The Astrophysical Journal Letters, 937:L33 (7pp), 2022 October 1 https://doi.org/10.3847/2041-8213/ac8e06

Rest-frame Near-infrared Sizes of Galaxies at Cosmic Noon: Objects in JWST's Mirror Are Smaller than They Appeared

Katherine A. Suess[1,2], Rachel Bezanson[3], Erica J. Nelson[4], David J. Setton[3], Sedona H. Price[5], Pieter van Dokkum[6], Gabriel Brammer[7], Ivo Labbé[8], Joel Leja[9,10,11], Tim B. Miller[6], Brant Robertson[1], Arjen van der Wel[12], John R. Weaver[13], and Katherine E. Whitaker[13,14]

[1] Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064 USA; suess@ucsc.edu
[2] Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA
[3] Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA
[4] Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80309, USA
[5] Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstr. 1, D-85748 Garching, Germany
[6] Astronomy Department, Yale University, 52 Hillhouse Ave, New Haven, CT 06511, USA
[7] Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark
[8] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, VIC 3122, Australia
[9] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[10] Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[11] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[12] Astronomical Observatory, Ghent University, Krijgslaan 281, Ghent, Belgium
[13] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[14] Cosmic Dawn Center (DAWN), Denmark

Received 2022 July 21; revised 2022 August 15; accepted 2022 August 21; published 2022 September 28

Abstract

Galaxy sizes and their evolution over cosmic time have been studied for decades and serve as key tests of galaxy formation models. However, at $z \gtrsim 1$ these studies have been limited by a lack of deep, high-resolution rest-frame infrared imaging that accurately traces stellar mass distributions. Here, we leverage the new capabilities of the James Webb Space Telescope (JWST) to measure the 4.4 μm sizes of ∼1000 galaxies with $\log M_*/M_\odot \geqslant 9$ and $1.0 \leqslant z \leqslant 2.5$ from public CEERS imaging in the Extended Groth Strip deep field. We compare the sizes of galaxies measured from NIRCam imaging at 4.4 μm ($\lambda_{\rm rest} \sim 1.6\,\mu$m) with sizes measured at 1.5 μm ($\lambda_{\rm rest} \sim 5500$ Å). We find that, on average, galaxy half-light radii are ∼9% smaller at 4.4 μm than 1.5 μm in this sample. This size difference is markedly stronger at higher stellar masses and redder rest-frame $V-J$ colors: galaxies with $M_* \sim 10^{11}\,M_\odot$ have 4.4 μm sizes that are ∼30% smaller than their 1.5 μm sizes. Our results indicate that galaxy mass profiles are significantly more compact than their rest-frame optical light profiles at cosmic noon, and demonstrate that spatial variations in age and attenuation are important, particularly for massive galaxies. The trend we find here impacts our understanding of the size growth and evolution of galaxies, and suggests that previous studies based on rest-frame optical light may not have captured the mass-weighted structural evolution of galaxies. This paper represents a first step toward a new understanding of the morphologies of early massive galaxies enabled by JWST's infrared window into the distant universe.

Unified Astronomy Thesaurus concepts: Galaxy evolution (594); Galaxy formation (595); Galaxy radii (617)

Supporting material: machine-readable table

1. Introduction

The growth and structural evolution of galaxies over cosmic time provides one of the strongest constraints on theoretical models of galaxy formation, as galaxy sizes are thought to reflect the growth of their host dark matter halos (e.g., Mo et al. 1998). However, measuring the sizes of distant galaxies is very difficult from the ground, as atmospheric seeing is roughly the same size as the half-light radii of galaxies beyond $z \sim 0.5$. Characterizing the size growth of galaxies was therefore one of the primary objectives of the Hubble Space Telescope (HST). One of the most exciting and unexpected discoveries from HST was that quiescent galaxies appeared to grow dramatically with time, more than doubling their sizes between $z \sim 2$ until the present day (e.g., Daddi et al. 2005; van Dokkum et al. 2008; Damjanov et al. 2009; van der Wel et al. 2014). This result provided evidence for the importance of gas-poor minor-merging in the growth of massive, quiescent galaxies (e.g., Bezanson et al. 2009; Hopkins et al. 2009; Naab et al. 2009; van de Sande et al. 2013). HST also revealed that high-redshift disks grow with time, as expected, but that the growth is slower than expected from basic halo growth models, perhaps implying evolving halo spin parameters (e.g., Somerville et al. 2008) or varying effects of feedback processes (e.g., Dutton & van den Bosch 2009). Over a decade of HST observations, the literature has coalesced around three basic rules of thumb: (1) more massive galaxies tend to be larger; (2) galaxies were smaller at cosmic noon ($z \sim 1$–2) than they are in the local universe; and (3) at fixed mass, star-forming galaxies are larger than their quiescent counterparts (e.g., Bell et al. 2012; Bruce et al. 2012; Newman et al. 2012; Barro et al. 2013; Cassata et al. 2013; Lang et al. 2014; van der Wel et al. 2014; Shibuya et al. 2015; van Dokkum et al. 2015; Mowla et al. 2019; Nedkova et al. 2021; Cutler et al. 2022).

However, these empirical structural measurements are fundamentally limited by the fact that even the longest wavelengths observed by HST (1.6 μm) correspond to rest-frame optical light at $z \sim 2$, and shift into the rest-frame ultraviolet at higher redshifts. Local galaxies are known to exhibit radial gradients in their mass-to-light (M/L) ratios caused by gradients in their stellar ages, metallicities, or dust attenuation (e.g., Saglia et al. 2000; La Barbera et al. 2005; Tortora et al. 2010; Greene et al. 2015; Woo & Ellison 2019; Bernardi et al. 2022). If these M/L gradients also exist at earlier time, then light-weighted size measurements at $z \sim 2$ could differ significantly from mass-weighted sizes. This effect is minimized in the rest-frame IR, where *K*-band light can be used as a reliable proxy for stellar mass; however, at bluer wavelengths the range in M/L varies by up to a factor of 10 (Bell & de Jong 2001). Given that the stellar populations within galaxies are often highly inhomogeneous, structures measured from rest-frame optical HST imaging may be fundamentally biased. These color gradients could be subtle, such as metallicity gradients in an elliptical galaxy, or extreme, such as a composite star-forming galaxy with a large bulge comprised of old stars.

Several studies have leveraged color information from multiband HST imaging to approximate mass-weighted structural measurements of distant galaxies (Szomoru et al. 2010; Suess et al. 2019a, 2019b; Mosleh et al. 2020; Miller et al. 2022). Szomoru et al. (2013) finds that luminosity-weighted sizes are indeed biased, as they are in the local universe, but that the effect is small and not strongly dependent on mass or redshift. However, the more recent studies find mass- and redshift-dependent trends, suggesting a less extreme redshift evolution for the structure of massive galaxies. These results call into question just how settled our understanding of galaxy size evolution really is. A more direct understanding of the mass distributions of galaxies requires high-resolution data in the rest-frame infrared: HST-based studies must use stellar population models to infer optical M/L and therefore depend on the assumptions that no optically thick dust is present in the galaxy and that our stellar population models are accurate. Even working at the resolution limit of HST's WFC3 instrument cannot solve these fundamental issues.

The launch and commissioning of the James Webb Space Telescope (JWST) enables a vast leap in resolution at infrared wavelengths, for the first time resolving the rest-frame near-infrared structures of galaxies at cosmic noon. Early imaging data in 4.4 μm from the NIRCam instrument provides nearly mass-weighted light maps of galaxies at $z \sim 1$, in principle enabling the study of the stellar mass distributions of galaxies rather than just their light distributions. In this paper we take a first step in this direction by analyzing images from the Cosmic Evolution Early Release Science (CEERS) program (PI: Finkelstein) to determine whether 4.4 μm sizes are more compact than those measured from rest-frame optical imaging.

This Letter is organized as follows: Section 2 describes the HST and JWST data used in this work and our structural measurements. In Section 3 we compare JWST/NIRCam galaxy sizes at 1.5 and 4.4 μm. Finally, in Section 4 we discuss the implications of these findings and speculate about the exciting structural evolution analysis that JWST will enable at cosmic noon and beyond. We adopt a standard ΛCDM cosmology throughout this paper, with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. All magnitudes are quoted in AB, and we use the term "size" to refer to the major-axis effective radius from a best-fitting Sérsic model to the galaxy light profile.

2. Data and Methods

The NIRCam F444W and F150W imaging was taken as part of the CEERS program (Finkelstein et al. 2017) in the AEGIS field.[15] The F444W observations we use for this project were taken on 2022 June 21–28 with total exposure times of ~1.6–6.3 hr per pointing, covering a total of ~40 sq $'$. We use mosaicked images and weight maps created following the procedure outlined in G. Brammer et al. 2022, (in preparation), based off of the public `grizli` software package (Brammer & Matharu 2021).[16]

We use stellar masses and photometric redshift estimates ("z_best") from the v4.1.5 3D-HST catalog (Brammer et al. 2012; Skelton et al. 2014; Momcheva et al. 2016). Undoubtedly these masses and redshifts will change for some galaxies in this sample given the new long-wavelength photometry now available in these fields; however, for this first-look paper our primary goal is to compare the sizes of these galaxies measured at 1.5 and 4.4 μm. At the most basic level, this test can be performed in on-sky units and is independent of stellar population fitting parameters.

Our sample consists of all galaxies at cosmic noon in the CEERS field that are both massive and sufficiently bright to allow straightforward structural measurements, corresponding to a selection of $1.0 \leqslant z \leqslant 2.5$, $\log M_*/\mathrm{M}_\odot \geqslant 9$, "use_phot" = 1 from the 3D-HST catalogs, and coverage in the initial CEERS imaging mosaic. Following van der Wel et al. (2012), we use an HST/F160W magnitude cut of 24.5 to ensure robust fits. These cuts result in a sample of 1179 galaxies, shown in Figure 1. The "use_phot" flag is to indicate photometry of reasonable quality, avoiding, e.g., stellar diffraction spikes and CCD defects, and thus does not bias the sample (Skelton et al. 2014).

We fit the sizes of all 1179 galaxies in our sample using the `GALFIT` software package (Peng et al. 2002). As inputs, `GALFIT` requires an image, a weight map, and a point-spread function (PSF). Because the centers of most stars are saturated in our CEERS mosaic, creation of an empirical PSF is impractical. We therefore use theoretical PSFs generated using the WebbPSF software (Perrin et al. 2014). By default, WebbPSF generates PSFs at a position angle of zero. Rotating these PSFs to match the position angle of the observations could introduce distortions unless the rotation is performed on an oversampled PSF. Therefore, we use WebbPSF to generate 9×-oversampled F444W and F150W PSFs assuming the same 0.″04 pixel scale of our mosaicked image. We then rotate the oversampled PSFs to the position angle of the CEERS exposures, convolve with a 9×9 square kernel, and downsample back to the 0.″04 mosaic pixel scale.

We choose an F444W cutout size of $100 \times r_{e,\mathrm{F160W}}$ as measured in the van der Wel et al. (2012) HST-based size catalog, enforcing a minimum cutout size of 80×80 pixels and a maximum cutout size of 200×200 pixels and adopting a size of 150×150 pixels for galaxies without existing HST structural measurements. We use standard `astropy` and `photutils` procedures to create a segmentation map of each

[15] All the JWST data used in this paper can be found in MAST, doi:10.17909/7v0n-6041.
[16] https://github.com/gbrammer/grizli

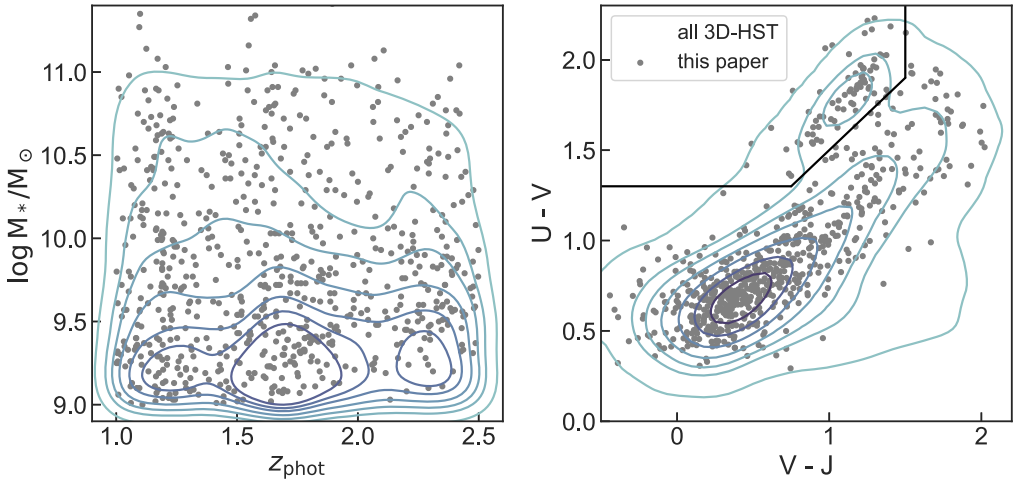

Figure 1. Galaxies in this paper (gray points) compared to all galaxies in the 3D-HST photometric catalog with $\log M_*/M_\odot > 9$ and $1.0 \leqslant z \leqslant 2.5$ (Skelton et al. 2014). Our targets follow a similar stellar mass and photometric redshifts distribution as the full 3D-HST catalog, showing that the subset is fairly representative of the full survey. Due to the limited area of the current CEERS observation, relatively few quiescent galaxies (indicated by the box in *UVJ* color–color space) appear in our sample.

Table 1
Morphological Measurements at 1.5 and 4.4 μm

ID	log$M_\star$	z	r_e (1.5 μm)	n (1.5 μm)	q (1.5 μm)	PA (1.5 μm)	r_e (4.4 μm)	n (4.4 μm)	q (4.4 μm)	PA (4.4 μm)
13225	9.31	2.48	0.56 ± 0.01	2.04 ± 0.07	0.48 ± 0.01	-11.14 ± 0.64	0.67 ± 0.01	1.84 ± 0.11	0.55 ± 0.01	-14.80 ± 1.14
13428	9.84	1.47	2.22 ± 0.03	0.91 ± 0.02	0.65 ± 0.01	14.34 ± 0.86	2.11 ± 0.01	0.73 ± 0.01	0.57 ± 0.00	24.23 ± 0.34
13486	9.70	1.27	3.68 ± 0.03	1.07 ± 0.01	0.29 ± 0.00	-24.58 ± 0.15	3.28 ± 0.01	1.05 ± 0.01	0.31 ± 0.00	-23.57 ± 0.11
13567	9.08	1.49	2.30 ± 0.02	0.44 ± 0.01	0.27 ± 0.00	-55.88 ± 0.23	2.29 ± 0.01	0.39 ± 0.01	0.31 ± 0.00	-56.20 ± 0.24
13598	9.86	1.27	4.39 ± 0.04	0.94 ± 0.01	0.54 ± 0.00	25.63 ± 0.40	3.04 ± 0.01	1.65 ± 0.01	0.45 ± 0.00	20.53 ± 0.12
13626	9.71	1.19	1.11 ± 0.02	2.90 ± 0.06	0.80 ± 0.01	-52.96 ± 1.42	0.60 ± 0.00	2.64 ± 0.04	0.75 ± 0.00	-59.93 ± 0.69
13647	9.05	1.44	2.37 ± 0.05	1.20 ± 0.03	0.22 ± 0.00	-10.21 ± 0.32	2.43 ± 0.02	0.71 ± 0.02	0.25 ± 0.00	-5.99 ± 0.23
13795	10.09	1.28	2.09 ± 0.02	1.60 ± 0.02	0.61 ± 0.00	-44.34 ± 0.43	1.63 ± 0.00	1.66 ± 0.01	0.61 ± 0.00	-47.61 ± 0.20
13842	10.36	1.47	2.81 ± 0.06	2.89 ± 0.05	0.86 ± 0.01	-26.65 ± 2.19	1.25 ± 0.01	1.72 ± 0.03	0.79 ± 0.01	-6.67 ± 1.03
13908	9.74	1.83	2.45 ± 0.02	0.53 ± 0.01	0.30 ± 0.00	20.52 ± 0.21	1.99 ± 0.01	0.79 ± 0.01	0.36 ± 0.00	18.97 ± 0.19
…	…	…	…	…	…	…	…	…	…	…

(This table is available in its entirety in machine-readable form.)

cutout to identify any additional galaxies in the image. We simultaneously model any galaxies that have magnitudes up to 2.5 mag fainter than the target galaxy and have centers within 2″ of the center of the cutout. We mask fainter or more distant galaxies. After performing segmentation, we estimate and subtract off a scalar local background correction using the `photutils` implementation of the SExtractor background subtraction scheme. Finally, we fit each background-subtracted cutout with `GALFIT` to determine the 4.4 μm size of these HST-selected galaxies. We repeat this procedure with F150W cutouts to measure 150 μm sizes. We remove from our sample 400 galaxies where `GALFIT` crashed or returned bad flags—typically, this occurs because one or more of the fit parameters reached our bounds on size or Sérsic index. We additionally flag and remove 15 galaxies where the best-fit `GALFIT` 1.5 or 4.4 μm magnitude differs by >2 mag from the magnitude as measured in our CEERS catalog. This offset typically occurs when our pipeline mistakenly fits a nearby bright galaxy instead of the fainter target galaxy. In total, we provide robust sizes for 703 of the 1179 of the galaxies in our parent sample. The fraction of galaxies successfully fit by our pipeline is slightly higher (∼60%) than the fraction of galaxies with no flags in the van der Wel et al. (2012) HST size catalog (∼50%).

We test the robustness of our measurement pipeline by comparing our measured 1.5 μm sizes to the 1.6 μm sizes measured by van der Wel et al. (2014). While the data depth, PSF size, and exact fitting details differ between these two sets of measurements, the underlying sizes should be relatively similar given the small difference in wavelength. We find that our 1.5 μm sizes are not systematically biased compared to the van der Wel et al. (2014) catalog, with an offset of <0.01 dex and a scatter of ∼0.12 dex. Given the significant differences between the two data sets and the analysis pipeline (e.g., source masking and background subtraction schemes), we believe this agreement is very robust. Based on this test, we adopt a systematic uncertainty on our size measurements.

Table 1 provides our morphological measurements at both 1.5 and 4.4 μm. ID numbers correspond to the v4.1 3D-HST catalog (Skelton et al. 2014). Error bars listed are from `GALFIT` and likely underestimated; systematic error bars on sizes are estimated to be ∼0.15 dex from mock recovery tests. We do not include integrated magnitudes in Table 1 as these are subject to still-evolving knowledge of exact photometric zero-points; our structural parameters and main results are not sensitive to any zero-point offsets.

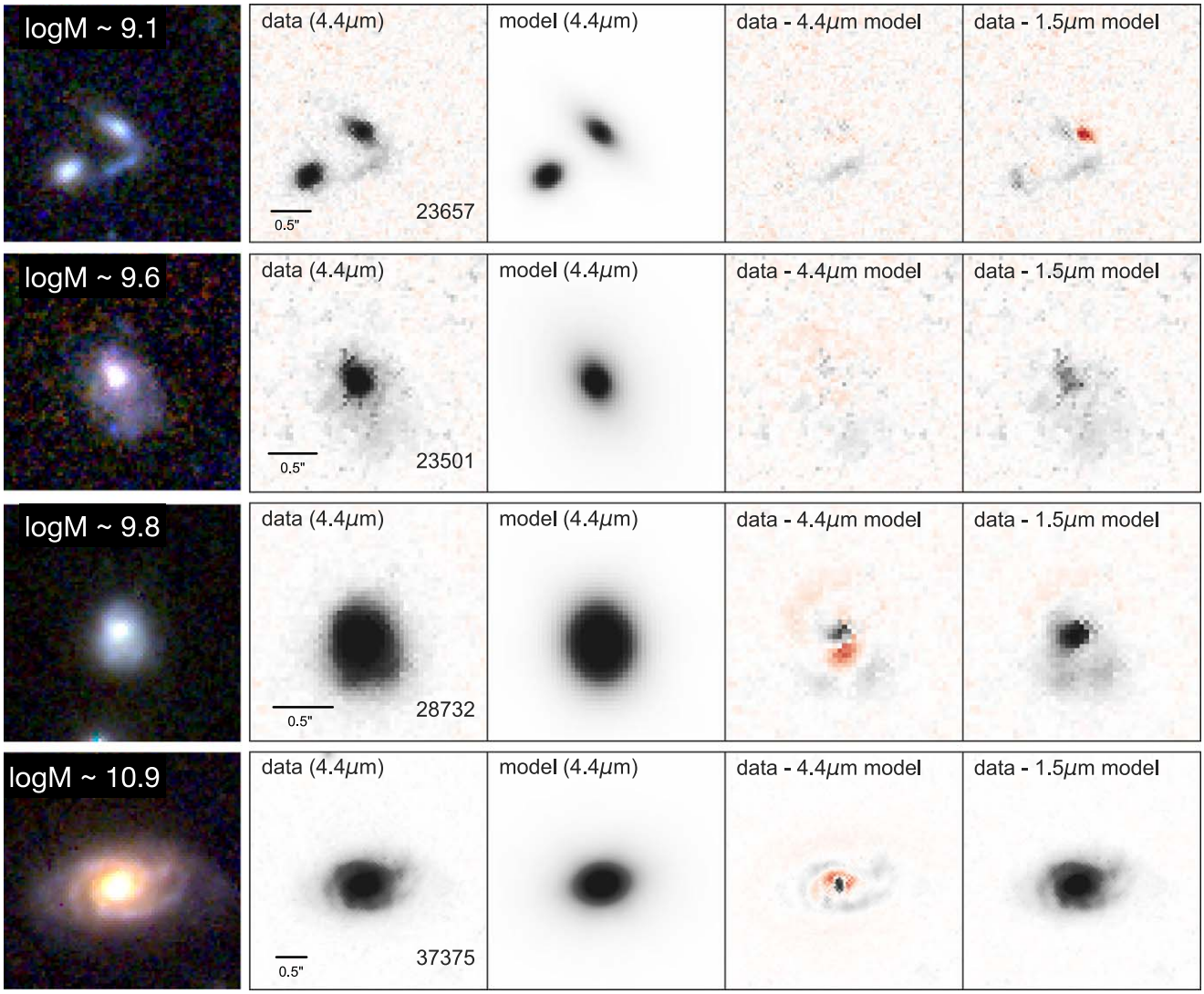

Figure 2. Example GALFIT fits for four objects in our sample, ordered from lowest to highest stellar mass. From top to bottom, the objects lie at redshifts of 1.9, 2.2, 1.2, and 1.5. The left column shows three-color images using JWST/NIRCam F150W, F277W, and F444W. The center three panels show the 4.4 μm data, our best-fit model, and the residual. For comparison, the right panel shows the difference between the 4.4 μm data and the best-fit 1.6 μm GALFIT model convolved with the 4.4 μm PSF. The data, model, and residuals are on the same symmetric colorbar for each object, with positive values in gray scale and negative values in red. Our fitting procedure is able to accurately reproduce galaxy cutouts for both isolated and crowded fields, but cannot capture complex morphologies such as spiral arms. The 1.6 μm model tends to underpredict the flux at the center of the galaxy; this is especially true for the higher-mass galaxies in the bottom rows of the plot.

Figure 2 shows several example fits in 4.4 μm. The left three columns show the 4.4 μm data cutout, model, and residual. The rightmost column highlights the differences in structures between the 4.4 and 1.5 μm imaging. We use GALFIT to generate a Sérsic model with the best-fit parameters of the 1.5 μm model, but convolve it with the 4.4 μm PSF. The figure shows the difference between the 4.4 data and this 1.5 μm fit. In all cases, we see residual flux remaining at the center of the galaxy that is not accounted for by the 1.5 μm model, indicating that 4.4 μm sizes are smaller than 1.5 μm sizes.

We note several caveats before moving on to discuss our results. First, we use a theoretical PSF rather than an empirical one. While this choice is necessary because the centers of stars are clipped/saturated in our mosaic, a theoretical PSF could introduce errors into the estimation of intrinsic galaxy shapes because the theoretical PSF is not a perfect representation of the true PSF in the images. We also do not account for any variations in the PSF across the field of view. Second, while we have subtracted off our best estimate of a scalar sky background, the backgrounds in this early NIRCam imaging retain some structure that may impact the modeling of the wings of the light distribution in galaxies. Third, as can be seen in Figure 2, many galaxies are not well approximated by the simple Sérsic profile with which we model them. Many exhibit, e.g., clumps and spiral arms that are not accounted for in our model and hence affect our parametric fit. One way to solve this in the future may be to move toward nonparametric models (e.g., Miller et al. 2022).

3. Results

In Figure 3, we compare JWST 4.4 μm (F444W) sizes to 1.5 μm (F150W) sizes, measured in arcseconds. These 1.5 μm sizes trace similar stellar populations as previous size measurements in HST/F160W (e.g., van der Wel et al. 2012). We stress that this size comparison does *not* depend

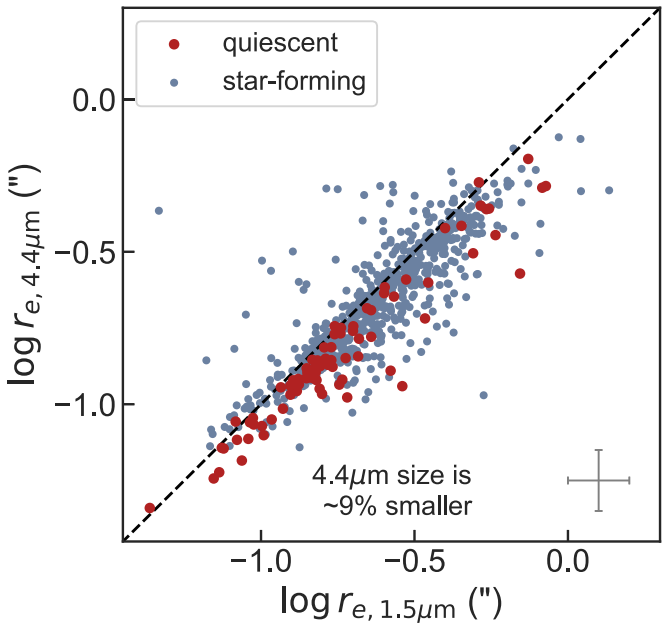

Figure 3. 4.4 μm size as a function of 1.5 μm size in arcseconds, colored by star-forming or quiescent based on rest-frame *UVJ* colors. Galaxies are slightly smaller in 4.4 μm than in 1.5 μm, the longest-wavelength filter previously accessible with HST. The lower right shows a typical error bar estimated from mock recovery tests (see text).

on stellar population modeling or photometric redshift estimates for these galaxies: we are simply comparing the on-sky extent of each galaxy. Points are colored by their designation as either star-forming or quiescent based on their rest-frame *UVJ* colors as measured in the 3D-HST catalog, using the same quiescent definition as Whitaker et al. (2012). Figure 3 shows that galaxies are systematically smaller in 4.4 μm than they appeared in 1.5 μm. Galaxies are 9% smaller on average at 4.4 μm than they appear in 1.5 μm. This systematic offset is a factor of ∼4 larger than the difference between our 1.5 μm sizes and the van der Wel et al. (2012) sizes, and is unlikely to be due to measurement error. With our newfound capacity to resolve the rest-frame near-infrared emission in galaxies at cosmic noon with JWST, we find that galaxies are more compact than they appeared when we could only observe their rest-frame optical emission with HST.

In Figure 4, we explore how the ratio of observed 4.4 μm size to 1.5 μm size changes with stellar mass and rest-frame $V-J$ color. Each panel shows a different redshift slice, because the 4.4 and 1.5 μm filters probe different rest-frame wavelengths in each bin, with rest-frame wavelength ranging between ∼1.3–2 μm for 4.4 μm and ∼4600–6600 Å for 1.5 μm. In all three redshift bins, we see a trend such that redder and more massive galaxies have stronger color gradients. While galaxies are color coded according to their star formation status, due to the limited area of our current study we do not have sufficient numbers of quiescent galaxies to fit the star-forming and quiescent trends separately. We perform a linear fit to the logarithmic size ratio for the entire population using `scipy`'s "curve_fit" function, estimating uncertainties in the best-fit parameters using bootstrapping. We find that the slope of the $\log(r_{e,4.4\,\mu m}/r_{e,1.5\,\mu m})-\log M_*/M_\odot$ trend is significant, $\sim -0.05 \pm 0.02$ in all three redshift bins. This mass effect is strong: at $10^9 M_\odot$, we find that galaxies are almost exactly the same size in 4.4 and 1.5 μm; by $10^{11} M_\odot$, galaxy 4.4 μm sizes are just ∼70% of their 1.5 μm sizes.

4. Discussion and Conclusions

In this Letter, we investigate the effect that new information from JWST—diffraction-limited space-based 4.4 μm imaging—has on the measured sizes of galaxies at cosmic noon. While the size–mass relation was thought to be relatively well understood over this mass and redshift regime (e.g., van der Wel et al. 2014; Mowla et al. 2019; Kawinwanichakij et al. 2021), recent work has called into question whether this view is biased by the rest-optical nature of HST observations (e.g., Suess et al. 2019a, 2019b; Mosleh et al. 2020; Suess et al. 2020; Miller et al. 2022). These new JWST observations allow us to directly measure the rest-frame infrared morphologies of galaxies—a reasonably reliable proxy for their stellar mass distribution—at cosmic noon for the first time, testing whether our previous picture of size evolution was biased by the available data.

We find that galaxies tend to be smaller at 4.4 μm than they are at 1.5 μm (Figure 3). This mirrors results in the local universe (e.g., Kelvin et al. 2012; Lange et al. 2015), and shows that galaxy structures were already complex 8–10 Gyr ago. Strikingly, the effect strongly depends on galaxy properties: more massive galaxies and galaxies with redder $V-J$ colors have stronger color gradients than their less massive, bluer counterparts. On average, $10^9 M_\odot$ galaxies have the same 4.4 and 1.5 μm sizes; by $10^{11} M_\odot$, 4.4 μm galaxy sizes are ∼30% smaller than their 1.5 μm sizes. These results qualitatively agree with previous studies which have used stellar population synthesis modeling to measure galaxy mass profiles: these studies have generally found that galaxy half-mass radii are smaller than half-light radii, with the effect becoming more important at higher stellar masses (Suess et al. 2019a; Miller et al. 2022). However, we emphasize the novel, model-independent nature of this result; all previous studies have relied on stellar population synthesis modeling to infer M/L ratios. This modeling is known to be sensitive to a host of systematic uncertainties due to, e.g., dust–metallicity–age degeneracies and even the initial mass function. The fact that galaxies do indeed appear smaller at 4 μm points to an exciting new era of stellar mass-weighted structural studies with JWST.

The differences we observe between 1.5 and 4.4 μm sizes call into question our understanding of the galaxy size–mass relation. Because more massive galaxies tend to have stronger 4.4/1.5 μm size differences, the inferred slope of the size–mass relation flattens: more massive galaxies may *not* be larger than less massive ones. If the strength of color gradients evolves with redshift (as suggested by Suess et al. 2019a, 2019b), then galaxies at cosmic noon may *not* be significantly smaller than their local counterparts. If color gradients differ between star-forming and quiescent galaxies—beyond the scope of our current limited sample size, but perhaps suggested by the trends with color we observe in Figure 4—then quiescent galaxies may *not* be smaller at fixed mass than star-forming galaxies. All three of our HST-based "rules of thumb" about galaxy sizes may change with JWST's new window into the infrared universe. While this first-look study focused only on the sizes of bright, HST-selected galaxies, our picture of the galaxy size–mass relation will undoubtedly change with JWST.

The implications of our observed size differences go beyond mitigating biases in our understanding of galaxy sizes: they

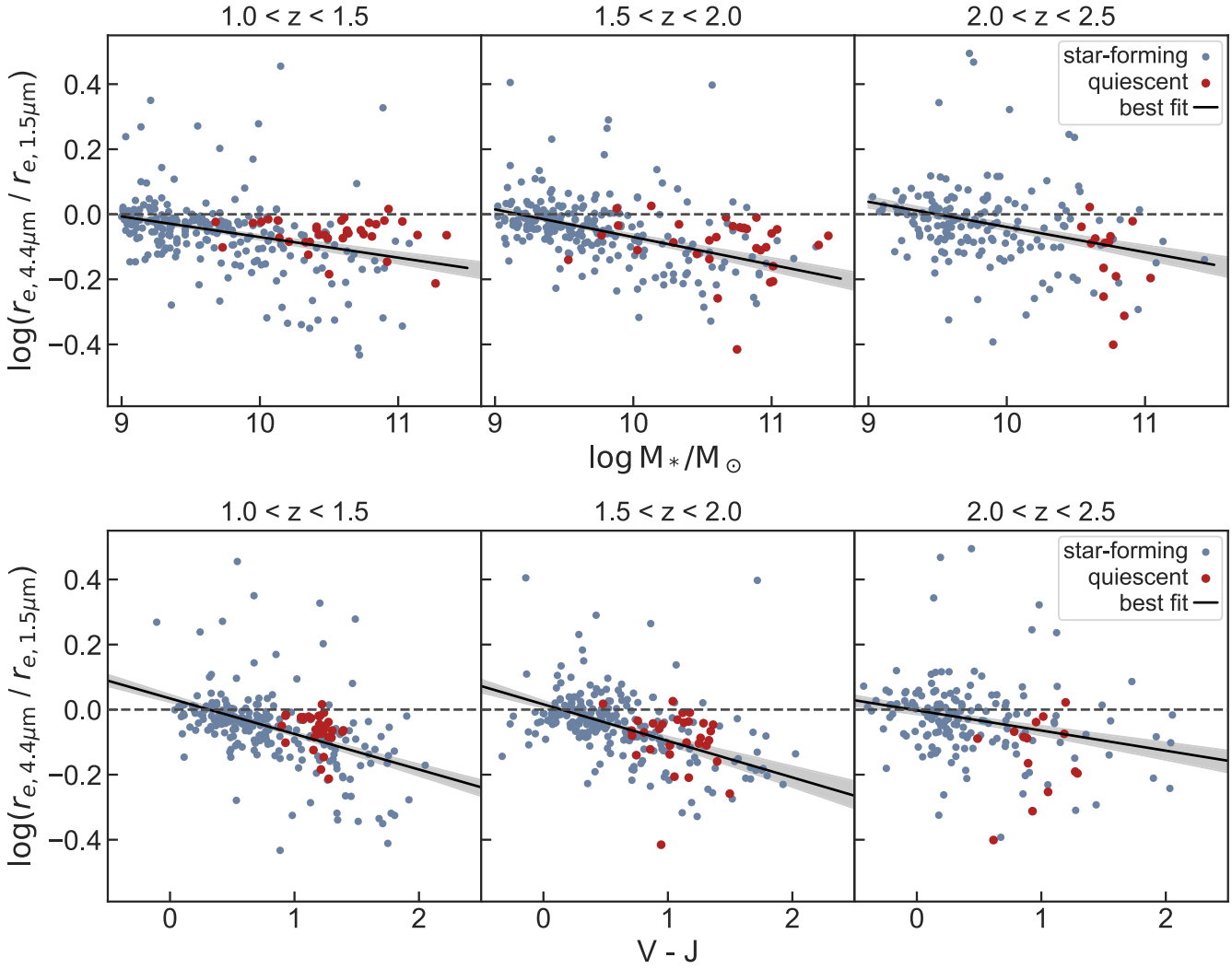

Figure 4. Estimate of color gradient strength as probed by the ratio of 4.4–1.5 μm sizes as a function of both stellar mass and rest-frame $V-J$ color in three redshift bins. Marker color indicates *UVJ*-selected star-forming and quiescent galaxies. We find significant trends with both mass and color, such that redder, more massive galaxies have larger differences between their 4.4 and 4.4 μm sizes. These color gradients indicate that more massive galaxies have increasingly redder centers compared to their outskirts.

allow us to understand *how* galaxies assemble their stellar mass. Differences in morphology at different rest-frame wavelengths can be mapped back to physical quantities—primarily radial variations in age and dust, but also of stellar metallicities. This means that the size differences seen in Figures 3 and 4 can be used to understand variations in stellar population properties of galaxies. Smaller 4.4 μm sizes than 1.5 μm sizes imply that stellar mass profiles are more compact than light profiles, indicating redder centers. These redder centers may be due to dust—previous work at cosmic noon has showed that galaxy centers tend to be more dust-obscured than their outskirts (e.g., Nelson et al. 2016), that more massive galaxies tend to be dustier (e.g., Whitaker et al. 2017), and that galaxies with redder $V-J$ colors are more likely to be edge-on disks with very obscured centers (e.g., Patel et al. 2012). Or, these redder centers may be due to older stellar ages—e.g., these massive disk galaxies may be in the process of assembling the bulge components that we see in massive galaxies in the local universe (e.g., Bezanson et al. 2009, 2011; Nelson et al. 2019; Tadaki et al. 2020). Our finding that even half-light radii, one of the most basic measures we have of galaxy morphologies, differ by up to ∼30% between 4.4 and 1.5 μm in massive galaxies indicates that our previous understanding of the true structures of massive galaxies at cosmic noon was incomplete. Although these differences may seem subtle, they could prove fundamental to our understanding of how galaxies quench and structurally transform. If light-weighted size estimates are indeed systematically biased and star-forming galaxies are closer in size to their quiescent counterparts at cosmic noon, this could alleviate the need for dramatic structural transformation (e.g., Zolotov et al. 2015) or careful progenitor and descendant matching (e.g., van Dokkum et al. 2015).

Moving forward, the remarkable public data sets gathered with JWST can be used to study *why* we see these differences in galaxy sizes across wavelengths. To date, quantitative measurements of age, dust, and metallicity gradients in galaxies have primarily been restricted to the local universe (e.g., Greene et al. 2012, 2015; Woo & Ellison 2019). New JWST multiband infrared imaging, along with a legacy of UV-optical imaging from HST, will allow us to extend spatially resolved stellar population fitting methodologies to quantify age and dust gradients in galaxies at cosmic noon, and to place strong constraints on their underlying stellar mass distributions.

Longer wavelength data with JWST/MIRI or ALMA may additionally help to constrain the sizes of massive star-forming galaxies at cosmic noon (e.g., Franco et al. 2020; Valentino et al. 2020; Gómez-Guijarro et al. 2022). These measurements can be directly compared to predictions from simulations (e.g., Wu et al. 2020; Pathak et al. 2021; Marshall et al. 2022) in order to gain a more complete picture of galaxy growth and assembly.

K.A.S. acknowledges the UCSC Chancellor's Postdoctoral Fellowship Program for support. B.E.R. acknowledges the use of the lux supercomputer at UC Santa Cruz, funded by NSF MRI grant AST 1828315, and support from NASA grants 80NSSC18K0563, 80NSSC22K0814, and the NIRCam science team. 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. E.J.N. acknowledges support from HST-AR-16146. R.B. acknowledges support from the Research Corporation for Scientific Advancement (RCSA) Cottrell Scholar Award ID No: 27587.

Software: astropy (Astropy Collaboration et al. 2013, 2018), GALFIT (Peng et al. 2002), grizli (Brammer & Matharu 2021), Seaborn (Waskom et al. 2017), WebbPSF (Perrin et al. 2014).

ORCID iDs

Katherine A. Suess https://orcid.org/0000-0002-1714-1905
Rachel Bezanson https://orcid.org/0000-0001-5063-8254
Erica J. Nelson https://orcid.org/0000-0002-7524-374X
David J. Setton https://orcid.org/0000-0003-4075-7393
Sedona H. Price https://orcid.org/0000-0002-0108-4176
Pieter van Dokkum https://orcid.org/0000-0002-8282-9888
Gabriel Brammer https://orcid.org/0000-0003-2680-005X
Ivo Labbé https://orcid.org/0000-0002-2057-5376
Joel Leja https://orcid.org/0000-0001-6755-1315
Tim B. Miller https://orcid.org/0000-0001-8367-6265
Brant Robertson https://orcid.org/0000-0002-4271-0364
Arjen van der Wel https://orcid.org/0000-0002-5027-0135
John R. Weaver https://orcid.org/0000-0003-1614-196X
Katherine E. Whitaker https://orcid.org/0000-0001-7160-3632

References

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E., et al. 2013, A&A, 558, A33
Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, ApJ, 765, 104
Bell, E. F., & de Jong, R. S. 2001, ApJ, 550, 212
Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, ApJ, 753, 167
Bernardi, M., Sheth, R. K., Dominguez Sanchez, H., et al. 2022, arXiv:2201.07810
Bezanson, R., van Dokkum, P. G., Franx, M., et al. 2011, ApJL, 737, L31
Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290
Brammer, G., & Matharu, J. 2021, gbrammer/grizli: Release 2021 v.1.3.2, Zenodo, doi:10.5281/zenodo.1146904
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Bruce, V. A., Dunlop, J. S., Cirasuolo, M., et al. 2012, MNRAS, 427, 1666
Cassata, P., Giavalisco, M., Williams, C. C., et al. 2013, ApJ, 775, 106
Cutler, S. E., Whitaker, K. E., Mowla, L. A., et al. 2022, ApJ, 925, 34
Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, ApJ, 695, 101
Dutton, A. A., & van den Bosch, F. C. 2009, MNRAS, 396, 141
Finkelstein, S. L., Dickinson, M., Ferguson, H. C., et al. 2017, (CEERS) Survey, JWST Proposal ID 1345. Cycle 0 Early Release Science, https://www.stsci.edu/jwst/phase2-public/1345.pdf
Franco, M., Elbaz, D., Zhou, L., et al. 2020, A&A, 643, A30
Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022, A&A, 658, A43
Greene, J. E., Janish, R., Ma, C.-P., et al. 2015, ApJ, 807, 11
Greene, J. E., Murphy, J. D., Comerford, J. M., Gebhardt, K., & Adams, J. J. 2012, ApJ, 750, 32
Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, MNRAS, 398, 898
Kawinwanichakij, L., Silverman, J. D., Ding, X., et al. 2021, ApJ, 921, 38
Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, MNRAS, 421, 1007
La Barbera, F., de Carvalho, R. R., Gal, R. R., et al. 2005, ApJL, 626, L19
Lang, P., Wuyts, S., Somerville, R. S., et al. 2014, ApJ, 788, 11
Lange, R., Driver, S. P., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2603
Marshall, M. A., Wilkins, S., Di Matteo, T., et al. 2022, MNRAS, 511, 5475
Miller, T. B., van Dokkum, P., & Mowla, L. 2022, arXiv:2207.05895
Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319
Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27
Mosleh, M., Hosseinnejad, S., Hosseini-ShahiSavandi, S. Z., & Tacchella, S. 2020, ApJ, 905, 170
Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, ApJ, 880, 57
Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJL, 699, L178
Nedkova, K. V., Häußler, B., Marchesini, D., et al. 2021, MNRAS, 506, 928
Nelson, E. J., Tadaki, K.-i., Tacconi, L. J., et al. 2019, ApJ, 870, 130
Nelson, E. J., van Dokkum, P. G., Momcheva, I. G., et al. 2016, ApJL, 817, L9
Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, ApJ, 746, 162
Patel, S. G., Holden, B. P., Kelson, D. D., et al. 2012, ApJL, 748, L27
Pathak, D., Belli, S., & Weinberger, R. 2021, ApJL, 916, L23
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Perrin, M. D., Maire, J., Ingraham, P., et al. 2014, Proc. SPIE, 9147, 91473J
Saglia, R. P., Maraston, C., Greggio, L., Bender, R., & Ziegler, B. 2000, A&A, 360, 911
Shibuya, T., Ouchi, M., & Harikane, Y. 2015, ApJS, 219, 15
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, MNRAS, 391, 481
Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019a, ApJ, 877, 103
Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019b, ApJL, 885, L22
Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2020, ApJL, 899, L26
Szomoru, D., Franx, M., van Dokkum, P. G., et al. 2010, ApJL, 714, L244
Szomoru, D., Franx, M., van Dokkum, P. G., et al. 2013, ApJ, 763, 73
Tadaki, K.-i., Belli, S., Burkert, A., et al. 2020, ApJ, 901, 74
Tortora, C., Napolitano, N. R., Cardone, V. F., et al. 2010, MNRAS, 407, 144
Valentino, F., Tanaka, M., Davidzon, I., et al. 2020, ApJ, 889, 93
van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85
van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJL, 677, L5
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
Waskom, M., Botvinnik, O., O'Kane, D., et al. 2017, Mwaskom/Seaborn: V0.8.1 (September 2017), Zenodo, doi:10.5281/zenodo.883859
Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, ApJ, 850, 208
Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29
Woo, J., & Ellison, S. L. 2019, MNRAS, 487, 1927
Wu, X., Davé, R., Tacchella, S., & Lotz, J. 2020, MNRAS, 494, 5636
Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327