JWST Reveals a Population of Ultralre, Flattened Galaxies at 2 ≤ z ≤ 6 Previously Missed by HST

Erica J. Nelson, Katherine A. Suess, Rachel Bezanson, Sedona H. Price, Pieter van Dokkum, Joel Leja, Bingjie Wang (王冰洁), Katherine E. Whitaker, Ivo Labbé, Laia Barrufet, Justus Gibson, Abigail I. Hartley, Benjamin D. Johnson, Kasper E. Heintz, Elijah Mathews, Tim B. Miller, Pascal A. Oesch, Lester Sandler, David J. Setton, Joshua S. Speagle (沈佳士), Sandro Tacchella, Ivo Labbé, Pieter van Dokkum, Laia Barrufet, Benjamin D. Johnson, Hannah Übler, Pascal A. Oesch, Abigail I. Hartley, Justus Gibson, Rachel Bezanson, Sedona H. Price, David J. Setton

© 2023. The Author. The Astrophysical Journal Letters, 948:L18 (10pp), 2023 May 10

OPEN ACCESS

Abstract

With just a month of data, JWST is already transforming our view of the universe, revealing and resolving starlight in unprecedented populations of galaxies. Although “HST-dark” galaxies have previously been detected at long wavelengths, these observations generally suffer from a lack of spatial resolution, which limits our ability to characterize their sizes and morphologies. Here we report on a first view of starlight from a subset of the HST-dark population that is bright with JWST/NIRCam (4.4 μm < 24.5 mag) and very faint or even invisible with HST (<1.6 μm). In this Letter we focus on a dramatic and unanticipated population of physically extended galaxies (z > 0.25). These 12 galaxies have photometric redshifts 2 < z < 6, high stellar masses M_⋆ ≳ 10^{10} M_☉, and significant dust-attenuated star formation. Surprisingly, the galaxies have elongated projected axis ratios at 4.4 μm, suggesting that the population is disk dominated or prolate and we hence refer to them as ultralre flattened objects. Most of the galaxies appear red at all radii, suggesting significant dust attenuation throughout. With R_e (F444W) ≈ 1–2 kpc, the galaxies are similar in size to compact massive galaxies at z ∼ 2 and the cores of massive galaxies and S0s at z ∼ 0. The stellar masses, sizes, and morphologies of the sample suggest that some could be progenitors of lenticular or fast-rotating galaxies in the local universe. The existence of this population suggests that our previous censuses of the universe may have missed massive, dusty edge-on disks, in addition to dust-obscured starbursts.

1. Introduction

1 Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA; erica.june.nelson@colorado.edu
2 Department of Astronomy and Astrophysics, University of Pennsylvania, Santa Cruz, CA 95064 USA
3 Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA
4 Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA
5 Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
6 Astronomy Department, Yale University, 52 Hillhouse Avenue, New Haven, CT 06511, USA
7 Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
8 Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA, USA
9 Department of Astronomy, University of Colorado, Boulder, CO 80309, USA
10 Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
11 Cosmic Dawn Center (DAWN), Denmark
12 Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, VIC 3122, Australia
13 Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1200 Versoix, Switzerland
14 Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark
15 Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
16 Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
17 Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK
18 David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St George Street, Toronto ON M5S 3H4, Canada
19 Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St George Street, Toronto, ON M5S 3H4, Canada
20 Department of Statistical Sciences, University of Toronto, 100 St George St, Toronto, ON M5S 3G3, Canada
21 National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

Received 2022 August 3; revised 2023 February 18; accepted 2023 March 5; published 2023 May 10

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Unified Astronomy Thesaurus concepts: Galaxy evolution (594); Galaxy structure (622); Galaxy formation (595)
Because these dusty galaxies are likely to be massive (e.g., Whitaker et al. 2017), and may also be the progenitors of today’s large elliptical galaxies (e.g., Toft et al. 2014), an ability to map and characterize their stellar properties is required to understand the growth of the most massive galaxies at cosmic noon. In particular, understanding their sizes and morphologies places constraints on their formation.

To date, detailed studies of the morphologies of these dust-obscured galaxies have been challenging as they require ≤0″2 spatial resolution. While this resolution is achievable with HST, these galaxies are HST-dark, making this definitionally not an option. While dusty star-forming galaxies have long been detectable at far-infrared wavelengths (e.g., Barger et al. 1998; Hughes et al. 1998; Coppin et al. 2006; Elbaz et al. 2011) and spectroscopically confirmed to reside at high redshift (e.g., Chapman et al. 2005; Pope et al. 2008), the >2″ resolution of infrared instruments has made it impossible to spatially resolve them in the infrared. This problem can be ameliorated somewhat in the submillimeter and radio where interferometry can be used to achieve the requisite spatial resolution. Studies resolving dusty star-forming galaxies have generally found compact submillimeter sizes (Barro et al. 2016; Hodge et al. 2016; Oteo et al. 2016; Rujopakarn et al. 2016; Fujimoto et al. 2017; Tadaki et al. 2017; Gullberg et al. 2019; Nelson et al. 2019; Tadaki et al. 2020), which have been interpreted as watching these galaxies in the process of building their bulges (e.g., Nelson et al. 2019; Hodge & da Cunha 2020; Tadaki et al. 2020).

However, due to the limited field of view of the Atacama Large Millimeter/submillimeter Array (ALMA) interferometer, these studies have generally been restricted to samples of galaxies already known with HST (e.g., Tadaki et al. 2020), relatively bright in submillimeter continuum (e.g., Walter et al. 2016; Cowie et al. 2018; Franco et al. 2018; Gómez-Guijarro et al. 2022) or lensed (e.g., Chen et al. 2014; Hsu et al. 2017). Perhaps more fundamentally, while millimeter interferometry provides constraints on the distribution of gas and dust, it does not tell us about the distribution of actual stellar mass. For this we need observations in the rest-frame near-infrared wavelengths covered by Spitzer whose >2″ spatial resolution is woefully inadequate for this purpose. To date, we have not had a telescope capable of resolving the distribution of stars in these dusty galaxies. The successful launch and commissioning of the James Webb Space Telescope (JWST) has changed the game; now, subarcsecond observations of these galaxies are possible out to 28 μm.

In this paper, we report the discovery of 26 HST-dark galaxies, 12 of which are extended, that are fully obscured at optical wavelengths but bright at 4.4 μm with JWST imaging from the Cosmic Evolution Early Release Science (CEERS) Survey. This 4.4 μm imaging traces rest-frame wavelengths of 1.1 μm (0.6 μm) at z = 3 (z = 6). These extended red star-forming galaxies are the main topic of this paper, which is organized as follows. Section 2 describes the observations and sample selection. In Section 3, we present our findings on the physical properties of this galaxy sample, including their redshifts and stellar populations. Section 4 contains our identification of the shapes and sizes of these galaxies, and we set forth our results on their observed color gradients in Section 4.3. We discuss the findings and conclusions of this study in Section 5.

In this paper, we assume the WMAP9 ΛCDM cosmology with Ω_m = 0.2865, Ω_L = 0.7135, and H_0 = 69.32 km s^{-1} Mpc^{-1} (Hinshaw et al. 2013). All magnitudes in this paper are expressed in the AB system (Oke 1974).

2. Data and Sample Selection

For this work, we use overlapping imaging in the AEGIS field from JWST via the CEERS program (Finkelstein et al. 2022) and HST via the CANDELS program (Grogin et al. 2011; Koekemoer et al. 2011). The early CEERS imaging covers a ~40 arcmin² portion of the AEGIS field and was taken in six broadband near-infrared filters (F115W, F150W, F200W, F277W, F356W, and F444W) and one medium band filter (F410M). All the JWST data used in this paper can be found in MAST doi:10.17909/tvn0-6041. Stage 2 of the JWST calibration pipeline (v1.5.2) produced flux-calibrated exposures, publicly available from the MAST archive. Further reduction, aligning, and coadding of the exposures was conducted using the public software package grizli (Brammer & Matharu 2021) and described fully in G. Brammer et al. 2023, (in preparation). Briefly, grizli masks imaging artifacts, subtracts an overall sky background, aligns the images to stars from the Gaia DR3 catalog, and projects images to a common pixel space using astrodrrizzle. Additional background structure was removed with a 5″ median filter after masking bright sources (see also, e.g., Labbe et al. 2022).

We detect sources and create segmentation maps on an inverse variance weighted combination of the F277W, F356W, and F444W images convolved with a 2.5 pixel FWHM Gaussian using standard astropy and photutils procedures. Photometry is done on all detected sources in 0″32 and 0″5 diameter circular apertures. These aperture fluxes are then scaled to total fluxes using the Kron autoscaling aperture measurement from the detection image plus a small additional correction based on the encircled energy from WebbPSF (Perrin et al. 2015). For additional details see Labbe et al. (2022).

When comparing between the JWST and HST color images, we quickly noticed a set of large JWST-bright galaxies that were invisible with HST. To collect this population, we select galaxies that are bright at the reddest wavelengths JWST/ NIRCam can detect and not detected at the <1.6 μm wavelengths previously visible with HST. Quantitatively, we identify HST-dark galaxies with total AB magnitudes of

1. F444W < 24.5 mag;
2. F150W > 25.5 mag;
3. F115W > 27 mag.

This selection identifies 26 optically faint galaxies, which we here refer to as HST-dark for conceptual consistency with previous literature. This is similar to the selection employed in Barrufet et al. (2023) of F160W–F444W > 2.3, F160W > 27 mag, and S/N > 5 in the F444W filter (which was in turn based on previous studies with Spitzer Caputi et al. 2012; Wang et al. 2016). The selected galaxies have a range of projected sizes, spanning 0″08–0″33 (see Figure 1). Visual inspection of the F150W–F444W red objects revealed a striking population of extended galaxies (see Figure 2). While extended light distributions are common for z~2–6 massive galaxies, existing size measurements for z~6 galaxies show very compact light distributions (e.g., Holwerda et al. 2015; Kawamata et al. 2018; Naidu et al. 2022; Yang et al. 2022).
We fit all galaxies that fall in our color selection with GALFIT (Peng et al. 2002, 2010) and focus on the more extended, applying a threshold of 0″25 to the GALFIT F444W half-light radii. Hence, in this paper, we highlight this previously unseen population of dusty galaxies at $2 < z < 6$ (the derivation of redshifts is described in the next section). A description of the properties of the full sample of HST-dark galaxies is presented in Barrufet et al. (2023). The primary difference between these two samples is the imposition of a size cut, which results in a lower median redshift and higher median mass. These requirements result in the sample of 12 galaxies shown in Figure 2. The selection is shown in Figure 1. These galaxies constitute some of the reddest, brightest sources, all undetected by HST.

3. Redshifts and Stellar Populations

We fit stellar population parameters using the Prospector Bayesian inference framework (Johnson et al. 2021), adopting the Prospector-$\alpha$ physical model (Leja et al. 2017, 2019). This models the star formation history by fitting for the mass formed in seven logarithmically spaced time bins, assuming a continuity prior that weights against large changes in star formation rate between bins. The stellar components are modeled with the MIST isochrones (Choi et al. 2016; Dotter 2016) in FSPS (Conroy & Gunn 2010), a Chabrier initial mass function, and a two-component dust model. The two-component dust model is particularly important for these red objects: it follows Charlot & Fall (2000) in modeling dust obscuration with a separate birth-cloud and diffuse screen, with the birth-cloud screen affecting only nebular emission and stars with age $< 10\, \text{Myr}$, with up to $A_V \sim 4$ allowed in each component. There is also a flexible attenuation curve following the Noll et al. (2009) prescription. Nebular emission is powered by the stellar ionizing continuum from the model (Byler et al. 2017). The fitting is sped up by a factor of $\sim 300$ using a neural net emulator which mimics stellar population synthesis models, dubbed parrot (Alsing et al. 2020; E. Mathews et al. 2023, in preparation). The parameters we report here are the median of the marginalized posterior probability function; the $1\sigma$ error bars are reported as the 84th–50th and 50th–16th interquartile ranges. We enforce a maximum signal-to-noise ratio of 20 in the observed photometry to allow for both unavoidable model-level errors and potential calibration uncertainties in the new JWST photometry.

Additionally, in order to avoid spurious high-mass, high-redshift solutions for these very red objects, we adopt a mass function prior for stellar mass, $P(\log M\big| z)$ (described in Wang et al. 2023). This is constructed by using the observed Leja et al. (2020) mass functions between $0.2 < z < 3$, inferred using the same Prospector-$\alpha$ stellar populations model. For $z < 0.2$ and $z > 3$, we adopt the nearest-neighbor solution, i.e., the $z = 0.2$ and $z = 3$ mass functions, as we do not yet have reliable high-resolution rest-frame optical selected mass functions at $z > 3$. In particular, this choice allows a conservatively high probability for yet-to-be-discovered populations of high-mass, high-redshift galaxies (hints of which have already been observed; see Labbe et al. 2022). We allow a solution of $0 < z < 12$ for all objects except 38029, which is allotted a
Figure 2. Three color images for all of the UFOs in F150W, F200W, and F444W filters, with a 1" bar indicated. Most of these galaxies are red throughout and consistently elongated. We note that a number of the objects include nearby sources, at least in projection. For two in particular, 8459 and 38,711, there are ambiguous patches with slightly bluer colors that may be associated with the primary objects, but may also be chance superpositions.
reduced range of $0 < z < 5$ for consistency with the EAzY solution (below).

We cross-check our redshifts and stellar masses by also fitting their photometry with the EAzY-py public photometric redshift code (Brammer et al. 2008). EAzY fits the observed spectral energy distribution (SED) with a set of templates, with the template weight and redshift as free parameters. Because the templates, generated with FSPS (Conroy & Gunn 2010), have associated stellar population properties like age, dust, star formation history, and $M_*/L$ ratio, EAzY can also be used to give an estimate of these properties. As a starting point, the default template set tweak_fspsf_QSF_12_v3 is used with one modification: an additional FSPS template is generated with an age of 1 Gyr and $A_V = 6$ attenuation to allow for more flexibility in reproducing extreme SED shapes.

The best-fit SEDs and observed photometry of our sample of F444W-selected galaxies are shown in Figure 3, the overall distribution of the sample in mass–redshift space is shown in the rightmost panel of Figure 4, and their properties are listed in Table 1. In general these are massive, dust-obscured, star-forming galaxies at a median redshift of $z = 3.5$ and span the range from $2.4 < z < 6.5$. They have a median log stellar mass of $\log M_*/M_\odot = 10.7$ from EAzY and 10.2 from parrot with a tail up to very massive galaxies with $\log (M_*) = 11.1$. With median dust attenuation values of $A_V \sim 2$ and star formation rates of SFR $\sim 40 M_\odot$ yr$^{-1}$ these are dusty star-forming galaxies as determined by their rest-optical and near-infrared light, but not extremely so, although we note these quantities are not well constrained. Compared to KIEROs and HIEROs (in Wang et al. 2012, 2016, respectively), the ultrafaint flattened objects (UFOs) in this paper have similar average redshifts; slightly lower average masses and 4.4/4.5 $\mu$m fluxes; and much lower average star formation rates. These are likely an extreme extension (in color and redshift) of the three red spirals found in the SMACS early release observations by Fudamoto et al. (2022).

Comparing redshifts from EAzY and parrot shows a median offset of only $\Delta z_{\text{EAzY}} – \Delta z_{\text{parrot}} = 0.04$, i.e., no significant systematic offsets in either direction. However, the median absolute offset is much larger: $|\Delta z_{\text{EAzY}} – \Delta z_{\text{parrot}}| = 0.4$. The stellar masses have a systematic offset of $\log M_*(\text{EAzY}) – \log M_*(\text{parrot}) = 0.25$ and a median absolute difference of 0.4 dex. The photometric redshifts and stellar populations of these objects are not terribly well constrained with existing data owing to their very red and often featureless SEDs. MIRI photometry and/or NIRSpec spectroscopy would significantly improve these constraints.

4. Resolved Galaxy Properties

We use the GALFIT software package (Peng et al. 2002, 2010) to fit the sizes and shapes of galaxies accounting for the point-spread function (PSF), following the procedure described in Suess et al. (2022). While the use of stars in the observed images to create empirical PSFs would be preferable, the centers of most stars in our CEERS mosaic are saturated or masked out, necessitating the use of a theoretical PSF. We generate theoretical PSFs with $9 \times 9$ oversampling on the same $0\"/04$ pixel scale as our image mosaic using the WebPSF software (Perrin et al. 2014). We then rotate these theoretical PSFs to match the position angle of the observations, convolve the $9 \times 9$ oversampled PSFs with a $9 \times 9$ square kernel, and downsample them to the pixel scale of the mosaic. The purpose of conducting the bulk of this procedure with an oversampled PSF image is to minimize distortions from the rotation algorithm.

To optimally utilize computational resources, GALFIT must be fed images that do not contain more than a handful of sources it must model simultaneously. It also needs a sufficiently large postage stamp to contain all of the light from the galaxy of interest. As a balance, we cut $80 \times 80$ pixel postage stamps of our galaxies out of their mosaics, corresponding to $>6r_e$ of all galaxies in our sample. We create a segmentation map of each postage stamp to identify all sources to be modeled or masked. Galaxies that have centers within $3\"$ of the target galaxy center and are less than 2.5 mag fainter are modeled simultaneously. Fainter and more distant galaxies are masked. With all sufficiently bright galaxies identified, we estimate and subtract the background in each stamp using the SEXtractor background algorithm as implemented in photutils. We run GALFIT on the F444W images, fitting Sérsic profiles to each unmasked galaxy.

4.1. Sizes

Figure 4(a) shows the (semimajor) effective radii at 4.4 $\mu$m of these galaxies versus their stellar mass. These galaxies have angular sizes of $0\"/25–0\"/51$ (the lower bound is introduced by selection), which is fairly typical for their fluxes (see Figure 1, second panel). Physically, they have sizes of 1 to 3 kpc. The black lines indicate the mass-weighted size–mass relation at $z \sim 2.25$ from (Suess et al. 2019, which builds on the light-weighted size–mass relations from Mowla et al. 2019 and van der Wel et al. 2014a). These mass-weighted sizes are consistent with sizes measured from F444W imaging (Suess et al. 2022). These galaxies lie just below the relation for star-forming galaxies. Given that galaxy sizes are expected to grow with redshift (e.g., Mo et al. 1998), the slight offset between the median sizes of the UFOs and the star-forming size–mass relation could easily be attributed to the higher median redshift of this sample. They are more extended than quiescent galaxies at $z \sim 2.25$; however, when compared to mass-weighted sizes of local early type galaxies (S0s and fast rotators) from MANGA Bernardi et al. (2023), their sizes could be consistent with the low-mass extension of the $z \sim 0$ relation.

With the same amount of dust, more compact galaxies will have higher dust column densities and hence more dust reddening. Thus, the naive intuition about galaxies that are bright in 4.4 $\mu$m, but so reddened as to be undetectable with HST may have been that they would be more compact than the population writ large (e.g., Nelson et al. 2014). However, as can be seen in both Figures 2 and 4, these galaxies are typical for their stellar masses. Interestingly, these galaxies have similar sizes to the far infrared (FIR) sizes of massive galaxies at $z \sim 2.5$ (e.g., submillimeter galaxies in Hodge et al. 2016 and mass-selected star-forming galaxies in Tadaki et al. 2020).

4.2. Axis Ratios

Figure 4(b) shows the distribution of projected axis ratios $q = b/a$ for the galaxies in our sample (black histogram), along with representative models of projected ellipsoids. Interestingly, all galaxies are flattened in projection, with axis ratios $0.24 < q < 0.65$ (see also Figure 2). This distribution is inconsistent with a population of round oblate spheroids (red line). With other effects (i.e., dust) being absent, one would expect randomly oriented 3D disks to exhibit flat distributions
orange solid line, not peaked ones at low $q$. This might suggest a selection bias against face-on galaxies (orange dashed line). If dust is well mixed, then our color selection might preferentially select edge-on disks with longer total optical path through the galaxy (e.g., Maller et al. 2009; Wild et al. 2011; Patel et al. 2012; Mowla et al. 2019), which could result in a preferentially peaked axis ratio distribution as opposed to the expected $\sim$flat distribution in axis ratio for disky systems. In

Figure 3. Spectral energy distributions (SEDs) spanning from 0.4 to (in some cases) 24 μm. Black points show observed JWST photometry, while gray points show HST and Spitzer photometry. Red shows the best-fit model from parrot. In each case, the SEDs are very red, attributable to a combination of redshift, dust, and age.
other words, face-on systems are more likely to be brighter at bluer wavelengths and hence may be excluded from our sample.

Another possibility is that the low-\(q\) values reflect a prolate population (blue line). Although such a model could also be consistent with the observed distribution, this would be at odds with the results for rest-frame optical axis ratio distributions of high-mass galaxies at slightly lower redshifts (e.g., van der Wel et al. 2014b; Zhang et al. 2019, 2022). The Sérsic indices of these systems are for the most part close to \(n = 1\), which could be consistent with a disk- or prolate-dominated population in F444W.

Although the intrinsic 3D shapes of galaxies cannot be measured directly, they can be statistically inferred based on the projected axis ratio distribution through statistical modeling (e.g., van der Wel et al. 2014b; Zhang et al. 2019). We follow the modeling procedure of Chang et al. (2013) and van der Wel et al. (2014b), adopting the median axis ratio uncertainty \(\delta q\) for our sample and using dynamic nested sampling (Dynesty, Speagle 2020) to determine the posterior distribution of the intrinsic galaxy geometries (i.e., \(E = 1 - (C/A)\) and \(T = (1 - (B/A)^2)/[1 - (C/A)^2]\) with intrinsic axes lengths \(A \geq B \geq C\), as in van der Wel et al. 2014b). For a free fit of \(\mu_E, \sigma_E, \mu_T, \sigma_T\), the fact that the observed distribution peaks at \(q \sim 0.5\) is best described by population with prolate geometries (see, e.g., van der Wel et al. 2014b; Zhang et al. 2022). Alternatively, if we assume these galaxies are axisymmetric \((T = 0)\), then the individual objects all have high probabilities \((P(\text{disk}) \gtrsim 50\%)\) of being disks (defined as \(C/A \leq 0.4\)). Such an axisymmetric disk geometry would be consistent with the possibility that these objects are dusty, highly inclined disks, which is what one might guess based on the images in Figure 2. Ultimately, this statistical fitting is inherently uncertain as the sample size is small and potentially dramatically biased due to the size and color selections.

Interestingly, submillimeter-selected galaxy samples appear to have similar axis ratio distributions, suggesting significant population overlap. These results have led to interpretations that those galaxies are either disk (Hodge et al. 2016) or triaxial systems (Gullberg et al. 2019). These previous HST studies of submillimeter bright galaxies have suffered from similar small number statistics to definitively characterize the 3D structures of heavily dust-obscured systems. However, further JWST imaging should continue to robustly detect starlight from dusty star-forming galaxies, ultimately building up much larger samples that will be better suited to answering these statistical questions.

4.3. Color Gradients

One of the most surprising and striking features of these galaxies is that they appear to be red throughout their disks. We quantify this by comparing their colors at different radii. To do this, we perform photometry with both 0.3 and 0.5 aperture. We then compare the central 0.3 and the 0.3–0.5 annulus

Figure 4. Physical Properties. Left: effective (half-light) semimajor radius in F444W vs. stellar mass. The gray points show the F444W sizes from Suess et al. (2022) for 1.0 < \(z\) < 2.5 galaxies at \(M_* > 10^9 M_\odot\), and the red symbols show UFOs, which are the subject of this paper. To guide the eye, the black lines show the relationship between half-mass–radius and stellar mass at \(z \sim 2\) for star-forming galaxies (upper relation) and quiescent galaxies (lower relation) (Suess et al. 2019), and the blue lines show the size–mass relation for S0s and fast rotators from Bernardi et al. (2023). Center: distribution of observed axis ratios and example population models with various intrinsic 3D shapes. Assuming axisymmetry, the UFOs could have intrinsic disk geometries, with the peaked distribution at \(q \sim 0.5\) caused by a preferential edge-on orientation for these objects. Without this assumption, prolate shapes also provide a reasonable fit to the data. Right: redshifts and stellar masses for the galaxies in our sample in red. Our galaxies have \(2 < z < 6\) and \(M_* \sim 10^{10} – 10^{11} M_\odot\).

Figure 5. Comparison of the F277W–F444W color in the central 0.3 (filled circles) and a 0.3–0.5 annulus (open circles) as a function of total F444W magnitude. Gray points show the colors of all galaxies with \(B-A-Y\) redshifts \(2 < z < 6\). Not only are the UFOs significantly redder than most galaxies at similar redshifts, they are red throughout their disks.
colors derived by subtracting the 0\degree 32 from the 0\degree 5 aperture fluxes. As shown in Figure 5, the 0\degree 32 and 0\degree 5–0\degree 32 colors are significantly redder than integrated galaxy colors of the full sample (black points). The central aperture has a median color $F277W-F444W = 1.22$ mag while the outer annulus has 1.05 mag and the median color gradient is 0.26 mag. The red colors in the outskirts of these systems is surprising and in contrast with a canonical view of massive disk galaxies with red bulges and blue disks. This difference may suggest that previous studies based on HST imaging would have excluded a fully optically-thick population included in this work.

### 5. Discussion and Conclusions

JWST is already allowing us to see the universe with new eyes, providing unprecedented spatial resolution in the infrared. With the JWST/NIRCam imaging released from the early CEERS program, we investigate the stellar structures of extremely red galaxies to which HST was previously blind, but are bright at 4 \mu m. This population includes 12 physically extended galaxies at $2 < z < 6$. While obtaining accurate photometric redshifts and stellar population models is difficult due to the very red light distributions of these galaxies, our sample appears to be massive ($M_\star \sim 10^{10}-10^{11}$), dusty, and star-forming, with potential contributions from older stellar populations. The discovery of these new objects at relatively late cosmic epochs—where we thought we had a reasonable census of the universe—highlights the incredible discovery space enabled by JWST.

Perhaps the most noteworthy result stems from the flattened shapes of these HST-dark galaxies. These massive, star-forming galaxies are the likely progenitors of today’s massive galaxies, which tend to be bulge/spheroid dominated. Disk-dominated objects at these cosmic times are expected to be gas-rich and gravitationally unstable, resulting in clump formation and migration to galactic centers, building central bulges (e.g., Ceverino et al. 2017). Finally, studies of the kinematics of star-forming galaxies show that they become progressively more dispersion-dominated and less rotation-dominated at earlier cosmic times (Gnerucci et al. 2011; Kassin et al. 2012; Wisnioski et al. 2015, 2019; Price et al. 2020). Taken together, the expectation may have been that the stellar bodies of these objects would already host significant bulges. This, however, is not what we observe in this sample. These objects have uniformly low axis ratios ($q \lesssim 0.6$) and Sérsic indices close to one, inconsistent with being significantly bulge-dominated/oblate at 4 \mu m. If we assume axisymmetry, these galaxies are likely to be disks. If not, the lack of high axis ratio objects could potentially suggest a prolate population. We think the former is more likely given our sample selection. Selecting things that are extended and with little to no blue light may result in not selecting face-on disks, which are less obscured than edge-on disks (e.g., Wild et al. 2011; Patel et al. 2012; Mowla et al. 2019).

Another surprising aspect of these objects is the spatial patterns in their colors. As a result of building up from the inside out (e.g., Nelson et al. 2012), we may expect age and hence color gradients in extended massive galaxies (e.g., Suess et al. 2019; Mosleh et al. 2020; Miller et al. 2023). We also expect the centers of galaxies to be more dust obscured than large radii, also contributing to color gradients (e.g., Nelson et al. 2016; Tacchella et al. 2018). Indeed, Suess et al. (2022) find smaller sizes in 4.4 than 1.5 \mu m, implying light that is more concentrated in the red than the blue, consistent with the physical picture of red bulges within blue disks. Hence, it came as a surprise that most of these objects are red throughout their disks. While two show some blue low-surface-brightness light at large radii, the bulk are red throughout, even in their outskirts. This suggests significant dust attenuation through the galaxies, not merely in their centers. The color gradients in our sample are consistent with the idea of looking through large dust columns at all radii in edge-on disks. A more in-depth discussion of the physical drivers of color gradients (or lack thereof) is presented in Miller et al. (2022).

Although this Letter focuses on a small sample of galaxies, this study emphasizes that JWST presents a new opportunity to connect the known populations of heavily dust-obscured star-forming galaxies that dominate the star formation history of the universe at cosmic noon (e.g., Barger et al. 2000; Wardlow et al. 2011; Casey et al. 2014; Zavala et al. 2021) with stellar mass-selected samples of galaxies across cosmic time. It will allow us to connect the distribution of existing stellar mass (e.g., with NIRCam at 4 \mu m in this study) with dust-obscured star formation (e.g., with MIRI). The full picture will extend early studies of the morphological evolution in massive galaxies (e.g., Hodge & da Cunha 2020). Furthermore, the less stochastically growing stellar components will likely help to test expectations of progenitor-descendant matching (e.g., Toft et al. 2014), in which uncertainties are currently dominated by the timescales of the submillimeter-bright phase.

| ID  | R.A. (deg) | Decl. (deg) | F115W (mag) | F150W (mag) | F444W (mag) | $r_e$ (\arcsec) | $z_{ext}$ | $z_{par,e}$ | log($M_e/\star$) ($M_\odot$) | log($M_{e,par,e}/\star$) ($M_\odot$) |
|-----|------------|-------------|-------------|-------------|-------------|----------------|----------|-----------|-----------------------------|-----------------------------|
| 1351 | 214.80996904 | 52.80974353 | 27.63 | 26.13 | 22.27 | 0.23 | 2.8 | 1.80 | 10.75 | 10.02 |
| 6898 | 214.90114317 | 52.83810470 | 27.06 | 25.73 | 22.16 | 0.31 | 2.7 | 1.80 | 10.71 | 10.01 |
| 8459 | 214.87066300 | 52.84611032 | 27.87 | 27.08 | 23.84 | 0.24 | 3.5 | 3.60 | 10.45 | 9.30 |
| 10275 | 214.85597860 | 52.85468575 | 31.36 | 26.68 | 22.97 | 0.36 | 3.6 | 4.10 | 11.01 | 10.40 |
| 20153 | 214.92575666 | 52.91852621 | 28.46 | 27.50 | 23.34 | 0.28 | 3.5 | 3.20 | 10.69 | 10.50 |
| 24554 | 214.97745753 | 52.95349457 | 27.33 | 26.23 | 22.96 | 0.40 | 2.9 | 2.80 | 10.70 | 10.70 |
| 31169 | 214.84026259 | 52.80111116 | 28.44 | 27.02 | 24.24 | 0.22 | 2.5 | 6.40 | 10.35 | 10.10 |
| 31961 | 214.76974546 | 52.81634720 | 28.04 | 27.07 | 23.23 | 0.51 | 3.6 | 3.70 | 10.74 | 10.50 |
| 38347 | 214.88067045 | 52.91296665 | 27.03 | 25.88 | 22.74 | 0.30 | 1.8 | 2.10 | 10.27 | 10.27 |
| 38711 | 214.93159121 | 52.92100046 | 29.70 | 28.28 | 24.47 | 0.31 | 4.4 | 2.90 | 10.15 | 9.80 |
| 39291 | 214.89186254 | 52.93388113 | 29.00 | 27.29 | 24.09 | 0.36 | 5.6 | 3.80 | 11.13 | 10.20 |
| 41736 | 215.02153540 | 52.99130527 | 28.53 | 25.95 | 21.93 | 0.27 | 2.9 | 2.70 | 11.00 | 10.20 |
We have mostly compared this sample of galaxies to massive disk galaxies with growing bulges; however, it is important to note that these galaxies could be the early progenitors of a very different population of quiescent galaxies in the local universe. Many quiescent, massive galaxies are either extremely flattened (lenticular) or significantly rotationally supported (fast rotators; Cappellari et al. 2011). We refer back to the the mass-weighted size–mass relation for S0 and fast-rotating galaxies from the MANGA survey (Bernardi et al. 2023), as shown in Figure 4(a). This small sample is insufficient to test whether continued star formation and self-similar growth would evolve the population onto the local relation. However, it is notable that, unlike the full population of star-forming galaxies at cosmic noon, this sample does not require significant structural evolution to reside within existing local descendants. Further studies with more precise redshifts and number densities will be necessary to perform a more careful progenitor-descendant linking.

Understanding the true nature of these ultrared, flattened galaxies will require significant additional data and many questions remain. First, more precise measurements of their redshifts will improve our ability to characterize their stellar populations. In this context, this avenue of study will benefit greatly from the spectroscopic capabilities of JWST and or ALMA. Furthermore, later this year, the CEERS program will cover the NIRCam footprint with MIRI imaging, which will further pin down the long-wavelength emission from these galaxies, breaking the degeneracy between age, dust, and redshift and tightening the confidence intervals. Second, determining whether the intrinsic shapes of these galaxies are disk dominated or prolate could be done with kinematic measurements using ionized gas with JWST/NIRSpec or molecular gas with ALMA. Third, it is possible that even more obscured stellar populations reside in these galaxies that have dust attenuations too high for even NIRCam to peer through (e.g., Arp 220 Scoville et al., 2017), which could be seen by mapping their FIR dust continuum with ALMA. Finally, if these objects exist in simulations of galaxy formation, simulations could be used to understand their formation, link to the overall galaxy population, and subsequent evolution.

E.J.N. acknowledges support from HST-AR-16146, K.A.S. acknowledges the UCSC Chancellor’s Postdoctoral Fellowship Program for support. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. Cloud-based data processing and file storage for this work is provided by the AWS Cloud Credits for Research program. R.B. acknowledges support from the Research Corporation for Scientific Advancement (RCSA) Cottrell Scholar Award ID No: 27587. H.U. gratefully acknowledges support by the Isaac Newton Trust and by the Kavli Foundation through a Newton-Kavli Junior Fellowship. This work is based in part on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with program DD-ERS#1345 (PI: Finkelstein).

Facilities: JWST(NIRCam), HST(WFC3).

Software: numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), matplotlib (Hunter 2007), astropy (Astropy Collaboration et al. 2013), grizli (Brammer & Matharu 2021), photutils (Bradley et al. 2020), astrodrizzle, GALFIT (Peng et al. 2002, 2010), Dynesty (Speagle 2020), Prospector (Johnson et al. 2021).

ORCID iDs
Erica J. Nelson https://orcid.org/0000-0002-7524-374X
Katherine A. Suess https://orcid.org/0000-0002-1714-1905
Rachel Bezanson https://orcid.org/0000-0001-5063-8254
Sedona H. Price https://orcid.org/0000-0002-0108-4176
Pieter van Dokkum https://orcid.org/0000-0002-8282-9888
Joel Leja https://orcid.org/0000-0001-6755-1315
Bingjie Wang (王冰洁) https://orcid.org/0000-0001-9269-5046
Katherine E. Whitaker https://orcid.org/0000-0001-7160-3632
Ivo Labbé https://orcid.org/0000-0002-2057-5376
Laia Barrufet https://orcid.org/0000-0001-6414-6185
Gabriel Brammer https://orcid.org/0000-0003-2680-005X
Daniel J. Eisenstein https://orcid.org/0000-0002-2929-3121
Justus Gibson https://orcid.org/0000-0003-1903-9813
Abigail I. Hartley https://orcid.org/0000-0002-5891-1603
Benjamin D. Johnson https://orcid.org/0000-0002-9280-7594
Kasper E. Heintz https://orcid.org/0000-0002-9389-7413
Eljah Mathews https://orcid.org/0000-0003-0384-0681
Tim B. Miller https://orcid.org/0000-0001-8367-6265
Pascal A. Oesch https://orcid.org/0000-0001-5851-6649
Lester Sandles https://orcid.org/0000-0001-9276-7062
David J. Setton https://orcid.org/0000-0003-4075-7393
Joshua S. Speagle (沈佳士) https://orcid.org/0000-0003-2573-9832
Sandro Tacchella https://orcid.org/0000-0002-8224-4505
Ken-ichiTadaki https://orcid.org/0000-0001-9728-8909
Hannah Übler https://orcid.org/0000-0003-4891-0794
John R. Weaver https://orcid.org/0000-0003-1614-196X

References
Aalsing, J., Peiris, H., Leja, J., et al. 2020, ApJS, 249, 5
Avila, R. J., Hack, W., Cara, M., et al. 2015, in ASP Conf. Ser. 495, Astronomical Data Analysis Software an Systems XXIV (ADASS XXIV), ed. A. R. Taylor & E. Rosolowsky (San Francisco, CA: ASP), 281
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Barger, A. J., Cowie, L. L., Blair, A. H., & Jones, L. H. 2022, ApJ, 934, 56
Barger, A. J., Cowie, L. L., & Richards, E. A. 2000, AJ, 119, 2092
Barger, A. J., Cowie, L. L., Sanders, D. B., et al. 1998, Natur, 394, 248
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2016, ApJL, 827, L32
Barrufet, L., Oesch, P. A., Weibel, A., et al. 2023, MNRAS, 522, 449
Bernardi, M., Sheth, R. K., Domínguez Sánchez, H., et al. 2023, MNRAS, 518, 3494
Bradley, L., Sipőcz, B., Robitaille, T., et al. 2020, astrophy/photutils: v1.0.0, Zenodo, doi:10.5281/zenodo.4044744
Brammer, G., & Matharu, J. 2021, grbammer/grizli: Release 2021, Zenodo, doi:10.5281/zenodo.1146904
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Byler, N., Dalcanton, J. J., Conroy, C., &Johnson, B. D. 2017, ApJ, 840, 44
Cappellari, M., Emel’yanov, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Caputi, K. I., Dunlop, J. S., McLure, R. J., et al. 2012, ApJL, 750, L20
Casey, C. M., Narayanan, D., & Cooray, A. 2014, PhR, 541, 45
Ceverino, D., Primack, J., Dekel, A., & Kassin, S. A. 2017, MNRAS, 467, 2664
Chang, Y.-Y., van der Wel, A., Rix, H.-W., et al. 2013, ApJ, 773, 149
Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718

9
