal, ID #6368
 Eisenstein D. J., et al., 2023, *arXiv e-prints*
 Evans I. N., et al., 2024, *The Astrophysical Journal Supplement Series*, 274, 22
 Finkelstein S. L., et al., 2025, *The Astrophysical Journal Letters*, 983, L4
 Fu S., et al., 2026, *Nature Astronomy*,
 Geris S., et al., 2026, *arXiv e-prints*, p. [arXiv:2606.21614](https://arxiv.org/abs/2606.21614)
 Greene J. E., et al., 2024, *The Astrophysical Journal*, 964, 39
 Harris C. R., et al., 2020, *Nature*, 585, 357
 Heintz K. E., et al., 2024, *arXiv e-prints*, p. [arXiv:2404.02211](https://arxiv.org/abs/2404.02211)
 Herrero-Carrión D., Spinoso D., Izquierdo-Villalba D., Su T., Bonoli S., Renard P., 2026, *Monthly Notices of the Royal Astronomical Society*, 547, 3478
 Hsiao T. Y.-Y., et al., 2023, *ApJ*, 949, L34
 Hunter J. D., 2007, *Computing in Science & Engineering*, 9, 90
 Hviding R. E., et al., 2025, *Astronomy & Astrophysics*, 702, A57
 Inayoshi K., 2025, *ApJ*, 988, L22
 Inayoshi K., Ho L. C., 2025, *arXiv e-prints*, p. [arXiv:2512.03130](https://arxiv.org/abs/2512.03130)
 Ji X., et al., 2026, *arXiv e-prints*, p. [arXiv:2604.03370](https://arxiv.org/abs/2604.03370)
 Kapoor S., et al., 2026, *arXiv e-prints*, p. [arXiv:2607.00084](https://arxiv.org/abs/2607.00084)
 Korber D., et al., 2026, *arXiv e-prints*, p. [arXiv:2604.08687](https://arxiv.org/abs/2604.08687)
 Labbé I., et al., 2023, *Nature*, 616, 266
 Loiacono F., et al., 2026, *arXiv e-prints*, p. [arXiv:2606.30253](https://arxiv.org/abs/2606.30253)
 Madau P., Maiolino R., 2026, *arXiv e-prints*, p. [arXiv:2602.22386](https://arxiv.org/abs/2602.22386)
 Madau P., Maiolino R., Scholtz J., D'Eugenio F., 2026, *arXiv e-prints*, p. [arXiv:2604.04216](https://arxiv.org/abs/2604.04216)
 Maiolino R., et al., 2025, *Monthly Notices of the Royal Astronomical Society*, 538, 1921
 Maseda M. V., et al., 2024, *Astronomy & Astrophysics*, 689, A73
 Matthee J., et al., 2024, *The Astrophysical Journal*, 963, 129
 Matthee J., et al., 2026, *arXiv e-prints*, p. [arXiv:2603.17667](https://arxiv.org/abs/2603.17667)
 McKinney W., 2010, in Proceedings of the 9th Python in Science Conference, pp 56–61, doi:10.25080/Majora-92bf1922-00a
 Naidu R. P., et al., 2025, *arXiv e-prints*, p. [arXiv:2503.16596](https://arxiv.org/abs/2503.16596)
 Papovich C., et al., 2026, *arXiv e-prints*, p. [arXiv:2608.04101](https://arxiv.org/abs/2608.04101)
 Pérez-González P. G., et al., 2024, *The Astrophysical Journal*, 968, 4
 Pérez-González P. G., et al., 2026, *arXiv e-prints*, p. [arXiv:2602.20247](https://arxiv.org/abs/2602.20247)
 Pollock C. L., et al., 2026, *A&A*, 708, A203
 Rinaldi P., et al., 2026, *arXiv e-prints*, p. [arXiv:2604.07138](https://arxiv.org/abs/2604.07138)
 Rusakov V., et al., 2026, *Nature*, 649, 574
 Sarrouh G., et al., 2025, *arXiv e-prints*
 Scholtz J., et al., 2026, *arXiv e-prints*, p. [arXiv:2603.22277](https://arxiv.org/abs/2603.22277)
 Shapley A. E., et al., 2025, *ApJ*, 980, 242
 Shen Y., et al., 2024, *arXiv e-prints*
 Sneppen A., et al., 2026a, *arXiv e-prints*, p. [arXiv:2601.18864](https://arxiv.org/abs/2601.18864)
 Sneppen A., Watson D., Matthews J. H., Nikopoulos G., 2026b, *arXiv e-prints*, p. [arXiv:2606.12509](https://arxiv.org/abs/2606.12509)
 Sun W. Q., et al., 2026, *The Open Journal of Astrophysics*, 9
 Treiber H., et al., 2025, *The Astrophysical Journal*, 984, 93
 Treu T., et al., 2022, *ApJ*, 935, 110
 Valentino F., et al., 2025, *A&A*, 699, A358

Virtanen P., et al., 2020, *Nature Methods*, 17, 261

Webb N., Traulsen I., et al., 2026, A&A submitted, Submitted to Astronomy & Astrophysics

Wu Q., Shen Y., 2022, *The Astrophysical Journal Supplement Series*, 263, 42

de Graaff A., et al., 2025a, *arXiv e-prints*, p. arXiv:2511.21820

de Graaff A., et al., 2025b, *Astronomy & Astrophysics*, 697, A189

7 ACKNOWLEDGEMENTS

LB and JSD acknowledge the support of the Royal Society through the award of a Royal Society University Research Professorship to JSD. LB and MSL thank Pedro Rodríguez Pascual (XMM-Newton Science Operations Centre, ESAC) for his assistance with the X-ray data. DHC acknowledges support from the Basque Government through the Programa Predoctoral de Formación de Personal Investigador No Doctor del departamento de Ciencia, Universidades e Innovación del Gobierno Vasco

This research has made use of data from the JWST/NIRSpec spectroscopic programmes contributing to the final sample, including JADES, CAPERS, RUBIES, UNCOVER, CEERS, NEXUS, CANUCS, GTO-WIDE, AURORA, GLASS-JWST, DeepDive, and other General Observer and Director’s Discretionary Time programmes. We acknowledge the principal investigators, teams, and staff involved in obtaining and processing these observations.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data products presented herein were retrieved from the Dawn JWST Archive (DJA), an initiative of the Cosmic Dawn Center (DAWN), which is funded by the Danish National Research Foundation under grant DNR140.

This research has made use of data obtained from the 5XMM serendipitous source catalogue compiled by the XMM-Newton Survey Science Centre, the XMM2ATHENA project, and in collaboration with the XMM-Newton Science Operations Centre. This research has also made use of data obtained from the Chandra Source Catalogue, provided by the Chandra X-ray Center (CXC) as part of the Chandra Data Archive.

APPENDIX A: ADDITIONAL SAMPLE PROPERTIES

This appendix provides supplementary information on the survey composition of our sample and illustrates the effects of the main selection criteria on the $H\alpha$ equivalent-width and optical-slope distributions.

Table A1 lists the contribution of each spectroscopic survey, the field observed and the number of sources to the final sample.

Figure A1 shows that the UV-slope criterion has little effect on the $H\alpha$ equivalent-width distribution, whereas the compactness cut preferentially selects sources with larger equivalent widths.

Figure A2 compares the selected populations with the DJA parent sample. Both LRDs and LBDs exhibit substantially larger $H\alpha$ equivalent widths than the parent population, with LBDs reaching the largest median value. The parent sample comprises approximately 24,000 sources with measurable $H\alpha$. Its distribution peaks around $EW_{\text{rest}}(H\alpha) \sim 200$, broadly consistent with the values expected for typical high-redshift star-forming galaxies. Nevertheless, given the heterogeneous selection of the surveys included in the DJA, the parent sample should not be regarded as fully representative of the general galaxy population. However, it broadly covers the $H\alpha$ equivalent-width range expected for typical galaxies.

Finally, Fig. A3 examines the β_{opt} distribution after applying all selection criteria except the Monte Carlo requirement that its sign be

Table A1. Spectroscopic surveys contributing to the final sample of 280 sources. The table lists the survey, field, and number of sources contributed by each programme. The sample combines spectroscopic observations from 19 surveys and programmes, with CAPERS and RUBIES providing the largest contributions.

Survey	Field	No.
CAPERS (Dickinson et al. 2024)	UDS/COSMOS	65
JADES (Eisenstein et al. 2023)	GOODS-N/S	54
RUBIES (de Graaff et al. 2025b)	EGS/UDS	43
UNCOVER (Bezanson et al. 2024)	Abell 2744	24
CEERS (Finkelstein et al. 2025)	EGS	18
NEXUS (Shen et al. 2024)	—	18
GTO-WIDE (Maseda et al. 2024)	GOODS-S/EGS/UDS	16
CANUCS (Sarrouh et al. 2025)	Cluster fields	12
DDT-6585	COSMOS	7
GO-4106	EGS	4
GO-2073	Quasar fields	4
MoM (Naidu et al. 2025)	—	4
GO-1433 (Hsiao et al. 2023)	MACS0647	3
GO-2198 (Barrufet et al. 2025)	GOODS-S	3
NIRSpec WIDE (Maseda et al. 2024)	COSMOS	1
AURORA (Shapley et al. 2025)	COSMOS/GOODS-N	1
GO-2767	RXJ2129	1
GLASS-JWST (Treu et al. 2022)	Abell 2744	1
DeepDive (Valentino et al. 2025)	EGS/UDS	1
Total		280

recovered with ≥ 95 per cent confidence. We fitted the 356 sources within $-6 \leq \beta_{\text{opt}} \leq 4$ with one- and two-component Gaussian-mixture models using `GaussianMixture` from `scikit-learn`. The two-component model is strongly preferred over a single Gaussian ($\Delta\text{BIC} = 28.65$), with components centred at $\beta_{\text{opt}} = -2.53$ and 0.44 . These components are clearly separated (Ashman’s $D = 3.08$), and a kernel-density estimate independently recovers peaks at $\beta_{\text{opt}} \approx -2.64$ and 0.32 . Thus, the bimodality is already present before imposing the Monte Carlo significance criterion.

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

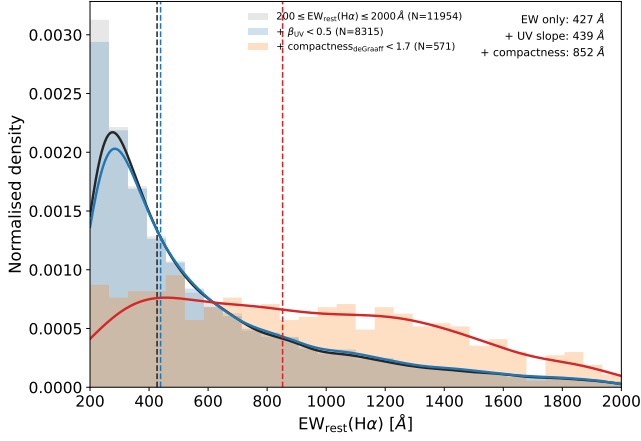


Figure A1. Normalised distributions of the rest-frame H α equivalent width after the sequential application of the selection criteria. The black distribution shows the 11,954 sources satisfying $200 \leq \text{EW}_{\text{rest}}(\text{H}\alpha) \leq 2000 \text{ \AA}$, while the blue distribution additionally imposes $\beta_{\text{UV}} < -0.5$, leaving 8,315 sources. The red distribution corresponds to the final sample after applying the compactness criterion, comprising 571 sources. Dashed vertical lines indicate the respective median equivalent widths. The UV-slope cut has little effect on the distribution, shifting the median from 427 to 439 $\text{\AA}$. By contrast, the compactness criterion substantially reduces the sample and shifts the median to 852 $\text{\AA}$.

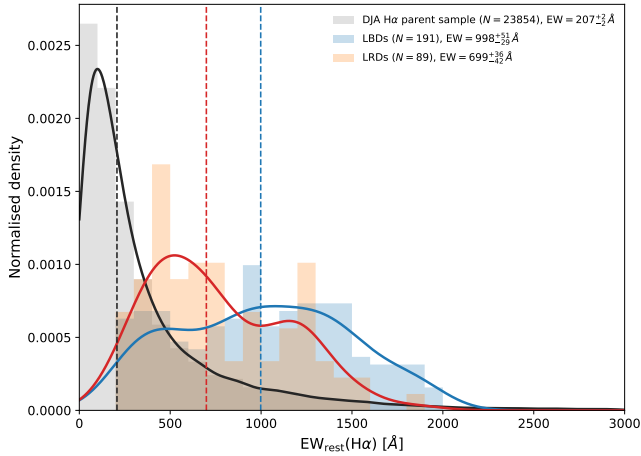


Figure A2. Normalised distributions of the rest-frame H α equivalent width for the DJA H α parent sample (black), LBDs (blue), and LRDs (red). The parent sample comprises 23,854 sources with successful H α fits and valid equivalent-width measurements within the displayed range, although its distribution may reflect the heterogeneous and predominantly high-redshift nature of the DJA archive. Dashed vertical lines indicate the median equivalent widths, with bootstrap uncertainties reported in the legend. Both selected populations are strongly shifted towards larger equivalent widths relative to the parent sample, whose median is $207^{+2}_{-2} \text{ \AA}$. LBDs exhibit a larger median equivalent width than LRDs, with $998^{+51}_{-29} \text{ \AA}$ and $699^{+36}_{-42} \text{ \AA}$, respectively.

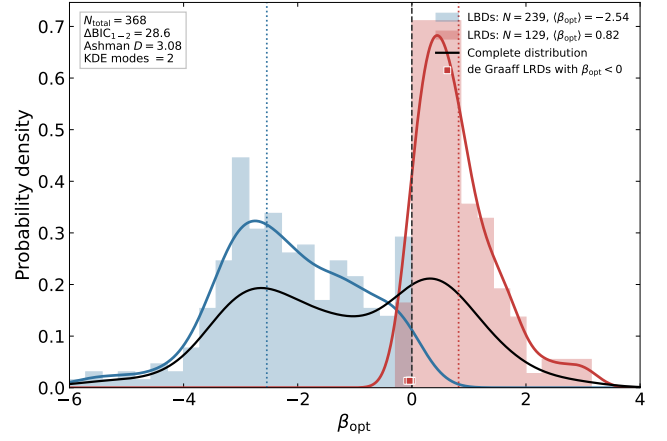


Figure A3. Distribution of the optical continuum slope, β_{opt} , after applying all selection criteria except the Monte Carlo significance requirement on β_{opt} . The normalised histograms and kernel-density estimates show LBDs (blue) and LRDs (red), while the black curve represents the complete sample. The vertical dotted lines mark the population means and the dashed line indicates $\beta_{\text{opt}} = 0$; red squares identify previously known de Graaff LRDs with $\beta_{\text{opt}} < 0$. A two-component Gaussian-mixture model is strongly preferred over a single Gaussian ($\Delta\text{BIC}_{1-2} = 28.6$), with well-separated components (Ashman's $D = 3.08$), while the KDE independently recovers two modes. These results indicate that the β_{opt} distribution is intrinsically bimodal before imposing its significance criterion.