

# Evidence for a Shallow Evolution in the Volume Densities of Massive Galaxies at $z=4-8$ from CEERS

Katherine Chworowsky<sup>1,35</sup>, Steven L. Finkelstein<sup>1</sup>, Michael Boylan-Kolchin<sup>1</sup>, Elizabeth J. McGrath<sup>2</sup>, Kartheik G. Iyer<sup>3</sup>, Casey Papovich<sup>4,5</sup>, Mark Dickinson<sup>6</sup>, Anthony J. Taylor<sup>1</sup>, L. Y. Aaron Yung<sup>7,8,36</sup>, Pablo Arrabal Haro<sup>6</sup>, Micaela B. Bagley<sup>1</sup>, Bren E. Backhaus<sup>9</sup>, Rachana Bhatawdekar<sup>10</sup>, Yingjie Cheng<sup>11</sup>, Nikko J. Cleri<sup>4,5</sup>, Justin W. Cole<sup>4,5</sup>, M. C. Cooper<sup>12</sup>, Luca Costantini<sup>13</sup>, Avishai Dekel<sup>14</sup>, Maximilien Franco<sup>15</sup>, Seiji Fujimoto<sup>16,17</sup>, Christopher C. Hayward<sup>18</sup>, Benne W. Holwerda<sup>19</sup>, Marc Huertas-Company<sup>20,21,22</sup>, Michaela Hirschmann<sup>23</sup>, Taylor A. Hutchison<sup>7,36</sup>, Anton M. Koekemoer<sup>8</sup>, Rebecca L. Larson<sup>24</sup>, Zhaozhou Li<sup>14</sup>, Arianna S. Long<sup>15,37</sup>, Ray A. Lucas<sup>8</sup>, Nor Pirzkal<sup>25</sup>, Giulia Rodighiero<sup>26,27</sup>, Rachel S. Somerville<sup>18</sup>, Brittany N. Vanderhoof<sup>28</sup>, Alexander de la Vega<sup>28</sup>, Stephen M. Wilkins<sup>29,30</sup>, Guang Yang<sup>31,32</sup>, and Jorge A. Zavala<sup>33,34</sup>

<sup>1</sup> Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

<sup>2</sup> Department of Physics and Astronomy, Colby College, Waterville, ME 04901, USA

<sup>3</sup> Dunlap Institute for Astronomy & Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada

<sup>4</sup> Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA

<sup>5</sup> George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA

<sup>6</sup> NSF's National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Ave., Tucson, AZ 85719, USA

<sup>7</sup> Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD 20771, USA

<sup>8</sup> Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD 21218, USA

<sup>9</sup> Department of Physics, 196 Auditorium Road, Unit 3046, University of Connecticut, Storrs, CT 06269, USA

<sup>10</sup> European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino Bajo del Castillo s/n, 28692 Villanueva de la Cañada, Madrid, Spain

<sup>11</sup> University of Massachusetts Amherst, 710 North Pleasant St., Amherst, MA 01003-9305, USA

<sup>12</sup> Department of Physics & Astronomy, University of California, Irvine, 4129 Reines Hall, Irvine, CA 92697, USA

<sup>13</sup> Centro de Astrobiología (CAB), CSIC-INTA, Ctra de Ajalvir km 4, Torrejón de Ardoz, 28850 Madrid, Spain

<sup>14</sup> Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

<sup>15</sup> Department of Astronomy, The University of Texas at Austin, Austin, TX, USA

<sup>16</sup> Cosmic Dawn Center (DAWN), Jagtvej 128, DK2200 Copenhagen N, Denmark

<sup>17</sup> Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK2100 Copenhagen Ø, Denmark

<sup>18</sup> Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

<sup>19</sup> Physics & Astronomy Department, University of Louisville, Louisville, KY 40292, USA

<sup>20</sup> Instituto de Astrofísica de Canarias, La Laguna, Tenerife, Spain

<sup>21</sup> Universidad de la Laguna, La Laguna, Tenerife, Spain

<sup>22</sup> Université Paris-Cité, LERMA—Observatoire de Paris, PSL, Paris, France

<sup>23</sup> Institute of Physics, Laboratory of Galaxy Evolution, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland

<sup>24</sup> School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Dr., Rochester, NY 14623, USA

<sup>25</sup> ESA/AURA Space Telescope Science Institute, USA

<sup>26</sup> Department of Physics and Astronomy, Università degli Studi di Padova, Vico del Osservatorio 3, I-35122, Padova, Italy

<sup>27</sup> INAF—Osservatorio Astronomico di Padova, Vico del Osservatorio 5, I-35122, Padova, Italy

<sup>28</sup> Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA

<sup>29</sup> Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

<sup>30</sup> Institute of Space Sciences and Astronomy, University of Malta, Msida MSD 2080, Malta

<sup>31</sup> Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

<sup>32</sup> SRON Netherlands Institute for Space Research, Postbus 800, 9700 AV Groningen, The Netherlands

<sup>33</sup> National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>34</sup> Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave., New York, NY 10010, USA

Received 2023 November 21; revised 2024 March 21; accepted 2024 March 22; published 2024 August 26

## Abstract

We analyze the evolution of massive ( $\log_{10}[M_*/M_\odot] > 10$ ) galaxies at  $z \sim 1-4$  selected from JWST Cosmic Evolution Early Release Survey (CEERS). We infer the physical properties of all galaxies in the CEERS NIRCam imaging through spectral energy distribution (SED) fitting with dense basis to select a sample of high-redshift massive galaxies. Where available we include constraints from additional CEERS observing modes, including 18 sources with MIRI photometric coverage, and 28 sources with spectroscopic confirmations from NIRSpec or NIRCam WFSS. We sample the recovered posteriors in stellar mass from SED fitting to infer the volume densities of massive galaxies across cosmic time, taking into consideration the potential for sample contamination by active galactic nuclei. We find that the evolving abundance of massive galaxies tracks expectations based on a constant baryon conversion efficiency in dark matter halos for  $z \sim 1-4$ . At higher redshifts, we observe an excess abundance

<sup>35</sup> NSF Graduate Fellow.

<sup>36</sup> NASA Postdoctoral Fellow.

<sup>37</sup> NASA Hubble Fellow.

of massive galaxies relative to this simple model, resulting in a shallower decline of observed volume densities of massive galaxies. These higher abundances can be explained by modest changes to star formation physics and/or the efficiencies with which star formation occurs in massive dark matter halos, and are not in tension with modern cosmology.

*Unified Astronomy Thesaurus concepts:* Galaxy evolution (594); Galaxy formation (595); High-redshift galaxies (734); Galactic and extragalactic astronomy (563)

*Materials only available in the online version of record:* machine-readable table

## 1. Introduction

Measurements of galaxy stellar masses can provide a powerful benchmark to compare to theoretical models and simulations. The ability of the Universe to create and form galaxies with large amounts of stellar mass is largely dependent on galaxy baryonic feedback processes, both internally and within a galaxy's environment. In particular, massive galaxies at high redshift provide interesting laboratories to study galaxy formation physics, as these extreme systems allow us to begin to understand how the Universe is able to rapidly build up large amounts of stellar mass in a short amount of time, and to constrain the feedback mechanisms that may be regulating these processes (e.g., Hayward et al. 2021).

Observationally, extensive efforts have been made in attempting to understand how galaxy stellar masses evolve over time, with many studies measuring galaxy stellar mass functions (SMFs) spanning cosmic time (e.g., Muzzin et al. 2013; Duncan et al. 2014; Kelvin et al. 2014; Tomczak et al. 2014; Moffett et al. 2016; Song et al. 2016; Wright et al. 2017; Bhatawdekar et al. 2019; Leja et al. 2020; Adams et al. 2021; McLeod et al. 2021; Stefanon et al. 2021; Weaver et al. 2022a). The stellar mass of a galaxy is typically inferred via modeling the rest-frame ultraviolet (UV) through near-infrared (NIR) spectral energy distribution (SED), which necessarily relies on assumptions about the galaxy's star formation history (SFH), metal enrichment history, dust attenuation, and stellar initial mass function (IMF). Prior to JWST, the Hubble Space Telescope (HST) provided the most accurate galaxy stellar mass estimates out to redshifts of  $z \sim 5$ , however, at increasingly high redshift, the wavelength coverage of HST ( $\sim 0.5$ – $1.8 \mu\text{m}$ ) probes an increasingly limited wavelength range of the SED, and only UV-luminous galaxies are detected. While the Spitzer Infrared Array Camera (IRAC) provided longer-wavelength space-based observations (e.g., Song et al. 2016; Bhatawdekar et al. 2019; Morishita 2021; Tacchella et al. 2022), the poor spatial resolution ( $\sim 2''$ ) and small collecting area of Spitzer IRAC with respect to HST meant that galaxy observations were often blended with neighboring sources, limiting the ability of observations to accurately attribute observed fluxes to their sources.

The launch of JWST (Gardner et al. 2023) enabled a huge leap forward in both the resolution and sensitivity of infrared imaging. With the longer-wavelength coverage of the JWST Near Infrared Camera (NIRcam; Rieke et al. 2023), observations of the rest-frame optical (including the SFH-sensitive 4000 Å break) are accessible out to  $z \sim 10$ , allowing for the measurement of accurate stellar masses out to redshifts that were previously inaccessible. Nearly 2 yr since the launch of JWST, the data collected have already begun to revolutionize our understanding of the Universe, with new discoveries answering long-standing questions in the field of extragalactic

science, while also providing unexpected results and generating new questions.

One of these unexpected early results is an apparent excess of UV-luminous  $z \gtrsim 8$  galaxies relative to many recent theoretical models (e.g., Castellano et al. 2022; Finkelstein et al. 2022, 2023, 2024; Naidu et al. 2022; Adams et al. 2024; Bouwens et al. 2023; Casey et al. 2024; Franco et al. 2023; Harikane et al. 2023; Leung et al. 2023). This population of galaxies may indicate evidence for different star formation and stellar feedback physics in the early Universe, leading to the earlier, more rapid formation of massive and luminous galaxies than previously thought possible (e.g., Dekel et al. 2023; Ferrara et al. 2023; Mason et al. 2023; Pallottini & Ferrara 2023; Shen et al. 2023; Yung et al. 2023a). While models shown tension with direct observables, simulation-inferred UV luminosities are very sensitive to the assumed IMFs, dust, and active galactic nucleus (AGN) contamination. Having reliable estimates for inferred stellar masses provides an alternative way of testing this putative tension with models. Excitingly, early JWST studies have claimed evidence for very massive galaxies (some exceeding  $\log_{10}(M_*/M_\odot) > 11$ ), further indicating the possibility of more efficient star formation activity at higher redshift than expected (e.g., Labb   et al. 2023; Xiao et al. 2023), and/or a revision to Lambda cold dark matter ( $\Lambda$ CDM; Boylan-Kolchin 2023).

To improve our understanding of the abundance of massive (defined here as  $\log_{10}[M_*/M_\odot] > 10$ ) galaxies and how this population evolves over cosmic time, we present a study of massive galaxies selected over the full the 88 arcmin<sup>2</sup> Cosmic Evolution Early Release Survey (CEERS) NIRCam field. We briefly introduce CEERS and the data reduction (Section 2.1) before discussing how the galaxies studied here were selected (Section 3), focusing on sources at  $z > 4$ . In selecting massive galaxies, we also note the interesting prominence of a population of galaxies arising from recent JWST observations: point-like sources with flat or blue rest-frame UV–optical spectra alongside increasing brightness at longer wavelengths, possibly hosting dust-reddened AGN (Barro et al. 2024; Greene et al. 2023; Kocevski et al. 2023; Labbe et al. 2023). We discuss how this category of sources may bias selection of massive galaxies (Section 3.5). We present how the cumulative number density of massive galaxies evolves across cosmic time in Section 4 and compare our findings to both previous studies (Section 5.1) and model predictions (Section 5.2). Finally, we discuss how the observed abundances of galaxies may indicate changes to the efficiency of star formation in galaxies (Section 5.3.1) or how the observed light from the galaxy is translated to mass (Section 5.3.2). Throughout this paper, we assume a cosmology of  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ , and  $\Omega_\Lambda = 0.7$ . When necessary, we assume a baryonic density parameter of  $\Omega_b h^2 = 0.02237$  (Planck Collaboration et al. 2016). We use a Chabrier (2003) IMF, and all magnitudes given are in the AB system (Oke & Gunn 1983).

## 2. Available Data

### 2.1. NIRCam and HST Catalog

We perform our analysis on the publicly released CEERS NIRCam mosaics (Data Releases 0.5 and 0.6, DR0.5 and DR0.6, respectively), which include imaging from 1 to 5  $\mu$ m in F115W, F150W, F200W, F277W, F356W, F410M, and F444W and cover 10 pointings in the Extended Groth Strip (Bagley et al. 2023). The CEERS observations were completed and reduced over two epochs. Epoch 1 NIRCam pointings 1, 2, 3, and 6 were obtained in parallel to the Mid-Infrared Instrument (MIRI) imaging in June 2022. These pointings are reduced with the JWST Calibration Pipeline (Bushouse et al. 2023) version 1.7.2 and Calibration Reference Data System (CRDS) pmap 0989, and are available from CEERS DR0.5. The epoch 2 observations released with CEERS DR0.6 include NIRCam imaging over pointings 4, 5, 7, 8, 9, and 10 obtained in 2022 December in parallel to NIRSpec microshutter assembly observations. Pointings 5, 7, 8, and 9 were also observed with NIRCam Wide Field Slitless Spectroscopy (WFSS) with MIRI imaging in parallel. These epoch 2 images were reduced with Pipeline version 1.8.5 and pmap 1023. The two CEERS NIRCam reductions follow identical procedures, with the only difference being in the Pipeline and CRDS versions. We note that the only NIRCam update between pmmaps 1089 and 1023 is to the distortion reference files, and so this change will not affect the noise properties, flat-fields, or flux calibration between CEERS DR0.5 and DR0.6. The CEERS NIRCam imaging reduction uses the JWST Calibration Pipeline, with custom modifications to correct for features and challenges in the data such as snowballs, wisps, and 1/f noise, and to improve the astrometric alignment. We direct the reader to Bagley et al. (2023) for further details on the reduction.

We make use of the photometric catalog from Finkelstein et al. (2024), which also includes the available HST imaging in this field from CEERS HST Data Release 1, including the Advanced Camera for Surveys F606W and F814W filters and WFC3 F105W, F125W, F140W, and F160W filters from HST programs including CANDELS (Grogin et al. 2011; Koekemoer et al. 2011). This catalog's methodology is similar to Finkelstein et al. (2023), with some updates to improve the accuracy of the color and total flux measurements. Flux uncertainties were calculated empirically via measurements of the flux spread in randomly placed apertures. We direct the reader to Finkelstein et al. (2024) for full details on all cataloging procedures.

### 2.2. MIRI

The CEERS survey includes MIRI imaging conducted over eight pointings, with four blue pointings (P3, P6, P7, and P9) providing deep imaging with F560W and F770W and four with contiguous wavelength coverage in F1000W, F1280W, F1500W, and F1800W (P1, P2, P5, and P8). Two of these pointings (P1 and P2) also include coverage in F770W and F2100W. These observations were performed in two epochs, in 2022 June and 2022 December. Pointings 1, 2, 3, and 6 were observed in epoch 1 as prime observations with parallel NIRCam imaging, and pointings 5, 7, 8, and 9 were observed in parallel to NIRCam WFSS. Details of the MIRI data reduction can be found in Yang et al. (2023).

Source Extractor v2.19.5 (Bertin & Arnouts 1996) was run on the blue pointings using F560W + F770W as detection

images, and using the point-spread function (PSF)-matched F560W and F770W images and their associated rms maps for flux measurements (Papovich et al. 2023). In the red MIRI pointings, photometry extraction was determined using TPhot v2.1 (Merlin et al. 2016) following Yang et al. (2021). TPhot was used for the red pointings instead of Source Extractor due to the large wavelength coverage of these pointings and to mitigate quality loss by PSF matching across filters.

### 2.3. Spectroscopy

CEERS also includes coverage with NIRSpec with the prism and F356W WFSS observations. The NIRSpec prism data processing is explained in detail in P. Arrabal Haro et al. (2024, in preparation). However, the main steps of the reduction follow those employed in Arrabal Haro et al. (2023), Fujimoto et al. (2023), Kocevski et al. (2023), and Larson et al. (2023), and are briefly summarized here. These NIRSpec data were reduced using JWST Calibration Pipeline v1.8.5 and CRDS mapping 1045. Custom parameters were utilized to improve correction of “snowball” events. The resulting count-rate maps are processed through stage two using `calwebb_spec2` pipeline modules, where the images of the three nods are combined to extract the 1D spectra.

The WFSS spectra were extracted following a simulation-based extraction method, which was described in Pirzkal et al. (2017). To summarize, we used the CEERS master mosaic and available photometry to duly simulate each individual WFSS observation and these were used to estimate and correct for spectral contamination in the 2D WFSS wavelength-calibrated spectra, before these were combined and used to produce a pair of independent (GRISM R and GRISM C) 1D spectra using optimal extraction. We used the latest, most up-to-date version of the NIRCAM WFSS calibration products, which were finalized in 2023 August. The estimated accuracy of the field dependence of the calibration is 0.2 pixels while the estimated uncertainty in the final wavelength calibration is 2.5  $\text{\AA}$ .

## 3. Selection of Massive Galaxies

Here, we detail our process of selecting a sample of massive galaxies. We first make use of `eazy` (Brammer et al. 2008) to determine photometric redshifts. We then use the derived redshift probability distribution functions (PDFs;  $\mathcal{P}(z)$ ) as a prior to dense basis (Iyer & Gawiser 2017; Iyer et al. 2019), which we use to perform Bayesian SED fitting to recover physical properties (i.e., stellar mass) and their corresponding posterior distributions.

### 3.1. Photometric Redshift Fitting

Photometric redshifts for the catalog were determined with `eazy`. We use the results from Finkelstein et al. (2024), briefly discussing the procedure here. `eazy` fits nonnegative linear combinations of stellar population synthesis templates through a user-defined grid of redshifts, and at each redshift finds the best-fitting synthetic template spectrum by minimizing  $\chi^2$ . `eazy` derives a  $\mathcal{P}(z)$  based on the quality of fit to the observed photometry. The templates used include the “tweak fspS QSf 12 v3” set of 12 FSPS (Conroy & Gunn 2010) templates recommended by the `eazy` documentation as well as an additional six templates designed to encompass bluer rest-frame UV colors expected from  $z > 9$  galaxies (Larson et al. 2022).

### 3.2. Spectral Energy Distribution Fitting

We perform SED fitting for all sources with signal-to-noise ratio in  $F444W > 3$  with the `dense basis` SED fitting code, which is designed to robustly recover SFH constraints from galaxy SEDs. `dense basis` uses a flexible nonparametric SFH represented by a Gaussian mixture model (Iyer et al. 2019), and stellar templates generated from `FSPS` (Conroy et al. 2009; Conroy & Gunn 2010), including implementation of nebular emission lines using `CLOUDY` (Byler et al. 2017; Ferland et al. 2017). For this work, we define three “shape” parameters that describe the SFH:  $t_{25}$ ,  $t_{50}$ , and  $t_{75}$  (requiring the recovered SFH of the galaxy to form “ $x$ ” fraction of its total mass by time  $t_x$ ). All sources were fit with `dense basis` assuming a Calzetti et al. (2000) dust law and a Chabrier (2003) IMF. We impose a uniform (flat) prior on the specific star formation rate (sSFR) with limits on the sSFR (sSFR  $\text{yr}^{-1} \in [-14, -7]$ ), an exponential prior on the dust attenuation over a wide range of values ( $A_V \in [0, 4]$ ), and a uniform (in log space) prior on the metallicity ( $Z/Z_\odot \in [0.01, 2.0]$ ). For redshift, `dense basis` is only able to take a prior in the form of a top-hat function. Therefore, in order to input the photometric redshift information as recovered by `eazy`, we modify the `eazy`  $\mathcal{P}(z)$  to a top-hat form by first fitting a Gaussian to the primary peak of the  $\mathcal{P}(z)$ , and taking the width of the top-hat function as the  $2\sigma$  width of the Gaussian, if  $\sigma > 1.5$ , or the  $5\sigma$  width of the Gaussian, if  $\sigma \leq 1.5$  (in the case of a very narrowly peaked Gaussian). For each source, we thus have a redshift prior based on the `eazy`-recovered redshift PDF.

While we perform SED fitting over the entire catalog and all available redshifts, as the emphasis of this work is on the evolution at  $z > 4$ , we further vet that higher-redshift sample to ensure accurate results. Therefore, we split our massive sample into two categories: (1) a low-redshift sample of 1796 sources with best-fit redshift  $z < 3.5$  and recovered median stellar masses of  $\log_{10}(M_\star/M_\odot) > 10$ , and (2) a parent high-redshift sample, with a best-fit photometric redshift of  $z > 3.5$  (from `dense basis`), at least 70% of the integrated  $\mathcal{P}(z)$  from `eazy` to be at  $z > 2.5$  (to reject objects with overly broad redshift distributions), and a less stringent mass cut of 2.5% of the posterior in stellar mass having  $\log_{10}(M_\star/M_\odot) > 10$ . The parent high-redshift sample is intentionally chosen to include sources which have median stellar masses below our intended mass limit, to ensure we perform visual inspection on all sources whose posteriors may include a massive ( $\log_{10}(M_\star/M_\odot) > 10$ ) solution to properly encompass the uncertainties from the mass estimates in computing the volume densities of these galaxies. We thus manually inspect 561 sources satisfying these criteria.

### 3.3. Sample Cuts

We remove a total of 297 sources from our sample due to unreliable photometry, e.g., too close to a detector edge, bad pixels, spurious sources from `Source Extractor` (Bertin & Arnouts 2010), and too close to a neighboring bright source; the majority (168/297) of the sources removed have  $z_{\text{phot}} > 8$ , therefore we are confident in their spurious nature. Two particularly high-redshift sources were determined to be incorrectly fit at high redshift (IDs 52092 and 75728 at  $z_{\text{phot}} \sim 15$  and  $z_{\text{phot}} \sim 12$ , respectively). Source ID 52092 has a low- $z$  solution with  $\Delta\chi^2 < 4$  from the high- $z$  fit, and source ID 75728 had significant flux detected below the putative Ly $\alpha$

break at its fitted redshift. Both sources were thus refit with `dense basis` at the low-redshift solution from `eazy`. After refitting, one source was removed from the sample for not satisfying our mass limit.

53 additional sources were “oversplit,” where multiple sources in the photometric catalog appear, based on visual inspection, to be part of a single large galaxy (this is because the `Source Extractor` run was optimized to find high-redshift, faint sources). For these galaxies, we further inspect the segmentation map to determine the neighboring components that are likely to be from the same galaxy. The photometry of each component was combined if the individual best-fitting redshift was within  $\Delta z \leq 0.5$ . The errors on flux are added in quadrature and the updated photometry was then refitted both with `eazy` and `dense basis` as described above. After correcting the photometry, 50/53 sources were still considered high redshift and massive.

This resulted in 261 sources with robust photometry which satisfied our selection. While the mass limit in this initial selection implies that many of these galaxies in fact have median stellar masses estimated to be below our intended mass limit, we draw stellar mass posteriors from this larger sample when measuring the volume densities of massive galaxies (details in Section 4.1), to account for the tails of their mass posteriors, which exceeds our mass limit of  $\log_{10}(M_\star/M_\odot) > 10$ . Figure 1 shows the redshift and stellar mass distributions recovered for the entire catalog. We expect our sample to be highly complete, even at high redshifts. With the typical magnitude of our  $z > 6$  sample at  $\sim 25$  mag in  $F444W$ , we expect  $> 90\%$  completeness based on completeness simulations run for the CEERS survey, as described in Finkelstein et al. (2023).

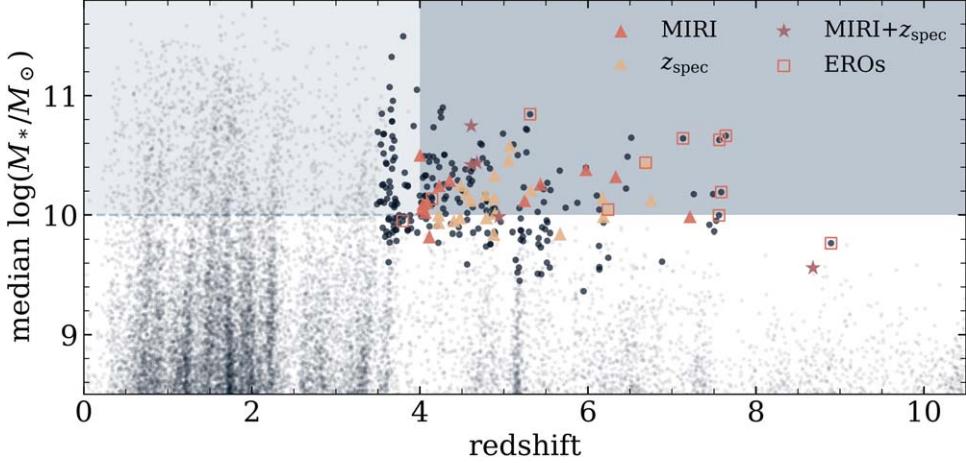
While we expect there to exist similar contaminants and spurious sources in the lower-redshift sample, we elect not to inspect the complete sample as this would be unfeasible for such a high number of sources. We do however inspect a random subsample of the low-redshift sample and found a very low spurious rate; therefore the expected contamination is not expected to contribute significantly to the derived statistics from the sample due to the large number of real sources. Therefore the following discussion primarily focuses on the visually inspected high-redshift subsample of 261 massive galaxies presented here.

### 3.4. Additional Data

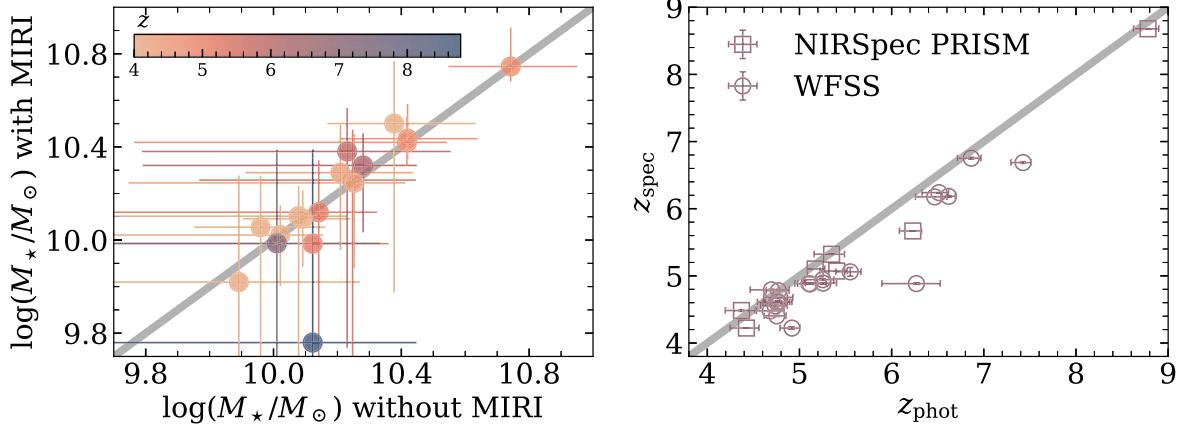
The CEERS field includes coverage by both MIRI imaging (Wright et al. 2023), NIRSpec spectroscopy (Böker et al. 2023), and NIRCam WFSS data (Rieke et al. 2023). NIRSpec allows for the detection of strong [O III] emission out to  $z \sim 8$ , while the CEERS WFSS observations with the F356W filter are sensitive to [O III] redshifts at  $5 < z < 7$ , allowing spectroscopic confirmation for sources with strong emission lines in our sample.

#### 3.4.1. Spectral Energy Distribution Fitting Including MIRI

18 high-redshift massive galaxies fall within the area covered by the CEERS MIRI imaging, extending the wavelength coverage for SED fitting to longer wavelengths; 16 were covered in the deeper blue pointings and two in the red pointings. We perform our SED fitting including MIRI constraints and rederive their physical parameters. We find that the stellar masses largely agree between SED fitting with and without MIRI observations within the errors; however, including the MIRI photometry



**Figure 1.** The distributions in redshift and stellar mass of the full CEERS catalog (Finkelstein et al. 2024), as measured by dense basis (with eazy priors on the dense basis-derived redshift). The darker shaded region denotes our redshift and mass ranges of interest, while the dark points indicate any source which has been visually inspected, the larger of which are sources whose median stellar mass and redshift satisfy our high-redshift mass criteria. We also show galaxies with MIRI coverage (red triangles) and those with spectroscopic confirmation (orange triangles). The galaxies marked with a red box are “extremely red objects” (EROs): F444W point sources with red F277W – F444W colors that are likely hosts to dust-reddened AGN (discussion in Section 3.5). We find 118 galaxies with median stellar mass and redshift values in our region of interest, including objects up to  $z \sim 7.5$ . Notably, we find no sources at  $\log_{10}(M_*/M_\odot) > 11$  or at  $z > 8$ , and the highest-redshift sources we do find are all classified as EROs; thus reliable mass measurements are not possible.



**Figure 2.** Left: estimated stellar masses for galaxies with and without MIRI photometry. With one exception (the Larson et al. 2023 galaxy, which also has spectroscopic confirmation), including MIRI photometric constraints does not significantly alter the recovered stellar masses (average  $\Delta M_* \lesssim 0.1$  dex), though it does reduce the median error by  $\sim 0.05$  dex. Right: comparison of photometric to measured spectroscopic redshifts for sources covered by NIRSpec (five sources) and WFSS (22 sources). Similar to previous analyses of JWST spectroscopy (e.g., Arrabal Haro et al. 2023; Fujimoto et al. 2023), we do observe a slight systematic offset between the photometric redshift estimates and the confirmed spectroscopic redshifts, with all but two galaxies having confirmed spectroscopic redshifts lower than the best-fit photometric estimates. However, we find general good agreement between  $z_{\text{phot}}$  and  $z_{\text{spec}}$ , with only four sources with  $\Delta z > 0.5$ , and no catastrophic outliers. We update our SED fitting with spectroscopic redshifts when available.

slightly decreases the stellar masses of the sample by  $\Delta M_* \sim 0.02$  dex (Figure 2).

One source (the Larson et al. 2023 galaxy) resulted in a decrease in the measured stellar mass of 0.5 dex, similar to the results found by Papovich et al. (2023). The inclusion of these longer-wavelength observations does provide marginally tighter constraints on the estimated stellar masses, reducing the median error bar by  $\sim 0.05$  dex.

### 3.4.2. Spectroscopic Redshifts

For this sample, five sources were observed with NIRSpec with the prism, and one object was also observed with the medium-resolution gratings. For all six sources, strong emission lines were observed, including [O III] and H $\alpha$ , providing clear spectroscopic redshift confirmations. One of these sources was previously presented in Larson et al. (2023), thus we use their measurements (the source with medium-

resolution grating observations). For the remaining five we measure spectroscopic redshifts by fitting Gaussian functions to the observed lines, converting to redshifts using the relevant vacuum rest-frame wavelength.

82 sources were covered by F356W WFSS observations. Of these 82 sources, 22 have clearly identifiable emission lines in both the row and column dispersions. We inspected each of the objects by eye and identified objects with obvious emission lines present in both the row and column dispersion data. We conservatively require detections in both observations to ensure no spurious lines were included; several sources contain partial observations in only one dispersion direction, thus future work may increase the yield of spectroscopic observations. We used a combination of the photometric redshifts and the presence or lack of a clear line doublet to classify the strong emission lines as the H $\alpha$  or [O III]  $\lambda\lambda 5007, 4959$  lines. For the H $\alpha$  objects, we fit a single Gaussian function with a local continuum offset to

**Table 1**  
Properties of the Massive Galaxy Sample

ID	R.A. (deg)	Decl. (deg)	Flux F277W ( $m_{AB}$ )	Stellar Mass $\log_{10}(M_*/M_\odot)$	$z_{\text{phot}}$	$z_{\text{spec}}$	MIRI
CEERS-42006	214.73721	52.72742	22.9	$10.90^{+0.25}_{-0.33}$	$4.27^{+0.25}_{-0.23}$	...	...
CEERS-93955	214.77138	52.74975	25.0	$10.86^{+0.32}_{-0.19}$	$4.25^{+0.50}_{-0.48}$	...	...
CEERS-37979 <sup>a</sup>	214.79537	52.78885	24.7	$10.84^{+0.10}_{-0.60}$	$5.31^{+0.10}_{-0.10}$	...	...
CEERS-45649	214.89649	52.87087	22.9	$10.82^{+0.16}_{-0.50}$	$4.25^{+0.16}_{-0.14}$	...	...
CEERS-80676	215.04413	52.89874	23.1	$10.75^{+0.17}_{-0.07}$	$4.77^{+0.10}_{-0.17}$	$4.611^{\text{c}}$	yes
CEERS-8028	215.03905	53.00278	23.9	$10.73^{+0.24}_{-0.19}$	$4.16^{+0.26}_{-0.30}$	...	...
CEERS-10933	214.90311	52.94573	23.5	$10.72^{+0.24}_{-0.43}$	$5.26^{+0.15}_{-0.17}$	...	...
CEERS-2167	215.01122	53.01374	23.6	$10.71^{+0.29}_{-0.78}$	$5.27^{+0.13}_{-0.17}$	...	...
CEERS-13082	214.91381	52.94294	23.6	$10.68^{+0.20}_{-0.38}$	$4.99^{+0.13}_{-0.17}$	...	...
CEERS-9317 <sup>a</sup>	214.98304	52.95601	25.7	$10.66^{+0.25}_{-0.22}$	$7.64^{+0.08}_{-0.17}$	...	...
CEERS-11949	214.89190	52.93386	25.1	$10.65^{+0.29}_{-0.29}$	$6.52^{+0.34}_{-0.55}$	...	...
CEERS-57442	214.88680	52.85538	26.0	$10.64^{+0.29}_{-0.31}$	$7.13^{+0.24}_{-0.36}$	...	...
CEERS-83784 <sup>a</sup>	215.08030	52.90790	25.5	$10.63^{+0.23}_{-0.29}$	$7.56^{+0.11}_{-0.15}$	...	...
CEERS-8825	215.01603	52.98242	23.5	$10.60^{+0.18}_{-0.35}$	$4.72^{+0.13}_{-0.16}$	...	...
CEERS-87742	214.94900	52.85175	24.6	$10.59^{+0.21}_{-0.24}$	$4.74^{+0.15}_{-0.22}$	...	...
CEERS-65868	215.08604	52.95223	23.5	$10.58^{+0.11}_{-0.89}$	$5.40^{+0.11}_{-0.18}$	$5.072^{\text{b}}$	...
CEERS-100439	214.86580	52.77020	23.9	$10.52^{+0.21}_{-0.22}$	$5.27^{+0.14}_{-0.19}$	...	...
CEERS-12787	214.85316	52.90160	23.0	$10.50^{+0.22}_{-0.77}$	$4.29^{+0.14}_{-0.18}$	...	...
CEERS-40706	214.76108	52.75069	24.1	$10.50^{+0.27}_{-0.73}$	$4.00^{+0.18}_{-0.19}$	...	yes
CEERS-26578	214.79037	52.84192	24.0	$10.49^{+0.24}_{-0.19}$	$4.11^{+0.22}_{-0.24}$	...	...
CEERS-78830	214.97479	52.86054	23.9	$10.49^{+0.25}_{-0.21}$	$4.05^{+0.25}_{-0.31}$	...	...
CEERS-24868	214.83842	52.88518	24.8	$10.49^{+0.23}_{-0.26}$	$6.47^{+0.15}_{-0.19}$	...	...
CEERS-95076	214.83275	52.78136	25.4	$10.49^{+0.30}_{-0.23}$	$4.16^{+0.57}_{-0.36}$	...	...
CEERS-6078	215.00289	52.98767	23.6	$10.47^{+0.15}_{-0.68}$	$4.75^{+0.14}_{-0.14}$	...	...
CEERS-68303	215.08717	52.94154	23.8	$10.46^{+0.13}_{-0.14}$	$5.55^{+0.12}_{-0.16}$	$5.060^{\text{c}}$	...
CEERS-4851	214.97421	52.97388	23.2	$10.45^{+0.16}_{-0.82}$	$4.20^{+0.14}_{-0.18}$	...	...
CEERS-14256	214.90682	52.93192	23.0	$10.44^{+0.24}_{-0.76}$	$4.11^{+0.12}_{-0.20}$	...	...
CEERS-79906	214.97836	52.85670	23.9	$10.44^{+0.24}_{-0.19}$	$4.64^{+0.14}_{-0.21}$	...	...
CEERS-44057 <sup>a</sup>	214.89224	52.87741	25.9	$10.44^{+0.20}_{-0.27}$	$7.43^{+0.09}_{-0.13}$	$6.687^{\text{c}}$	...
CEERS-84799	214.94344	52.86410	24.1	$10.44^{+0.15}_{-0.09}$	$4.79^{+0.14}_{-0.17}$	$4.680^{\text{c}}$	yes
CEERS-98459	214.85289	52.77394	24.3	$10.43^{+0.25}_{-0.25}$	$4.80^{+0.21}_{-0.36}$	...	...
CEERS-31016	214.80816	52.83222	24.6	$10.43^{+0.19}_{-0.19}$	$4.38^{+0.17}_{-0.24}$	...	...
CEERS-78348	215.04260	52.91196	23.6	$10.42^{+0.11}_{-0.10}$	$4.77^{+0.15}_{-0.14}$	$4.611^{\text{c}}$	yes
CEERS-73986	214.97734	52.88961	23.5	$10.40^{+0.14}_{-0.78}$	$4.23^{+0.16}_{-0.18}$	...	...
CEERS-80453	215.00767	52.87410	24.7	$10.40^{+0.24}_{-0.26}$	$5.98^{+0.27}_{-0.28}$	...	...
CEERS-19333	214.94405	52.92974	23.8	$10.39^{+0.22}_{-0.35}$	$4.63^{+0.18}_{-0.23}$	...	...
CEERS-75554	214.98134	52.88257	24.5	$10.39^{+0.24}_{-0.23}$	$4.83^{+0.21}_{-0.29}$	...	...
CEERS-13769	214.91046	52.93917	23.7	$10.38^{+0.16}_{-0.73}$	$4.41^{+0.14}_{-0.17}$	...	...
CEERS-9235	215.03880	52.99599	23.6	$10.38^{+0.16}_{-0.78}$	$4.54^{+0.14}_{-0.17}$	...	...

**Notes.** Table reporting the redshifts and stellar masses of  $z > 4$  massive galaxy candidates, sorted by mass. Sources with auxiliary data are marked accordingly, the full table can be found in the electronic version of this paper.

<sup>a</sup> EROs.

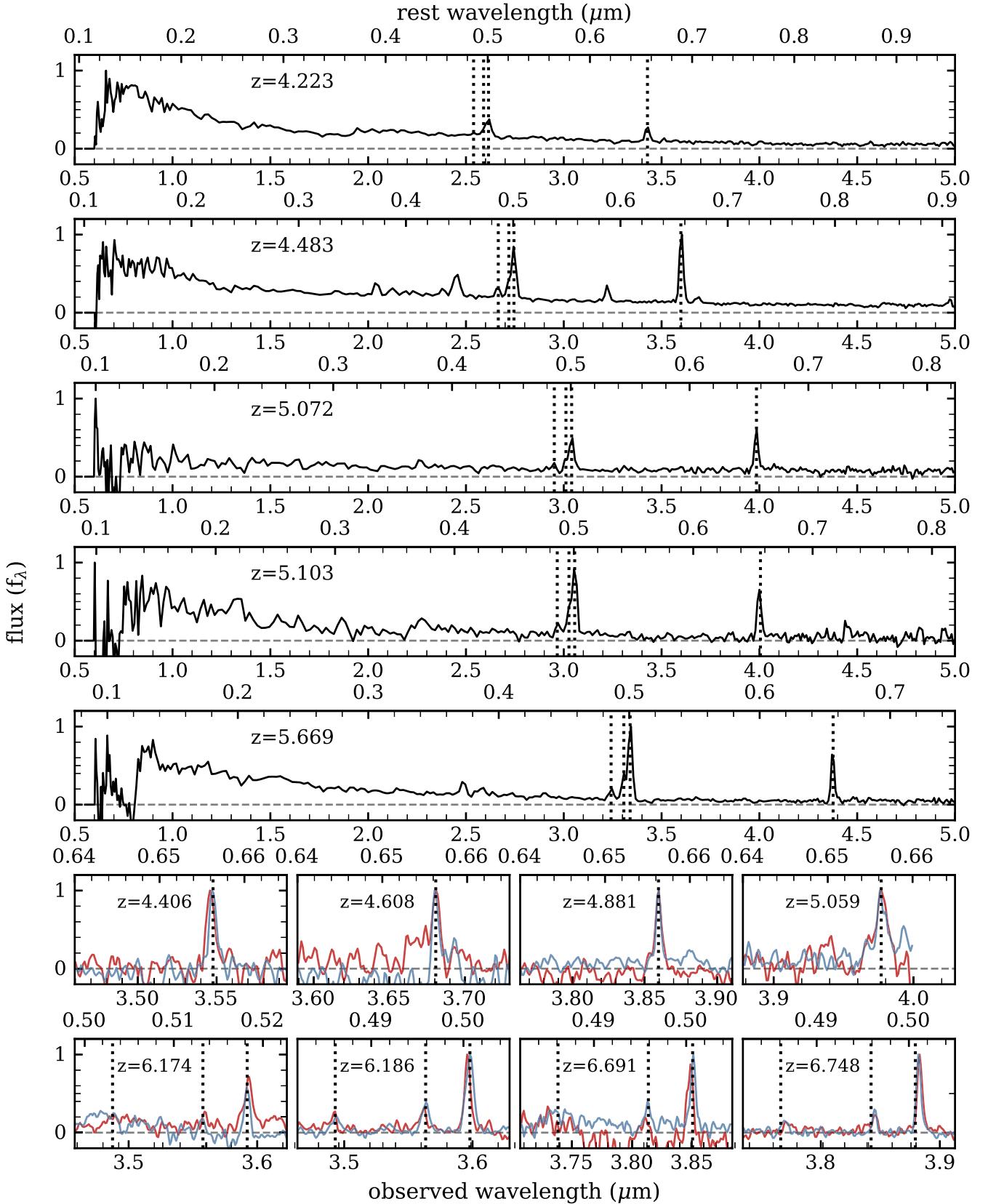
<sup>b</sup> NIRSpec spectroscopic redshift.

<sup>c</sup> NIRCam WFSS spectroscopic redshift.

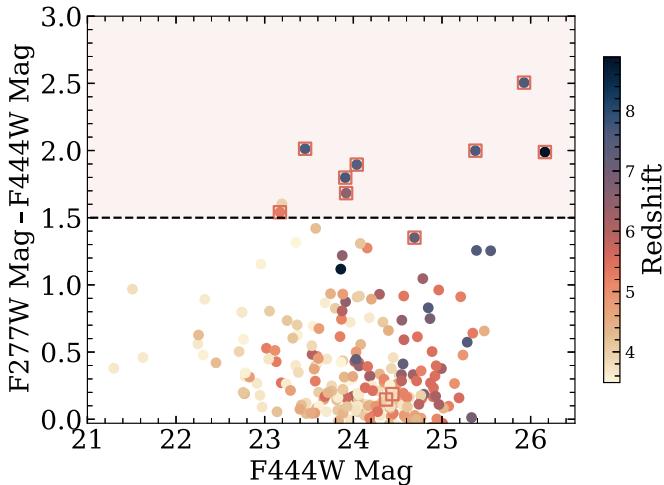
(This table is available in its entirety in machine-readable form in the [online article](#).)

the emission line, and for the [O III] objects we fit two Gaussians with a fixed 2.98:1 amplitude ratio, a fixed line-center ratio of 5006.843:4958.911, a common Gaussian width, and a local continuum offset. Using a Monte Carlo technique, we fit the appropriate model to each object 1000 times, and varied the flux data of the object spectrum using the spectral error data between each run. We use the median line centers and their standard deviations from these Monte Carlo runs to determine spectroscopic redshifts with uncertainties for each object. We report these values in Table 1.

For the majority of galaxies with spectroscopic coverage, we find remarkable agreement between the photometric and spectroscopic redshifts, with four sources with  $\Delta z > 0.5$  and only one source with  $\Delta z > 1$  between the photometric and spectroscopic redshifts (Figure 2, right panel). Notably, we find no catastrophic outliers, with all galaxies remaining in our sample with their spectroscopic redshifts, providing confidence in the validity of the full sample of massive galaxies. We show observed spectra for the five NIRSpec prism sources and eight example NIRCam WFSS detections in Figure 3, all showing



**Figure 3.** Examples of spectroscopic observations of sources within our sample of  $\log_{10}(M_*/M_\odot) > 10$  galaxies. The top five panels show spectra for the five sources with NIRSpec PRISM observations (one additional source, presented in Larson et al. 2023, has a NIRSpec grating redshift). Each galaxy shows clearly detected [O III] doublet + H $\beta$  and H $\alpha$  emission at  $\lambda_{\text{rest}} = 5007 \text{ \AA}$  and  $6563 \text{ \AA}$ , respectively. The bottom eight panels show line detections made with the GRISM, with the top row showing H $\alpha$  and the bottom row shows the [O III] doublet + H $\beta$ . The red (blue) shows the row (column) dispersion. Here we smooth the GRISM data using a Gaussian filter with a width of one spectral element for clarity. 28 sources have clear emission lines providing robust spectroscopic redshift confirmation of many of the sources in our sample.



**Figure 4.** Color–color space of EROs (Barro et al. 2024) colored by redshift, where sources with red squares are selected as ERO based on visual inspection. These sources are bright and point-source-like in F444W, exhibiting blue rest-frame UV colors with redder rest-frame optical colors. Due to the brightness of these sources at  $\lambda_{\text{observed}} > 3 \mu\text{m}$ , SED fits generally recover large stellar masses. However, these galaxies have photometric colors consistent (and degenerate) with different combinations of dust-obscured massive star-forming galaxies, less massive non-dust-obscured galaxies, and dust-reddened AGN. Due to these degeneracies, the estimated stellar masses can vary significantly, up to  $\sim 3$  dex, depending on the assumed galaxy components. Therefore, we remove these sources from our fiducial analysis; however, we still show cumulative number densities found inclusive of these objects for comparison (Figure 5).

clear emission line detections. We again perform our SED fitting with this updated redshift information.

### 3.5. Extremely Red Objects

One of the early JWST discoveries is the prominence of a new class of sources: high-redshift point-like sources that are blue or flat in the short-wavelength channels in JWST but show extreme reddening toward longer wavelengths. Barro et al. (2024) analyzed these sources and found that the most likely explanation of the photometric colors are various combinations of dust-reddened AGN with bluer stellar components. Indeed, early JWST observations seem to indicate that accreting supermassive black holes are relatively common at  $z > 5$  (e.g., Furtak et al. 2023; Harikane et al. 2023; Juodžbalis et al. 2023; Kocevski et al. 2023; Labbe et al. 2023; Larson et al. 2023; Leung et al. 2023; Maiolino et al. 2023; Matthee et al. 2023; Scholtz et al. 2023). This interpretation is supported by a number of such sources with confirmed broad-line AGN (Greene et al. 2023; Kocevski et al. 2023), though weaker than expected long-wavelength MIRI fluxes may indicate this population also includes non-AGN (Williams et al. 2023).

With only photometric colors available, it is extremely difficult to accurately determine the light contributed by the AGN component of these galaxies, making photometric stellar mass estimates for these sources extremely uncertain (Barro et al. 2024; Kocevski et al. 2023). These degeneracies mean that estimated stellar masses may vary significantly, up to  $\sim 3$  dex, depending on the assumed galaxy components (Barro et al. 2024). These uncertainties in stellar mass would bias our results in determining the cumulative number densities of massive galaxies observed; therefore we remove these sources from our fiducial sample via visual inspection for point-like sources in F444W. We show the color space of our sources in

Figure 4 in comparison to the color selection presented by Barro et al. (2024), who have found that a straightforward color cut of  $F277W - F444W > 1.5$  can effectively identify sources dominated by potential AGN emission. We find nine sources which satisfy this color cut, many of which are fit as higher-redshift sources ( $z > 7$ ). We find that all but one source is a point source in F444W and remove it from our ERO sample. We further add into our sample three point-like sources not identified from the color–color cut.

With this in mind, we perform the following analysis on two samples, one inclusive of all sources determined to have stellar masses of  $> 10^{10} M_{\odot}$ , and the same sample, but with the red sources described here removed.

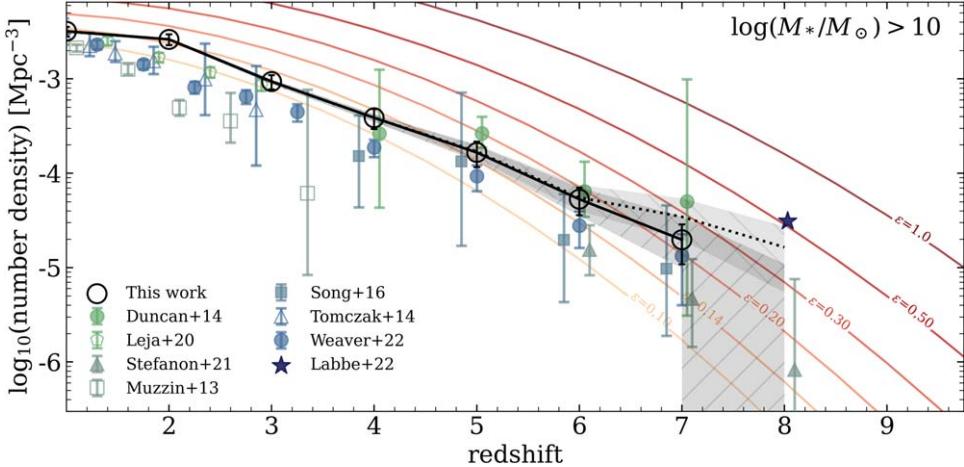
In summation, we use SED fitting to identify galaxies with  $\log_{10}(M_{\star}/M_{\odot}) > 10$  at  $z = 1\text{--}9$  in the CEERS survey. We visually inspect sources in the high-redshift sample; 28 of these galaxies have spectroscopic redshifts. We also identify 11 extremely reddened point-like objects in F444W and explore the implications of the following results both with and without these EROs included in the sample.

## 4. Results

### 4.1. Cumulative Number Densities of Massive Galaxies

To determine the cumulative number densities of massive galaxies across cosmic time, we separate our sample into redshift bins of  $\Delta z = 1$ . At  $z < 4$ , we take the median fitted physical parameters for each source recovered from dense basis. Above  $z > 4$  we draw from the recovered posterior distribution for each individual (visually inspected) source. For each source, we conduct a draw from the posterior SED to obtain a stellar mass and redshift (stellar mass and redshift sampled as a pair as they are covariant), and assign the galaxy to a redshift bin if the drawn stellar mass is greater than our intended mass limit ( $\log_{10}(M_{\star}/M_{\odot}) > 10$ ). We perform this sampling over 500 draws to recover the number of massive galaxies across the redshift bins. For each draw, we further sample from the Poisson distribution of the number of galaxies in each redshift bin. Combined, these provide uncertainties in the number densities inclusive of the spread in redshift, stellar mass, and Poisson uncertainties. To account for cosmic variance, or the variance that arises from the effects of large-scale structure, we use the approach presented in Yung et al. (2023b) and calculate the variance from 2500 random samples of a  $6' \times 14.7'$  elongated field from five of the  $2 \text{ deg}^2$  light cones presented in that work. In Figure 5, we report the median as the number densities and show the 68% confidence limit as the shaded region, where the cosmic variance is shown as error bars.

We show in Figure 5 the cumulative number densities of  $\log_{10}(M_{\star}/M_{\odot}) > 10$  galaxies as recovered from this analysis in both our entire galaxy sample, and the same sample with red point sources removed (Section 3.5) in the CEERS field, as a function of redshift. For comparison, we also plot curves of cumulative number densities of  $\log_{10}(M_{\star}/M_{\odot}) > 10$  galaxies given various values of constant baryon conversion efficiency, or the rate of conversion from baryonic to stellar mass using halo mass functions from the Bolshoi–Planck and MultiDark–Planck  $\Lambda$ CDM cosmological simulations (Rodríguez-Puebla et al. 2016). Assuming a cosmic baryon fraction  $f_{\text{baryon}} = 0.16$ , which is consistent with the value established by observations of the cosmic microwave background (Spergel et al. 2003;



**Figure 5.** The redshift evolution of massive ( $\log_{10}(M_*/M_\odot) > 10$ ) galaxies in CEERS. The black solid (dotted) line shows the volume density for  $\log_{10}(M_*/M_\odot) > 10$  galaxies exclusive (inclusive) of EROs, with the darker (lighter) shaded region showing the 68% confidence interval (CI) inclusive of sampling the redshift and stellar mass posteriors as well as the Poisson error; the error bars show the error from cosmic variance. We show a variety of pre-JWST results (green–blue points) and a comparison to the massive galaxies presented in Labb   et al. (2023), all converted to a Chabrier IMF. The orange-red curves show the expected cumulative number densities above our mass threshold for a given baryon conversion efficiency ( $\epsilon$ ) calculated using the Rodr  guez-Puebla et al. (2016) halo mass function. Our results are systematically higher than previous studies though often within the uncertainties, while following a similar evolution to higher redshifts (details in Section 5). We find that the volume density of massive galaxies follows that expected given a constant baryon conversion efficiency of  $\epsilon \sim 0.14$  up to  $z \sim 5$ , before which the volume densities begin to exceed the predicted values.

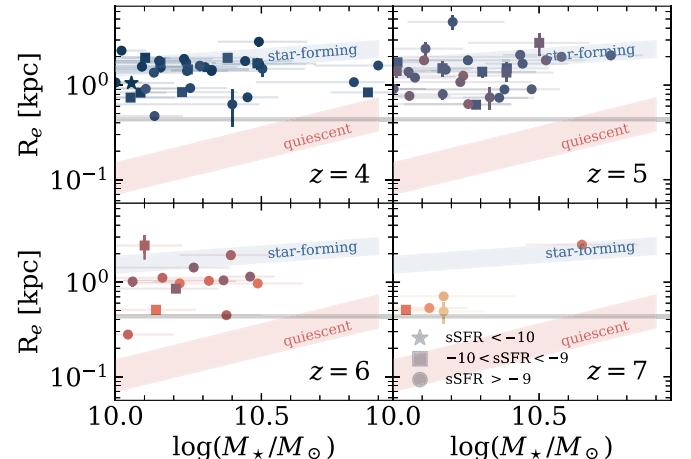
Jarosik et al. 2011), we convert the underlying halo mass function to an SMF. We then integrate the converted SMF of  $\log_{10}(M_*/M_\odot) > 10$  sources assuming varying values of the efficiency,  $\epsilon$ .

In CEERS, we see that above a redshift of  $z \sim 4$ , while the number density of massive galaxies continues to decrease with increasing redshift, there is an increased abundance over that expected with a constant baryon conversion efficiency as we progress to higher redshifts. This increased abundance could be explained by a difference in star formation physics in the early Universe, resulting in an increased global baryon conversion efficiency and a nonstandard galaxy IMF, discussed in Section 5.

#### 4.2. Sizes of Massive Galaxies

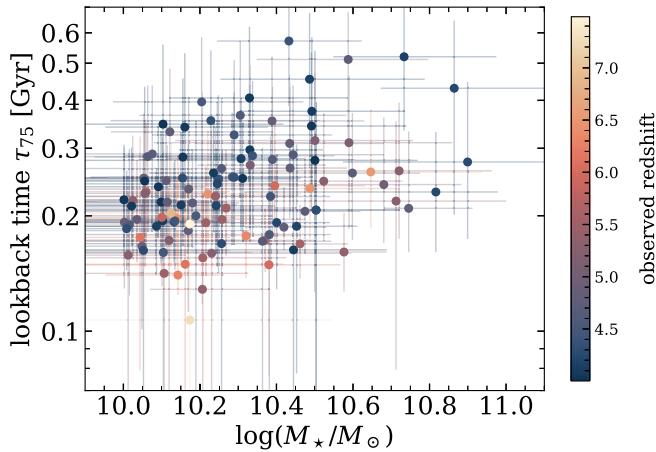
In Figure 6, we present the size–mass distribution of our massive sample from  $4 < z < 8$ . Sizes were measured with GALFIT and are effective semimajor axes measured at rest-frame  $0.5 \mu\text{m}$ . Details of the GALFIT fitting will be provided in a future paper (E. J. McGrath et al. 2024, in preparation). Briefly, sources were fit using empirical PSFs (Finkelstein et al. 2024), and the ERR (error) array was used as the input sigma image, which includes noise elements from the detector, sky, and source. Neighboring galaxies were either fit simultaneously or masked during fitting, depending on brightness and distance from the primary source. Only galaxies whose best-fit models yielded magnitudes consistent with our photometry and which did not reach a constraint limit are shown (83/118 sources). We also show the best-fit size–mass relations from van der Wel et al. (2014) extrapolated to this redshift range for comparison.

The slope in the size–mass plane for our massive galaxies sample is consistent with the results for star-forming galaxies at lower redshift, but when we extrapolate the evolving intercept of the size–mass relation from lower redshift to  $z > 4$ , we find that galaxies at  $4 < z < 8$  are smaller on average by a factor of 1.5–2 than predicted from previous work (van der Wel et al. 2014). This is consistent with the results of Ormerod et al. (2023) and Ward et al. (2023), who also found evolution



**Figure 6.** The size–mass distribution of  $z > 4$  massive galaxies presented here. We show the effective semimajor axes measured at rest-frame  $0.5 \mu\text{m}$  (measured in F277W at  $z = 4$  and F444W at  $z = 7$ ). The points are colored by redshift and the shape indicates the 10 Myr averaged sSFR values from dense basis. The best-fit size–mass relations for quiescent and star-forming galaxies from van der Wel et al. (2014) extrapolated to  $z = 4–8$  are also shown for comparison. The gray line shows the size range of a point source from  $z = 4$  to 7. We find that our sample is primarily star forming and the slope of our sample in the size–mass plane is consistent with the results for star-forming galaxies at lower redshifts, however, they are smaller by a factor of  $\sim 2$  than predicted when the relation is extrapolated out to higher redshifts.

toward smaller sizes at higher redshifts. Furthermore, we see evidence for continued redshift evolution, with  $z = 6–7$  galaxies smaller at fixed mass than  $z = 4–5$  galaxies by a factor of 1.67, implying that the processes responsible for driving this evolution are already in place at very early times in the Universe. Evolution toward smaller sizes at higher redshifts would imply that the gas densities in these galaxies should also evolve to higher values at higher redshifts. Simulations predict that higher gas densities should give rise to a greater fraction of baryons converted into stars (Fukushima & Yajima 2021). Thus our observation of smaller sizes at higher redshift is



**Figure 7.** The estimated formation time of 75% of the stellar mass (in terms of lookback time,  $\tau_{75}$ ) for each galaxy in our massive, high-redshift sample. The points are colored by redshift, where we bin galaxies into bins of stellar mass, with each bin having an equal number of sources. The open black squares show the median of each bin, with the error bars denoting the 18% and 84% spread. There is a slight trend with  $\tau_{75}$  with respect to the observed stellar mass, with galaxies of larger mass forming the bulk of their stars at an earlier time, but this is not significant with the constraints on these quantities from the current data. Due to the difficulty constraining SFHs with only photometry, we note there exist large errors associated with  $\tau_{75}$ . SFHs are often difficult to constrain based on photometry alone, with follow-up spectroscopy often necessary to properly disentangle the underlying stellar populations. Furthermore, at very high redshifts, our sample is small (of order a few galaxies at  $z > 7$ ); larger surveys will be able to provide the sample sizes needed to properly infer trends with redshift.

consistent with an evolution in the efficiency of gas conversion, as discussed in the previous subsection.

#### 4.3. Formation Timescales

SFHs, or the SFR of a galaxy across its lifetime, probes galaxy formation and evolution. By studying the SFHs of massive galaxies in particular, we can attempt to understand how these galaxies manage to form surprisingly large amounts of stellar mass over relatively short timescales, and provide valuable insights into the early stages of galaxy formation and crucial benchmarks for galaxy formation theory.

From dense basis, we estimate the times at which each galaxy formed 25%, 50%, and 75% of its total stellar mass at the time of observation and the associated spread in each timescale. Figure 7 shows the distribution of  $\tau_{75}$  (the difference between the time of observation compared to the time when the galaxies formed 75% of their stellar mass) for our sample with median  $\log_{10}(M_*/M_\odot) > 10$  and  $z > 4$  with EROs removed (118 sources). We separate our sample into five bins of an equal number of galaxies across stellar mass and show the median  $\tau_{75}$  and 68% spread in each bin. We notice that there may be a slight dependence on  $\tau_{75}$  at which 75% of the observed stellar mass formed with respect to current observed stellar mass, similar to the results found by Estrada-Carpenter et al. (2020); however the spread and errors are large, making this relationship difficult to robustly constrain.

We show some example fitted SFHs in Figure 8. With dense basis, we recover a diversity of SFH shapes, including periods of bursts and rising/falling SFHs. We note that the posterior SFHs often have a large spread in the recovered shapes, with some sources exhibiting unconstrained SFHs, and others which are more constrained. Without spectral information, it is difficult to disentangle the emission/absorption lines from the continuum

for each galaxy; therefore the exact stellar populations also become difficult to fully constrain, particularly for older stellar populations which are often outshined by younger more massive stars (Conroy 2013; Giménez-Arteaga et al. 2023; Papovich et al. 2023; Song et al. 2023; though as shown in Figure 2 we do not find evidence that the stellar masses of our objects are underestimated when MIRI photometry is not available). While recent bursts of star formation activity can be indicated by photometry, spectroscopic observations are necessary to robustly constrain the full mass assembly histories of these galaxies (Estrada-Carpenter et al. 2020, 2020). Furthermore, larger samples of massive galaxies may be able to recover more robust trends with star formation timescales. While only with available photometry, the SFHs are difficult to constrain, follow-up detailed spectroscopic observations may help disentangle possible SFH models that occur in the early Universe, enabling further understanding at the timescales and the onset of star formation of massive galaxies.

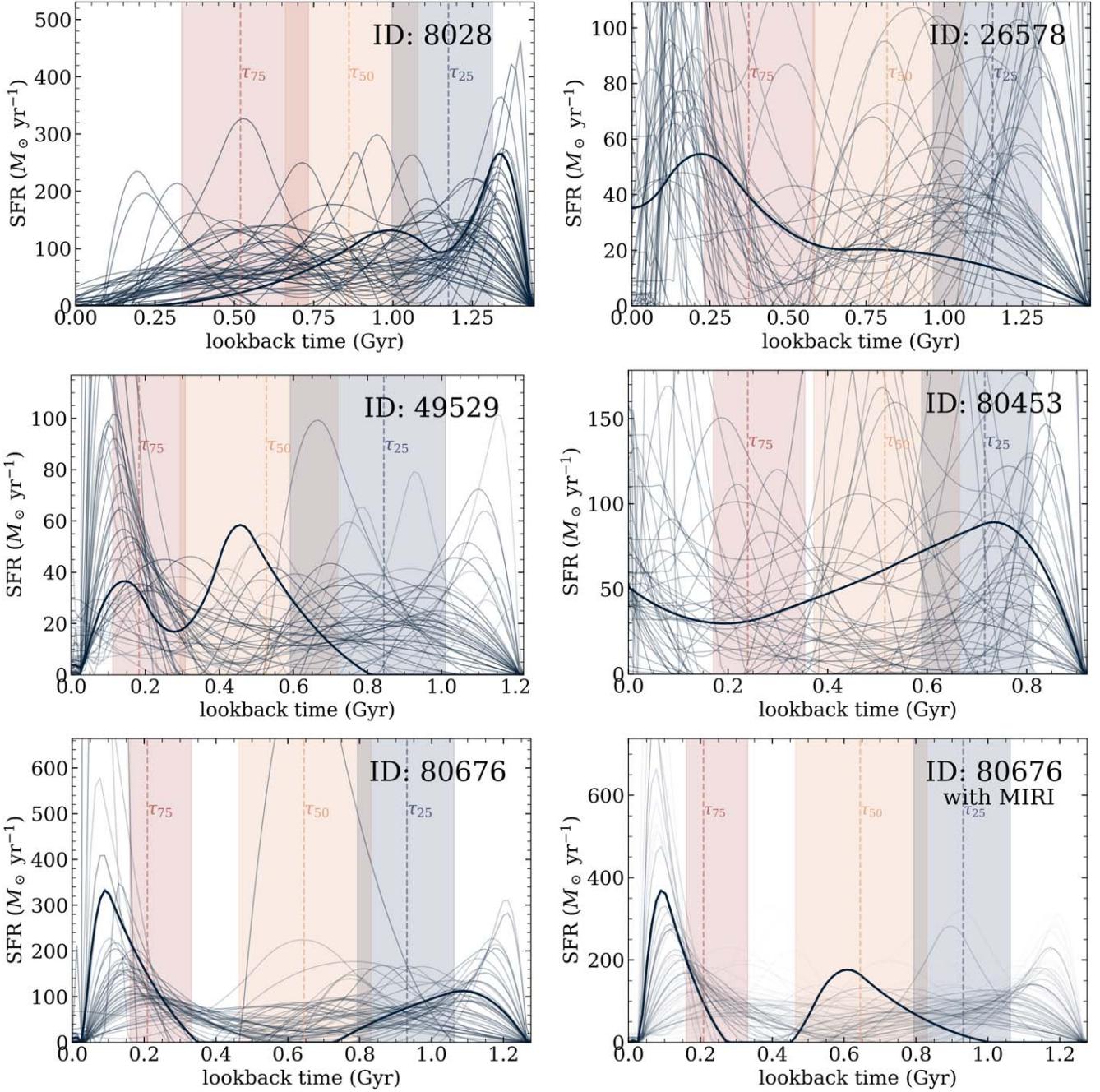
## 5. Discussion

### 5.1. Comparison to Literature Results

Here we compare our measured cumulative number densities to those inferred from previously published SMFs, converting all results to a Chabrier IMF where needed (Figure 5). At lower redshifts, ( $z < 4$ ), we compare to results from Tomczak et al. (2014), who surveyed an area of 316 arcmin<sup>2</sup> distributed over three independent fields, using observations from the FourStar Galaxy Evolution Survey. Our results show slightly higher number densities that agree within the quoted uncertainties. We also show results from the  $K_s$ -selected catalog of the 1.62 deg<sup>2</sup> COSMOS/UltraVISTA field from  $0.2 < z < 4.0$  from Muzzin et al. (2013). Our results show higher number densities; however, we note that beyond  $z \geq 2.0$ , the analysis from Muzzin et al. (2013) is limited to galaxies with  $\log_{10}(M_*/M_\odot) > 10.51$ , reducing its measured number densities. We also show recent work from Leja et al. (2020), who compute an updated SMF from SED fitting with nonparametric SFHs of  $\sim 10^5$  galaxies in the 3D-HST and COSMOS-2015 surveys to infer updated stellar masses, and construct a continuity model that directly fits for the redshift evolution of the SMF.

At higher redshifts, our results are overall higher than the SMFs from Great Observatories Origins Deep Survey Reionization Era Wide-area Treasury from Spitzer program (Stefanon et al. 2021). We are within the error of the SMF presented by Song et al. (2016), calculated from a rest-frame UV-selected sample of  $\sim 4500$  galaxies, over the  $\sim 280$  arcmin<sup>2</sup> CANDELS/GOODS HUDF fields, and Duncan et al. (2014) who measure the SMF in the CANDELS GOODS-South field from  $4 < z < 7$ . Weaver et al. (2022a) studied the galaxy SMF from redshifts  $0.2 < z \leq 7.5$ , leveraging the 1.27 deg<sup>2</sup> coverage of the COSMOS2020 catalog (Weaver et al. 2022b). We find overall higher number densities across all redshifts, while following the same general trend as we push to higher redshifts.

While the resulting cumulative number densities in this work are generally higher than in previous works, our results make use of deeper data covering both the rest-frame UV and optical at these redshifts. The improvement in angular resolution compared to Spitzer/IRAC largely corrects the source blending seen with IRAC, resulting in higher completeness. We note that in particular, the number density inclusive of EROs we find here (if extrapolated) agrees within the uncertainties with



**Figure 8.** This figure set contains example plots of the recovered SFHs from dense basis of  $\log(M_*/M_\odot) > 10$  and  $z > 4$  galaxies. The blue (orange, red) panels show the time at which the galaxy formed 25% (50%, 75%) of the observed stellar mass. The starting time is defined as the start of the Universe. The black line shows the best-fitting SFH, and the thinner blue lines show other draws for the SFH; the opacity scales with the  $\chi^2$  of each fit, where darker represent a lower  $\chi^2$  (better fit). The bottom two panels show the added constraints that including MIRI observations provide. The SFHs of high- $z$  massive galaxies are varied, and in certain cases, difficult to tightly constrain when determined from photometry alone.

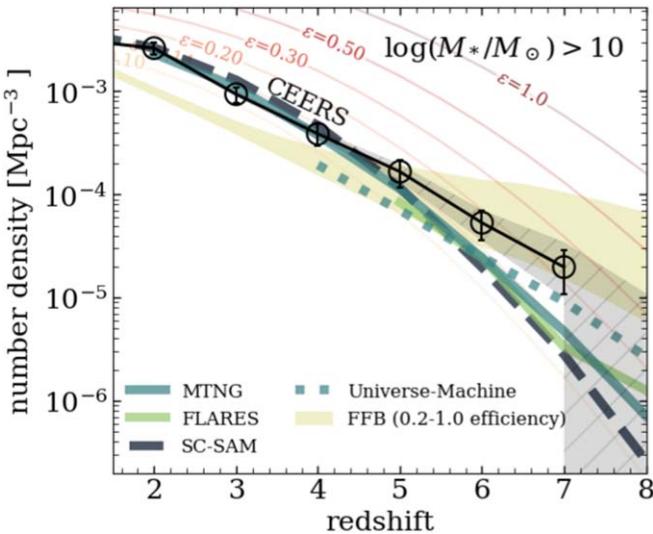
previous studies of massive galaxies in the CEERS field by Labb   et al. (2023).

### 5.2. Comparisons to Models

We show in Figure 9 comparisons between the number densities of massive galaxies in CEERS and a variety of models, including simulations Millennium-TNG (MTNG; Nelson et al. 2021; Kannan et al. 2023; Pakmor et al. 2023) and FLARES (Lovell et al. 2021; Wilkins et al. 2023), the empirical model Universe Machine (Behroozi et al. 2019), semianalytic models (SAMs) including the Santa Cruz SAM

(SC-SAM; Somerville et al. 2015; Yung et al. 2019a, 2019b) and DRAGONS (Mutch et al. 2016; Poole et al. 2016), which combines the hydrodynamical model Illustris-TNG (Pillepich et al. 2017) with the dark matter-only Millennium simulation (Lemson & the Virgo Consortium 2006) through semianalytic galaxy formation models, and the Feedback-Free Starbursts (FFB) model (Dekel et al. 2023).

While the models agree well at low redshift, they tend to diverge at higher redshifts due to a combination of effects from physical assumptions and the volume probed. Below a redshift of  $z \sim 5$  (i.e.,  $z < 5$ ), our results agree well with SC-SAM,



**Figure 9.** Comparison of the number density of  $\log_{10}(M_*/M_\odot) > 10$  galaxies found in CEERS vs. models. Here we show the sampled number densities with EROs removed (details in Section 3.5), the gray shaded region in our 95% CI. We note that at lower redshifts ( $z < 4$ ), the models seem to agree well; however, they diverge at higher redshifts. Overall, we find good agreement with DRAGONS, MTNG, and SC-SAM at  $z < 4.5$ . At  $z > 5$  we find a higher median number density than predicted by all models shown aside from FFB, where our results are best fit by the FFB predictions with  $\epsilon_{\max} = 0.3$  at  $z > 6$ .

MTNG, and DRAGONS; however we find higher abundances of massive galaxies than that predicted by FFB, which is made by construction to converge with the Universe Machine at low redshifts. At  $z \gtrsim 5$ , we find a higher median cumulative number density than almost all of the models that we compared with except for FFB, which argues for high star formation efficiency due to FFB in massive galaxies at cosmic dawn, where the gas density in star-forming clouds is above a critical value (Dekel et al. 2023). The shaded area corresponds to a range of maximum efficiency from  $\epsilon_{\max} = 0.2$  (bottom) to 1.0 (top). Our results at  $z > 6$  most closely follow the FFB model with  $\epsilon_{\max} \sim 0.3$ .

### 5.3. Evolution of Massive Galaxies

JWST is revealing a high abundance of bright galaxies in the early Universe (e.g., Castellano et al. 2022; Naidu et al. 2022; Adams et al. 2024; Bouwens et al. 2023; Casey et al. 2024; Finkelstein et al. 2023; Franco et al. 2023; Harikane et al. 2023; Leung et al. 2023). These observations may be explained by changing the physical processes that govern the relationship between observed galaxy light and the host dark matter halo mass. In this analysis we find a similar trend of a high abundance of massive galaxies in the early Universe. Here, we discuss two possible physical processes that can explain this excess: (1) a higher baryon conversion efficiency at high redshifts and (2) a changing mass-to-light ratio, where the assumptions made in converting the observed light to stellar mass grounded in the local Universe no longer hold at high redshifts. Other physical processes may also be invoked to explain excesses in UV-luminous and massive galaxies, such as modifying cosmology, varying dust attenuation, and/or levels of stochasticity of star formation, though we note that the latter two are encompassed in the performed SED modeling and should not effect the recovered stellar masses significantly; we refer the reader to the works by Boylan-Kolchin (2023),

Dressler et al. (2024), Ferrara et al. (2023), Finkelstein et al. (2024), Maio & Viel (2023), Mason et al. (2023), Pallottini & Ferrara (2023), Parashari & Laha (2023), Shen et al. (2023), and Sun et al. (2023) for further discussion on these particular topics.

#### 5.3.1. A Redshift Dependence in the Global Baryon Conversion Efficiency

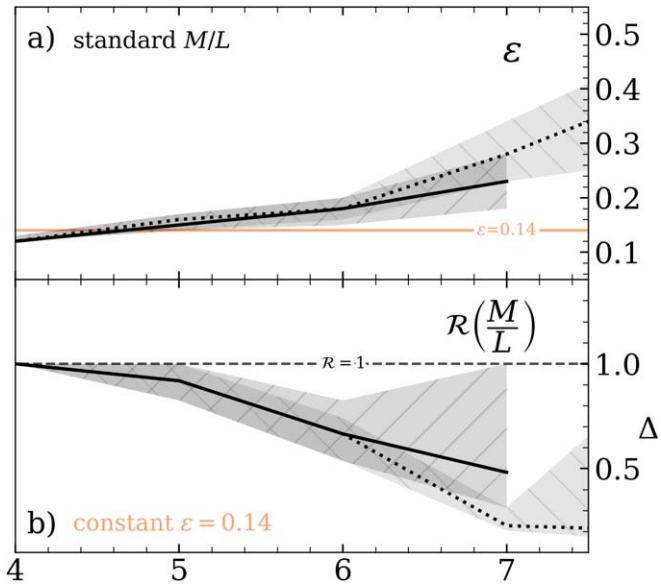
The baryon conversion efficiency ( $\epsilon$ ) describes how efficiently the baryonic gas accreted onto dark matter halos is converted into stars. It is defined as:

$$M_* = \epsilon f_{\text{baryon}} M_{\text{halo}}, \quad (1)$$

where  $M_*$  is the stellar mass,  $f_{\text{baryon}}$  is the fraction of baryons available, and  $M_{\text{halo}}$  is the mass of the underlying dark matter halo. This efficiency, as defined here, is averaged over the SFH of the galaxy, and is determined to strongly correlate with both halo mass as well as the stellar mass of the central galaxy (Behroozi et al. 2010; Moster et al. 2010). In the local Universe, low values of  $\epsilon$  ( $\lesssim 0.15$ ) are inferred (Bregman 2007; Zhang et al. 2022). Energetic feedback from a combination of effects including stellar winds, supernovae, AGN, mergers/shock heating, and morphological quenching are thought to quench and prevent large amounts of efficient star formation from occurring in the local Universe (e.g., Kauffmann & Charlot 1998; Grimes et al. 2009; Martig et al. 2009; Somerville & Davé 2015; Naab & Ostriker 2017; Kondapally et al. 2023, etc.). Dekel et al. (2023) have argued that at extremely high redshifts ( $z \gtrsim 10$ ),  $\epsilon$  may approach unity because in the early Universe, the combination of high ISM density and low metallicities implies that molecular cloud freefall times are shorter than the time it takes for stars to develop winds and supernovae. When they explode in higher-density gas, supernova explosions become less efficient at heating and driving winds, perhaps causing weaker stellar feedback in the early Universe (Walch et al. 2015).

To explore how the baryon conversion efficiency may have had to evolve in order to match the observed number densities for each redshift bin, we find the cumulative number density for each value of  $\epsilon$  that most closely matches the observed cumulative number density in CEERS. From this, we show how the baryon conversion efficiency would evolve with redshift to explain the surplus of massive galaxies in the early Universe. We find that, to match the observed number of  $\log_{10}(M_*/M_\odot) > 10$  galaxies in this study, the efficiency remains relatively constant ( $\epsilon \sim 0.14$ ) from  $z = 1$  to  $z \sim 4$ , at which point it steadily increases, reaching  $\epsilon \sim 0.3$  at  $z \sim 7$ , a factor of 3 more efficient than the average efficiency expected in the local Universe (Bregman 2007; Zhang et al. 2022; Figure 10, panel (a)).

If the light observed from EROs is indeed primarily contributed by their stellar components, and thus the estimated masses are robust, we show that there would need to be a dramatic increase in  $\epsilon$ , particularly at  $z > 7$ , where they represent a significant portion (5/13) of massive galaxies in our sample. Including these sources the baryon conversion efficiency would have to have reached a value of 0.4 by  $z \sim 7$ . At even higher redshifts Dekel et al. (2023) predict an even higher efficiency value (close to unity at  $z \sim 10$ ). We do not probe to such high redshifts in this work due to our mass limit



**Figure 10.** (a) The evolution in baryon conversion efficiency given a constant mass-to-light ratio ( $M/L$ ) needed to match our observations of massive galaxies (Section 5.3.1). (b) The factor change in the mass-to-light ratio required to match the number densities of  $\log_{10}(M_*/M_\odot) > 10$  expected, assuming a constant efficiency of  $\epsilon = 0.14$ . We find that the observed cumulative number densities in CEERS follow that expected from a constant baryon conversion efficiency of  $\sim 0.14$  up to a redshift of  $z \sim 4$ , before which we find an excess of  $\log_{10}(M_*/M_\odot) > 10$  galaxies, potentially indicating that the global baryon conversion efficiency is higher *or* that the mass-to-light ratio is lower in the early Universe than at lower redshift.

combined with the cosmic volume probed (the area coverage of CEERS is not expected to find rare, extremely massive high-redshift sources; Santini et al. 2021).

Our inference of  $\epsilon$  as defined in Equation (1) requires knowledge of the underlying dark matter halo mass function. We use mass functions obtained directly from large cosmological simulations (Bolshoi–Planck and MultiDark–Planck) run with cosmological parameters that are consistent with the Planck cosmology (see Section 4.1). If the true halo mass function deviates from these simulation results, the inferred star formation efficiencies will change accordingly. However, this would only affect our results at a quantitative rather than qualitative level unless the background cosmology itself deviates from the flat Planck cosmology. For a recent detailed investigation of halo mass functions at high redshift, see Yung et al. (2023a).

### 5.3.2. Changing Mass-to-light Ratio

Another explanation for the implied excess of massive galaxies may be due to an incorrect assumption about the mass-to-light ratio at high redshifts. Determination of the SMF is a cornerstone in astrophysics, for the stellar mass distribution determines the evolution, surface brightness, chemical enrichment, and baryonic content of galaxies. SED fitting codes estimate the mass of a galaxy from its observed brightness, dependent on an assumed IMF, or the assumption of the abundance of stars at any given mass populating the galaxy. Various functional forms of the IMF are widely used (e.g., Salpeter 1955; Scalo & Miller 1979; Kroupa & Boily 2002; Chabrier 2003). The determination of the IMF is not direct, but instead is dependent on the transformation of the observed light from objects into mass reliant on theories of stellar evolution.

Current IMFs are calibrated on a Galactic scale, where individual stars are resolvable, while the IMF characteristics of early star formation at large redshift remain undetermined. However, changes in the characteristic stellar mass for an IMF may be expected in the early Universe (Sneppen et al. 2022; Steinhardt et al. 2023). The IMF depends upon several properties of the star-forming clouds, with the high temperatures and low gas metallicities in the early Universe motivating the formation of more massive stars. This can lead to a decrease in the mass-to-light ratio by a factor of several (e.g., Bromm et al. 1999; Raiter et al. 2010), which could also explain the high observed abundances of UV-bright galaxies at  $z > 10$  from recent JWST observations (Finkelstein et al. 2023; Harikane et al. 2023; Yung et al. 2023a).

In Figure 10, we show how the masses measured from our SED fitting would need to be adjusted in order for the cumulative number densities of massive galaxies observed to match those expected from a constant baryon conversion efficiency ( $\epsilon = 0.14$ ). Our reported stellar masses are determined using a Chabrier (2003) IMF, thus the change in the mass-to-light ratio shown here is the deviation from such. To determine this, we repeat the procedure as detailed in Section 4.1, drawing from the posteriors of the SED fits, and bin them according to mass and redshift. For each redshift bin, we then divide all masses by some value  $\delta$  (starting at 1.0 with increments of 0.0001) until the number of galaxies with  $\log_{10}(M_*/M_\odot) > 10$  is within 10% of that which is expected given an assumed constant baryon conversion efficiency of  $\epsilon = 0.14$  (Figure 10, panel (b)). This analysis shows that the mass-to-light ratio should be  $\sim 2 \times$  lower by  $z = 6.5$  and  $\sim 3 \times$  lower by  $z = 7$ . These ratios are consistent with those expected of top-heavy IMFs; e.g., the results from Yung et al. (2023a, 2024) are SAM predictions based on the GUREFT simulation suite, which show that a  $3 \times$  decrease in the mass-to-light ratio can be invoked to match the excess of UV-luminous galaxies and match the observed UV luminosity function at  $z \sim 11$ .

## 6. Conclusion

In this work, we study the evolution of massive ( $\log_{10}(M_*/M_\odot) > 10$ ) galaxies in the CEERS survey. The launch of JWST has brought with it an unparalleled ability to measure stellar masses at redshifts beyond  $z > 5$ , which was historically limited due to the abilities of previous telescope facilities to constrain rest-frame optical emission.

To select massive galaxies, we perform SED fitting with eazy to obtain photometric redshift probabilities and with dense basis to fit flexible nonparametric SFHs and recover posteriors on stellar masses. We perform visual inspection of all high-redshift ( $z > 3.5$ ) sources for which the 97th percentile posterior in stellar mass is  $\log_{10}(M_*/M_\odot) > 10$  (561 sources), allowing us to further vet the photometry and their resulting fits. We remove a total of 300 sources from our final high-redshift sample and note that this significant number of contaminant sources speaks to the necessity of careful selection for these rare galaxies. We remove likely AGN contaminants from our sample, as the stellar mass estimates cannot be accurately determined without further information on the light contribution from the central AGN. We then sample the posterior fit for each galaxy and take the median  $\pm 1\sigma$  spread and arrive at a cumulative number density inclusive of errors from SED fitting, Poisson, and cosmic variance. From this

analysis, we find that the volume densities of massive galaxies in CEERS follow that expected from a constant baryon conversion ratio of  $\epsilon \sim 0.14$  up to  $z \sim 5$ , before which the volume densities exceed a constant  $\epsilon$  and appear to indicate higher  $\epsilon$  at higher redshifts.

We find that the number of  $z > 4$  massive galaxies in CEERS exceeds most theoretical predictions, particularly at redshifts above  $z \sim 5$ . While we do find a higher than expected cumulative number density at  $z > 4$ , our findings are entirely consistent with the  $\Lambda$ CDM cosmological model if we posit an increased efficiency of conversion of baryons into stars in massive dark matter halos with increasing redshift. We show that a redshift-dependent global baryon conversion efficiency or a change to the mass-to-light ratios of galaxies' stellar populations, e.g., due to variations in the IMF, can both reproduce the number of observed massive galaxies at high redshifts. To disentangle these scenarios, further observations to measure galaxy clustering may provide insights into star formation efficiencies (Muñoz et al. 2023) while deep spectroscopic observations to detect very massive stars via strong hydrogen and helium recombination lines alongside a lack of heavy metal features to provide constraints on assumed IMFs (Trinca et al. 2024; Venditti et al. 2023; Yung et al. 2023a).

### Acknowledgments

K.C. and S.L.F. acknowledge support from the National Science Foundation through grant AST-2009905 and from NASA through STScI award JWST-ERS-1345. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE 2137420. M.B.K. acknowledges support from NSF CAREER award AST-1752913, NSF grants AST-1910346 and AST-2108962, NASA grant 80NSSC22K0827, and HST-AR-15809, HST-GO-15658, HST-GO-15901, HST-GO-15902, HST-AR-16159, HST-GO-16226, HST-GO-16686, HST-AR-17028, and HST-AR-17043 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. Some of the data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via DOI:10.17909/z7p0-8481.

*Facilities:* HST (STIS) and JWST (STScI).

*Software:* astropy (Astropy Collaboration et al. 2013, 2018, 2022), EAZY (Brammer et al. 2008), Source Extractor (Bertin & Arnouts 1996), Dense Basis (Iyer et al. 2019; Iyer & Gawiser 2017), and Colormaps (Crameri 2023).

### ORCID iDs

Katherine Chworowsky <https://orcid.org/0000-0003-4922-0613>

Steven L. Finkelstein <https://orcid.org/0000-0001-8519-1130>

Michael Boylan-Kolchin <https://orcid.org/0000-0002-9604-343X>

Elizabeth J. McGrath <https://orcid.org/0000-0001-8688-2443>

Kartheik G. Iyer <https://orcid.org/0000-0001-9298-3523>

Casey Papovich <https://orcid.org/0000-0001-7503-8482>

Mark Dickinson <https://orcid.org/0000-0001-5414-5131>

Anthony J. Taylor <https://orcid.org/0000-0003-1282-7454>

L. Y. Aaron Yung <https://orcid.org/0000-0003-3466-035X>

Pablo Arrabal Haro <https://orcid.org/0000-0002-7959-8783>

Micaela B. Bagley <https://orcid.org/0000-0002-9921-9218>

Bren E. Backhaus <https://orcid.org/0000-0001-8534-7502>

Rachana Bhatawdekar <https://orcid.org/0000-0003-0883-2226>

Yingjie Cheng <https://orcid.org/0000-0001-8551-071X>

Nikko J. Cleri <https://orcid.org/0000-0001-7151-009X>

Justin W. Cole <https://orcid.org/0000-0002-6348-1900>

M. C. Cooper <https://orcid.org/0000-0003-1371-6019>

Luca Costantin <https://orcid.org/0000-0001-6820-0015>

Avishai Dekel <https://orcid.org/0000-0003-4174-0374>

Maximilien Franco <https://orcid.org/0000-0002-3560-8599>

Seiji Fujimoto <https://orcid.org/0000-0001-7201-5066>

Christopher C. Hayward <https://orcid.org/0000-0003-4073-3236>

Benne W. Holwerda <https://orcid.org/0000-0002-4884-6756>

Marc Huertas-Company <https://orcid.org/0000-0002-1416-8483>

Michaela Hirschmann <https://orcid.org/0000-0002-3301-3321>

Taylor A. Hutchison <https://orcid.org/0000-0001-6251-4988>

Anton M. Koekemoer <https://orcid.org/0000-0002-6610-2048>

Rebecca L. Larson <https://orcid.org/0000-0003-2366-8858>

Zhaozhou Li <https://orcid.org/0000-0001-7890-4964>

Arianna S. Long <https://orcid.org/0000-0002-7530-8857>

Ray A. Lucas <https://orcid.org/0000-0003-1581-7825>

Nor Pirzkal <https://orcid.org/0000-0003-3382-5941>

Giulia Rodighiero <https://orcid.org/0000-0002-9415-2296>

Rachel S. Somerville <https://orcid.org/0000-0002-6748-6821>

Brittany N. Vanderhoof <https://orcid.org/0000-0002-8163-0172>

Alexander de la Vega <https://orcid.org/0000-0002-6219-5558>

Stephen M. Wilkins <https://orcid.org/0000-0003-3903-6935>

Guang Yang <https://orcid.org/0000-0001-8835-7722>

Jorge A. Zavala <https://orcid.org/0000-0002-7051-1100>

### References

Adams, N. J., Bowler, R. A. A., Jarvis, M. J., Häußler, B., & Lagos, C. D. P. 2021, *MNRAS*, **506**, 4933

Adams, N. J., Conselice, C. J., Austin, D., et al. 2024, *ApJ*, **965**, 169

Arrabal Haro, P., Dickinson, M., Finkelstein, S. L., et al. 2023, *ApJL*, **951**, L22

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, **935**, 167

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33

Bagley, M. B., Finkelstein, S. L., Koekemoer, A. M., et al. 2023, *ApJL*, **946**, L12

Barro, G., Perez-Gonzalez, P. G., Kocevski, D. D., et al. 2024, *ApJ*, **963**, 128

Behroozi, P., Wechsler, R. H., Hearn, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, *ApJ*, **717**, 379

Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393

Bertin, E., & Arnouts, S. 2010, SExtractor: Source Extractor, *Astrophysics Source Code Library*, ascl:1010.064

Bhatawdekar, R., Conselice, C. J., Margalef-Bentabol, B., & Duncan, K. 2019, *MNRAS*, **486**, 3805

Böker, T., Beck, T. L., Birkmann, S. M., et al. 2023, *PASP*, **135**, 038001

Bouwens, R. J., Stefanon, M., Brammer, G., et al. 2023, *MNRAS*, **523**, 1036

Boylan-Kolchin, M. 2023, *NatAs*, **7**, 731

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503

Bregman, J. N. 2007, *ARA&A*, **45**, 221

Bromm, V., Coppi, P. S., & Larson, R. B. 1999, *ApJ*, **527**, L5

Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2023, JWST Calibration Pipeline, v1.10.2, Zenodo, doi:10.5281/zenodo.7829329

Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, *ApJ*, **840**, 44

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682

Casey, C. M., Akins, H. B., Shuntov, M., et al. 2024, *ApJ*, **965**, 98

Castellano, M., Fontana, A., Treu, T., et al. 2022, *ApJL*, **938**, L15

Chabrier, G. 2003, *PASP*, **115**, 763

Conroy, C. 2013, *ARA&A*, **51**, 393

Conroy, C., & Gunn, J. E. 2010, FSPS: Flexible Stellar Population Synthesis, Astrophysics Source Code Library, ascl:1010.043

Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, **699**, 486

Cramer, F. 2023, Scientific colour maps, v8.0.0, Zenodo, doi:10.5281/zenodo.8035877

Dekel, A., Sarkar, K. C., Birnboim, Y., Mandelker, N., & Li, Z. 2023, *MNRAS*, **523**, 3201

Dressler, A., Rieke, M., Eisenstein, D., et al. 2024, *ApJ*, **964**, 150

Duncan, K., Conselice, C. J., Mortlock, A., et al. 2014, *MNRAS*, **444**, 2960

Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2020, *ApJ*, **898**, 171

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *RMxAA*, **53**, 385

Ferrara, A., Pallottini, A., & Dayal, P. 2023, *MNRAS*, **522**, 3986

Finkelstein, S. L., Bagley, M. B., Arrabal Haro, P., et al. 2022, *ApJL*, **940**, L55

Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, *ApJL*, **946**, L13

Finkelstein, S. L., Leung, G. C. K., Bagley, M. B., et al. 2024, *ApJL*, **969**, L2

Franco, M., Akins, H. B., Casey, C. M., et al. 2023, arXiv:2308.00751

Fujimoto, S., Arrabal Haro, P., Dickinson, M., et al. 2023, *ApJL*, **949**, L25

Fukushima, H., & Yajima, H. 2021, *MNRAS*, **506**, 5512

Furtak, L. J., Labbé, I., Zitrin, A., et al. 2023, *Nature*, **628**, 57

Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, *PASP*, **135**, 068001

Giménez-Arteaga, C., Oesch, P. A., Brammer, G. B., et al. 2023, *ApJ*, **948**, 126

Greene, J. E., Labbe, I., Goulding, A. D., et al. 2023, *ApJ*, **964**, 39

Grimes, J. P., Heckman, T., Aloisi, A., et al. 2009, *ApJS*, **181**, 272

Grogan, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35

Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, *ApJ*, **959**, 39

Hayward, C. C., Sparre, M., Chapman, S. C., et al. 2021, *MNRAS*, **502**, 2922

Iyer, K., & Gawiser, E. 2017, *ApJ*, **838**, 127

Iyer, K. G., Gawiser, E., Faber, S. M., et al. 2019, *ApJ*, **879**, 116

Jarosik, N., Bennett, C. L., Dunkley, J., et al. 2011, *ApJS*, **192**, 14

Juodžbalis, I., Conselice, C. J., Singh, M., et al. 2023, *MNRAS*, **525**, 1353

Kannan, R., Springel, V., Hernquist, L., et al. 2023, *MNRAS*, **524**, 2594

Kauffmann, G., & Charlot, S. 1998, *MNRAS*, **294**, 705

Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, *MNRAS*, **444**, 1647

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

Kondapally, R., Best, P. N., Raouf, M., et al. 2023, *MNRAS*, **523**, 5292

Kroupa, P., & Boily, C. M. 2002, *MNRAS*, **336**, 1188

Labbé, I., van Dokkum, P., Nelson, E., et al. 2023, *Natur*, **616**, 266

Labbe, I., Greene, J. E., Bezançon, R., et al. 2023, arXiv:2306.07320

Larson, R. L., Hutchison, T. A., Bagley, M., et al. 2022, *ApJ*, **958**, 141

Larson, R. L., Finkelstein, S. L., Kocevski, D. D., et al. 2023, *ApJL*, **953**, L29

Leja, J., Speagle, J. S., Johnson, B. D., et al. 2020, *ApJ*, **893**, 111

Lemon, G. & the Virgo Consortium 2006, arXiv:astro-ph/0608019

Leung, G. C. K., Bagley, M. B., Finkelstein, S. L., et al. 2023, *ApJL*, **954**, L46

Lovell, C. C., Vijayan, A. P., Thomas, P. A., et al. 2021, *MNRAS*, **500**, 2127

Maio, U., & Viel, M. 2023, *A&A*, **672**, A71

Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2023, arXiv:2308.01230

Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, *ApJ*, **707**, 250

Mason, C. A., Trenti, M., & Treu, T. 2023, *MNRAS*, **521**, 497

Matthee, J., Naidu, R. P., Brammer, G., et al. 2023, *ApJ*, **963**, 129

McLeod, D. J., McLure, R. J., Dunlop, J. S., et al. 2021, *MNRAS*, **503**, 4413

Merlin, E., Bourne, N., Castellano, M., et al. 2016, *A&A*, **595**, A97

Moffett, A. J., Lange, R., Driver, S. P., et al. 2016, *MNRAS*, **462**, 4336

Morishita, T. 2021, *ApJS*, **253**, 4

Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, *ApJ*, **710**, 903

Muñoz, J. B., Mirocha, J., Furlanetto, S., & Sabti, N. 2023, *MNRAS: Letters*, **526**, L47

Mutch, S. J., Geil, P. M., Poole, G. B., et al. 2016, *MNRAS*, **462**, 250

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18

Naab, T., & Ostriker, J. P. 2017, *ARA&A*, **55**, 59

Naidu, R. P., Oesch, P. A., van Dokkum, P., et al. 2022, *ApJL*, **940**, L14

Nelson, D., Springel, V., Pillepich, A., et al. 2021, *ComAC*, **6**, 2

Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713

Ormerod, K., Conselice, C. J., Adams, N. J., et al. 2023, *MNRAS*, **527**, 6110

Pakmor, R., Springel, V., Coles, J. P., et al. 2023, *MNRAS*, **524**, 2539

Pallottini, A., & Ferrara, A. 2023, *A&A*, **677**, L4

Papovich, C., Cole, J. W., Yang, G., et al. 2023, *ApJL*, **949**, L18

Parashari, P., & Laha, R. 2023, *MNRAS*, **526**, L63

Pillepich, A., Springel, V., Nelson, D., et al. 2017, *MNRAS*, **473**, 4077

Pirzkal, N., Malhotra, S., Ryan, R. E., et al. 2017, *ApJ*, **846**, 84

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, **594**, A13

Poole, G. B., Angel, P. W., Mutch, S. J., et al. 2016, *MNRAS*, **459**, 3025

Raiter, A., Schaerer, D., & Fosbury, R. A. 2010, *A&A*, **523**, A64

Rieke, M. J., Kelly, D. M., Misselt, K., et al. 2023, *PASP*, **135**, 028001

Rodríguez-Puebla, A., Behroozi, P., Primack, J., et al. 2016, *MNRAS*, **462**, 893

Salpeter, E. E. 1955, *ApJ*, **121**, 161

Santini, P., Castellano, M., Merlin, E., et al. 2021, *A&A*, **652**, A30

Scalo, J. M., & Miller, G. E. 1979, *ApJ*, **233**, 596

Scholtz, J., Maiolino, R., D'Eugenio, F., et al. 2023, arXiv:2311.18731

Shen, X., Vogelsberger, M., Boylan-Kolchin, M., Tacchella, S., & Kannan, R. 2023, *MNRAS*, **525**, 3254

Sneppen, A., Steinhardt, C. L., Hensley, H., et al. 2022, *ApJ*, **931**, 57

Somerville, R. S., & Davé, R. 2015, *ARA&A*, **53**, 51

Somerville, R. S., Popping, G., & Trager, S. C. 2015, *MNRAS*, **453**, 4337

Song, J., Fang, G., Lin, Z., Gu, Y., & Kong, X. 2023, *ApJ*, **958**, 82

Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, *ApJ*, **825**, 5

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, *ApJS*, **148**, 175

Stefanon, M., Bouwens, R. J., Labbé, I., et al. 2021, *ApJ*, **922**, 29

Steinhardt, C. L., Kokorev, V., Rusakov, V., Garcia, E., & Sneppen, A. 2023, *ApJL*, **951**, L40

Sun, G., Faucher-Giguère, C.-A., Hayward, C. C., et al. 2023, *ApJ*, **955**, 35

Tacchella, S., Finkelstein, S. L., Bagley, M., et al. 2022, *ApJ*, **927**, 170

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, *ApJ*, **783**, 85

Trinca, A., Schneider, R., Valiante, R., et al. 2024, *MNRAS*, **529**, 3563

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, **788**, 28

Venditti, A., Graziani, L., Schneider, R., et al. 2023, *MNRAS*, **522**, 3809

Walch, S., Girichidis, P., Naab, T., et al. 2015, *MNRAS*, **454**, 238

Ward, E. M., de la Vega, A., Mobasher, B., et al. 2023, *ApJ*, **962**, 176

Weaver, J. R., Davidzon, I., Toft, S., et al. 2022a, *A&A*, **677**, 184

Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022b, *ApJS*, **258**, 11

Wilkins, S. M., Vijayan, A. P., Lovell, C. C., et al. 2023, *MNRAS*, **519**, 3118

Williams, C. C., Alberts, S., Ji, Z., et al. 2023, *ApJ*, **968**, 34

Wright, A. H., Robotham, A. S. G., Driver, S. P., et al. 2017, *MNRAS*, **470**, 283

Wright, G. S., Rieke, G. H., Glasse, A., et al. 2023, *PASP*, **135**, 048003

Xiao, M., Oesch, P., Elbaz, D., et al. 2023, arXiv:2309.02492

Yang, G., Papovich, C., Bagley, M. B., et al. 2021, *ApJ*, **908**, 144

Yang, G., Papovich, C., Bagley, M., et al. 2023, *ApJL*, **956**, A12

Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Popping, G., & Davé, R. 2019a, *MNRAS*, **483**, 2983

Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Wilkins, S. M., & Gardner, J. P. 2023a, *MNRAS*, **527**, 5929

Yung, L. Y. A., Somerville, R. S., Nguyen, T., et al. 2024, *MNRAS*, **530**, 4868

Yung, L. Y. A., Somerville, R. S., Popping, G., et al. 2019b, *MNRAS*, **490**, 2855

Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., et al. 2023b, *MNRAS*, **519**, 1578

Zhang, Z., Wang, H., Luo, W., et al. 2022, *A&A*, **663**, A85