The Astrophysical Journal Letters, 989:L50 (18pp), 2025 August 20 https://doi.org/10.3847/2041-8213/adf73d


# Lonely Little Red Dots: Challenges to the Active Galactic Nucleus Nature of Little Red Dots through Their Clustering and Spectral Energy Distributions

María Carranza-Escudero[1], Christopher J. Conselice[1], Nathan Adams[1], Thomas Harvey[1], Duncan Austin[1], Peter Behroozi[2], Leonardo Ferreira[3], Katherine Ormerod[1,4], Qiao Duan[1], James Trussler[5], Qiong Li[1], Lewi Westcott[1], Rogier A. Windhorst[6], Dan Coe[7], Seth H. Cohen[6], Cheng Cheng[8,9], Simon P. Driver[10], Brenda Frye[11], Lukas J. Furtak[12], Norman A. Grogin[13], Nimish P. Hathi[13], Rolf A. Jansen[6], Anton M. Koekemoer[13], Madeline A. Marshall[14,15], Rosalia O'Brien[6], Norbert Pirzkal[13], Maria Polletta[16], Aaron Robotham[10], Michael J. Rutkowski[17], Jake Summers[6], Stephen M. Wilkins[18], Christopher N. A. Willmer[19], Haojing Yan[20], and Adi Zitrin[12]

[1] Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK; maria.carranzaescudero@ox.physics.ac.uk
[2] Steward Observatory and Department of Astronomy, University of Arizona, Tuscon, AZ 85721, USA
[3] Department of Physics & Astronomy, University of Victoria, Finnerty Road, Victoria, BC V8P 1A1, Canada
[4] Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool, L3 5RF, UK
[5] Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA
[6] School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA
[7] AURA for the European Space Agency (ESA), Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[8] Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100101, People's Republic of China
[9] 2 CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
[10] International Centre for Radio Astronomy Research (ICRAR) and the International Space Centre (ISC), The University of Western Australia, M468, 35 Stirling Highway, Crawley, WA 6009, Australia
[11] University of Arizona, Department of Astronomy/Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721, USA
[12] Department of Physics, Ben-Gurion University of the Negev, P.O. Box 653, Be'er-Sheva 84105, Israel
[13] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[14] National Research Council of Canada, Herzberg Astronomy & Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
[15] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[16] INAF—Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, Via A. Corti 12, I-20133 Milano, Italy
[17] Minnesota State University–Mankato, Department of Physics & Astronomy, Trafton Science Center North 141, Mankato, MN 56001, USA
[18] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton, BN1 9QH, UK
[19] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0009, USA
[20] Department of Physics and Astronomy, University of Missouri, Columbia, MO 65211, USA
*Received 2025 June 4; revised 2025 July 25; accepted 2025 July 30; published 2025 August 19*

## Abstract

Observations with the James Webb Space Telescope reveal a previously unseen population of compact red objects, known as "little red dots" (LRDs). We study a new photometrically selected sample of 124 LRDs in the redshift range $z \sim 3$–10 selected from Near Infrared Camera coverage of the Cosmic Evolution Early Release Science Survey (CEERS), North Ecliptic Pole Time Domain Field (NEP-TDF), James Webb Space Telescope Advanced Deep Extragalactic Survey (JADES), and JEMS. For JADES, the NEP-TDF, and CEERS, we compare spectral energy distribution (SED) models with and without active galactic nucleus (AGN) components and analyze the impact of an AGN component on the goodness of fit using the Bayesian information criterion (BIC). We find that while the $\chi^2$ of the majority of models containing AGN components is improved compared to models without AGN components, we show that the BIC suggests that models without AGN are a more appropriate fit to LRD SEDs, especially when MIRI data are available. We also measure LRD clustering in the CEERS field, JADES field, and NEP-TDF, where we compare the spatial distribution of LRDs and galaxies with Kolmogorov–Smirnov tests of equality of distribution. We find that the neighbourhood of LRDs tends to be less dense compared to galaxies at all selections and masses and at similar redshifts. We further measure upper limit estimates for the halo masses of LRDs using abundance matching. While the population of LRDs could be a mixture of several different inherent populations, as a whole, it does appear that these systems are mostly hosting compact galaxies or star clusters in formation.

*Unified Astronomy Thesaurus concepts:* Galaxies (573); AGN host galaxies (2017); Active galactic nuclei (16); High-redshift galaxies (734); James Webb Space Telescope (2291)

## 1. Introduction

The discovery of a mysterious set of red objects exhibiting a V-shaped continuum and point-source morphology has received a large amount of attention due to their puzzling nature (e.g., L. J. Furtak et al. 2023, 2024; D. D. Kocevski et al. 2023; J. Matthee et al. 2024a; I. Labbe et al. 2025). These so-called little red dots (LRDs) prove challenging to interpret and understand due to the similarity of their spectral energy distributions (SEDs) to both stellar populations and dust-reddened active galactic nuclei (AGN). LRDs are typically found around redshift $z \sim 5$ (I. Labbe et al. 2025). These LRDs are a unique high-redshift population whose nature is still very much uncertain.

The SEDs of LRDs do not resemble any typical SEDs and seem to be composed of a blue component starting at roughly 3600 Å and a red component at longer wavelengths (D. J. Setton et al. 2024). Investigations into the blue rest-ultraviolet (UV) of LRD SEDs (H. B. Akins et al. 2023; P. G. Pérez-González et al. 2024; I. Labbe et al. 2025) suggest, for example, that their origin could be unobscured star formation (SF) or, alternatively, light scattered from an AGN accretion disk. Attempts have been made to constrain the origin of the red rest-frame optical SEDs, some making use of Mid-Infrared Instrument (MIRI) data, in which few LRDs are detected. The red component is often attributed to warm dust heated by some of the following scenarios. Using MIRI data, P. G. Pérez-González et al. (2024) find that the red optical and near-infrared data fit an obscured accretion disk, but stellar-dominated (often starburst) models provide an even better fit. C. C. Williams et al. (2024) find that data using the Hubble Space Telescope (HST) and Near Infrared Camera (NIRCam), but not MIRI and Atacama Large Millimeter/submillimeter Array (ALMA) data, result in larger masses and SF rates than when MIRI+ALMA data are used. Z. Li et al. (2024) find no need for stellar emission and scattered AGN light and instead fit LRD SEDs with AGN embedded in an extended dusty medium and a relatively gray extinction curve.

Employing spectroscopic observations of LRDs, J. E. Greene et al. (2024) find that ∼75% of their LRD sample exhibits broad-line H$\alpha$ emission, suggesting that these are dust-reddened AGN. Supermassive black hole (SMBH) mass estimates from the broadness of LRD H$\alpha$ and H$\beta$ lines give rise to a range of masses ($10^5$–$10^9$ $M_\odot$; D. D. Kocevski et al. 2023; L. J. Furtak et al. 2024; J. E. Greene et al. 2024). The abundance of LRDs would suggest a much higher density of AGN (V. Kokorev et al. 2023; H. B. Akins et al. 2024; D. D. Kocevski et al. 2025) than predicted by ground-based surveys (Y. Matsuoka et al. 2018; M. Niida et al. 2020; W. He et al. 2024), with implications for Lyman continuum radiation and reionization (e.g., A. Grazian et al. 2024; P. Madau et al. 2024). Theories discussing "black hole stars" (A. de Graaff et al. 2025; R. P. Naidu et al. 2025) or hyperdense gas cocoons surrounding smaller black holes also exist (K. Inayoshi & R. Maiolino 2025), with early ALMA results providing upper mass limits that are in line with this (C. M. Casey et al. 2025).

Investigations into the local Universe ($z < 0.1$) demonstrate evidence for a strong evolutionary relationship between SMBHs and their host galaxies (J. Hu 2008; K. Gültekin et al. 2009; J. Kormendy & L. C. Ho 2013). As most LRDs lack detectable extended components, it is possible to place upper limits on stellar mass based on their maximum physical size. Using the empirical relationship between the SMBH-to-host-galaxy mass ratio presented in J. Kormendy & L. C. Ho (2013) and then comparing it to LRD SMBH mass estimates from spectroscopy (V. Kokorev et al. 2023; J. E. Greene et al. 2024) reveals an SMBH-to-host-galaxy mass ratio that is startlingly higher for LRDs and, in the case of a lensed LRD, is perhaps unphysically high with a very high broad-line width of $\sim$2000 km s$^{-1}$ (L. J. Furtak et al. 2024).

Other challenges to the AGN interpretation of LRDs exist. For example, M. Kokubo & Y. Harikane (2024) report that LRDs do not exhibit the typical photometric variability associated with standard AGNs. However, M. Kokubo & Y. Harikane (2024) suggest that this could be the result of intrinsically nonvariable AGN or the dominance of AGN emission through scattering. Z. Zhang et al. (2024) also find that LRDs do not usually display strong photometric variability. Another challenge is that unlike what is expected of type I AGN, the majority of LRDs are nondetected in X-rays (R. Maiolino et al. 2024; M. Yue et al. 2024). A sample of LRDs exhibiting broad-line H$\alpha$ emission studied by T. T. Ananna et al. (2024) shows a stacked signal of only $\sim$2.6$\sigma$. Theories on the lack of X-ray detection of AGN are discussed in R. Maiolino et al. (2024), who suggest that the lack of X-ray detections is due to a mixture of scenarios including X-ray absorption, intrinsic X-ray weakness, and non-AGN sources. J. F. W. Baggen et al. (2024) find an alternative explanation to the broad lines found in LRDs, suggesting that they are indicative of a brief phase in which galaxies have high central densities.

In this work, we select and present a sample of 124 LRDs using our own photometric selection based on previous LRD selection criteria. Our sample spans the redshift range $z \sim 3$–10 in the fields covered by the Cosmic Evolution Early Release Science Survey (CEERS; M. B. Bagley et al. 2023), the North Ecliptic Pole Time Domain Field (NEP-TDF; R. A. Windhorst et al. 2022) survey, and the James Webb Space Telescope (JWST) Advanced Deep Extragalactic Survey (JADES; M. J. Rieke et al. 2023). Most LRDs appear around $z \sim 5$ (I. Labbe et al. 2025), although some can reach photometric redshifts of $z > 9$ (G. C. K. Leung et al. 2024). In fact, the highest-redshift LRD presented in our sample has $z = 10.4^{+0.6}_{-1.3}$.

In this Letter, we investigate SED modeling with models containing an AGN and those with no AGN component, and we compare the best-fit $\chi^2$ statistics and Bayesian information criterion (BIC) of these to determine which models are most suitable. We briefly discuss the differences in stellar mass and dust for AGN and non-AGN models. We also carry out a study of the local environment of LRDs and compare it to the local environment of galaxies using Kolmogorov–Smirnov (K-S) tests. Finally, we create estimates for the upper limits of halo masses for our LRD sample using abundance matching.

Section 2 describes the imaging and data reduction of the fields used in this work. The selection criteria for LRDs are described in Section 3.1. In Section 4.2, we analyze the local environment of LRDs and compare this to the local environment of galaxies at similar redshifts. We investigate SED modeling and how the inclusion of AGN components affects fitting in Section 4.1.

Throughout this work, unless stated otherwise, we assume a standard cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. All magnitudes listed follow the AB magnitude system (J. B. Oke 1974; J. B. Oke & J. E. Gunn 1983).

## 2. Imaging and Data Reduction

For this study, we use JWST NIRCam imaging of CEERS, the NEP-TDF survey, and JADES. To cover objects at lower redshifts ($z \sim 4.5$–6), we make use of data sets from the HST. For CEERS, we use the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; N. A. Grogin et al. 2011; A. M. Koekemoer et al. 2011) imaging, specifically the Extended Groth Strip (EGS; M. Davis et al. 2007). For NEP-TDF, we incorporate HST Advanced Camera for Surveys (ACS) Wide Field Channel imaging from programs GO-15278 (PI: R. Jansen) and GO-16252/16793 (PIs: R. Jansen & N. Grogin). To cover JADES, which lies on the Great Observatories Origins Deep Survey South (GOODS-S)

footprint, we make use of HST data from the most recent mosaic (v2.5) from the Hubble Legacy Fields team (G. Illingworth et al. 2016; K. E. Whitaker et al. 2019). This section describes the details of the observations and data reduction used, as well as source identification and extraction. A more detailed overview of this can be found in C. J. Conselice et al. (2024).

### 2.1. CEERS JWST NIRCam and HST Imaging

The JWST/NIRCam observations of CEERS (ID: 1345; PI: S. Finkelstein; see also M. B. Bagley et al. 2023) consist of 10 pointings covering 66.40 arcmin$^2$ in the EGS field. The observations cover seven photometric bands: F115W, F150W, F200W, F277W, F356W, F410M, and F444W. We process CEERS data independently as described in Section 2.5. Further details on the calibration process and data products can be found in N. J. Adams et al. (2024) and C. J. Conselice et al. (2024).

Due to the lack of F090W imaging, we include HST CANDELS imaging (N. A. Grogin et al. 2011; A. M. Koekemoer et al. 2011) of the F606W and F814W filters to cover a bluer wavelength range. This imaging was reduced by the CANDELS team and is aligned using the Gaia EDR3 (A. G. A. Brown et al. 2021). The average depth is 28.5 mag and 28.3 mag for the F606W and F814W filters, respectively.

### 2.2. NEP-TDF JWST NIRCam and HST Imaging

NEP-TDF is part of the JWST Prime Extragalactic Areas for Reionization and Lensing Science (PEARLS) project (R. A. Windhorst et al. 2022; J. M. Diego et al. 2023; B. L. Frye et al. 2023). NEP-TDF observations have eight pointings covered by eight filters: F090W, F115W, F150W, F200W, F277W, F356W, F410M, and F444W. We process the data as described in Section 2.5. This is described in more depth in N. J. Adams et al. (2024). The total area is 57.32 arcmin$^2$, with a resolution of $0\farcs03$ pixel$^{-1}$.

To cover bluer wavelengths, we incorporate HST observations that utilize the F606W filter (R. O'Brien et al. 2024). These observations were obtained through the GO-15278 (PI: R. Jansen) and GO-16252/16793 (PIs: R. Jansen & N. Grogin) programs from 2017 October 1 through 2022 October 31.

### 2.3. JADES JWST NIRCam and HST Imaging

In this Letter, we also use JADES DR1 (M. J. Rieke et al. 2023), covering 22.98 arcmin$^2$ in the GOODS-S field footprint (PID: 1180; PI: D. Eisenstein; D. J. Eisenstein et al. 2023). JADES consists of six overlapping pointings of nine filters: F090W, F115W, F150W, F200W, F277W, F335M, F356W, F410M, and F444W. We reduce these data using our own pipeline once again to ensure consistency with other fields as described in Section 2.5 and in greater detail in N. J. Adams et al. (2024).

We include HST/ACS data of the F435W, F606W, F775W, and F814W filters. This imaging is derived from the v2.5 GOODS-S mosaic from the Hubble Legacy Fields team (G. Illingworth et al. 2016; K. E. Whitaker et al. 2019).

### 2.4. MIRI Imaging

We make use of the Systematic Mid-infrared Instrument Legacy Extragalactic Survey (SMILES; PID 1207; PI: G. Rieke; S. Alberts et al. 2024) coverage of JADES, which has 15 pointings in the F560W, F770W, F1000W, F1280W, F1500W, F1800W, F2100W, and F2550W filters. The total area covered by SMILES is $\sim$34 arcmin$^2$. The public data release can be found on the MAST website.[21] Details of the reduction process and alignment can be found in S. Alberts et al. (2024).

### 2.5. Reduction Process

We process all uncalibrated lower-level JWST NIRCam data products with a modified version of the STScI JWST Pipeline v1.8.2 (H. Bushouse et al. 2022) and use Calibration Reference Data System v1084 for the most up-to-date NIRCam calibration files at the time of writing. After running stage 1 of the JWST pipeline, we subtract templates of "wisps," large-scale artifacts affecting F150W and F200W imaging (N. J. Adams et al. 2024). After stage 2 of the pipeline, we apply a $1/f$ noise correction derived by Chris Willott.[22] We then perform background subtraction on each NIRCam frame before continuing to stage 3, after which we align the final images. The final resolution of the drizzled images is $0\farcs03$ pixel$^{-1}$.

### 2.6. Source Extraction

To carry out source identification and extraction, we make use of `SExtractor` (E. Bertin & S. Arnouts 1996). We take the inverse variance weighted stack of the F277W, F356W, and F444W bands and run this in dual-image mode to detect and select objects. We carry out forced aperture photometry for multiband measurements. Photometry is calculated within circular apertures with $0\farcs32$ diameters, chosen to enclose the central and brightest 70%–80% of the flux of a point source and yet small enough to avoid contamination. We include an aperture correction derived from simulated `WebbPSF` point-spread functions for each band used (M. D. Perrin et al. 2012, 2014).

To avoid underestimating the photometric errors, we use the $5\sigma$ local depth as the error. Local depth is calculated using `galfind`[23] by placing apertures with an approximately uniform number density of around $\sim$2500 arcmin$^{-2}$ in the "empty" regions of the images, where "empty" refers to an aperture where no preexisting source is in the image. Overlapping apertures are not permitted to ensure that these provide independent measurements of the background. To calculate the photometric error for each source, we use the normalized mean absolute deviation (NMAD; D. C. Hoaglin et al. 1983) of the nearest 200 apertures, making the depth resistant to potential outliers.

For each field, we carefully mask areas of the images affected by defects such as diffraction spikes and snowballs. The apertures are placed at least 30 pixels from the mask. More details of this process can be found in our EPOCHS Paper I (C. J. Conselice et al. 2024) and Paper II (N. J. Adams et al. 2024).

[21] https://archive.stsci.edu/hlsp/smiles
[22] https://github.com/chriswillott/jwst
[23] https://github.com/duncanaustin98/galfind

### 2.7. Photometric Redshifts

We use the EAZY photometric redshift code (G. B. Brammer et al. 2008) to calculate both the photometric redshift probability distribution and the most likely photometric redshift. We include templates from R. L. Larson et al. (2023), which expand the default template sets that use 12 templates generated with the Flexible Stellar Population Synthesis code (C. Conroy & J. E. Gunn 2010). We use and test the SED templates used in K. N. Hainline et al. (2024a) but find that the R. L. Larson et al. (2023) templates match the current spectroscopic results somewhat more closely.

To determine the quality of our photometric redshifts, we employ the outlier fraction $\eta$ and the NMAD. The outlier fraction is defined as the fraction of photometric redshifts that differ from spectroscopic redshifts by more than 15% and is given by

$$\eta = \frac{N_{115} + N_{85}}{N_{\rm total}}, \tag{1}$$

where $N_{115}$ represents the number of points above the line $z_{\rm phot} = 1.15 \times z_{\rm spec}$ and $N_{85}$ represents the number of points below the line $z_{\rm phot} = 0.85 \times z_{\rm spec}$.

The NMAD quantifies the dispersion in the redshift variances and is normalized. It is defined as

$$\sigma_{\rm NMAD} = 1.48 \times \text{median} \left| \frac{z_{\rm spec} - z_{\rm phot}}{1 + z_{\rm spec}} \right|. \tag{2}$$

Note that the factor of 1.48 in Equation (2) normalizes the expectation value of the NMAD to be equivalent to the standard deviation of a normal distribution.

We match objects to published spectroscopic redshifts, including those from the JADES DR3 (F. D'Eugenio et al. 2024) release, spectra and redshifts from the EGS region from CEERS (P. Arrabal Haro et al. 2023a) and a follow-up DDT program (P. Arrabal Haro et al. 2023b, PID 2750), and PID 2565 (K. Glazebrook et al. 2024). To increase our sample for the calibration of photometric redshifts, we also include spectroscopic redshifts for the GLASS–z12 object (M. Castellano et al. 2024) and results from the MACS–0416 field (Z. Ma et al. 2024) and the SMACS–0723 ERO program (K. M. Pontoppidan et al. 2022). We compare these spectroscopic redshifts to our calculated photometric redshifts and find that for redshifts $z > 6.5$, the NMAD value is 0.021. We find our outlier fraction to be $\eta = 9/86$, or $\sim$10%. These measures indicate a high-quality photometric redshift sample. The lack of F115W for the SMACS–0723 cluster makes some redshifts uncertain at $7.5 < z < 9.5$. When SMACS–0723 is omitted from the redshift sample, the fraction of outliers drops to $\eta = 6/86$, or $\sim$7%. Further details can be found in N. J. Adams et al. (2024) and C. J. Conselice et al. (2024).

### 2.8. LRD NIRSpec Spectra

We use the spectra of 26 LRDs (see Section 3.6) found in the Cosmic Dawn Center (DAWN) JWST Archive[24] (DJA; G. Brammer 2023; K. E. Heintz et al. 2024). The majority of these spectra are taken as part of the RUBIES program (GO-4233; PI: A. de Graaff; A. de Graaff et al. 2024). The remaining spectra are from the NIRSpec WIDE GTO Survey (GTO-1211 to 1215; PI: M. Maseda; M. V. Maseda et al. 2024), CEERS, and JADES DR1 (A. J. Bunker et al. 2024). The reduction process of these spectra is described in A. de Graaff et al. (2024) and K. E. Heintz et al. (2024).

[24] https://dawn-cph.github.io/dja/

## 3. Methodology

### 3.1. LRD Sample Selection

LRDs were selected following the general previous criteria used by V. Kokorev et al. (2024), which aims to identify compact sources with a red rest-frame optical continuum and blue rest-frame UV light. We require objects to be strongly detected in the F444W band by applying an F444W $< 27.7$ AB mag cut and a $>14\sigma$ detection. Using the same redness and significance criteria as V. Kokorev et al. (2024), we select red objects under the criteria red1 or red2, where

$$\begin{aligned} \text{red1} = {} & \text{F115W} - \text{F150W} < 0.8 \text{ and} \\ & \text{F200W} - \text{F277W} > 0.7 \text{ and} \\ & \text{F200W} - \text{F356W} > 1.0, \end{aligned} \tag{3}$$

$$\begin{aligned} \text{red2} = {} & \text{F150W} - \text{F200W} < 0.8 \text{ and} \\ & \text{F277W} - \text{F356W} > 0.6 \text{ and} \\ & \text{F277W} - \text{F444W} > 0.7, \end{aligned} \tag{4}$$

but only where one band of each color cut is $>2\sigma$ detected and the other is $>3\sigma$ detected.

These two sets of redness criteria target two redshift bins. We find that LRDs that meet the red1 criteria typically have $z \lesssim 6$, while those that meet the red2 criteria have $z \gtrsim 5$. We find $\sim$80% of our LRDs in red1 and $\sim$40% of our LRDs in red2. Around 10% of the LRDs are present in both red1 and red2.

To ensure the compactness of objects in the LRD sample, we require the condition

$$\text{compact} = f_{\rm F444W}(0\farcs5)/f_{\rm F444W}(0\farcs32) < 1.4. \tag{5}$$

This is different from other works on LRDs, as we use aperture sizes of 0.″5 and 0.″32, as opposed to the more commonly used 0.″4 and 0.″2, which require a ratio of $<1.7$ (e.g., J. E. Greene et al. 2024; V. Kokorev et al. 2024).

To reduce the amount of contaminants with strong emission lines rather than a red continuum, we add the color cut criterion

$$\text{F200W} - \text{F410M} > 0.9 \tag{6}$$

when F410M photometry is available. To choose this color cut, we compare objects with emission lines in the F200W or F410M bands to those without emission lines in either band. This is similar to D. D. Kocevski et al. (2025), who define a limit of F277W $-$ F410M $> 0.42$ instead. We find a surplus of objects below this color cut in the latter category. This removed a further 23 objects ($\sim$14%) from the sample. Finally, we also remove objects that appear to be diffuse or hot pixels by eye.

To increase the size of our clustering samples, we attempted to select a sample of LRDs from the COSMOS-Web field. However, we find that the relatively small number of bands means that the criteria could not be adjusted to avoid selecting a very different population. We also find that less than 20% of our sample in CEERS, NEP-TDF, and JADES overlaps with the color cut criterion in H. B. Akins et al. (2024). We remark that this criterion is meant to select only the subset of red

objects known as extremely red objects. For the reasons outlined above, we decide not to include COSMOS-Web in our LRD sample.

### 3.2. Brown Dwarfs

It is possible that some compact red sources are in fact brown dwarfs rather than LRDs (D. Langeroodi & J. Hjorth 2023). In fact, the fraction of brown dwarfs found in some samples is between up to ∼25% (D. Langeroodi & J. Hjorth 2023) and ∼5% (D. D. Kocevski et al. 2025). It may not be possible to rule out brown dwarfs in JWST data based on their sizes (B. W. Holwerda et al. 2024). To remove brown dwarfs, we fit the SEDs of the sample using brown dwarf templates from the Sonora Bobcat (M. S. Marley et al. 2021) and Sonora Cholla (K. N. Hainline et al. 2024b) models. We consider an object in the sample to be a brown dwarf if $\chi^2 < 20$ for the best-fitting brown dwarf model. Following the suggestion of J. E. Greene et al. (2024), we investigate using an additional color cut to remove potential brown dwarfs from LRD samples. The additional color cut is given by

$$(\text{bd\_removal}) = \text{F115W} - \text{F200W} > -0.5. \quad (7)$$

We compare this cut to the cuts made using our Sonora Bobcat and Sonora Cholla fits. In comparison to the objects selected as brown dwarfs by our $\chi^2 < 20$ cut, we find that the cut in Equation (7) removes eight contaminant brown dwarfs and six LRDs but misses one brown dwarf. We conclude that this color cut is effective in producing a mostly clean sample. Ultimately, we remove nine brown dwarfs using our own $\chi^2 < 20$ brown dwarf cut to remove this form of contaminant while maximizing our sample size.

### 3.3. Final LRD Sample

The number of objects in our final sample is 63 LRDs in the CEERS field, 42 in the NEP-TDF, and 19 in the JADES field, totaling 124 LRDs. One LRD in JADES is outside the SMILES footprint and thus has no MIRI data. We match 33 out of the 44 LRDs in CEERS found by V. Kokorev et al. (2024). We also match 30 of the 64 LRDs in CEERS and 12 of the 46 LRDs in JADES found by D. D. Kocevski et al. (2025). Our sample of LRDs in the NEP-TDF is the first to be published.

### 3.4. Robust 4 < z < 9 Galaxy Sample

To investigate the local environment of LRDs, we compare the clustering of LRDs and galaxies. To create a sample of comparison galaxies to use in our clustering studies, we follow the method of Q. Li et al. (2025). Limited to the range $4 < z < 9$ due to the available photometric bands, we use the following criteria (N. J. Adams et al. 2024; C. J. Conselice et al. 2024) to identify robust $4 < z < 9$ galaxies.

1. The first and second bands redward of the break are $\geqslant 5\sigma$ detected and any other bands redward of the break are $\geqslant 2\sigma$ detected to ensure a strong Lyman-break detection.
2. We also require any bands blueward of the Lyman break to not be $\geqslant 3\sigma$ detected.
3. The majority of the redshift probability density function $P(z)$ must be located inside the primary peak, achieved with the criteria $\int_{0.9\times z_{\rm phot}}^{1.10\times z_{\rm phot}} P(z)dz \geqslant 0.6$.

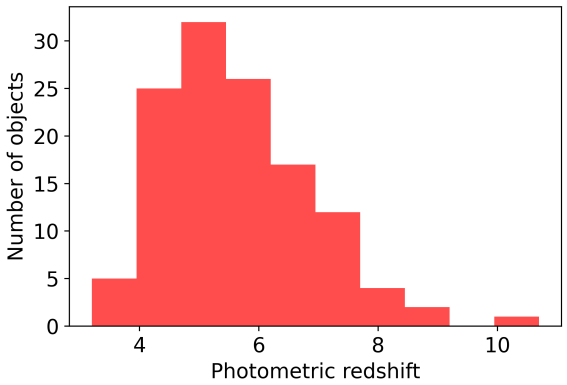


**Figure 1.** The redshift distribution for our LRD sample. The distribution peaks around $z \sim 5$ and contains most LRDs in the range $4 \lesssim z \lesssim 6$, similar to D. D. Kocevski et al. (2023), V. Kokorev et al. (2024), and I. Labbe et al. (2025).

4. If a secondary peak exists, we require it to be less than 50% of the higher-probability $z$ solution, so $P(z_{\rm sec}) < 0.5 \times P(z_{\rm phot})$.
5. For a best-fitting SED, we require $\chi^2_{\rm red} < 3$ for the SED fit to be considered robust.

The number of objects in the final galaxy sample is 3685. We later put galaxies into redshift bins for clustering analysis and restrict our galaxy sample to $4.75 < z < 8.25$ as described in Section 4.2.3. We note that the majority of LRDs selected by the criteria in Section 3.1 do not meet the criteria described above.

### 3.5. Redshift and Number Distribution of LRDs

To investigate the quality of our photometric redshifts for LRDs, we look for matches of our LRDs in the DJA (G. Brammer 2023; K. E. Heintz et al. 2024) for spectra and find matches for 26 out of 124 in total. We investigate the quality of the photometric redshifts of our sample of LRDs using the same method as in Section 2.7, focusing on the grade 3 redshift estimates, where grade 3 refers to spectra whose fits have been visually checked. Of the 26 matched LRDs, 24 have grade 3 redshift estimates. The outlier fraction of redshifts for our LRD sample is $\eta = 9/24$, or ∼40%. The NMAD for this sample is 0.112. This is noticeably poorer than for our parent sample (Section 2.7). This difference in quality is likely tied to our current lack of understanding of the physics involved in LRDs. Parameter estimates produced by various SED modeling codes may make nonrepresentative assumptions about LRDs, meaning that photometric redshift estimates are less accurate.

Most LRDs in our sample have redshift $4 \lesssim z \lesssim 6$, similar to D. D. Kocevski et al. (2023), V. Kokorev et al. (2024), and I. Labbe et al. (2025), spanning a total range of $3.5 < z < 10.4$ as shown in Figure 1. To calculate the number density of LRDs, we split our sample into redshift bins of size $\Delta z = 1$ from $z = 3$ to $z = 11$ as shown in Figure 2 and assuming a Poissonian error on the count of LRDs per bin. We find that the number density of our LRD sample in most of these bins is $\sim 10^{-5}$ cMpc$^{-3}$, in agreement with E. Pizzati et al. (2024). For the bins $8 < z < 9$ and $10 < z < 11$, which contain only one LRD each, the number density drops to $\sim 10^{-6}$ cMpc$^{-3}$.

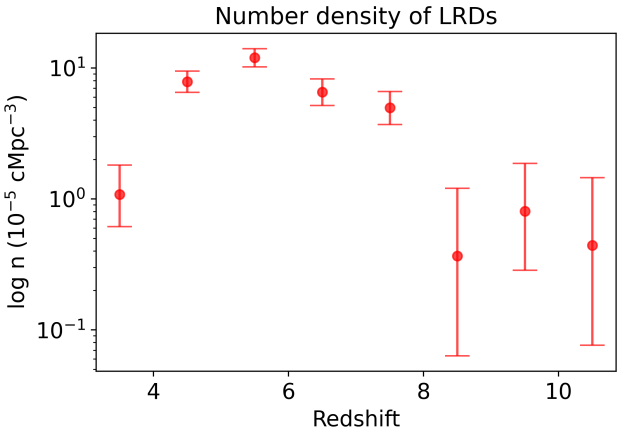


**Figure 2.** The number density evolution of our LRD sample. The errors on the number density are calculated assuming a Poissonian error on the count of LRDs in each bin.

### *3.6. Broad Lines in LRDs*

While the presence of somewhat broad H$\alpha$ lines (2000 km s$^{-1}$ > FWHM > 1000 km s$^{-1}$) could be due to other processes, broader lines (2000 km s$^{-1}$ > FWHM) are indicative of broad-line AGN (M. Habouzit 2025). For this reason, we investigate the broad-line fraction of our sample. We use the spectra of our 26 LRDs found in the DJA to search for broad H$\alpha$ lines (FWHM > 1000 km s$^{-1}$). If there are multiple spectra of one object taken with different resolution gratings, we select the spectrum with the highest-resolution grating and reject objects that only have PRISM spectra covering the wavelength range for H$\alpha$, as the PRISM resolution is generally not high enough to resolve broad lines. This leaves 19 spectra. To determine the presence of broad H$\alpha$ lines, we fit the spectra using a model with a single-component Gaussian curve and a model with a double-component Gaussian curve. We fix the center of the previously mentioned Gaussian curves for simplicity. To ensure that absorption features do not alter our results, we also use a version of the double-component model that models an absorption feature as a negative Gaussian curve of variable center. While it is possible that some spectra could contain [N II] emission, this was not obvious upon inspection and thus was not included in the models.

## 4. Results

To determine whether or not an H$\alpha$ line is broad, we compare the fits for the three models. To select the best model, we use a $\chi^2$ difference test (C. Werner & K. Schermelleh-Engel 2010), where we take the difference of the values of $\chi^2$ and compare this to the critical $\chi^2$ for the corresponding number of degrees of freedom at the 95th percentile. If the double-component Gaussian model (with or without absorption) is selected, we check the standard deviation of the two components to ensure that they are not too similar and are in line with the expected values for broad and narrow lines. We take the expected values for the full width at half-maximum (FWHM), ∼5000 km s$^{-1}$ to ∼1000 km s$^{-1}$ and ∼800 km s$^{-1}$ to ∼100 km s$^{-1}$, respectively (J. F. W. Baggen et al. 2024; J. E. Greene et al. 2024; X. Ji et al. 2025). Out of 19 LRD spectra, 3 have H$\alpha$ lines that are redshifted out of the range of both gratings and the PRISM, leaving 16 spectra that can be fit.

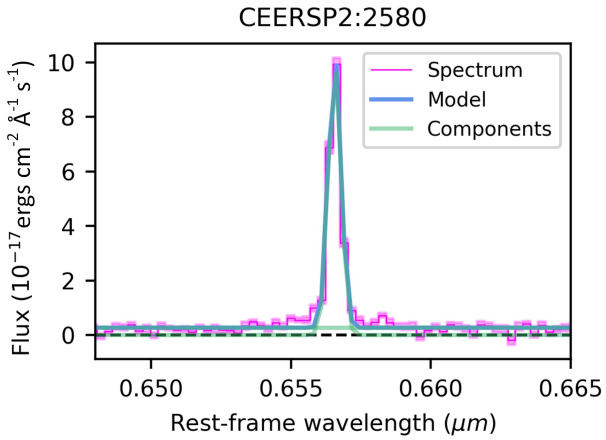


**Figure 3.** Spectrum and model of H$\alpha$ line for CEERSP2:2580. The best model selected for this LRD is a single-component Gaussian with an FWHM of ∼240 km s$^{-1}$.

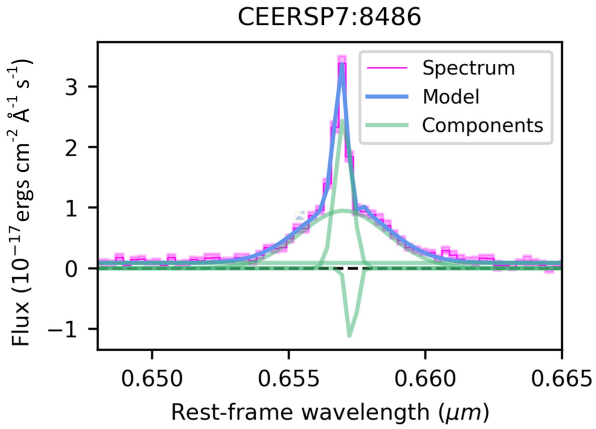


**Figure 4.** Spectrum and model of H$\alpha$ line for CEERSP7:8486. The best model selected is a double-component Gaussian with an absorption feature. The FWHM of the components is ∼1980 km s$^{-1}$ and ∼560 km s$^{-1}$ for the broad and narrow components, respectively. We note that the supposed detection of an absorption feature may instead be due to the presence of [N II] emission creating a small secondary peak. This was the only spectrum that displayed a possible hint of [N II] emission.

Of these 16, 14 are better fit by a double-component Gaussian with or without an absorption feature and 2 are better fit by a single narrow Gaussian component. One of the two single narrow lines found is shown in Figure 3. We provide two examples of the fits of the double-component Gaussian models with absorption features in Figures 4 and 5. The distribution of the FWHM of the broad-line fits is shown in Figure 6. In summary, we find that ∼80% ± 10% of LRDs with grating spectra have broad lines, in approximate agreement with J. E. Greene et al. (2024; ∼75%).

### *4.1. SED Modeling*

We investigate the composition of LRDs through SED modeling. We compare the results of models with and without AGN components for LRDs selected from CEERS, NEP-TDF, and JADES.

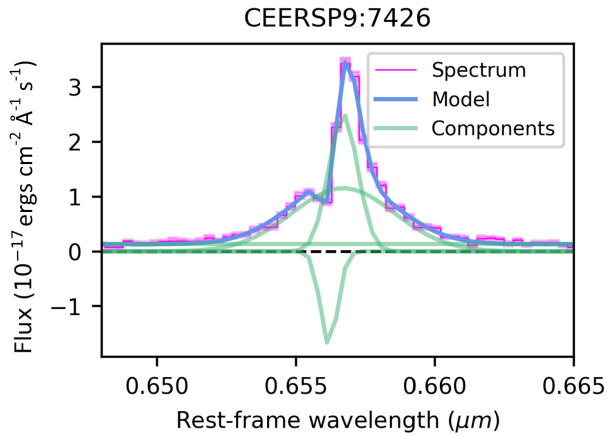


**Figure 5.** Spectrum and model of H$\alpha$ line for CEERSP9:7426. The best model selected is a double-component Gaussian with an absorption feature. The FWHM of the components is ∼1700 km s$^{-1}$ and ∼320 km s$^{-1}$ for the broad and narrow components, respectively.

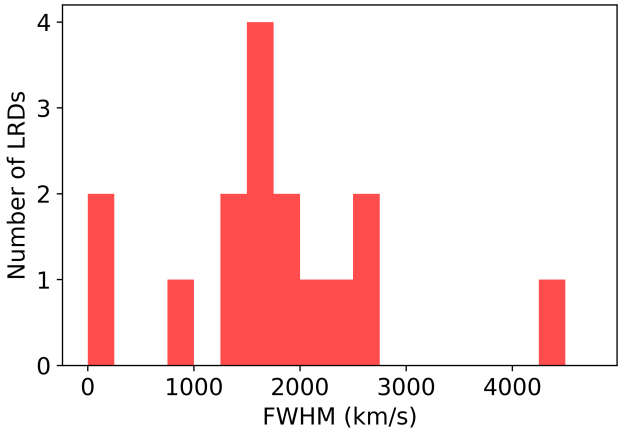


**Figure 6.** Distribution of FWHM in km s$^{-1}$ of the broadest component of each H$\alpha$ line for LRDs.

#### 4.1.1. Fitting Code

We analyze SEDs using CIGALE v2022.1 (D. Burgarella et al. 2005; S. Noll et al. 2009; M. Boquien et al. 2019), which models the spectrum of galaxies between the far-UV and radio. CIGALE builds composite models using templates that describe stellar populations with a flexible SF history (SFH), emission from ionized gas, AGN emission, dust emission attenuation, and nebular emission. To search for redshifts, CIGALE also has a photometric redshift mode. To determine whether CIGALE gives us different photometric redshift estimates, we run a subset of LRDs with and without MIRI data with CIGALE in photometric redshift mode and then compare these to an identical run with fixed photometric redshifts produced by EAZY. The resulting models are not significantly affected by the redshift setting used. Due to our large sample and the resulting computation time, we keep our photometric redshifts as a fixed variable. This is also consistent with how we have carried out similar analyses with the EPOCHS sample (e.g., C. J. Conselice et al. 2024).

The sfhdelayed SED module models a standard delayed $\tau$ SF model. We choose this module as it includes an optional exponential burst. We use a similar method to E. Durodola et al. (2024) to select the range of main stellar ages for the module, where we use the range of redshifts of our LRD sample to determine the range of possible main stellar ages. We make use of the bc03 module (G. Bruzual & S. Charlot 2003) to model a simple stellar population and set the initial mass function (IMF) to G. Chabrier (2003). Unlike the m2005 module (C. Maraston 2005), the alternative stellar population model available in CIGALE, the bc03 module can be combined with the nebular module. The choice of IMF is known to have a strong impact on stellar mass estimates. We select a G. Chabrier (2003) IMF as it includes fewer low-mass stars than a Salpeter IMF, reducing stellar mass, and seems to be somewhat more in line with available data (M. Cappellari et al. 2006; P. G. van Dokkum 2008).

To model nebular emission, we employ nebular, whose nebular templates are based on A. K. Inoue (2011). To model dust attenuation, we use dustatt_modified_starburst, which follows the D. Calzetti et al. (2000) starburst attenuation curve. We select this dust attenuation module as it allows the $E(B-V)$ of the continuum to be varied with one factor. We briefly investigated dust attenuation with higher $E(B-V)$ values but found that this significantly worsens the quality of the fits. For this reason, we do not change the standard CIGALE values for dustatt_modified_starburst. To model dust emission we apply the dl2014 module that is based on the B. T. Draine & A. Li (2007) models and updated in B. T. Draine et al. (2014). This is one of the more up-to-date dust emission modules and focuses on dust in the galaxies. We select this module as it does not contain an AGN component, which allows an AGN to be modeled as a separate module.

Finally, to model an AGN component, we use the clumpy AGN skirtor2016 model (M. Stalevski et al. 2012, 2016) SED module, which has controls for the gradient of dust density with both angle and radius, and the fraction of the total dust mass contained in clumps. The skirtor2016 module also includes parameters for the opening angle of the torus, the edge-on optical depth at 9.7 $\mu$m, and polar dust extinction and allows for both type I and type II AGN. However, we find that allowing type II AGN causes CIGALE to select the lowest AGN fraction and the highest polar dust extinction that we allow. For this reason, as well as because of the high broad-line fraction of our LRD sample, we model AGN of type I only. Note that these models are for those AGN with central tori, and thus other physical models of AGN, including those that are not understood or fully modeled, may give different SED shapes and forms.

To create fits, we combine the sfhdelayed, bc03, nebular, dustatt_modified_starburst, and dl2014 SED modules. We run the fitting procedure twice, once with skirtor2016 to model AGN and once without. Following the example of E. Durodola et al. (2024), we only allow a small set of variables to vary. Our choice of variables we vary is listed in Table 1. Any other variables are kept the same as the default single value given by CIGALE.

Due to the shallow nature of MIRI, many objects that are >5$\sigma$ detected in NIRCam data are not detected in MIRI data. HST data also contain nondetections due to the Lyman break of objects in our sample. To create our CIGALE fits, we treat any bands with <5$\sigma$ detections as upper limits.

**Table 1**
List of Variables Used in CIGALE Modeling

| Module | Parameter | Values |
|---|---|---|
| SFH | Main population *e*-folding time | 100, 400, 800, 1000, 2000, 3000, 4000, 5000 Myr |
| sfhdelayed | Main population age | 100, 200, 300, 400, 500, 750, 1000, 1250, 1500, 1750 Myr |
| | Burst population *e*-folding time | 10, 50, 100 Myr |
| | Burst population age | 10, 30, 50 Myr |
| | Burst fraction | 0, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5 |
| Stellar population | IMF | Chabrier |
| bc03 | … | … |
| AGN emission | AGN contribution | 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99 |
| skirtor2016 | Viewing angle | 30° |
| | Extinction in polar direction | 0.1, 0.5, 1, 2, 3, 4, 5, 6 [$E(B-V)$] |

**Note.** Any variables that are not listed were kept as the defaults given by CIGALE.

#### 4.1.2. AGN versus Non-AGN Model Composition

We briefly investigate the composition of the models created by CIGALE, examples of which are shown in Figures 7 and 8. A comparison of the stellar mass of the AGN and non-AGN models is shown in Figure 9. All of our LRDs, save two in the JADES and two in the CEERS fields, produce a higher stellar mass for non-AGN models. A total of 61 LRDs (∼50%) have a mass difference less than $\log_{10}(M_{*\mathrm{non-AGN}}/M_{\odot}) - \log_{10}(M_{*\mathrm{AGN}}/M_{\odot}) = 0.1$. This includes all except two JADES LRDs with MIRI data. The single LRD in JADES without MIRI data has a higher stellar mass when using an AGN component. This could be partly due to the greater depth of the NIRCam data of JADES, but the reason for this outlier is not clear.

The dust extinction of AGN models is shown in Figure 10. Most AGN models have a dust extinction at extremes of the allowed values with $E(B-V) = 0.1$, 0.5, 1, 2, or 6, with only three LRD in CEERS having $E(B-V) = 3$. No LRDs have $E(B-V) = 4$ or 5. We note that allowing even higher values of $E(B-V)$ typically results in LRDs with $E(B-V)$ moving to these higher values. For some of the LRDs, the unphysically high dust masses might discount torus-surrounded AGN.

We compare the stellar mass in LRDs with the fraction of AGN IR luminosity to total IR luminosity in Figure 11. We find no correlation between the fraction of AGN IR luminosity and the stellar mass. Most fractions of AGN produce similar ranges of stellar masses for LRDs. However, the highest fraction of AGN IR luminosity ($f_{\mathrm{AGN}} = 0.99$) produces a larger range of stellar masses than other fractions.

We note that the $\chi^2_r$ given by CIGALE is not necessarily the most suitable value of $\chi^2_r$, as it is calculated using the number of photometric bands available to determine the number of degrees of freedom. There is also some difficulty associated with determining the number of orthogonal parameters. For this reason, we utilize the BIC to determine whether AGN or non-AGN models are more suitable, rather than comparing the $\chi^2_r$ of the models.

#### 4.1.3. BIC

Due to the difficulty in determining the number of orthogonal degrees of freedom of the CIGALE models, we make use of the BIC to determine whether adding the skirtor2016 SED module results in the overfitting of LRD SEDs. The BIC is used in model selection and penalizes models based on the number of free parameters used. When comparing BICs between models, a lower BIC is generally preferred. For two given models, a difference >3.2 between the BICs indicates a substantial preference for the model with the lower BIC, and a difference >10 indicates a very strong preference (R. E. Kass & A. E. Raftery 1995). The BIC is given by

$$\mathrm{BIC} = k \log N + \chi^2, \tag{8}$$

where $k$ is the number of free parameters, $N$ is the number of photometric bands, and $\chi^2$ is defined as usual. This particular expression for the BIC relies on the assumption that the errors are independent of each other and distributed as identical Gaussian distributions. The AGN model we use has eight free parameters, while the non-AGN model has five free parameters. Out of our sample of 124 LRDs, we find that 93 LRDs have a lower $\chi^2$ for their AGN models (Figure 12). However, 84 out of 124 LRDs (68%) have a lower BIC for the non-AGN model than for the corresponding AGN model (Figure 13).

Most $\chi^2$ values are higher for the non-AGN models, and when the $\chi^2$ is higher for the AGN model, it is only slightly higher than the $\chi^2$ value for the corresponding non-AGN model (Figure 12). There is also a clear preference for models with higher fractions of AGN IR luminosity, where $\Delta\chi^2$ tends to a larger range of more negative values.

Around ∼70%, or 84 out of 124 pairs, of AGN and non-AGN models have a positive ΔBIC value, with most of the values for fields with MIRI data around ∼9 (Figure 13). As the BIC penalizes models with more parameters, positive values of ΔBIC suggest that most LRDs may be overfit by AGN components. Of these 84, 72 of them show substantial evidence of overfitting with $\Delta\mathrm{BIC} > 3.2$, and eight of these show very strong evidence with $\Delta\mathrm{BIC} > 10$. Notably, these eight are all in JADES. Generally, CEERS and NEP-TDF LRDs tend to have a lower ΔBIC than JADES LRDs. This is to be expected, as an AGN component is expected to have a larger impact when fitting MIRI data than NIRCam data, thus broadening the gap between BIC values when MIRI data are used. Most of the values for CEERS and NEP-TDF LRDs with lower AGN fractions are $\Delta\mathrm{BIC} \simeq 6$.

We briefly investigate if there exists a relation between the broadness of the Hα line in LRDs and the BICs of our AGN fits. We may expect the very broadest of lines to be a sign of AGN and so expect that these would be better fit by SED models with an AGN component. However, no relationship or cutoff is found between the FWHMs of the broadest line component and either the ΔBICs or BICs of the AGN fits.

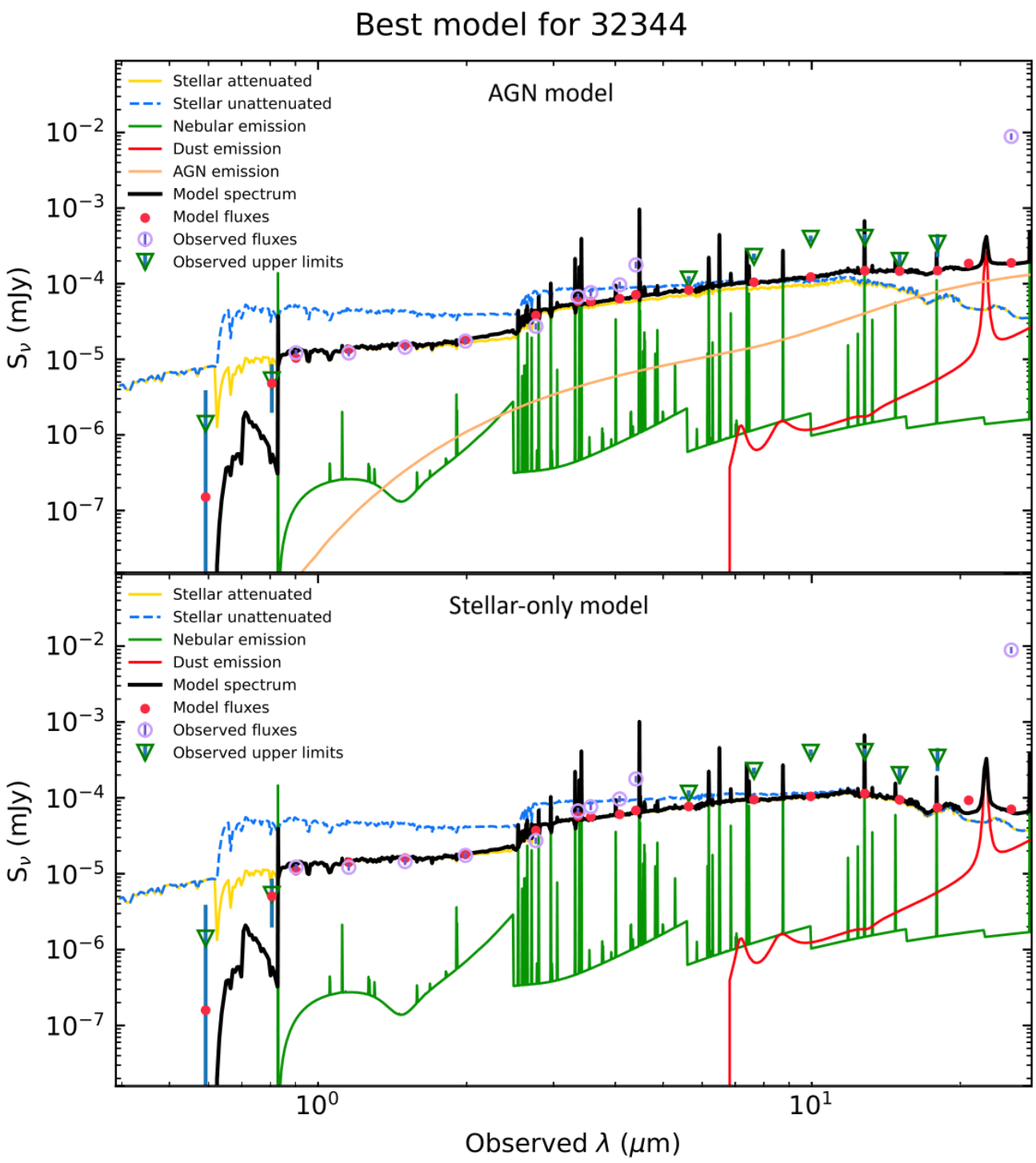


**Figure 7.** AGN (top) and non-AGN (bottom) CIGALE models for JADES:32344 ($z = 5.82$), which has MIRI coverage. Note that six of the longest-wavelength model flux points corresponding to MIRI data are treated as upper limits, while the F2550W band is treated as an observed flux. Both models produce a stellar mass of $\sim 10^{9.3}\ M_\odot$. The AGN IR luminosity fraction for this LRD is $f_{\rm AGN} = 0.9$.

### 4.2. Clustering of LRDs

#### 4.2.1. Local Environment

To study the local environment of galaxies and LRDs, we follow Q. Li et al. (2025) and use the nearest-neighbor method. We make use of the $k$-dimensional tree data structure to search for the nearest neighbors of objects and determine the separation between them. We use the term "objects" to refer to both LRDs and galaxies. We use the term "galaxy" to exclude all LRDs to avoid confusion. For any given object, we search for a nearest neighbor in the set of all objects, rather than searching for the nearest neighbor in the subset of like objects. We constrict our nearest-neighbor search to a specific maximum redshift offset $\Delta z$ relative to the object in question.

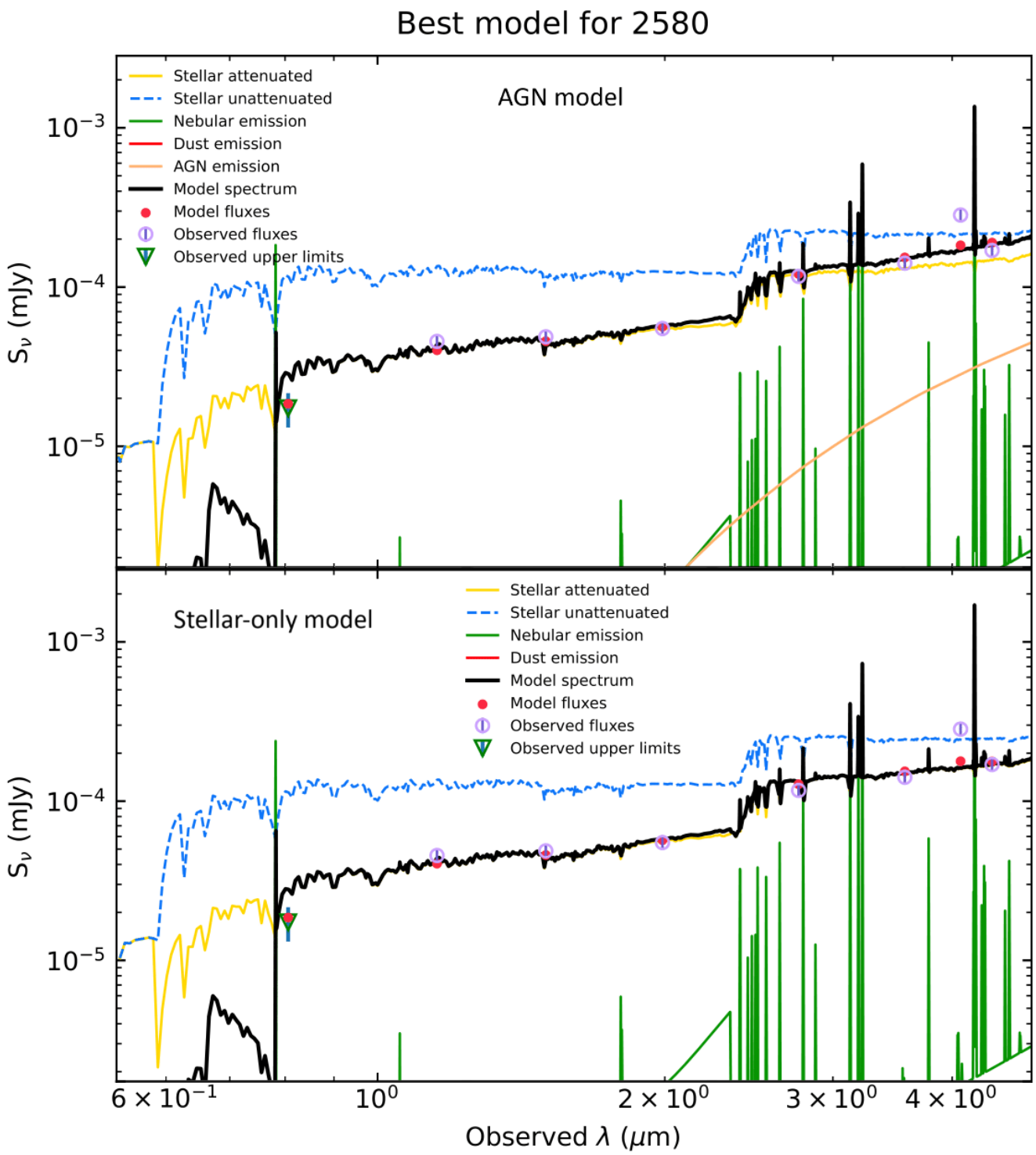


**Figure 8.** AGN (top) and non-AGN (bottom) CIGALE models for CEERS:2580, which has redshift $z = 5.43$.

We define a redshift offset mask as

$$\mathcal{M}_{\rm redshift} = \begin{cases} 1 & \text{if } \Delta z < 0.2 \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

The value 0.2 is chosen to ensure that $\Delta z$ is larger than the NMAD of the photometric redshifts for LRDs, which is found to be 0.112 (see Section 3.5). Any objects falling outside of this mask relative to the object in question are not included in the nearest-neighbor search.

#### 4.2.2. Impact of Image Depth and Borders

As the nearest neighbor(s) of objects close to the edges of an image may fall outside the image itself, the local density of these objects is usually underestimated. While this should have

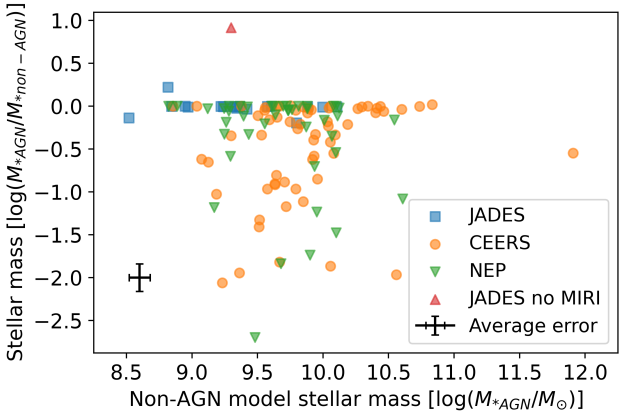


**Figure 9.** The stellar mass calculated by CIGALE for non-AGN models compared to the difference in stellar mass for AGN and non-AGN models, expressed logarithmically. The average error is included in the bottom left corner. The stellar mass of the non-AGN model is higher for all but two LRDs in JADES and two LRDs in CEERS.

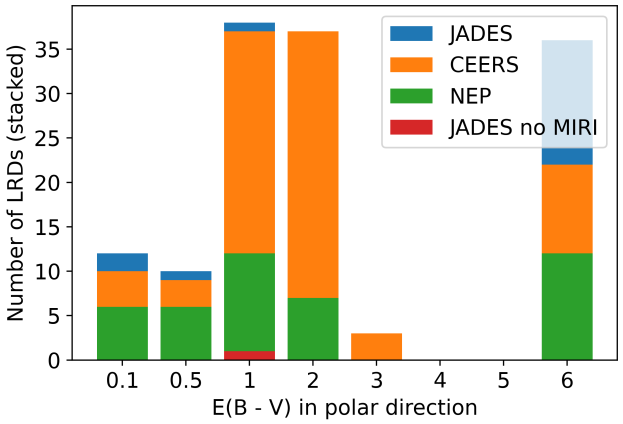


**Figure 10.** Stack plot of the dust extinction calculated by CIGALE for AGN models. Most AGN models are fit by a dust extinction close to the extremes (0.1 and 6) of the allowed values.

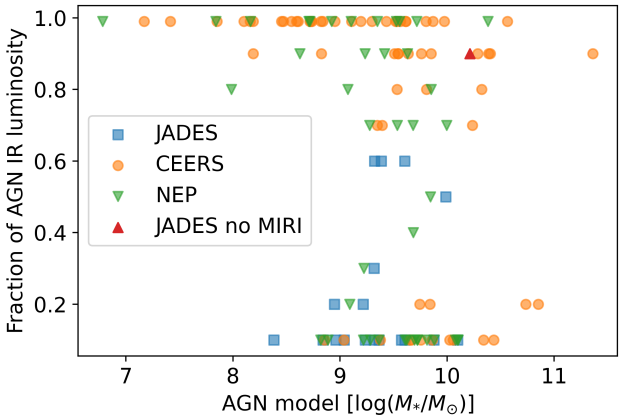


**Figure 11.** Comparison of the stellar mass with the fraction of AGN IR luminosity to total IR luminosity $f_{\rm AGN}$. The fraction $f_{\rm AGN}$ is quantized because of the choice of values allowed. We find no correlation between the fraction of AGN and stellar mass.

little impact on the comparison between LRDs and galaxies, accounting for image edges will give a somewhat more complete measure of the local environment around LRDs. To

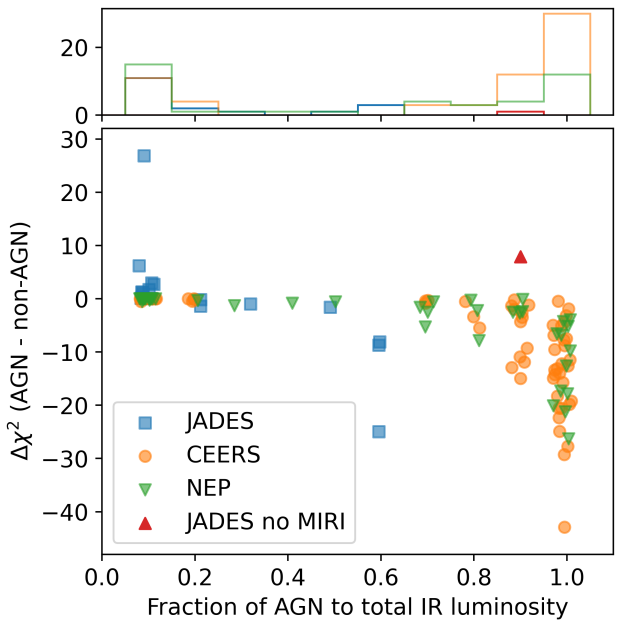


**Figure 12.** Comparison of $\Delta\chi^2$ across CEERS, NEP-TDF, and JADES with the fraction of AGN IR luminosity to total IR luminosity. Note that a jitter has been added to distinguish data points. The red symbol is the LRD in JADES without MIRI data. Note that some values of $\chi^2$ remain effectively the same. The LRDs with the highest AGN fractions tend to have a considerably improved $\chi^2$ compared to non-AGN models.

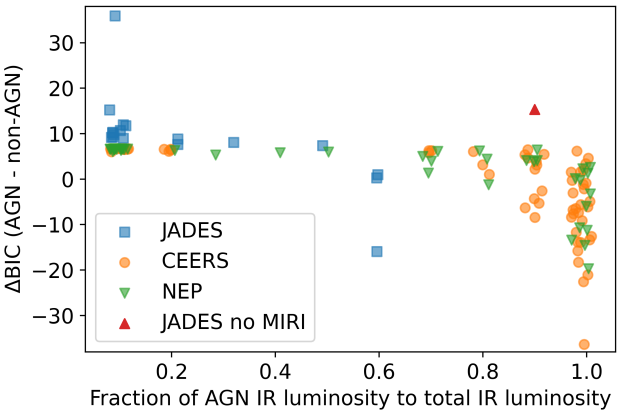


**Figure 13.** Comparison of the BIC with the fraction of AGN IR luminosity to total IR luminosity. Note that a jitter has been added to distinguish data points. Around ∼70% of the pairs of AGN and non-AGN models have a positive ΔBIC value. For LRDs with MIRI data, this is ∼9. In comparison, CEERS and NEP-TDF LRDs tend to have a lower ΔBIC than JADES LRDs. Most of the values for CEERS and NEP-TDF LRDs with lower AGN fractions are ∼6.

reduce the impact of image borders, we simply exclude objects that are expected to be significantly affected. We calculate the shortest distance to the boundary for each object, $d_{\rm edge}$. We define a mask for objects that are affected by the boundary:

$$\mathcal{M}_{\rm edge} = \begin{cases} 0 & \text{if } d_{\rm edge} < 1\,\text{cMpc} \\ 1 & \text{otherwise.} \end{cases} \quad (10)$$

Objects falling within this mask are not included in the sample.

**Table 2**
List of Redshift Bins and Number of Objects for Galaxies and LRDs per Field and in Total

| Redshift | Number of LRDs | | | | Number of Galaxies | | | |
|---|---|---|---|---|---|---|---|---|
| | CEERS | NEP-TDF | JADES | Total | CEERS | NEP-TDF | JADES | Total |
| $4.75 < z < 6.5$ | 33 | 24 | 10 | 67 | 840 | 869 | 498 | 2207 |
| $6.5 < z < 8.25$ | 10 | 10 | 4 | 24 | 283 | 273 | 253 | 809 |

**Note.** The bins were chosen to maximize the sample sizes while ensuring similar redshift distributions so as to avoid skewing $\Sigma_5$ measurements and distributions.

The local galaxy density in units of galaxies per Mpc$^2$, $\Sigma_n$, is then given by

$$\Sigma_n = \frac{n}{\pi d_n^2}, \tag{11}$$

where $d_n$ is the projected distance to the $n$th nearest neighbor in Mpc. While $\Sigma_5$ is typically used to study galaxy clusters (P. A. A. Lopes et al. 2016), we opt for $n = 5$ following the method given by Q. Li et al. (2025) and use it to compare the local density surrounding LRDs to the local density of galaxies, and we do not require $n$ to be less than the number of objects in a cluster.

We combine the data from the three fields and divide the samples of galaxies and LRDs into the redshift bins $4.75 < z < 6.5$ and $6.5 < z < 8.25$, described in Table 2. We choose these bins to include the largest possible sample of LRDs and to ensure that the distributions of LRDs and galaxies are similar to each other, as seen in Figure 14, thus minimizing biases to our nearest-neighbor search. Ideally, stellar masses should be matched to avoid further biases (J. Matthee et al. 2024b), but due to the difficulty in determining the stellar masses of LRDs through SED fitting and the necessity of choosing between AGN and non-AGN models, this is not possible. We find the $\langle\Sigma_5\rangle$ of non-LRD galaxies and LRDs to be $14.91^{+0.79}_{-0.65}$ cMpc$^{-2}$ and $9.56^{+1.51}_{-1.38}$ cMpc$^{-2}$, respectively, at $4.75 < z < 6.5$, and for $6.5 < z < 8.25$, $7.80^{+4.40}_{-4.56}$ cMpc$^{-2}$ and $4.65^{+1.63}_{-1.83}$ cMpc$^{-2}$, respectively. We show the $\Sigma_5$ distribution of both categories in Figure 15. We also calculate and plot $\langle\Sigma_5\rangle$ for random points with the same redshift distribution as LRDs and the same number density as galaxies and LRDs combined. The $\langle\Sigma_5\rangle$ for random points are 10.51 cMpc$^{-2}$ and 3.48 cMpc$^{-2}$ at redshift $4.75 < z < 6.5$ and $6.5 < z < 8.25$, respectively. For LRDs, this distribution is high-end deficient and significantly more common than the galaxy distribution at the very lowest $\Sigma_5$ values.

The fields used in this study are not completely homogeneous in depth, which in theory could reduce the $\langle\Sigma_5\rangle$ by a small amount compared to completely homogeneous fields. However, the magnitude limit for our galaxy selection is significantly brighter than the brightest variation of the depths across each of the three fields, which makes this variation unimportant for our analysis.

#### 4.2.3. Distribution Testing

We apply a K-S test using `Scipy` to test and compare the distributions of $\Sigma_5$ of LRDs and non-LRD galaxies. K-S tests can be used to determine whether one sample came from a given probability distribution or, in the case of a two-sample K-S test, whether two samples came from the same parent distribution. Our choice of redshift bins $4.75 < z < 6.5$ and $6.5 < z < 8.25$ shows the similarity in redshift distributions between LRDs and galaxies. This should ensure that the K-S tests are not biased by differing redshift distributions.

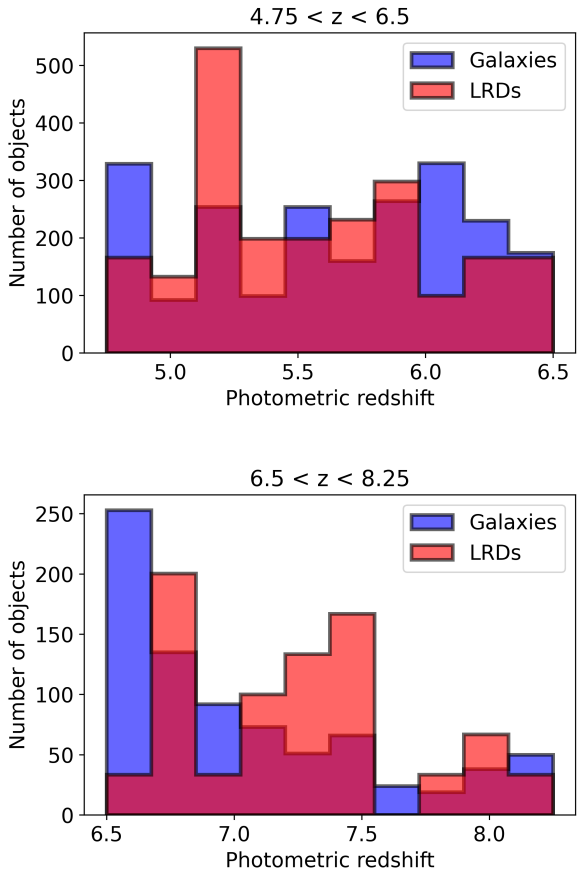


**Figure 14.** Distributions of LRD and galaxy redshifts for the $4.75 < z < 6.5$ and $6.5 < z < 8.25$ bins. Note that the LRD histograms are weighted to be normalized compared to the galaxy histograms for ease of visual comparison. The redshift distribution of LRDs and galaxies is roughly similar, especially for the $4.75 < z < 6.5$ bin.

For each redshift bin, we run a two-sample K-S test on the $\Sigma_5$ distribution of LRDs and of galaxies. If a $p$-value smaller than the default value of 0.003 is calculated, corresponding to a $3\sigma$ detection, then we reject the null hypothesis that the distributions of $\Sigma_5$ of LRDs and galaxies originate from the same parent distribution. The K-S tests give a $p$-value of 0.044 for $4.75 < z < 6.5$ and 0.014 for $6.5 < z < 8.25$; thus, we cannot reject the null hypothesis at a $3\sigma$ level. Regardless, this result is a tentative indication at the $2\sigma$ level that LRDs and galaxies do not have the same distribution.

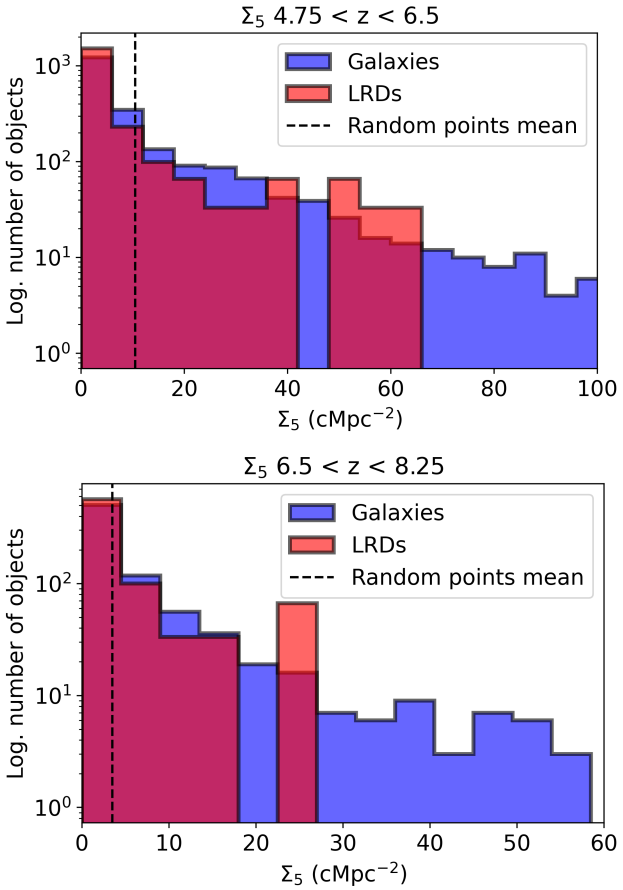


**Figure 15.** Comparison of $\Sigma_5$ for galaxies and LRDs for redshift bin $4.75 < z < 6.5$ (top) and $6.5 < z < 8.25$ (bottom). The black dashed line represents the $\langle\Sigma_5\rangle$ of random points. Note that the LRD histograms are weighted to be normalized compared to the galaxy histograms and have a logarithmic $y$-axis. LRDs occupy a smaller range of lower densities than galaxies do and are not found in dense environments.

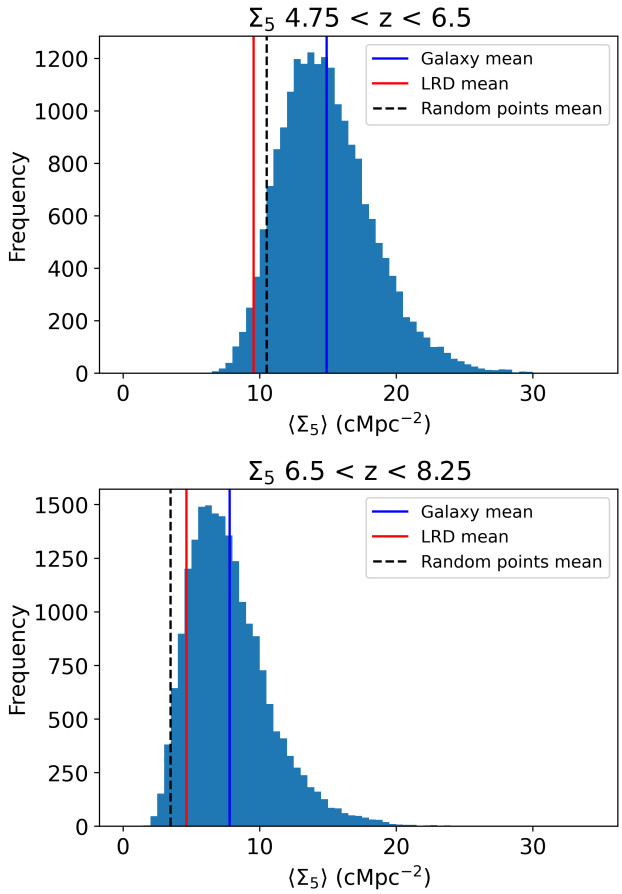


**Figure 16.** The $\langle\Sigma_5\rangle$ of each run of a random galaxy sample for redshift bin $4.75 < z < 6.5$ (top) and $6.5 < z < 8.25$ (bottom). The $\langle\Sigma_5\rangle$ for our LRD and galaxy samples are shown in red and blue, respectively, with the $\langle\Sigma_5\rangle$ of random points shown with a dashed black line for comparison.

There is a considerably larger proportion of LRDs with the lowest $\Sigma_5$ values ($\lesssim$5 cMpc$^{-2}$) when compared to galaxies. To investigate the impact of this on the strength of K-S test results, we randomly remove galaxies in all but the lowest $\Sigma_5$ bin shown in Figure 15 until the proportion of galaxies matches that of LRDs in the lowest $\Sigma_5$ bin and rerun the K-S tests. The resulting $p$-values are 0.84 for $4.75 < z < 6.5$ and 0.03 for $6.5 < z < 8.25$. In the lowest redshift bin, the $p$-value is far higher than its original value. This means the difference in distribution, at least at lower redshift, is largely due to the large proportion of LRDs in the very lowest-density environments.

To further investigate this difference, we run additional $K$-sample Darling–Anderson (D-A) tests, which determine whether a set of samples originates from a given population. The results are somewhat strong, with the lower redshift and higher redshift bins producing $p$-values of 0.024 and 0.018 and statistics of 2.79 and 3.11, respectively, pointing to a low likelihood that LRDs and galaxies have the same distribution. This, alongside the K-S tests, indicates a possibly significant difference in distribution of LRDs and galaxies.

To confirm that the above results are not due to the small size of our LRD sample or unknown effects, we randomly select a number of galaxies per redshift bin equal to the number of LRDs in each bin, so as to mimic any effects arising from the small size of the LRD sample. We then perform the same K-S test on the random galaxy sample and the parent galaxy sample and calculate the $\langle\Sigma_5\rangle$ of the random galaxy sample. This is repeated 20,000 times, the results of which are shown in Figure 16. As the random galaxy sample originates from the parent galaxy sample, we expect significant results with low $p$-values to appear very infrequently. Ideally, the frequency should be the same percentage as the corresponding threshold $p$-value, such that a $p$-value of 0.05 only appears in $\sim$5% of runs. The K-S tests on the redshift bins $4.75 < z < 6.5$ and $6.5 < z < 8.25$ produce a $p$-value of less than 0.05 only in $\lesssim$4% of the runs, suggesting that the K-S tests are behaving as expected. The distribution of the $\langle\Sigma_5\rangle$ of the random galaxy sample shows that the vast majority of random samples produce a $\langle\Sigma_5\rangle$ that is higher compared to the $\langle\Sigma_5\rangle$ of LRDs.

### 4.3. Halo Masses

Employing halo mass functions based on work by J. Tinker et al. (2008) and P. S. Behroozi et al. (2013), we estimate the halo masses of LRDs. We use the abundance matching approach to measure these halo masses within a Planck cosmology (Planck Collaboration et al. 2016). The halo masses we calculate are upper limits of the halo masses of these systems. We employ this approach as an experiment, under the hypothesis that the LRDs are galaxies and not AGN. This allows us to assume that these systems are not variable, with a

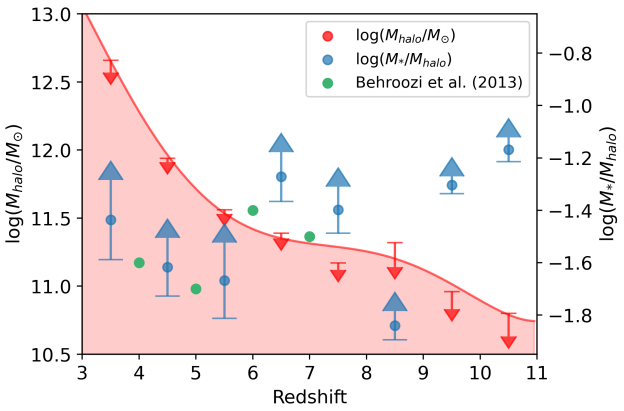


**Figure 17.** Halo mass from abundance matching (red) over the redshift range $3 < z < 11$ with redshift bin size $\Delta z = 1$. The SHMR (blue) is calculated from stellar masses extracted by CIGALE for non-AGN models. The SHMR holds relatively constant throughout the bins. The red shaded area is included for visualization and is an estimated fit of our data. Using this estimate fit, we include approximate points (green) for the corresponding SHMR of our halo mass at each redshift available in P. S. Behroozi et al. (2013).

short lifetime, but that they are galaxies of a similar semihomogeneous population.

The result of this is experiment is shown in Figure 17. What we can see is that the average halo mass grows considerably over time, such that under this hypothesis an LRD grows its halo mass by almost a factor of ∼100. This is a significant increase in the halo masses for these systems. When we examine the growth of stellar masses inferred for these systems in the non-AGN model fits, we see that they also grow with time for an LRD selected sample at a similar rate. It is of course not at all clear or obvious that LRDs at high redshift are the same as those we see at lower redshift, or even for that matter how long the LRD phase lasts within these systems.

We compare the average stellar mass estimates of non-AGN models given by CIGALE against the halo masses given by the halo mass function, finding that the mean stellar-to-halo mass ratio (SHMR) varies around $\sim 10^{-1.4}$. As the halo masses are upper limits, we expect the SHMRs to be lower limits. The value $10^{-1.4}$ is among the peak values for lower redshifts around $0 < z < 4$ found by simulations (C. A. Correa & J. Schaye 2020; G. Girelli et al. 2020). We create an approximate fit to our data using a polynomial of degree 5 to find estimates for halo mass at integer values of redshift. Using these estimates of halo mass at integer redshifts, we extract corresponding SHMRs for each redshift from P. S. Behroozi et al. (2013). These corresponding SHMRs, shown in Figure 17, are slightly lower than ours but seem to agree with our results.

If LRDs are hosted in dark matter halos, their clustering can place constraints on their host halo masses and duty cycles. For random points at $4.75 < z < 6.5$ and $6.5 < z < 8.25$, we find $\langle \Sigma_5 \rangle = 10.51\,\mathrm{cMpc}^{-2}$ and $3.48\,\mathrm{cMpc}^{-2}$, respectively. Remarkably, random points have a slightly higher average $\Sigma_5$ in the lower redshift bin than what we find for LRDs ($\langle \Sigma_5 \rangle \sim 9.6\,\mathrm{cMpc}^{-2}$). The natural implication is that LRDs at lower redshifts cannot form in dense regions—i.e., there is a physical effect associated with dense regions (such as ionizing radiation or metal pollution) that inhibits LRDs from forming there. Even for higher redshifts, the excess clustering above random implies a low bias relative to galaxies,

$$\frac{b_{\mathrm{LRDs}}}{b_{\mathrm{galaxies}}} \approx \frac{\Sigma_{5,\mathrm{LRDs}} - \Sigma_{5,\mathrm{Random}}}{\Sigma_{5,\mathrm{galaxies}} - \Sigma_{5,\mathrm{Random}}} \sim 0.5, \tag{12}$$

where the approximation comes because $\Sigma_5$ is an integral over slightly different physical scales for the different populations. From the cumulative number densities, the typical host halos of our galaxy sample (for $M_* > 10^{9.5}\ M_\odot$) in the higher redshift bin ($6.5 < z < 8.25$) would have $M_h > 10^{11.4}\ M_\odot$, with a mean bias of 10.8 (J. L. Tinker et al. 2010), naively implying that the masses of the LRD hosts (with a mean bias of half as much) would be $M_h \sim 10^{10.1}\ M_\odot$, which would typically host $10^7\ M_\odot$ galaxies (P. Behroozi et al. 2019). Given the exceptionally low clustering in the low redshift bin, we view it as plausible that there is an isolation effect that also prevents LRDs from forming in the densest regions in the higher redshift bin, and so this halo mass estimate should be viewed as a lower limit.

## 5. Discussion

### 5.1. SED Modeling and BIC Results

Most LRDs in our and other samples appear to have broad H$\alpha$ lines (Section 3.6). Most works interpret these broad lines as indicative of AGN (J. E. Greene et al. 2024; J. Matthee et al. 2024a; D. D. Kocevski et al. 2025), which we consider in our SED fits. J. F. W. Baggen et al. (2024) provide an alternative explanation to AGN, in which the broad lines are instead the result of a short-lived phase of galaxy evolution and reflect the kinematics of extremely high densities of such compact galaxies. Alternatively, V. Rusakov et al. (2025) suggest that LRDs may be intrinsically narrow-line AGN in which the broadening of lines is partially due to electron scattering through a cocoon of ionized gas, potentially originating from mild outflows due to feedback from a burst of SF or, more simply, accretion onto an SMBH. In the weak AGN or SF scenarios, it is possible that LRDs still contribute to reionization given that their main LyC escape and SF-driven outflow is directed away from the observer. LRDs would therefore not be a new class of objects but rather compact dust star-formers with dust obscuration with relatively unobscured counterparts with visible rest-frame UV. Given the mass measurements and limits based on clustering, it might be the case that these systems are forming star clusters that would not be expected to have a high clustering. As our mass limits are $10^7\ M_\odot$, which is about 10 times higher than globular clusters today, it is likely that these are perhaps the more commonly observed super star clusters that likely dissolve at later times (e.g., Y. Guo et al. 2018).

Of the 31 LRDs with a better $\chi^2$ for the non-AGN models, the values of $\Delta\chi^2 = \chi^2_{\mathrm{AGN}} - \chi^2_{\mathrm{non-AGN}}$ are usually small, with a typical value of $\Delta\chi^2 = 1$, suggesting that only minor improvements to fits are made when using an AGN component. Model SEDs of JADES 32344 and CEERS 1919 are shown in Figures 7 and 8, respectively. Of those that are not improved by including an AGN component, 12 are from JADES, 7 from CEERS, and 12 from NEP-TDF. However, the BIC calculated for the models suggests that adding the skirtor2016 SED module may lead to overfitting, as the BIC for AGN models is higher than that for the non-AGN models for 84 out of 124 LRDs in this sample (Figure 13) and

more than half of LRDs showing substantial evidence ($\Delta$BIC $> 3.2$) pointing toward non-AGN models. The $\Delta$BIC is especially high in the presence of MIRI data, where an AGN component is expected to dominate if present. LRDs with MIRI data are much more likely to show very strong evidence preferring the non-AGN model. Before completely ruling out the possibility of LRDs hosting AGN, we note that it may be possible to model the SED of these objects with two AGN components, such as a reddened component and a scattered component with different normalization and extinction. For example, I. Labbe et al. (2024) use a composite AGN SED model to fit the SED of most optically luminous LRDs found to date and find a very strong fit. This implies that if AGN are found in LRDs, they cannot have a dusty torus as modeled by `CIGALE`.

There is also a marked difference between the BICs calculated for LRDs in NEP-TDF and CEERS and the LRDs in JADES as seen in Figure 13. For CEERS and NEP-TDF, which use HST and NIRCam data only, the BICs of AGN models are typically $\sim$30, while for non-AGN models, they are $\sim$25. For JADES, these values are $\sim$110 and $\sim$100, respectively. All of the LRDs found in JADES have a lower $\chi^2$ for the non-AGN model, likely due to the presence of MIRI data. The BICs suggest that non-AGN models are more suitable, especially when MIRI data are used. A further effect of MIRI data, as seen in Figure 9, is that the difference in stellar mass between AGN and non-AGN models is much reduced compared to models with only NIRCam+HST data.

The SED models presented in this Letter are not necessarily representative of the true composition of LRDs. This is in part due to the poorly understood nature of LRDs. Despite this, we explore the model compositions and attempt to draw some conclusions. We find that AGN models without MIRI data tend to have high ($f_{\rm AGN} \sim$ 0.7–0.99) AGN-to-total-IR luminosity fractions (Figures 12 and 13), while those with MIRI data tend to have a lower AGN fraction ($f_{\rm AGN} = 0.1$). There tends to be a more negative and greater range of $\Delta\chi^2$ associated with a larger fraction of AGN. Furthermore, a higher AGN fraction is also associated with a more negative and greater range of $\Delta$BIC. Both the BIC and $\chi^2$ of AGN versus non-AGN models with MIRI data indicate a better fit for the non-AGN models.

The presence of an AGN component sometimes leads to a significantly lower stellar mass, which is up to 2 dex lower, as seen in Figure 9. This somewhat alleviates the anomalously high central stellar mass densities implied for LRDs (H. B. Akins et al. 2024). Some model stellar masses ($\sim$50%) are largely unaffected by the presence of an AGN component. In comparison, G. C. K. Leung et al. (2024) model LRDs as galaxies using `Bagpipes` (A. C. Carnall et al. 2018) and create AGN models using the `qsogen` code from M. J. Temple et al. (2021), and they find that fitting as a galaxy results in stellar masses that are $\sim$2 dex higher than models with AGN. We note that stellar masses for non-AGN models are typically lower and similar to AGN models for LRDs with HST+NIRCam+MIRI data. This is of particular interest, as it suggests that the problem of overmassive stellar masses of LRDs (e.g., V. Kokorev et al. 2023; C. A. Guia et al. 2024) may be at least partially tackled when MIRI data are taken into consideration.

There is no clear correlation visible in Figure 11 between the AGN fraction and the stellar mass. Most fractions of AGN produce a similar range of stellar masses for LRDs, but the highest fraction ($f_{\rm AGN} = 0.99$) produces the largest range of stellar masses.

### 5.2. Environment Results

The relatively large size of the redshift offset mask used to determine the nearest neighbors ($\Delta z < 0.2$; Section 4.2.1) reduces the impact on the $\langle\Sigma_5\rangle$ results of the somewhat poorer quality of photometric redshifts for LRDs compared to the rest of our sample (Section 3.5). However, this will still have an impact on the strength of the results that is difficult to assess.

The results of both the K-S and D-A tests suggest that the lower $\langle\Sigma_5\rangle$ result of LRDs is not due to a smaller sample size but rather due to a real difference between the distributions of LRDs and non-LRD galaxies. LRDs have a tendency to be found in less dense environments. Moreover, the range of $\Sigma_5$ of LRDs is also reduced compared to galaxies, as shown in Figure 15, although this could be in part caused by the smaller sample size of LRDs. Only $\sim$4% of the 20,000 randomly selected galaxy samples produce a $p$-value less than 0.05, suggesting that K-S tests are an effective method in determining that the difference in distributions of the $\Sigma_5$ of LRDs and galaxies is not coincidental or due to small sample size. The $\langle\Sigma_5\rangle$ of the random galaxy samples are also typically significantly higher than the $\langle\Sigma_5\rangle$ of our LRD sample, further suggesting that the lower $\Sigma_5$ of LRDs is not an accidental or random result.

We attempt to interpret the above results for LRDs as due to stellar mass by carrying out further K-S tests on various mass bins. However, a lower $\Sigma_5$ is not necessarily a sign of a lower or higher stellar mass than the comparison galaxies, as we find that the stellar mass of galaxies has a very weak if any correlation with density, in agreement with Q. Li et al. (2025). To make any estimates constraining halo mass, two-point clustering studies are required (A. Durkalec et al. 2015).

An alternative explanation to the lower-density environments in which LRDs are found could be that higher-density environments speed up the LRD evolutionary phase, meaning that any LRDs that were in higher-density environments have evolved past the LRD stage within the redshift range of this study. For example, T. Morishita et al. (2025) find an increased fraction of galaxies with weak H$\alpha$+[N II] emission lines in two overdensities of galaxies at $z \sim 5.7$ compared to field samples at similar redshifts. These weak emission line galaxies also exhibit a strong continuum break at 4000 Å, which is associated with evolved stellar populations. This evidence is consistent with the suggestion that high-density environments accelerate the evolution of galaxies. Some works studying overdensities find that LRDs tend to be isolated. For example, Y. Fudamoto et al. (2025) find an LRD candidate in a lower surface density area of an overdensity. J. B. Champagne et al. (2025) study a protocluster and find that objects meeting photometric LRD criteria (J. E. Greene et al. 2024) line the edge of the protocluster, and they argue that galaxy evolution occurs "inside-out" in dense environments. To extract further information, a more general study of the makeup of populations found around the edges of overdensities may be necessary.

### 5.3. Halo Mass and SHMR Results

As our halo masses are upper limits, our SHMRs are expected to be lower limits. To estimate a lower limit on halo masses, we assume that the SHMRs cannot exceed the maximum at $z = 0$, which is for the most massive systems. This value is ∼20% of the baryon fraction (17%; P. S. Behroozi et al. 2013), or approximately $10^{-1.5}$. The inferred SHMR is greater than this in four redshift bins, suggesting one of a few scenarios. This includes the idea that all massive halos host LRDs, which might be the case if our LRDs are, for example, forming star clusters. Alternatively, a large fraction of the light detected from LRDs must originate from AGN to increase their inferred measured stellar masses. However, we have shown that this is unlikely to be the case for the majority of LRDs given their SED fits.

The remarkably low clustering of LRDs compared to both galaxies and random points suggests a unique formation mechanism, which prevents LRDs from forming in the densest regions. This is opposite to most other traditional formation mechanisms, for example, the formation of galaxies and black holes, which is expected to be easier in high-density regions. This also poses a challenge to interpretation: all existing templates for galaxies and black holes are for objects that, at high redshifts, are easier to form in dense regions and so may not be a good description of the physics occurring in LRDs. At least, however, we can be confident that LRDs do not look like traditional templates for AGN.

It is of course interesting to consider the objects that may be preferentially formed in lower-density regions. One example is direct-collapse black holes (V. Bromm & A. Loeb 2003) and their potential precursors of supermassive stars (M. C. Begelman 2010), which need an ionizing radiation source to prevent gas fragmentation but require extremely low metallicity for the same reason and so may be anticorrelated with large-scale structure by $z \sim 5$. A more exotic example is gas clouds that are supported by dark matter annihilation—and hence do not need a separate ionizing radiation source—but nonetheless require low metallicity to avoid fragmentation (N. Banik et al. 2019). Key observational signatures of both such object classes are (as is the case for LRDs) lower formation rates once the intergalactic medium has been metal-enriched (especially at $z < 3$; A. De Cia et al. 2018) and lower clustering compared to galaxies. However, we note that initial low metallicities during formation may be obscured by later gas accretion and SF by the time the objects are bright enough to be observed as LRDs. We remain agnostic as to whether either (or neither) of these two can explain some of the LRD population, but the clustering measured in this work strongly motivates spectroscopic follow-up studies to understand the nature of these mysterious objects.

## 6. Conclusion

The LRDs, discovered with the JWST, are some of the most mysterious objects yet found in the very distant Universe. These systems are defined by three major characteristics: very red SEDs and spectra, very compact to unresolved structures, and very bright. In fact, when they were discovered, it was thought that the large brightness of these systems made their inferred stellar masses larger than is acceptable within ΛCDM cosmology (e.g., I. Labbe et al. 2025). Thus, there is an opportunity and importance in uncovering the nature of these systems.

In this Letter, we select a sample of a total of 124 LRDs from the CEERS, NEP-TDF, and JADES fields. Of those with grating spectra, 14 out of 16 LRDs show evidence of broad lines (FWHM 1000 km $s^{-1}$), a similar fraction to J. E. Greene et al. (2024). For the range $3 < z < 8$, we find a number density of $\sim 10^{-5}$ $cMpc^{-3}$, in line with E. Pizzati et al. (2024), which peaks at $\sim 10^{-4}$ $cMpc^{-3}$ at $5 < z < 6$.

We investigate and compare SED models for LRDs with and without AGN components. To determine whether adding an AGN component to an SED model of an LRD results in overfitting, we compare the $\chi^2$ and BIC of models with and without AGN. We find that while the $\chi^2$ of ∼75% of the LRDs in our sample suggests that SED models containing AGN are more suitable, the BICs calculated for our sample suggest that an AGN component often results in overfitting the SEDs of LRDs. This is particularly the case when using MIRI data, as we find that the presence of MIRI data results in a greater difference in the BICs of AGN and non-AGN models. AGN models with MIRI data also tend to have lower AGN fractions, while non-AGN models with MIRI data tend to have lower stellar masses.

Comparing the spatial distributions of LRDs to non-LRD galaxies, we find that the $\Sigma_5$ of LRDs tends to be lower than that of galaxies. We employ a K-S test to determine whether this result is significant and find that the redshift bins $4.75 < z < 6.5$ and $6.5 < z < 8.25$ produce a $p$-value of 0.044 and 0.014, respectively. Both $p$-values are relatively low, giving a tentative indication that the $\Sigma_5$ distributions of general galaxies and LRDs are different. Overall, this suggests that LRDs are typically found in lower-density environments than galaxies.

We calculate upper limit estimates for the halo masses and SHMR for our LRD sample using halo mass functions (J. Tinker et al. 2008; P. S. Behroozi et al. 2013) and `CIGALE`. We find that the inferred SHMR is relatively constant (∼0.03) over the range $3 < z < 11$. Our halo mass and SHMR values agree with simulation data over the range $4 < z < 8$ and are comparable with the peak SHMR values found by simulations for lower redshifts. The SHMR is greater than the maximum SHMR at $z = 0$ for four redshift bins. Because of this, we expect that either all massive halos host LRDs, in which case we would expect LRDs to be found in pairs, or AGN provide a large fraction of the light from LRDs.

The low $\langle \Sigma_5 \rangle$ of LRDs at redshift $4.75 < z < 6.5$ implies that LRDs are not formed in dense regions. In fact, even at higher redshifts ($6.5 < z < 8.25$), the low $\langle \Sigma_5 \rangle$ values suggest that LRD host halos have masses of $M_h \sim 10^{11.4}\ M_\odot$. These halos usually host galaxies with stellar masses of $10^7\ M_\odot$. Alternatively, LRDs may only be able to form in lower-density regions, in which case the halo mass calculated would be a lower limit.

The population of LRDs of the early Universe will continue to surprise for the foreseeable future. This work already uncovers some curious properties, such as the low clustering of these objects. By further examining clustering, it may be possible to place further constraints on the characteristics of LRDs.

## Acknowledgments

We thank the JADES, PEARLS, and CEERS teams for their work in designing and preparing these public and GTO observations and the STScI staff that carried them out. We acknowledge support from the ERC Advanced Investigator grant EPOCHS (788113), as well three studentships from the STFC. R.A.W., S.H.C., and R.A.J. acknowledge support from NASA JWST Interdisciplinary Scientist grants NAG5-12460, NNX14AN10G, and 80NSSC18K0200 from GSFC. L.F. acknowledges financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brazil (CAPES) in the form of a PhD studentship. This work is based on observations made with the NASA/ESA Hubble Space Telescope (HST) and NASA/ESA/CSA James Webb Space Telescope (JWST) obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST and NAS 5-26555 for HST. The observations used in this work are associated with JWST programs 1345, 1180, 1176, and 2738. The data described here may be obtained from the MAST archive at data set doi: 10.17909/5h64-g193.

Some of the data products presented herein were retrieved from the Dawn JWST Archive (DJA). DJA is an initiative of the Cosmic Dawn Center (DAWN), which is funded by the Danish National Research Foundation under grant DNRF140.

M.M.C.E. would like to thank Charlie Ellis for helping her with the presentation of the SED plots in this work. This work made use of the following Python libraries: Astropy (The Astropy Collaboration et al. 2022), Scipy (P. Virtanen et al. 2020), and Matplotlib (J. D. Hunter 2007). J.A.A.T. acknowledges support from the Simons Foundation and JWST program 3215. Support for program 3215 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 NAS5-03127. A.Z. acknowledges support by grant No. 2020750 from the United States–Israel Binational Science Foundation (BSF), grant No. 2109066 from the United States National Science Foundation (NSF), and Israel Science Foundation grant No. 864/23.

## ORCID iDs

María Carranza-Escudero https://orcid.org/0009-0002-0274-2676
Christopher J. Conselice https://orcid.org/0000-0003-1949-7638
Nathan Adams https://orcid.org/0000-0003-4875-6272
Thomas Harvey https://orcid.org/0000-0002-4130-636X
Duncan Austin https://orcid.org/0000-0003-0519-9445
Peter Behroozi https://orcid.org/0000-0002-2517-6446
Leonardo Ferreira https://orcid.org/0000-0002-8919-079X
Katherine Ormerod https://orcid.org/0000-0003-2000-3420
Qiao Duan https://orcid.org/0009-0009-8105-4564
James Trussler https://orcid.org/0000-0002-9081-2111
Qiong Li https://orcid.org/0000-0002-3119-9003
Lewi Westcott https://orcid.org/0009-0008-8642-5275
Rogier A. Windhorst https://orcid.org/0000-0001-8156-6281
Dan Coe https://orcid.org/0000-0001-7410-7669
Seth H. Cohen https://orcid.org/0000-0003-3329-1337
Cheng Cheng https://orcid.org/0000-0003-0202-0534
Simon P. Driver https://orcid.org/0000-0001-9491-7327
Brenda Frye https://orcid.org/0000-0003-1625-8009
Lukas J. Furtak https://orcid.org/0000-0001-6278-032X
Norman A. Grogin https://orcid.org/0000-0001-9440-8872
Nimish P. Hathi https://orcid.org/0000-0001-6145-5090
Rolf A. Jansen https://orcid.org/0000-0003-1268-5230
Anton M. Koekemoer https://orcid.org/0000-0002-6610-2048
Madeline A. Marshall https://orcid.org/0000-0001-6434-7845
Rosalia O'Brien https://orcid.org/0000-0003-3351-0878
Norbert Pirzkal https://orcid.org/0000-0003-3382-5941
Maria Polletta https://orcid.org/0000-0001-7411-5386
Aaron Robotham https://orcid.org/0000-0003-0429-3579
Michael J. Rutkowski https://orcid.org/0000-0001-7016-5220
Jake Summers https://orcid.org/0000-0002-7265-7920
Stephen M. Wilkins https://orcid.org/0000-0003-3903-6935
Christopher N. A. Willmer https://orcid.org/0000-0001-9262-9997
Haojing Yan https://orcid.org/0000-0001-7592-7714
Adi Zitrin https://orcid.org/0000-0002-0350-4488

## References

Adams, N. J., Conselice, C. J., Austin, D., et al. 2024, ApJ, 965, 169
Akins, H. B., Casey, C. M., Allen, N., et al. 2023, ApJ, 956, 61
Akins, H. B., Casey, C. M., Lambrides, E., et al. 2024, arXiv:2406.10341
Alberts, S., Lyu, J., Shivaei, I., et al. 2024, ApJ, 976, 224
Ananna, T. T., Bogdán, Á., Kovács, O. E., Natarajan, P., & Hickox, R. C. 2024, ApJL, 969, L18
Arrabal Haro, P., Dickinson, M., Finkelstein, S. L., et al. 2023a, ApJL, 951, L22
Arrabal Haro, P., Dickinson, M., Finkelstein, S. L., et al. 2023b, Natur, 622, 707
Baggen, J. F. W., van Dokkum, P., Brammer, G., et al. 2024, ApJL, 977, L13
Bagley, M. B., Finkelstein, S. L., Koekemoer, A. M., et al. 2023, ApJL, 946, L12
Banik, N., Tan, J. C., & Monaco, P. 2019, MNRAS, 483, 3592
Begelman, M. C. 2010, MNRAS, 402, 673
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Brammer, G. 2023, NIRSpec Analyis Tools, v0.6.17, Zenodo, doi:10.5281/zenodo.8319596
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bromm, V., & Loeb, A. 2003, ApJ, 596, 34
Brown, A. G. A., Vallenari, A., Prusti, T., et al. 2021, A&A, 650, C3
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bunker, A. J., Cameron, A. J., Curtis-Lake, E., et al. 2024, A&A, 690, A288
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2022, JWST Calibration Pipeline, v1.8.2, Zenodo, doi:10.5281/zenodo.7325378
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Casey, C. M., Akins, H. B., Finkelstein, S. L., et al. 2025, arXiv:2505.18873
Castellano, M., Napolitano, L., Fontana, A., et al. 2024, ApJ, 972, 143
Chabrier, G. 2003, PASP, 115, 763
Champagne, J. B., Wang, F., Yang, J., et al. 2025, ApJ, 981, 114
Conroy, C., & Gunn, J. E. 2010, FSPS: Flexible Stellar Population Synthesis, Astrophysics Source Code Library, ascl:1010.043
Conselice, C. J., Adams, N., Harvey, T., et al. 2025, ApJ, 983, 30

Correa, C. A., & Schaye, J. 2020, MNRAS, 499, 3578
Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJL, 660, L1
De Cia, A., Ledoux, C., Petitjean, P., & Savaglio, S. 2018, A&A, 611, A76
de Graaff, A., Brammer, G., Weibel, A., et al. 2025, A&A, 697, 21
de Graaff, A., Rix, H.-W., Naidu, R. P., et al. 2025, arXiv:2503.16600
D'Eugenio, F., Cameron, A. J., Scholtz, J., et al. 2025, ApJS, 277, 4
Diego, J. M., Meena, A. K., Adams, N. J., et al. 2023, A&A, 672, A3
Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172
Draine, B. T., & Li, A. 2007, ApJ, 657, 810
Durkalec, A., Le Fèvre, O., Pollo, A., et al. 2015, A&A, 583, A128
Durodola, E., Pacucci, F., & Hickox, R. C. 2025, ApJ, 985, 169
Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, arXiv:2306.02465
Frye, B. L., Pascale, M., Foo, N., et al. 2023, ApJ, 952, 81
Fudamoto, Y., Helton, J. M., Lin, X., et al. 2025, arXiv:2503.15597
Furtak, L. J., Labbé, I., Zitrin, A., et al. 2024, Natur, 628, 57
Furtak, L. J., Zitrin, A., Plat, A., et al. 2023, ApJ, 952, 142
Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, A&A, 634, A135
Glazebrook, K., Nanayakkara, T., Schreiber, C., et al. 2024, Natur, 628, 277
Grazian, A., Giallongo, E., Boutsia, K., et al. 2024, ApJ, 974, 84
Greene, J. E., Labbe, I., Goulding, A. D., et al. 2024, ApJ, 964, 39
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Guia, C. A., Pacucci, F., & Kocevski, D. D. 2024, RNAAS, 8, 207
Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, ApJ, 698, 198
Guo, Y., Rafelski, M., Bell, E. F., et al. 2018, ApJ, 853, 108
Habouzit, M. 2025, MNRAS, 537, 2323
Hainline, K. N., Helton, J. M., Johnson, B. D., et al. 2024b, ApJ, 964, 66
Hainline, K. N., Johnson, B. D., Robertson, B., et al. 2024a, ApJ, 964, 71
He, W., Akiyama, M., Enoki, M., et al. 2024, ApJ, 962, 152
Heintz, K. E., Watson, D., Brammer, G., et al. 2024, Sci, 384, 890
Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding Robust and Exploratory Data Analysis (New York: Wiley)
Holwerda, B. W., Hsu, C.-C., Hathi, N., et al. 2024, MNRAS, 529, 1067
Hu, J. 2008, MNRAS, 386, 2242
Hunter, J. D. 2007, CSE, 9, 90
Illingworth, G., Magee, D., Bouwens, R., et al. 2016, arXiv:1606.00841
Inayoshi, K., & Maiolino, R. 2025, ApJL, 980, L27
Inoue, A. K. 2011, MNRAS, 415, 2920
Ji, X., Maiolino, R., Übler, H., et al. 2025, arXiv:2501.13082
Kass, R. E., & Raftery, A. E. 1995, JASA, 90, 773
Kocevski, D. D., Finkelstein, S. L., Barro, G., et al. 2025, ApJ, 986, 126
Kocevski, D. D., Onoue, M., Inayoshi, K., et al. 2023, ApJL, 954, L4
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kokorev, V., Caputi, K. I., Greene, J. E., et al. 2024, ApJ, 968, 38
Kokorev, V., Fujimoto, S., Labbe, I., et al. 2023, ApJL, 957, L7
Kokubo, M., & Harikane, Y. 2024, arXiv:2407.04777
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Labbe, I., Greene, J. E., & Bezanson, R. 2025, ApJ, 978, 92
Labbe, I., Greene, J. E., Matthee, J., et al. 2024, arXiv:2412.04557
Langeroodi, D., & Hjorth, J. 2023, ApJL, 957, L27
Larson, R. L., Hutchison, T. A., Bagley, M., et al. 2023, ApJ, 958, 141
Leung, G. C. K., Finkelstein, S. L., Pérez-González, P. G., et al. 2024, arXiv:2411.12005
Li, Q., Conselice, C. J., Sarron, F., et al. 2025, MNRAS, 539, 1796
Li, Z., Inayoshi, K., Chen, K., Ichikawa, K., & Ho, L. C. 2025, ApJ, 980, 39
Lopes, P. A. A., Rembold, S. B., Ribeiro, A. L. B., Nascimento, R. S., & Vajgel, B. 2016, MNRAS, 461, 2559
Ma, Z., Sun, B., Cheng, C., et al. 2024, ApJ, 975, 87
Madau, P., Giallongo, E., Grazian, A., & Haardt, F. 2024, ApJ, 971, 75
Maiolino, R., Risaliti, G., Signorini, M., et al. 2025, MNRAS, 538, 1921
Maraston, C. 2005, MNRAS, 362, 799
Marley, M. S., Saumon, D., Visscher, C., et al. 2021, ApJ, 920, 85
Maseda, M. V., de Graaff, A., Franx, M., et al. 2024, A&A, 689, A73
Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, ApJ, 869, 150
Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, ApJ, 963, 129
Matthee, J., Naidu, R. P., Kotiwale, G., et al. 2025, ApJ, 988, 246
Morishita, T., Liu, Z., Stiavelli, M., et al. 2025, ApJ, 982, 153
Naidu, R. P., Matthee, J., Katz, H., et al. 2025, arXiv:2503.16596
Niida, M., Nagao, T., Ikeda, H., et al. 2020, ApJ, 904, 89
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
O'Brien, R., Jansen, R. A., Grogin, N. A., et al. 2024, ApJS, 272, 19
Oke, J. B. 1974, ApJS, 27, 21
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Pérez-González, P. G., Barro, G., Rieke, G. H., et al. 2024, ApJ, 968, 4
Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, Proc. SPIE, 9143, 1174
Perrin, M. D., Soummer, R., Elliott, E. M., Lallo, M. D., & Sivaramakrishnan, A. 2012, Proc. SPIE, 8442, 84423D
Pizzati, E., Hennawi, J. F., Schaye, J., et al. 2025, MNRAS, 539, 2910
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Pontoppidan, K. M., Barrientes, J., Blome, C., et al. 2022, ApJL, 936, L14
Rieke, M. J., Robertson, B., Tacchella, S., et al. 2023, ApJS, 269, 16
Rusakov, V., Watson, D., Nikopoulos, G. P., et al. 2025, arXiv:2503.16595
Setton, D. J., Greene, J. E., de Graaff, A., et al. 2024, arXiv:2411.03424
Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popović, L. č. 2012, MNRAS, 420, 2756
Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288
Temple, M. J., Hewett, P. C., & Banerji, M. 2021, MNRAS, 508, 737
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
The Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
van Dokkum, P. G. 2008, ApJ, 674, 29
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Werner, C., & Schermelleh-Engel, K. 2010, Deciding Between Competing Models: Chi-Square Difference Tests. Introduction to structural equation modeling with LISREL, 1
Whitaker, K. E., Ashas, M., Illingworth, G., et al. 2019, ApJS, 244, 16
Williams, C. C., Alberts, S., Ji, Z., et al. 2024, ApJ, 968, 34
Windhorst, R. A., Cohen, S. H., Jansen, R. A., et al. 2022, AJ, 165, 13
Yue, M., Eilers, A.-C., Ananna, T. T., et al. 2024, ApJL, 974, L26
Zhang, Z., Jiang, L., Liu, W., & Ho, L. C. 2025, ApJ, 985, 119