

EPOCHS VI: the size and shape evolution of galaxies since $z \sim 8$ with *JWST* Observations

K. Ormerod¹,¹★ C. J. Conselice,¹ N. J. Adams¹,¹ T. Harvey¹,¹ D. Austin,¹ J. Trussler,¹ L. Ferreira,² J. Caruana,^{3,4} G. Lucatelli,¹ Q. Li¹ and W. J. Roper⁵

¹Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK

²Department of Physics & Astronomy, University of Victoria, Finnerty Road, Victoria, British Columbia, V8P 1A1, Canada

³Department of Physics, University of Malta, Msida MSD 2080, Malta

⁴Institute of Space Sciences & Astronomy, University of Malta, Msida MSD 2080, Malta

⁵Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

Accepted 2023 November 17. Received 2023 November 17; in original form 2023 September 11

ABSTRACT

We present the results of a size and structural analysis of 1395 galaxies at $0.5 \leq z \leq 8$ with stellar masses $\log(M_*/M_\odot) > 9.5$ within the *James Webb Space Telescope* Public CEERS field that overlaps with the *Hubble Space Telescope* Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey EGS observations. We use GALFIT to fit single Sérsic models to the rest-frame optical profile of our galaxies, which is a mass-selected sample complete to our redshift and mass limit. Our primary result is that at fixed rest-frame wavelength and stellar mass, galaxies get progressively smaller, evolving as $\sim(1+z)^{-0.71 \pm 0.19}$ up to $z \sim 8$. We discover that the vast majority of massive galaxies at high redshifts have low Sérsic indices, thus do not contain steep, concentrated light profiles. Additionally, we explore the evolution of the size–stellar mass relationship, finding a correlation such that more massive systems are larger up to $z \sim 3$. This relationship breaks down at $z > 3$, where we find that galaxies are of similar sizes, regardless of their star formation rates and Sérsic index, varying little with mass. We show that galaxies are more compact at redder wavelengths, independent of sSFR or stellar mass up to $z \sim 3$. We demonstrate the size evolution of galaxies continues up to $z \sim 8$, showing that the process or causes for this evolution is active at early times. We discuss these results in terms of ideas behind galaxy formation and evolution at early epochs, such as their importance in tracing processes driving size evolution, including minor mergers and active galactic nuclei activity.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: structure.

1 INTRODUCTION

Ever since the discovery of galaxies, their extended nature has been a clear indicator of their different properties from unresolved sources such as stars. These resolved properties have been of interest for many years, and in many ways are the oldest studied properties of galaxies (e.g. Hubble 1926; Buitrago et al. 2008; Conselice 2014; van der Wel et al. 2014; Kartaltepe et al. 2022; Suess et al. 2022; Ferreira et al. 2022a, b). As far back as the 18th century, the Herschels catalogued what we now know to be extragalactic objects, commenting on their appearance (Herschel 1786). Following up on this, Lord Rosse discovered spiral structures in galaxies, through his observations of M51 and other nearby galaxies (Rosse 1850). After inferring the distances to these objects, photography allowed Hubble and his immediate successors to develop the dominant morphological classification scheme we use today – the Hubble Sequence – which classifies extragalactic objects as spiral, elliptical, or irregular (Hubble 1926). This has continued until this day, with *James Webb Space Telescope* (*JWST*) observations setting us on a

new pathway towards understanding the structures and morphologies of the very first galaxies (e.g. Whitney et al. 2021; Kartaltepe et al. 2022; Suess et al. 2022; Ferreira et al. 2022a, b; Huertas-Company et al. 2023; Jacobs et al. 2023; Morishita et al. 2023; Ono et al. 2023; Tacchella et al. 2023).

The study of galaxy structure and morphology are amongst the oldest subjects within the extragalactic field, and continue to hold significant importance in our quest to understand the evolutionary processes of galaxies over cosmic time, however, our knowledge about these features remains limited at $z > 3$. Tracking the changes in the structural properties of galaxies from the era of early galaxy formation until the present-day provides valuable insights into the processes of galaxy evolution. There have been major efforts over the past 30 yr to study morphology and structure of distant galaxies with the *Hubble Space Telescope* (*HST*; Buitrago et al. 2008; Delgado-Serrano et al. 2010; Conselice 2014), where the rest-frame optical properties of galaxies up to $z \sim 3$ can be studied and examined. These *HST* observations have shown us that galaxies appear to become progressively more irregular and peculiar at higher redshifts (e.g. Conselice 2003; Lotz, Primack & Madau 2004; Delgado-Serrano et al. 2010; Mortlock et al. 2013; Conselice 2014; Schawinski et al. 2014; Whitney et al. 2021).

* E-mail: katherineormerod@gmail.com

However, the red Hubble filters are limited, and do not allow us to measure or observe the rest-frame optical light of galaxies back to within the first few Gyr of the universe. In fact, the reddest *HST* filter, *F160W* on Hubble’s Wide Field Camera 3 (WFC3), only probes the rest-frame visible light of galaxies up to $z \sim 2.8$, but galaxies exist at much higher redshifts, and this paper aims to probe their evolution (e.g. Atek et al. 2022; Castellano et al. 2022; Donnan et al. 2022; Naidu et al. 2022; Yan et al. 2022; Austin et al. 2023; Curtis-Lake et al. 2023; Finkelstein et al. 2023; Harikane et al. 2023; Adams et al. 2023b). *JWST* is the best approach for measuring galaxies in the rest-frame optical where the effects of dust are limited. We can see why this longer wavelength data is needed through the use of cosmological simulations. These simulations predict that at higher redshift, galaxies in the absence of the effects of dust are intrinsically more compact resulting in a negative far-UV size–luminosity relation. However, when simulated with dust, the bright cores of these galaxies are attenuated, increasing the observed half-light radius. This brings the FUV size–luminosity relations in line with observations. Simulations, in agreement with observations, have also shown that galaxy sizes typically increase with increasing stellar mass, and show that compact galaxies may grow in size due to mergers or renewed star formation, with high-redshift stars moving outwards (e.g. Furlong et al. 2017; Marshall et al. 2022; Roper et al. 2022).

The relatively recently launched *James Webb Space Telescope* allows us to obtain the same type of data with the Near Infrared Camera (NIRCam) probing rest-frame optical light as far out as $z \sim 9$, with filters reaching up to $\sim 4.4 \mu\text{m}$. The superior resolution of *JWST* and the extended wavelengths of its filters allow us to examine galaxy structure in significantly greater detail than with *HST* (Ferreira et al. 2022b). Early *JWST* observations reveal morphological and structural features of galaxies at $1.5 < z < 8$ that were not possible to discern fully with *HST*, thus resulting in the re-classification of many galaxies previously believed to have peculiar morphologies (e.g. Ferreira et al. 2020). A major discovery has been that galaxies appear morphologically much more disc-like than previously thought (e.g. Ferreira et al. 2022a). While these early *JWST* papers show that galaxies at $z > 2$ are different than we thought on the basis of *HST* imaging, a significant amount of quantitative analysis is still needed.

As such, in this paper, we present an analysis of the sizes and Sérsic indices of a mass complete sample of massive galaxies with stellar masses $\log(M_*/M_\odot) > 9.5$, for which we can now quantitatively measure rest-frame structure up to $z \sim 8$ with *JWST*. Whilst previously we could also measure galaxy sizes and morphologies with *HST*, these are often unreliable due to image fidelity and the range of wavelengths we were able to probe (Ferreira et al. 2022b). Our results at $0.5 < z < 8$ allow us to probe deeper and at higher redshifts than previous studies.

We present a quantitative analysis of measured galaxy shapes, based on the Sérsic index, n , obtained from Sérsic profile fitting, and size measurements, based on half-light radii measurements, to determine the evolution of galaxy structure over most of cosmic time. Whilst the Sérsic profile is one of the simplest forms that can be fit to a light profile, this initial analysis is important as it enables a continuation of work completed with *HST*, and allows a direct comparison of the results at different redshift ranges. It should be noted that a Sérsic profile will not model the full detail of some galaxies, although the galaxies within our sample are not resolved enough to see the very core regions or to show kinks that are seen in surface brightness profiles of very nearby galaxies. In this paper, we present an initial analysis of Sérsic profile fitting with *JWST*, to answer how galaxy sizes and their overall shapes change for a mass-

Table 1. The 5σ depths of the *HST* photometric data covering the EGS, for full details see Stefanon et al. (2017).

Filter name	Depth
<i>HST/ACS F606W</i>	28.8
<i>HST/ACS F814W</i>	28.2
<i>HST/WFC3 F125W</i>	27.6
<i>HST/WFC3 F140W</i>	26.8
<i>HST/WFC3 F160W</i>	27.6

complete sample, and complete further analysis on sub-samples based upon two properties: specific star formation rate (sSFR), and Sérsic indices. We separate our sample into sub-samples based on sSFR to investigate the evolution of Sérsic index in the more star-forming and passive populations. We also separate our sample into sub-samples with a high and low Sérsic index, with a separating value of $n = 2$, as a proxy for elliptical- and disc-like galaxies, and to investigate their sizes across cosmic time. This differs from previous work which has focused on either just the highest redshift galaxies (e.g. Ono et al. 2023) or those which are passive (e.g. Ito et al. 2023). The evolution of these properties within a stellar mass selection is a key observable for galaxy evolution as well as an important way to trace processes that drive galaxy size evolution in massive galaxies, including galaxy minor mergers and active galactic nuclei (AGNs) activity (e.g. Bluck et al. 2012).

The structure of this paper is as follows: Section 2 discusses the data used and the data-reduction process, Section 3 discusses properties of the galaxies within our sample. The fitting process is explained in Section 4, and we discuss how we select a sample of robust morphological fits in Section 5. We present our results, along with a comparison to simulations to confirm that our findings are not the result of redshift effects in Section 6. We assume a standard Lambda cold dark matter cosmology throughout of $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Where we reference galaxy ‘sizes’, we are referring to the half-light radii of our objects. All magnitudes are given in the AB system (Oke 1974; Oke & Gunn 1983).

2 DATA

We use *JWST* NIRCam imaging (Rieke et al. 2022) to analyse the light profiles of a large sample of high-redshift galaxies in the *F115W*, *F150W*, and *F200W* short-wavelength (SW) bands, and *F277W*, *F356W*, *F410M*, and *F444W* long-wavelength (LW) bands. The Cosmic Evolution Early Release Science (CEERS, PID: 1345, PI: S. Finkelstein) Survey (Finkelstein et al. 2017, 2022; Bagley et al. 2023) is one of 13 *JWST* ERS programmes, with the goal of examining galaxy formation at $0.5 < z < 10$ and perhaps beyond. The galaxies analysed in this work are within the CEERS NIRCam footprint, and within the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) observations (Koekemoer et al. 2011). This is important as it allows us to use both the *JWST* data as well as the deep data from *HST*’s WFC3 and Advanced Camera for Surveys (ACS), which also aids the determination of photometric redshifts. As such, the photometric redshifts, stellar masses, and star formation rates used in this paper are based on the original CANDELS + GOODS WFC3/ACS imaging and data, *Spitzer*/IRAC S-CANDELS observations (Ashby et al. 2015), and Canada–France–Hawaii Telescope ground-based observations (Stefanon et al. 2017). A summary of *HST* filters used and their 5σ depths is shown in Table 1.

Table 2. Average 5σ depths in our reduced CEERS images, for point sources in 0.32 arcsec diameter apertures.

Filter	Depth
<i>F115W</i>	28.75
<i>F150W</i>	28.60
<i>F200W</i>	28.80
<i>F277W</i>	28.95
<i>F356W</i>	29.05
<i>F410M</i>	28.35
<i>F444W</i>	28.60

Notes. This aperture size is chosen as it encloses the central 70–80 per cent of the flux of a point source, and is small enough to avoid contamination. Depths are calculated by placing random apertures in regions of the image that are empty based on the final segmentation maps.

The CANDELS survey was designed to investigate galaxy evolution and the birth of black holes at $1.5 < z < 8$, and consists of multiwavelength observations in five fields. The CANDELS/DEEP survey covers 125 arcmin^2 within GOODS-N and GOODS-S, and the remainder consists of the CANDELS/Wide survey, covering three additional fields (Extended Groth Strip, COSMOS, Ultra-Deep Survey), covering a total of 800 arcmin^2 across all fields (Grogin et al. 2011). The primary sample we select originates from these deep observations within the EGS, where there is overlap with the CEERS *JWST* NIRCcam data, and whose analysis is described in detail in Duncan et al. (2019).

2.1 CEERS data reduction

We process the *JWST* data products on this field using a modified version of the official *JWST* pipeline, explained in depth in Adams et al. (2023a, 2023b). We use the standard *JWST* pipeline (pipeline version 1.8.2 and Calibration Reference Data System v0995), with some minor modifications. Between Stage 1 and Stage 2, we subtract templates of ‘wisp’ artefacts from the *F150W* and *F200W* data.¹ After Stage 2 of the pipeline, we apply a correction for 1/F noise, derived by Chris Willott.² We extract the sky subtraction step from Stage 3 of the pipeline and run this on each NIRCcam frame independently. We then align calibrated imaging for each exposure to GAIA Data Release 3 (Gaia Collaboration et al. 2021), using `tweakreg` from the DRIZZLEPAC PYTHON³ package. We finally pixel-match the final mosaics using `astropy reproject`.⁴ The final resolution of our drizzled images is 0.03 arcsec per pixel. The total unmasked area of *JWST* images used in this paper is 64.15 arcmin^2 , and the average depths of each filter are listed in Table 2.

3 GALAXY PROPERTIES

3.1 Photometric redshifts, stellar masses, and star formation rates

We begin our analysis with a catalogue of 1649 massive objects with $\log(M_*/M_\odot) > 9.5$, which have photometric redshifts and physical properties calculated in previous works. The photometric redshifts used in this paper are calculated in Duncan et al. (2019), using

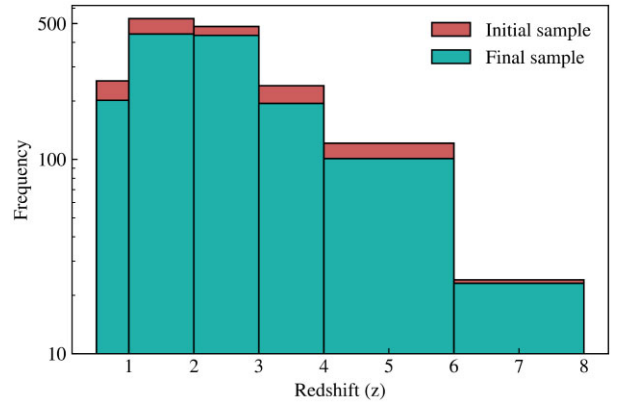


Figure 1. Histogram showing the distribution of photometric redshifts within our redshift bins. The ‘Initial sample’ is the sample of 1649 high-mass galaxies we begin the fitting process with, and the ‘Final sample’ represents the galaxies that are selected as good fits in Section 5.

the EAZY (Brammer, van Dokkum & Coppi 2008) photometric redshift code, with three separate template sets fitted to the observed photometry. These templates include zero-point offsets, which alter the input fluxes, and fix additional wavelength-dependent errors (see Duncan et al. 2018a, b for full details). A Gaussian process code (Almosallam, Jarvis & Roberts 2016) is used to measure further empirical estimates, using a subset of the photometric bands. Individual redshift posteriors are calculated, and all measurements are combined in a statistical framework via a Bayesian combination to give a final estimate of redshift. From a comparison with spectroscopic redshifts, for sources where spectroscopic measurements are available, these photometric redshifts are seen to have a high degree of accuracy; for full details, see section 2.4 of Duncan et al. (2019). We show the photometric redshift distribution of our sample in Fig. 1 for both our initial and final samples. We note that there is a significant decrease in the number of galaxies in the highest redshift bin, although this is not unexpected due to the catalogue being compiled from previous *HST* works, thus *HST* dark galaxies are not included in this sample.

The stellar masses we use are those measured in Duncan et al. (2014, 2019), using a custom template-fitting code (see Section 4 of Duncan et al. 2014). With this custom spectral energy distribution (SED) fitting code, the stellar mass is measured at all redshifts within the photometric redshift-fitting range. The masses also have a ‘template error function’, described in Brammer, van Dokkum & Coppi (2008), accounting for uncertainties driven by the template set and wavelength effects. These stellar mass measurements assume a BC03 stellar population synthesis model (Bruzual & Charlot 2003), with a wide range of stellar population parameters, and a Chabrier (2003) initial mass function (IMF). The star formation histories used within these fits follow the form $\text{SFR} \propto e^{-t/\tau}$, with time-scales of $|\tau| = 0.25, 0.5, 1, 2.5, 5, 10$, where negative values of τ represent exponentially increasing histories. A short burst model is also used ($\tau = 0.05$), as well as continuous star formation models ($\tau = 1/H_0$). Nebular emission is also included in the model SEDs assuming an escape fraction of $f_{\text{esc}} = 0.2$. In order to ensure that our stellar masses do not suffer from systematic biases, they are compared to stellar masses calculated independently within the CANDELS collaboration (Santini et al. 2015). While there is some scatter between the two mass estimates, there is no significant bias (see section 2.5 of Duncan et al. 2019 for a detailed discussion). We aim to calculate masses using *JWST* data for galaxies at $z > 4.5$ in Harvey et al. (in preparation).

¹<https://jwst-docs.stsci.edu/jwst-near-infrared-camera/nircam-instrument-features-and-caveats/nircam-claws-and-wisps>

²<https://github.com/chriswillott/jwst>

³<https://github.com/spacetelescope/drizzlepac>

⁴<https://reproject.readthedocs.io/en/stable/>

Star formation rates for our sample of galaxies are calculated using the UV slope (β) of the SED, which gives a measure of the dust attenuation within the galaxy. We aim to measure this with *JWST* within this EPOCHS paper series (Austin et al., in preparation). From this, we correct for dust and obtain the total star formation rates for our galaxies, which agree well with star formation rates derived directly from SED fitting (Duncan et al. 2014, 2019). We avoid degeneracies between the stellar masses and star formation rates by deriving the star formation rates using this separate method. The SFR and stellar mass estimates used in this work are calculated in Duncan et al. (2014, 2019), and are compared to several other works within the CANDELS collaboration which find no significant biases. The observed scatter found is due to the photo- z estimates used not being exactly the same in each paper, and is most significant at $\log(M_*/M_\odot) < 9$, which is not within the mass range of this paper.

3.2 Visual classifications

For 470 objects within our sample, we use visual classifications from Ferreira et al. (2022a) as part of our analysis. The categories we use are as follows:

- (i) *Discs*: Sources with a resolved disc with an outer area of lower surface brightness, that regularly increases towards the centre of the galaxy.
- (ii) *Spheroids*: Resolved sources that are symmetrical, with a centrally concentrated, smooth light profile, that are round or elliptical.
- (iii) *Peculiar*: Resolved sources with a disturbed morphology, which dominates any smooth components.
- (iv) *Other*: Mainly made up of sources classified as ‘ambiguous’, due to the classifiers not reaching a majority agreement on the classification of a source. This category also contains sources that are classified as point sources due to an angular size smaller than the full-width half maximum (FWHM) of the point spread function (PSF), or clear spikes that are consistent with point sources, and any sources that were unable to be classified due to faintness or image issues.

4 MORPHOLOGICAL FITTING

We use GALFIT version 3.0.5 (Peng et al. 2002, 2010) to fit a single Sérsic light profile to each galaxy. Ultimately, the overall goal with *JWST* is to measure the light profiles in more detail, such as obtaining bulge to disc ratios and other features (Margalef-Bentabol et al. 2016). However, it is important to first determine structural properties using single profile fitting, and assess how they characterize the data (e.g. van der Wel et al. 2012, 2014; Suess et al. 2022).

GALFIT is a least-squares-fitting algorithm which finds the optimum solution to the surface brightness profiles for galaxies through using a Levenberg–Marquardt algorithm. GALFIT uses the reduced chi-squared, χ^2_v , to determine goodness-of-fit and finds the best-fitting model through χ^2_v minimization. The χ^2_v is given by

$$\chi^2_v = \frac{1}{N_{\text{DOF}}} \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} \frac{(f_{\text{data}}(x, y) - f_{\text{model}}(x, y))^2}{\sigma(x, y)^2} \quad (1)$$

summed over n_x and n_y pixels, and where N_{DOF} is the number of degrees of freedom. As seen in equation 1, GALFIT requires a data image from which the galaxy surface brightness is measured, $f_{\text{data}}(x, y)$ and a sigma image, $\sigma(x, y)$, giving the relative error at each position within the image, which are then used to calculate the model image, $f_{\text{model}}(x, y)$.

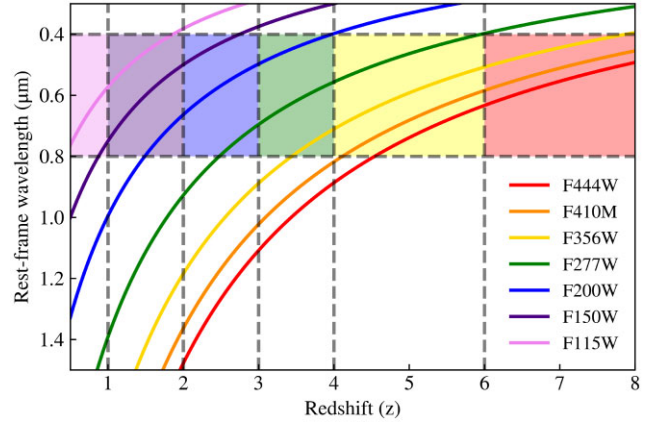


Figure 2. Plot showing the rest-frame wavelengths at given redshifts, for all filters used within the *JWST* CEERS NIRCcam observations, and used within this paper. The shaded regions show the selected filter we use to observe sources in the rest-frame optical at the given redshift. The grey-dashed lines show where the filter we use to provide a rest-frame optical view of galaxies at different redshift changes.

We run GALFIT for all available filters, but only report results here for the filters that best match the rest-frame optical wavelength of the source. This minimizes, or even eliminates, the effect of morphological k -correction, as the qualitative and quantitative structure of galaxies changes as a function of wavelength (Taylor-Mager et al. 2007), which can result in significant structural changes between rest-frame UV and rest-frame optical images. This also enables us to limit the CMOD effect (Papaderos, Östlin & Breda 2023), by probing rest-frame optical wavelengths, which is possible up to $z \sim 8$ with *JWST*. The band selected at a given redshift is shown in Fig. 2. The figure shows which filter we use within different redshift ranges, and what rest-frame wavelength we probe within that filter at that redshift. As can be seen, we are always probing the rest-frame optical at wavelengths redder than the Balmer break at all epochs in which we view our galaxy sample.

The Sérsic profile we use has the form

$$I(R) = I_e \exp \left\{ -b_n \left[\left(\frac{R}{R_e} \right)^{1/n} - 1 \right] \right\}, \quad (2)$$

where $I(R)$ is the intensity at a distance R from the centre of the galaxy, R_e is the half-light radius of galaxy (the radius where 50 per cent of the total luminosity is enclosed), I_e is the intensity at the half-light radius, n is the Sérsic index, which controls the shape of the light profile of the galaxy (Sérsic 1963; Ciotti 1991; Caon, Capaccioli & D’Onofrio 1993), and b_n can be approximated as $b(n) \approx 2n - \frac{1}{3} + \frac{4}{405n} + \frac{46}{25515n^2}$ (Ciotti & Bertin 1999). GALFIT gives us a best-fitting value for each of these terms. The errors on these values are also calculated through this method, and a full description of the GALFIT error calculation can be found in Peng et al. (2002).

4.1 GALFIT pipeline

We use a custom pipeline for single-component Sérsic fits with GALFIT. The process is as follows:

- (i) *Source detection*: We use SExtractor (Bertin & Arnouts 1996) to detect sources within the *F444W* images for each CEERS pointing, following the parameters and method in Adams et al. (2023b). These catalogues are then cross-matched within 1 arcsec to the catalogues created from the analysis completed in Duncan

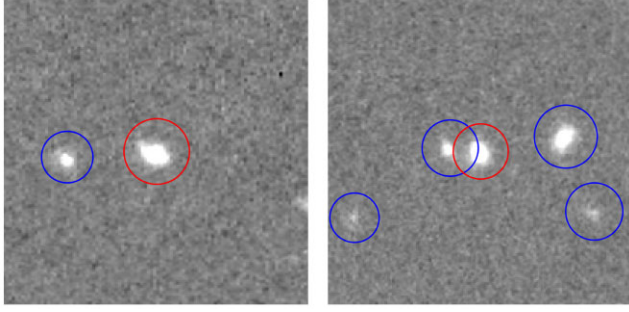


Figure 3. Plots showing the Kron radii (semimajor axis of the Kron ellipse), where the red radius in each image is that of the primary source, and the blue radii are those of the neighbouring sources. We give several scenarios for how these systems would be found. Left panel: No other sources would be fit simultaneously to the primary, and the pixels belonging to all other objects according to the SExtractor segmentation maps are masked. Right panel: The source where the Kron radius overlaps that of the primary source are simultaneously fit in this instance, and all other sources would be masked. Cutout sizes shown are 6×6 arcsec.

et al. (2019) to create our final catalogues for each pointing. The average separation between both catalogues is ~ 0.15 arcsec.

(ii) *Cutout and mask creation:* We create 200×200 pixel (6×6 arcsec) cutouts of each source, in order to ensure the entire surface brightness profile of the galaxy is enclosed within the cutout, along with that of any neighbours that may need to be modelled simultaneously. We use the SExtractor segmentation maps to make masks, creating the same 200×200 pixel cutout of the segmentation map, and then masking the necessary objects.

In order to create masks for each object, and select which neighbouring objects must be masked, we use the Kron ellipses (Kron 1980), as defined by SExtractor and plot circular apertures with a radius equal to the semimajor axis of the Kron ellipse. We do this to select galaxies that are sufficiently close enough for their surface brightness profile to interfere with that of the target, ‘primary’, object and must be fit simultaneously. If any neighbouring galaxy has an overlapping Kron aperture with that of the primary object, the neighbouring galaxy is deemed to be sufficiently close that it must be modelled alongside the primary galaxy, to account for both light profiles. We do this for as many neighbouring objects as necessary. Objects that do not have an overlapping Kron aperture are far enough away that they do not also need to be fit, and therefore are masked instead, primarily to save computational time. Through visual inspection of fits and residuals of fits from the data, we conclude that this criterion is good for determining when to fit neighbouring objects. The pixels that are masked are those that the SExtractor segmentation maps assign to each source. An example of these selection criteria is shown in Fig. 3.

GALFIT also requires a sigma image to give relative weight to the pixels in the image during the fit. As the input sigma image, we use the ‘ERR’ extension of the images, which is a measure of the noise of the image, and this is created using the same method as the object cutout images.

(iii) *Input parameters:* In order to fit a single Sérsic profile, initial estimates of parameters must be provided to GALFIT. The input parameters that GALFIT requires estimates for are x and y image coordinates, total magnitude, half-light radius, axis ratio, position angle, and Sérsic index. Similarly to Kartaltepe et al. (2022), we use the SExtractor catalogue for our initial parameter estimates. As these catalogues do not contain values of Sérsic index, we estimate this as $n = 1$ initially and allow the fitting process to determine the

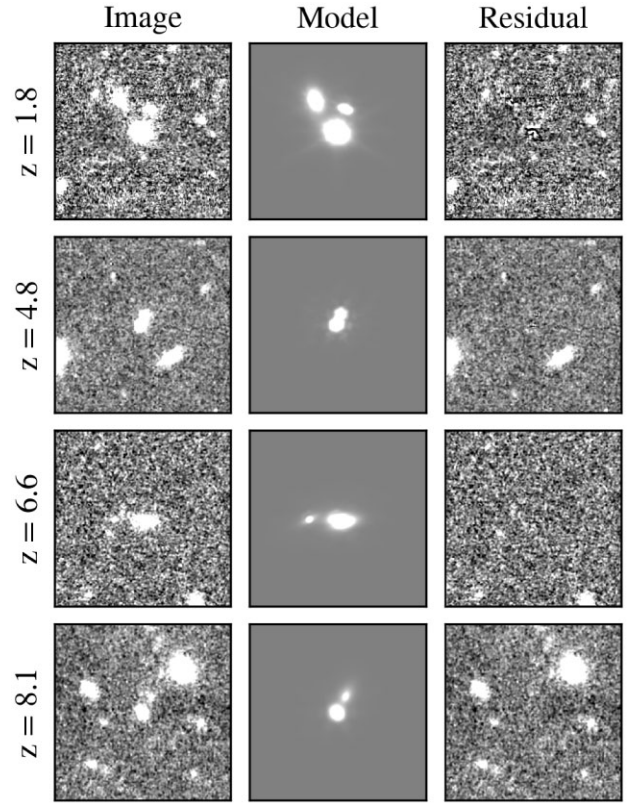


Figure 4. Example fits showing the data image, model image, and the residual image (data–model). Each cutout is 6×6 arcsec. The redshift of each galaxy is shown to the left of the images.

best-fitting Sérsic index. We find that using different values of n as the initial estimate have virtually no effect on the output parameters. We only apply constraints to the image position of the sources within ± 2 pixels to ensure the correct source is being fit.

(iv) *PSF:* GALFIT requires the appropriate PSF for each filter, which is obtained using WebbPSF (Perrin et al. 2014), and resampled to our pixel scale. We experimented with different PSFs created through this method and find that the results do not significantly change.

Although the central position of the source is constrained, all other parameters are allowed to vary freely, and a selection process is used to select good fits with physical parameters after fitting is complete, as the overuse of constraints can lead to GALFIT converging on unphysical results. Example fits can be seen in Fig. 4.

4.2 Comparison with IMFIT

IMFIT is an alternative light profile-fitting program which uses Levenberg–Marquardt, Nelder–Mead, and differential evolution algorithms to find the best-fitting parameters (Erwin 2015). In order to test the robustness of our method, we present a comparison of best-fitting half-light radii and Sérsic indices for a representative sample of 146 objects in CEERS Pointing 1, which have GALFIT fits that meet the selection criteria explained in Section 5 and have a stellar mass of $\log(M_*/M_\odot) > 9.5$. We run IMFIT using the same input parameters and initial guesses as those used for GALFIT and again use the Levenberg–Marquardt algorithm for consistency. The results of this comparison are shown in Fig. 5a and b, which plot the half-light radii and Sérsic indices measured for these 146 objects

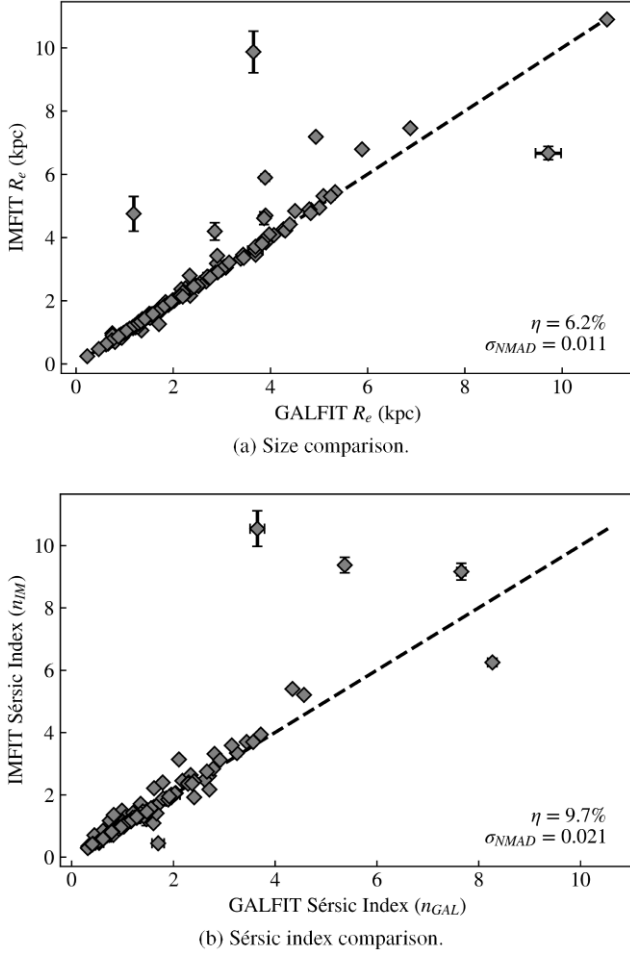


Figure 5. A comparison of two measures – size (top panel) and Sérsic index (bottom panel) – as obtained with GALFIT and IMFIT. The black line shows the one-to-one relation. Size and Sérsic index are generally in good agreement, with more variation at larger values, in particular, with the Sérsic index. We show the outlier rate (η) and NMAD in the bottom right corner of each figure.

using GALFIT and IMFIT, showing a good agreement between the measured values.

We use two numerical values to provide a further indicator of reliability, and measure these for the sample of 146 objects used across the comparison. First, we use the outlier rate, defined as the fraction of radii/Sérsic indices obtained by IMFIT that disagrees with the radii/Sérsic indices obtained by GALFIT by more than 15 per cent in $(1 + x)$, where x is the measured quantity. Secondly, we use the Normalized Median Absolute Deviation (NMAD) (Hoaglin, Mosteller & Tukey 1983), which is defined as $1.48 \times \text{median}[|\Delta x|/(1 + x)]$, where $\Delta x = x_{\text{galfit}} - x_{\text{imfit}}$. The NMAD is a measure of the spread of the IMFIT measurements around the GALFIT measurements, maintaining its reliability when outliers are present.

For the half-light radius, we find an outlier rate of 6.2 per cent and NMAD of 0.012, and for Sérsic index we find an outlier rate of 9.7 per cent and NMAD of 0.021, showing a slightly better agreement between codes for half-light radius than Sérsic index. We also find a better agreement with Sérsic index at lower values of Sérsic index. We see greater disagreements for a few objects at higher Sérsic index n due to larger contrast which exists at higher values of n . A slight change in the fitting will provide a larger change in n when n is larger.

Overall, however, we find a good agreement between these different codes and use the GALFIT results throughout the rest of this paper.

5 SAMPLE SELECTION

We select massive galaxies with stellar masses $\log(M_*/M_\odot) > 9.5$, and then make further selections based on the goodness of fit achieved by GALFIT for each object. We do not make selections based on the χ^2_v obtained, which is only used by GALFIT to determine when it has reached the best fitting. The χ^2_v is provided for the fit as a whole, so in cases where multiple objects are modelled simultaneously, this is not a measure of the goodness-of-fit for only the target object. Instead, we make our own selections based on the output parameters and the RFF, which we discuss in the following sections.

5.1 GALFIT parameters

In order to remove extreme cases, a fit must meet all of the following criteria:

- (i) The half-light radius must be within $0.01 < R_e \text{ (pixels)} < 100$. This ensures that fits where GALFIT has reached the minimum size possible or where the model would be larger than the cutout size are excluded.
- (ii) The fit Sérsic index lies within the range $0.05 < n < 10$.
- (iii) The fit axis ratio must be $(b/a) > 0.01$, removing unphysical models. This is particularly prevalent in faint sources, where GALFIT sometimes converges upon a ‘bad’ fit with a small axis ratio (van der Wel et al. 2012). We do not make selections based upon GALFIT magnitude, and all models are fit by GALFIT with a rest-frame optical magnitude < 28 .

Where neighbouring objects are being simultaneously modelled, these criteria are only applied to the best-fitting parameters of the central object. Where GALFIT does not converge, and gives no best-fitting parameters, or the best-fitting parameters do not meet the above criteria, we reject the fit. We start the fitting process with a sample of 1649 galaxies with $\log(M_*/M_\odot) > 9.5$, and through our selection criteria, we reject 192 galaxies, for which we repeat the fitting process with Sérsic index held at a value of $n = 1$. The ‘Fixed Sérsic’ fits must then meet the above criteria for all other parameters, and out of these 192 galaxies, we still reject 93 galaxy fits, thus our sample contains 99 galaxies with a Sérsic index fixed at $n = 1$, and 1457 objects with a free value of n at this stage, with a total sample size of 1556 galaxies.

5.2 Residual flux fraction

We calculate the RFF for our fits that met the previous criteria in Section 5.1. The RFF is a measure of the signal in the residual image that cannot be explained by background fluctuations (Hoyos et al. 2012). As in Margalef-Bentabol et al. (2016), we define this as

$$\text{RFF} = \frac{\sum_{(j,k) \in A} |I_{j,k} - I_{j,k}^{\text{GALFIT}}| - 0.8 \sum_{(j,k) \in A} \sigma_{B,j,k}}{\text{FLUX_AUTO}}, \quad (3)$$

where I is the NIRCcam image of the galaxy, I^{GALFIT} is the model image created by GALFIT, σ_B is the background RMS image, and FLUX_AUTO is the flux of the galaxy calculated by SExtractor, all of which are in the rest-frame optical filter of the object. The factor of 0.8 in the numerator ensures that the expected value of the RFF is 0 for a Gaussian noise error image (Hoyos et al. 2011). We calculate the RFF within the Kron radius of the galaxy, where we

define the Kron radius as the semimajor axis of the Kron ellipse. Calculating the RFF over a large radius leads to the RFF decaying to zero, where the outer areas can dominate the calculation, even if there is a complex residual at the centre of the image.

We calculate the background term following the method used in Margalef-Bentabol et al. (2016), where we assume that

$$\sum_{(j,k) \in A} \sigma_{B,j,k} = N \langle \sigma_B \rangle, \quad (4)$$

where $\langle \sigma_B \rangle$ is the mean value of the background sigma for the whole image. We calculate this by placing apertures on blank areas of sky in the image ‘ERR’ extension that we use for creating the GALFIT sigma images (see Section 4.1), and calculating the mean value of these regions. The value of N is the number of pixels within the radius that we are using for the RFF calculation.

In order to remove any remaining objects that are poorly fit with large residuals, we complete visual checks to select an appropriate RFF cutoff value. As a result, we select objects with an RFF value below 0.5. This enables us to remove objects where the light is either very overaccounted or underaccounted for, yet also allows for features that are not modelled precisely (due to features such as spiral arms and bars not being accounted for in single Sérsic fits) but where the measured properties are otherwise reasonable. Through RFF measurements, we reject a further 161 fits due to large residuals. This results in a final sample of 1395 robust galaxy fits, of which 1313 (94.1 per cent) were fit with a free value of n , and 82 (5.9 per cent) were fit with a fixed value of $n = 1$, which we further analyse in Section 6. The rejected fits are mostly comprised of lower mass galaxies, although there are fits rejected at all masses within our sample, and we do not see a change in RFF as a function of redshift.

5.3 Final sample

We begin the fitting process with a sample of 1649 galaxies, and after our quality cuts we recover 1395 galaxies for our final sample. This recovery rate of 84.6 per cent is higher than comparable analyses, such as Suess et al. (2022) (~ 60 per cent), although we repeat the fitting process with a fixed Sérsic index where the first fitting procedure has failed. However, fits with a fixed Sérsic index only account for 5.9 per cent of our sample. We note that we have not made any selections based upon the magnitude of the GALFIT model, as the input and output magnitudes are in good agreement. In the rest-frame optical, 3.66 per cent of the GALFIT best-fitting models have a half-light radius in pixels that is smaller than the FWHM of the PSF, thus almost all of our sources are resolved. This sample is used throughout this paper to carry out our analyses of the evolution of galaxy size and structure.

6 RESULTS

In the following sections, we describe the results of our analysis. This includes examining the sizes of our systems, as well as the overall shapes of these galaxies based on their Sérsic indices and how these evolve with time. Where we state results for passive and star-forming galaxies, whereby these are defined by having a sSFR below or above the midpoint of the distribution of values within the redshift bin. We calculate the sSFR using the star formation rates and stellar masses from Duncan et al. (2014, 2019) (see Section 3.1 for full details). We then define a galaxy with an sSFR greater (lower) than the median sSFR within the redshift bin, to be a star-forming (quiescent) galaxy. While we use the half-light radius as a measure of size throughout this paper, it is worth noting that when we compared our trends to

those obtained with another size measure, such as R_{90} , our findings do not change. The use of R_e also allows for a continuation of work completed with *HST*, and a wide comparison of the following results to previous work in this area.

6.1 Half-light radii and sérsic indices

We measure the half-light radii of our objects using GALFIT, and convert the values from pixels to their physical half-light radii in kpc. Fig. 6 shows the size evolution with redshift for our sample of 1395 galaxies. We use the radius from the filter nearest to the rest-frame optical for each object, and find that the sizes are well fit by the power-law relation:

$$\langle R_e \rangle = 4.50 \pm 1.32(1+z)^{-0.71 \pm 0.19}. \quad (5)$$

This is such that the average sizes of our sample, in terms of effective radii ($\langle R_e \rangle$), become progressively smaller at increasing redshifts. This trend for galaxies at a given mass selection to become smaller at higher redshifts had been known to exist at $z < 3$ for many years (e.g. Trujillo et al. 2007; Buitrago et al. 2008; van der Wel et al. 2012), yet this is the first time this has been shown using *JWST* observations for similar types of studies. We also compare our power-law function to those derived in comparable studies, shown in Fig. 6. We note that the curves presented in Buitrago et al. (2008) are normalized with respect to SDSS data, thus we perform an arbitrary normalization of these curves to align them with the scale employed. We also extrapolate all curves to cover our entire redshift range. We compare to a range of individual points and power-law curves.

It is important to note that the evolution of galaxy size and Sérsic index are highly dependent on the redshift ranges studied and the stellar mass and/or magnitude ranges which are included in the analysis. For example, if we study just the supermassive galaxies at $\log(M_*/M_\odot) > 11$, we would observe that systems have a stronger evolution than at lower masses (e.g. Buitrago et al. 2008, 2013). As such it is important to establish a clear basis for comparison, as no previous study has measured galaxies using the exact same approach as in this study. The points from van der Wel et al. (2014) are for galaxies with $\log(M_*/M_\odot) \sim 10.75$, Bridge et al. (2019) and Kubo et al. (2017) are at $\log(M_*/M_\odot) \sim 10$, and Yang et al. (2022) select bright objects based on their magnitudes in the *F444W* band. The curves given are for a mass selection of $\log(M_*/M_\odot) > 9$ at $3 \leq z \leq 6$ (Costantin et al. 2023), $\log(M_*/M_\odot) > 11$ at $1.7 < z < 4$ (Buitrago et al. 2008), and for a number density-based selection at $z < 2$ (van Dokkum et al. 2010). We extrapolate these curves to cover our entire redshift range.

Fig. 6 shows that there is a steep evolution for galaxies within the mass range employed in this paper. We find, as previously mentioned, that the size evolution of galaxies follows a power law with the shape $\sim (1+z)^{-0.71}$, which is less steep than previous results when comparing the evolution up to $z \sim 3$. This is partially due to the fact that we are observing galaxies at higher redshifts where the size evolution tapers off and does not continue as steeply at higher z . It is important to keep in mind that redshift (z) values do not scale linearly with time, and there is much more time at a given δz at low redshift than at the higher redshifts. Another reason for the difference, can be seen at the lowest redshifts, where our galaxies are on average smaller than the previous work. This is likely due to us using a lower mass cut to define our sample of galaxies, resulting in on average smaller systems. This is consistent with findings from simulations, where it has been shown that galaxy size correlates with stellar mass, thus resulting in lower mass samples having smaller sizes, on average (Furlong et al. 2017; Ma et al. 2018).

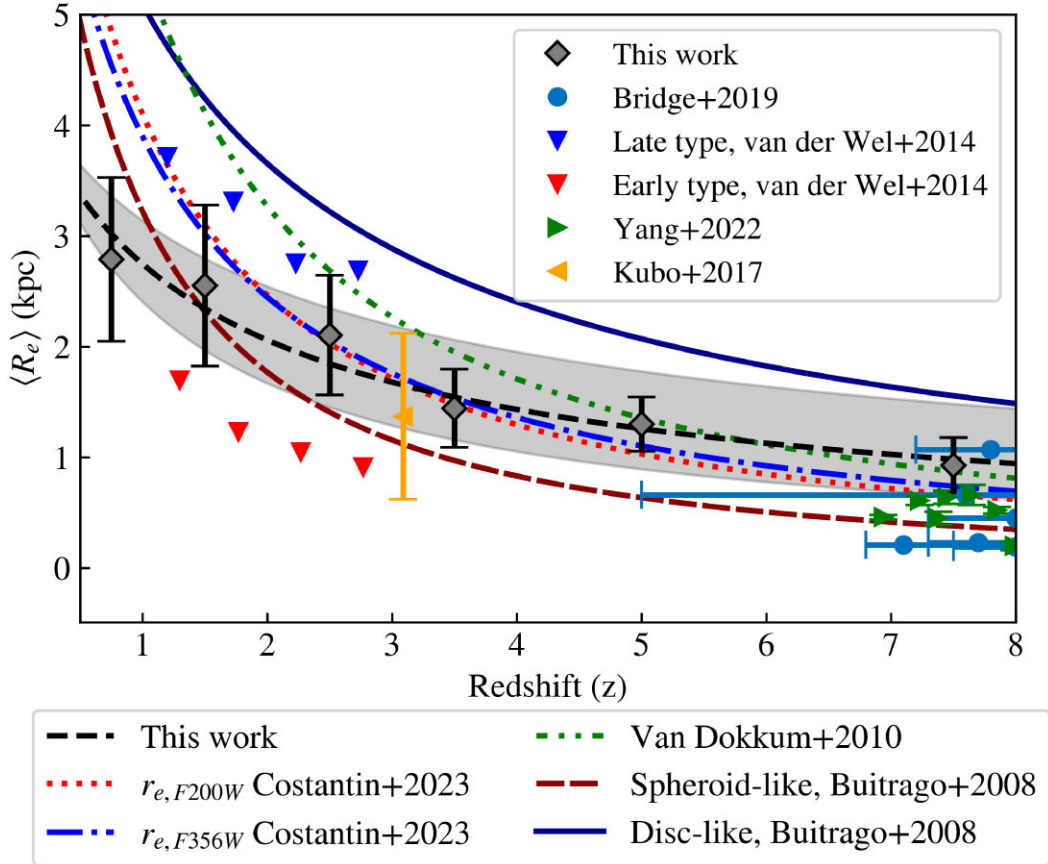


Figure 6. Power-law fit showing the size evolution of all galaxies within our sample, compared to results from previous work. The black-dashed line is of the form $R_e(\text{kpc}) = 4.50 \pm 1.32(1+z)^{-0.71 \pm 0.19}$, with the grey, shaded area showing the error on the power-law fit. The grey diamond points are the median galaxy sizes in each redshift bin, and error bars are 1σ in length. The previous work shown in the Figure uses *HST* data, except for Yang et al. (2022), which uses *JWST* data, van Dokkum et al. (2010) which uses NOAO/YaleNEWFIRM Medium Band Survey Data, and Costantin et al. (2023) which uses the TNG50 simulation to produce mock CEERS observations. We note that we only plot redshift errors for Bridge et al. (2019), as radius errors are not provided.

We also investigate the difference in the size–redshift relation for populations of galaxies with high and low Sérsic indices, defined as $n > 2$ and $n < 2$, respectively, although using $n = 2.5$ to separate the sample produces effectively the same results. We do this to determine how the size evolution depends on the shape of the profile, with galaxies at $n > 2$ possibly more like the massive galaxies we see in the local universe and those with $n < 2$ possibly progenitors of disc galaxies or those undergoing mergers.

As can be seen in Fig. 7, at low redshift, the galaxy populations have a clearly different size–redshift relation, but at redshifts higher than $z \sim 3$, the relations show a greater similarity, suggesting that effective radius is less dependent upon the Sérsic index at high redshift compared to lower redshifts. What this means is that galaxies do not differentiate between overall morphology, as measured by the Sérsic index, until around $z \sim 3$, consistent with findings at $z < 5$, where star-forming galaxies exhibit inside out growth (Roper et al. 2023). This suggests that this aspect of the ‘Hubble Sequence’ was in place by at least $z \sim 3$, with a disc-like ($n < 2$) and elliptical-like ($n > 2$) population clearly defined.

As shown in Figs 7 and 6, we also find that galaxies at higher redshift are almost all compact objects, regardless of their Sérsic index, and are smaller than their low-redshift counterparts of similar mass, in agreement with previous studies and simulations (e.g. Trujillo et al. 2007; Buitrago et al. 2008; van der Wel et al. 2014; Costantin et al. 2023). Furthermore, Fig. 7 shows that the sizes of

the high and low Sérsic index populations evolve differently at $z < 3$, with the objects with lower Sérsic indices following a steeper size–redshift relation, suggesting a difference in structural growth mechanisms in each population, likely due to the onset of inside out growth in these systems with lower Sérsic indices. Overall, we also show that the compact systems of the early universe are not representative of the galaxy population today, suggesting a strong size evolution must occur, continuing until the present-day. This growth is such that we find an increase of a factor of three from $z \sim 7$ to ~ 1 , with roughly a doubling of size from $z \sim 7$ to ~ 3 , a time period of ~ 1.4 Gyr. The relation is also still evolving at the highest redshifts, confirming that galaxy evolution as well as evolution in structure were already taking place in the first Gyr since the big bang.

For those objects fit with a free Sérsic index, we also investigate the evolution of Sérsic index with redshift, as shown in Fig. 8. Fig. 8 shows that for all free-fit n galaxies within our sample, Sérsic index decreases with increasing redshift, suggesting a higher proportion of disc-type galaxies in the early universe. This is in agreement with findings shown in Robertson et al. (2023), where it is shown disc candidates appear at high redshifts of $z \sim 5$ with $\langle n \rangle = 1.04$. Whilst these objects have Sérsic indices similar to modern pure disc galaxies, this does not necessarily imply that these are rotating disks (e.g. Buitrago et al. 2014). This is in contrast to previous findings using *HST* that there were fewer disc galaxies at $z > 1.5$,

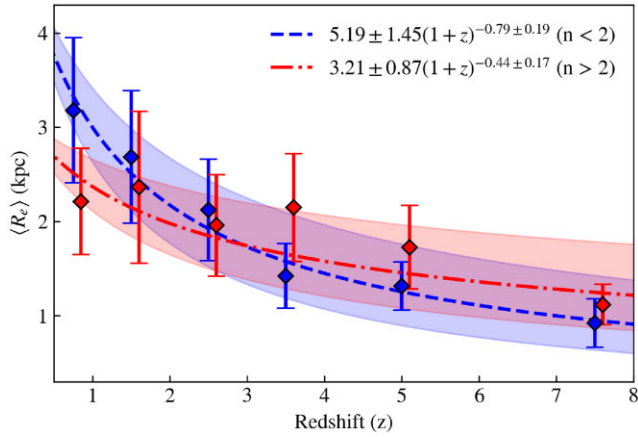


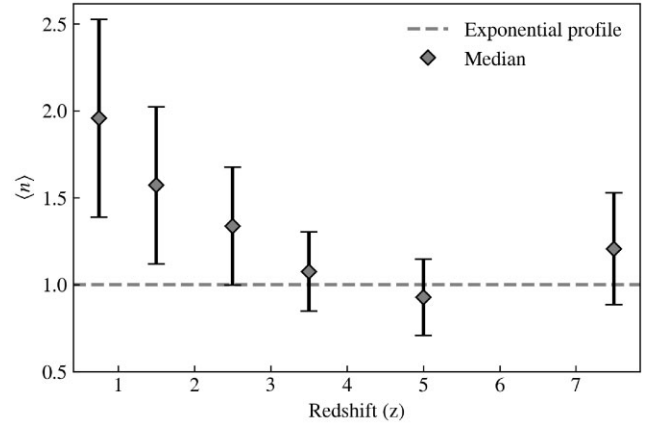
Figure 7. Power-law fit for galaxies separated into two groups by Sérsic index as defined at $n = 2$, where $n < 2$ represents disc-like galaxies, and $n > 2$ represents elliptical-like galaxies. The sizes of these objects mostly diverge at the lowest redshifts. The diamond points are the median galaxy sizes in each redshift bin, and the error bars are 1σ in length. The shaded region around each line represents the error on the power-law fit.

although the lower resolution of *HST* lead to misclassification of galaxies (Ferreira et al. 2022b), with the improved depth and quality of *JWST* imaging when compared to *HST* imaging revealing disc morphologies previously obscured due to noise (Robertson et al. 2023). Exploring this complicated subject is beyond the scope of this paper. The slight increase in Sérsic index at $z \sim 7.5$ could be due to an increase of spheroid galaxies reported in Ferreira et al. (2022b), particularly in the star-forming population (see Fig. 8b), as spheroid galaxies have been found to account for a higher proportion of the sSFR budget (Ferreira et al. 2022a) at high redshifts, but could also simply be due to random chance and increased errors at higher redshifts. We further discuss size and Sérsic index changes as a function of morphology in Section 6.4.

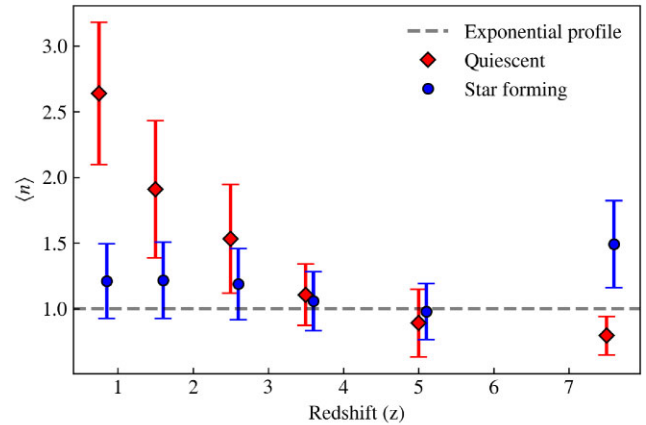
In Fig. 8b, we investigate Sérsic indices for passive and star-forming populations, and find that the star-forming galaxies have Sérsic indices around $n \sim 1$, suggesting that most star formation takes place within disc galaxies, which is also the case when identifying these systems through visual means (Ferreira et al. 2022b). This is also in agreement with measurements of star-forming galaxies, which are found to be typically discy with $n \sim 1.2$ (Paulino-Afonso et al. 2016). The slight deviation from this trend at the highest redshifts could be due to stars forming in young, compact sources, before evolving into disc-like galaxies. The populations start to show differing trends at $z \sim 3$, again hinting at the establishment of a bifurcation of the galaxy population in structure around this time. This shows that the star forming and passive galaxy populations possibly evolve through different mechanisms to create differing physical properties at later times, such as inside out star formation in star-forming galaxies, and stellar migration in passive systems.

6.2 Galaxy size–mass distribution

We plot the galaxy size–mass relation in different redshift bins, as shown in Fig. 9. For each redshift bin, we separate galaxies into either a star-forming or passive population with respect to the average value in the given redshift bin, using the median sSFR of each bin as the separating value. This mid-point is determined by plotting a histogram of all the sSFR values for galaxies within a given redshift bin. The median value is then measured, and galaxies which are lower



(a) Sérsic - redshift distribution for all galaxies within our sample.



(b) Sérsic - redshift distribution for passive and star-forming galaxies.

Figure 8. The Sérsic index – redshift relations for all galaxies (top panel), and passive and star-forming populations, relative to the average sSFR in each redshift bin (bottom panel). Large diamond and circle points are median values of each redshift bin, with error bars 1σ in length. The grey dashed line at $n = 1$ represents the special case of the exponential disc profile.

than this we call ‘passive’, and those above this ‘active’, relative to the average sSFR in the given redshift bin. This is the same method we used previously for separating star-forming from non-star-forming galaxies when investigating trends with size and Sérsic index.

In general, we find that galaxy size increases with stellar mass, in both the quiescent and star-forming populations up to $z \sim 6$, although this tends to be more obvious for the star-forming galaxies. We also see that the star-forming galaxies at the lower redshifts are nearly always larger at a given mass than the passive galaxies, although this tends to break down at the lower masses.

We also find that at redshifts higher than $z > 3$, the sizes of the quiescent and star-forming populations are statistically the same, suggesting that at a fixed mass at high redshift, quiescent galaxies may not be smaller than star-forming galaxies, as predicted in Suess et al. (2022), likely due to transient quiescence driven by high-redshift AGNs activity (Lovell et al. 2023). This means that whatever is differentiating the sizes of galaxies as a function of shape or star formation rate does not come into play until past redshift $z \sim 3$. This implies that the physical processes of growing galaxies is truncated for the passive galaxies, even if these systems still acquire stellar mass throughout their history. We discuss possible reasons for this later in the discussion section.

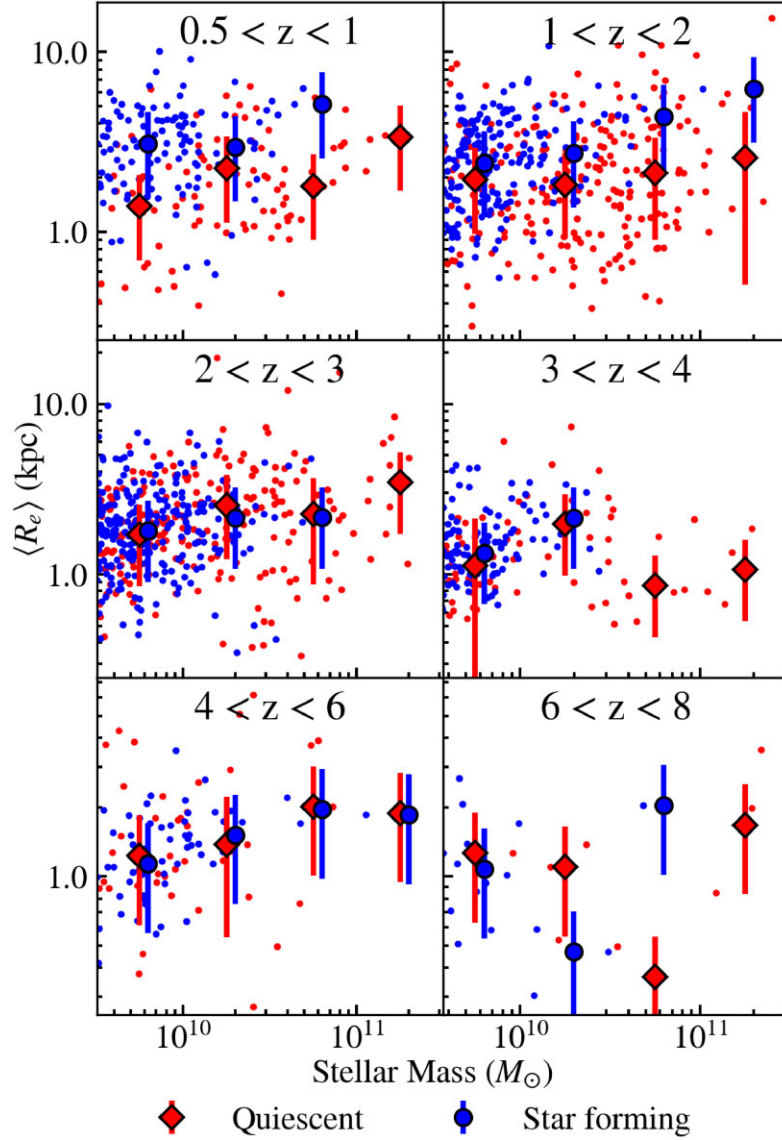


Figure 9. The size–stellar mass distribution of ‘passive’ and ‘star forming’ galaxies, relative to the mid-point of the sSFR in each bin, as explained in Section 6.2. The larger points represent the median values in mass bins. We use a 50 per cent error floor, with error bars one standard deviation in length, or representing a 50 per cent error on the median, in cases where the error floor is larger.

6.3 Changes in effective radii as a function of observed wavelength

The changes in a galaxy’s appearance and size as a function of wavelength can provide many clues to the formation history and stellar populations of modern galaxies (e.g. Windhorst et al. 2002; Taylor-Mager et al. 2007; Mager et al. 2018; Suess et al. 2022). An example of this is that if a galaxy forms inside-out, with the older stellar populations in the centre of the galaxy, then most likely these systems would appear larger at shorter wavelengths, and vice versa if formation occurred via outside-in.

As such, we measure half-light radii in all available wavelengths for our sources, and use these sizes to probe the size evolution of our galaxy sample as observed at different wavelengths. As seen in Fig. 10, average galaxy sizes become increasingly more compact and smaller when observed at longer wavelengths, for objects with both high and low Sérsic indices at $z < 3$. We also see again, that those galaxies with higher Sérsic indices are smaller on average at all

wavelengths to those systems with lower indices. This is an indication that galaxy sizes are larger at shorter wavelengths where bluer and young light is probed, due to the formation mechanisms for these galaxies. This would be such that the outer parts of these systems consist of younger stars, compared to their inner portions made of older stars. More detailed analysis of the colour gradients and star formation gradients of these galaxies would answer this question. We also note that dust attenuation can increase the observed half-light radius (e.g. Marshall et al. 2022; Popping et al. 2022; Roper et al. 2022), although this would require a more in depth observational analysis.

At $z < 3$, we also find that galaxies with higher Sérsic indices have smaller radii in both mass bins, with this effect being more noticeable in the highest mass bin. However, we do not see the same effect past $z \sim 3$, where we see a much flatter relation, suggesting that galaxies at high redshift are forming stars throughout the entire galaxy, with blue and red light emitted throughout. However, when

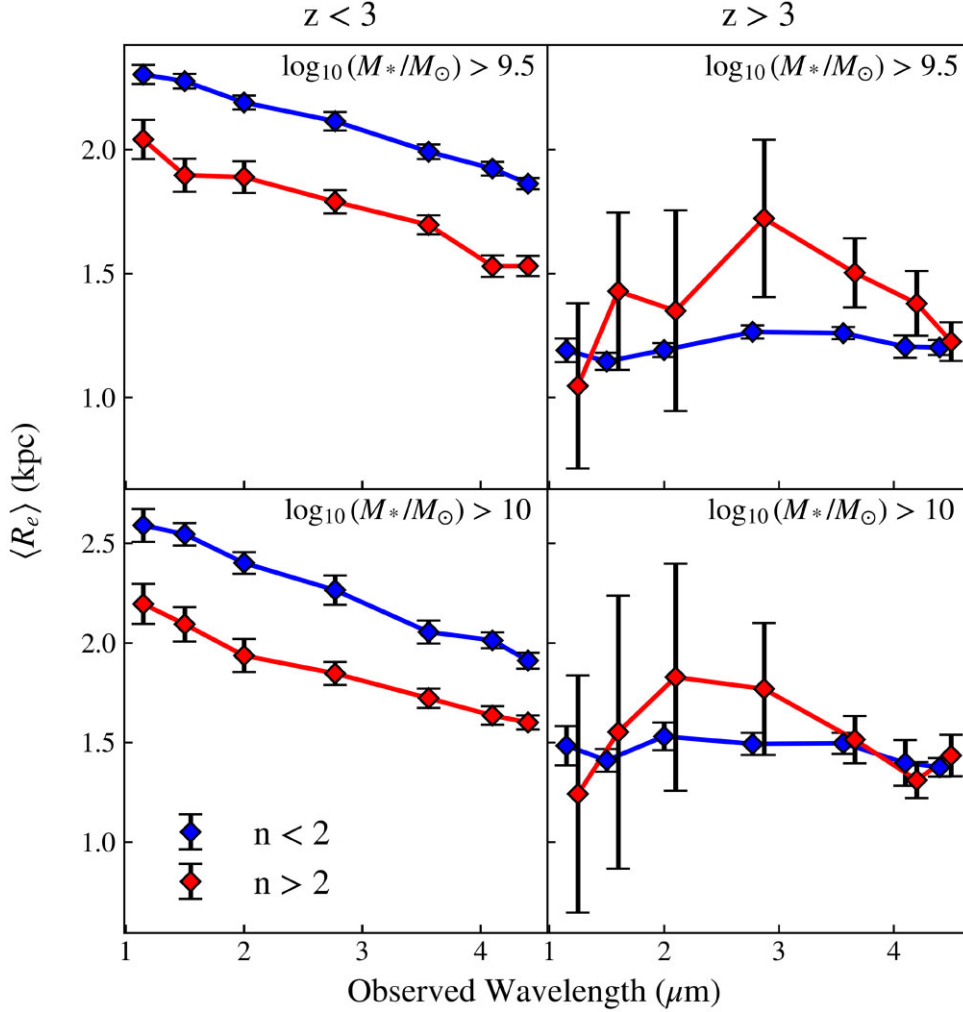


Figure 10. The half-light radius as a function of wavelength. These plots show the radius evolution for $z < 3$ (left panels), and $z > 3$ (right panels) galaxies. In each redshift bin, we show the evolution for low-mass and high-mass galaxies, with the galaxies separated into high-Sérsic (red) and low-Sérsic (blue) index populations. These figures show that up to $z \sim 3$ (left panels), galaxies are more compact when measured in redder filters, regardless of selection method. Past $z = 3$ (right panels), the relation is flatter, with the galaxies having $n < 2$ being smaller at this stage.

we divide the higher redshift galaxies into different bins, we obtain a noisier trend, and thus we hesitate to draw any further conclusions from this observed trend. For objects with Sérsic indices of $n > 2$, there is much more scatter within the relation, although there are larger errors, due to the smaller number of galaxies with high Sérsic index at $z > 3$. Analysing this in smaller bins, such as $\Delta z = 1$, did not yield different conclusions.

To compare with previous *JWST* work on similar questions, we also compare the sizes of our galaxies measured in the *F444W* ($4.4 \mu\text{m}$) band to the *F150W* ($1.5 \mu\text{m}$) band sizes, as shown in Fig. 11. This allows us to probe the rest-frame optical, observed in *F150W*, and compare with the rest-frame near-infrared, observed with *F444W*. This comparison is quoted in arcseconds, as a comparison of the on-sky sizes. We find that for galaxies at cosmic noon ($1 \leq z \leq 2.5$), the $4.4\text{-}\mu\text{m}$ sizes are 11.4 ± 1.28 per cent smaller than the $1.5\text{-}\mu\text{m}$ sizes on average, in agreement with the ~ 9 per cent difference found in Suess et al. (2022). Taking sizes measured in the near-infrared (observed with $4.4 \mu\text{m}$ at this redshift range) to be a reasonable proxy for stellar mass distributions, this shows that the stellar mass profiles of galaxies are smaller and more compact than their star-forming ‘light’ profiles. We do not find that

the galaxies outside of cosmic noon show the same effect, as the *F150W* and *F444W* filters no longer correspond to the rest-frame optical and rest-frame infrared of these galaxies. As we have shown that the smaller appearance in *F444W* is indeed due to the rest-frame infrared profile being smaller, thus showing the mass profile of the galaxy is smaller than the light profile, we investigate how this varies with the stellar masses of galaxies. This effect is dependent on the mass of the galaxy, as shown in Fig. 12, where we examine the size difference at cosmic noon in two mass bins. We find that for galaxies with $9.5 \leq \log(M_*/M_\odot) < 10$, the *F444W* sizes are 5.82 ± 1.76 per cent smaller than their *F150W* sizes. This increases with increasing mass, with sizes being 17.9 ± 1.80 per cent smaller for objects with $10 \leq \log(M_*/M_\odot)$. This implies that, on average, we see a greater difference in size between different wavelengths for the highest mass galaxies.

6.4 Correlation with visual morphology

For 470 galaxies within our sample, we use visual morphological classifications from Ferreira et al. (2022b), as defined in Section 3.2 to determine how the properties and relations we have found are

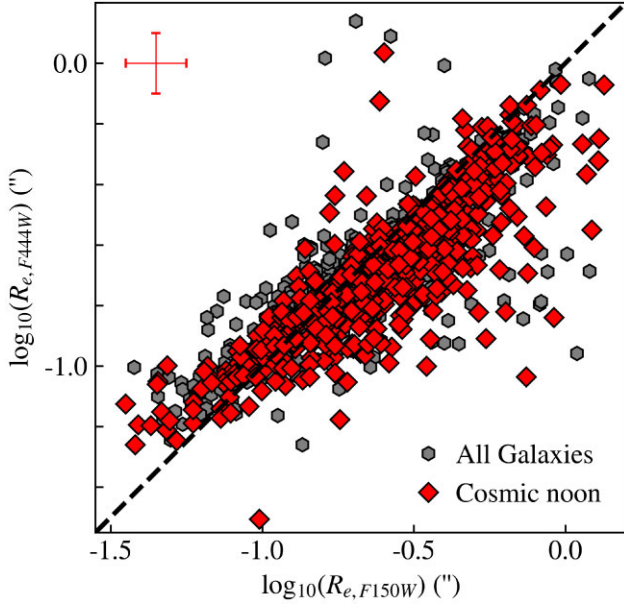


Figure 11. A comparison of sizes measured in the $F444W$ and $F150W$ filters, showing that $F444W$ sizes are smaller than those in $F150W$. This implies that galaxies measured in the near-infrared with *JWST* are more compact than rest-frame optical sizes previously measured with *HST*. Galaxies at cosmic noon ($1 \leq z \leq 2.5$) are shown as red diamonds whilst the grey hexagons represent all galaxies within our sample that do not fall within cosmic noon. These latter galaxies do not show the same effect. The black, dashed line is the one-to-one relation. The error bar represents a typical error of 0.2 dex on these measurements.

determined by overall morphology. This is a small sample of galaxies, but represents one of the first times that we can examine these measured properties with visual morphologies. Using these classifications, we present an analysis of size and Sérsic index as a function of morphology, as shown in Fig. 13.

Fig. 13a shows that for all galaxy types, radius decreases with increasing redshift, but at all redshifts, spheroid-type galaxies are the smallest. Fig. 13b shows how Sérsic index varies with redshift for all galaxy types. Disc-type galaxies have $n \sim 1$ as expected, with ‘peculiar’ and ‘other’ galaxies showing a slight decrease with redshift. Spheroid galaxies have the highest Sérsic index at all redshifts, significantly above other galaxy types, as found in Kartaltepe et al. (2022). The small sizes of these galaxies combined with their high Sérsic index is a clear indicator of their compact, concentrated nature.

6.5 Comparison with image simulations

One major issue with a study such as this, which deals with imaging and the analysis of structure at galaxies at vastly different redshifts, is the fact that the surface brightness measurements of galaxies declines as $(1+z)^4$, and this can produce significant changes in the way that structure for distant galaxies would be imaged by a telescope. In fact, it is clear that galaxy structure can, in principle, change substantially and that many galaxies are potentially being missed at the highest redshifts (Conselice 2003; Whitney et al. 2020, 2021). Therefore, it is very important that we carry out simulations to determine if the trend we see in this paper, namely that galaxies get progressively smaller up to $z \sim 7$, is due to a real evolution or due to galaxies appearing smaller and fainter at higher redshifts. Previous work using *HST* shows that whilst we are likely missing galaxies at the highest redshifts, we can

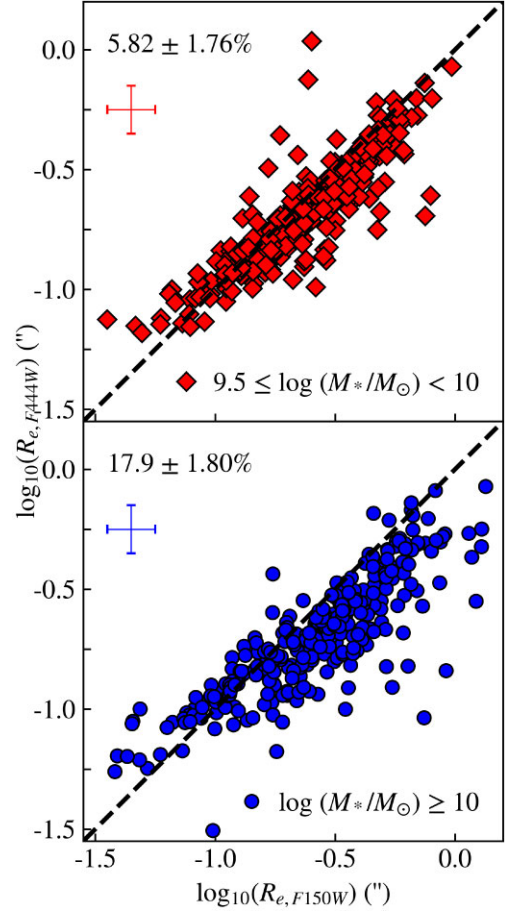


Figure 12. Size comparison for different mass bins at cosmic noon. We find an increasing size difference with increased stellar mass. The black-dashed line is the one-to-one relation. The percentages in the upper left corner are the average size difference for each redshift bin. The error bar represents a typical error of 0.2 dex.

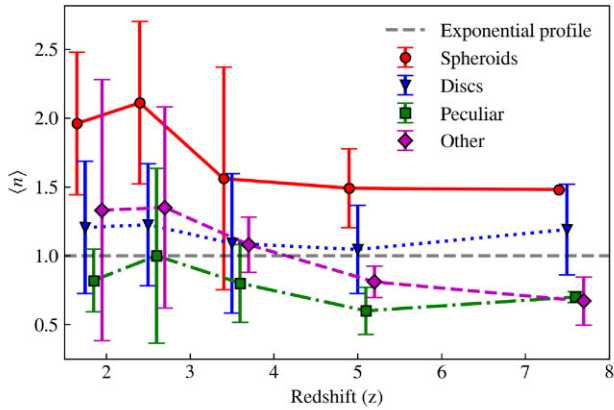
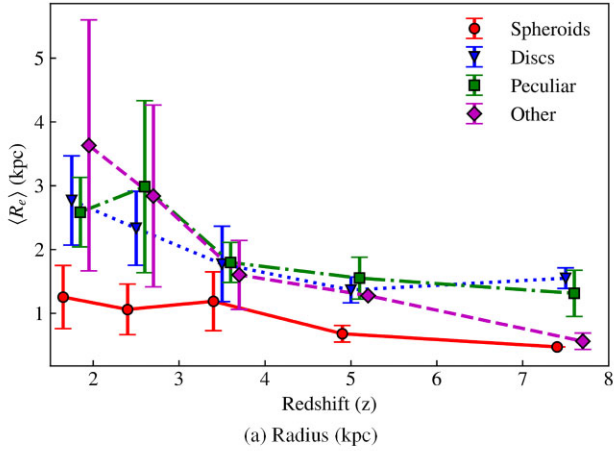
still measure accurately their structural parameters (Whitney et al. 2020, 2021).

To understand this issue in depth for *JWST* data, we take a sample of 186 low-redshift galaxies at redshifts $0.5 < z < 1$, and create simulated images of these galaxies at higher redshifts, in intervals of $\Delta z = 0.5$, up to $z = 7.5$, including all known cosmological effects. A representative sample of galaxies is used, with Sérsic indices ranging from $0.06 < n < 7.58$. This analysis is done in order to separate real evolution effects from redshift effects. To do this simulation, we use the redshifting code AREIA.⁵ We give here a brief overview of the steps taken are described below, for a more detailed discussion, we refer the reader to Tohill et al. (2021) and Whitney et al. (2021).

First, the source is extracted from the original stamp by measuring a segmentation map with GALCLEAN.⁶ Then, the image is geometrically re-binned from the source redshift to the target redshift based on the standard cosmology, preserving its flux. This is done to ensure that higher redshift sources have the appropriate geometric scaling due to the adopted cosmology. The resulting image from the rebinning has its flux scaled due to the cosmological dimming effect. Furthermore, shot noise is sampled from the source new light distribution, which

⁵<https://github.com/astroferreira/areia>

⁶<https://github.com/astroferreira/galclean>



(b) Sérsic index (n). The grey dashed line represents the $n = 1$ profile of an exponential disc.

Figure 13. Size and Sérsic index evolution as a function of visually determined morphology from Ferreira et al. (2022b). At all redshifts, we find that spheroid galaxies have the smallest radius, and the highest Sérsic index, displaying the compact, concentrated nature of these objects.

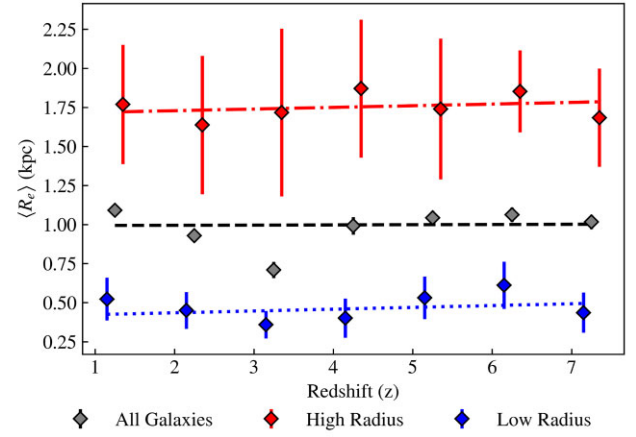
is then convolved with the target *JWST* PSF of the rest-frame filter in the target redshift, to mimic an observation in the rest-frame optical. Then, the final redshifted source is placed on a random real CEERS background to mimic a real observation. This results in a sample of 2418 simulated images.

Although AREIA allows the user to include size corrections and brightness corrections to mimic redshift evolution, we keep all intrinsic properties of the galaxies constant at each redshift, simulating only observational effects.

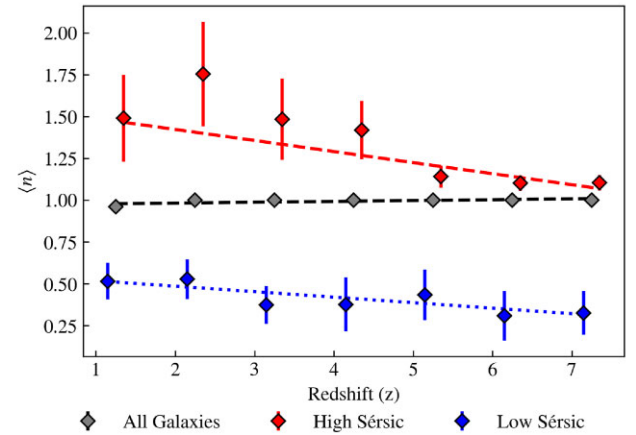
Following the same method described in Section 4.1, we measure the sizes and Sérsic indices of our new simulated galaxies. We follow the same selection method, described in Section 5 to ensure robust results. For our Sérsic index analysis, we further select objects with a GALFIT measured magnitude < 30 .

We also analyse the trends for objects with values above and below the overall median, indicated by the ‘high’ and ‘low’ radius and Sérsic groups.

We use the *emcee* package to obtain the lines of best fit (Foreman-Mackey et al. 2013). As shown in Fig. 14, we find that our sizes and Sérsic indices are best fit by linear fits, with gradients and errors stated in Table 3. We find our results are consistent with ‘flat slopes’ – that is we find no change with redshift for the measured sizes and Sérsic indices within these simulations. This shows that the simulated



(a) Median simulated galaxy sizes at each redshift. The grey points and line show this change for the simulated galaxies for the total simulated sample, whilst the red and blue are for those galaxies that are larger and smaller, respectively, than the median radius.



(b) Median Sérsic index of simulated galaxies at each redshift. The lines are the same as explained in the plot of effective radius with redshift.

Figure 14. Sizes and Sérsic indices for our simulated galaxies in redshift bins. The error bars represent the standard error of the median for each bin. We divide these samples into different sub-types to determine how different selections would evolve differently within these simulations.

Table 3. Best-fitting results for linear fits to all galaxies, and high and low subsets for both radius and Sérsic index.

Radius subset	Gradient	Sérsic subset	Gradient
All galaxies	0.001 ± 0.137	All galaxies	0.005 ± 0.139
High radius	0.009 ± 0.159	High Sérsic	-0.067 ± 0.148
Low radius	0.010 ± 0.138	Low Sérsic	-0.032 ± 0.149
(a) Radius (kpc)		(b) Sérsic index (n)	

Notes. The slopes are consistent with being flat – with little change with redshift, showing that our main findings are due to evolutionary effects, not redshift effects.

galaxies continue to have very similar size measurements at different redshifts, meaning that in the absence of evolution we would expect the same galaxies to have the same effective radii and Sérsic indices measured at all redshifts. This confirms that our method recovers the same result regardless of the redshift in which the galaxy is observed.

This is vastly important for this work, and shows that the trends obtained in Section 6, are due to a change in galaxy properties and populations with increasing redshift, not because the galaxies appear smaller at higher redshifts due to cosmological or redshift effects.

7 DISCUSSION

The results in this paper reveal a myriad of new observational facts about the sizes and shape evolution of galaxies up to $z \sim 8$. Our core result is that galaxies become progressively smaller at fixed stellar mass at higher redshifts. Whilst this was known for some time up to $z \sim 3$ (Trujillo et al. 2007; Buitrago et al. 2008; van Dokkum et al. 2010; van der Wel et al. 2012), *JWST* is now allowing an analysis of this in the rest-frame optical light at higher redshifts than previously possible. In fact, we find that from $z \sim 7$ to ~ 3 , galaxies with larger masses roughly double in size. This evolution is not as dramatic as is seen for the highest mass galaxies at $z < 3$ (e.g. Buitrago et al. 2008) and we will require that more area be covered before we have statistics to probe these very high-mass galaxies at such high redshift, as very few are imaged due to the limited numbers and sky coverage with existing and reliable *JWST* data.

Another major result is that we find very little variation in the size distribution and size evolution for our sample at $z > 3$ irrespective of how we divide the sample. This is true for different Sérsic cuts, as well as for different cuts in the sSFR for these galaxies. One of the main signatures of the formation of the Hubble sequence is a bifurcation in galaxies into morphologically distinct populations of star-forming and relatively passive systems (often simply divided into discs and ellipticals). Whilst we are not seeing the entire formation of this Hubble sequence at $z \sim 3$, it is clear that the major bifurcation begins at this epoch and increasingly differentiates itself. This is likely due to different formation mechanisms coming into play at $z < 3$ that were not present at the higher redshifts. This is likely something to do with mergers and feedback from either AGN or star formation (e.g. Bluck et al. 2012), particularly inside out star formation, where star-forming galaxies begin to grow more rapidly than their passive counterparts.

While we do not see much difference between galaxies at $z > 3$, we do find that galaxies have a well-established difference in size and Sérsic index as a function of the wavelength of observation. This is such that galaxies are more compact in redder wavelengths, showing that the outer parts are made up of more recent stars and star formation events. This could be a sign that galaxies are forming inside-out and that we are witnessing the formation of bulges and the cores of giant galaxies at these early times that are growing outward from minor mergers and/or accreted gas in star formation (e.g. Tacchella et al. 2015a, b, 2018; Nelson et al. 2016, 2019; Wilman et al. 2020; Matharu et al. 2022; Roper et al. 2023). This is consistent with the formation of bulges and disks occurring gradually at about $z \sim 3$ and at lower redshifts (e.g. Margalef-Bentabol et al. 2016).

We also find that there is a correlation between stellar mass and size for galaxies – for both galaxies that are relatively more star forming and those that are more passive – up to $z \sim 3$. Our *JWST* results allow us to accurately probe these mass ranges, even at the higher redshifts where *HST* has had a difficult time resolving these systems. These results provide another indication that something is regulating the sizes and masses of galaxies, and that this appears to be more present at $z < 3$ than at earlier times. When we compare with simulations, such as the TNG50 simulation at $3 < z < 6$ (see Fig. 6) we find that there is a good agreement. The causes of this change around $z \sim 3$, and the physics behind the mergers that produce these size increases within these simulations are still to be determined.

7.1 Caveats and future work

This study’s limitations, which future research could address, include the relatively small number of galaxies in the higher redshift categories. This limitation stems from our reliance on catalogues developed from *HST*. While these catalogues offer accurate photometric redshifts for lower z galaxies, they inevitably miss some fainter galaxies at higher redshifts. Future research that concentrates exclusively on high-redshift galaxies, using *JWST* data, should not face this limitation. Additionally, as more deep, high- z observations are made, the growing number of known distant galaxies will further mitigate this issue.

We also note that our comparisons of the passive and star-forming populations in this work are relative to the average sSFR in each redshift bin. While this is clear throughout the analysis, we feel that this could be improved upon when greater populations of each are known, or with studies focusing entirely on quiescent or star-forming galaxies, with stricter definitions of each. We also note that the average sSFR in each bin varies, but we chose this method as using a fixed value to define the populations in each redshift bin resulted in either the low- or high- z bins being entirely one population, thus we chose to present a comparison of the more passive and more star-forming population at each redshift. For example, using $\log_{10}(\text{sSFR}) = -9$ as a constant value resulted in the lowest redshift bin being made up of mostly passive systems in a way that did not allow for a comparison to be made.

8 SUMMARY AND CONCLUSIONS

We present an analysis of 1395 carefully selected massive galaxies with stellar masses $\log(M_*/M_\odot) > 9.5$, within the CEERS and CANDELS fields between $z = 0.5$ and 8, using light profile fitting. Our galaxy sample is taken from the CANDELS field to enable the use of robust masses, redshifts, and star formation rates from optical to NIR data, which is necessary to obtain accurate measurements for galaxies over our large redshift range. In this paper, we fit single Sérsic profiles to our galaxies and analyse their effective radius and Sérsic index to probe the evolution of these properties through cosmic time. We also probe the variation in size and shape as a dependence on stellar mass and wavelength. The trends we find are robust to redshift effects as we show through simulations of placing low-redshift galaxies at high redshift that parameters would not change simply due to being at higher redshifts as imaged with *JWST*.

To carry out our analysis, we use a custom-built GALFIT pipeline, fitting a single Sérsic fit, and fitting neighbouring galaxies where appropriate. We verify this method via a comparison with another galaxy profile fitting code, IMFIT and find that our results are generally in good agreement between these two codes, and thus reliably measured. We then analyse the evolution of effective radius and Sérsic index, for our sample as a whole, and for quiescent/star-forming population, and high/low Sérsic index populations. We verify that our results are due to evolutionary effects rather than redshifting effects, by fitting the same galaxies redshifted to higher redshifts using the exact same process we apply on the real galaxies. We find ‘flat’ relations between measured parameters and redshift, thus confirming the robustness of our method and its ability to recover the correct measurements regardless of the distance to the objects.

Our main findings are as follows:

(i) Sérsic indices decrease on average with increasing redshift, suggesting a higher proportion of ‘disc-like’ galaxies in the early universe. We find that passive and star-forming populations of galaxies show a different evolution of Sérsic index with redshift, with

star-forming galaxies hovering around value of $n = 1$, suggesting that most star formation occurs within disc-like galaxies, at least in terms of structure. We cannot, however, rule out that some of these galaxies are involved in mergers. In principle, this confirms what has been found when classifying galaxies visually (Ferreira et al. 2022b).

(ii) In general, more massive galaxies have a larger effective radius up to at least $z \sim 3$ compared to lower mass galaxies. At redshifts higher than $z \sim 3$, at a fixed stellar mass, star-forming and quiescent galaxies have very similar sizes, suggesting that quiescent galaxies may not be smaller than star-forming galaxies at fixed stellar mass at high redshift, which has been a finding for almost 20 yr at $z < 3$ (e.g. Buitrago et al. 2008).

(iii) We find that galaxies appear more compact and smaller when observed in redder filters, and demonstrate the mass dependence of this effect, where more massive galaxies are more compact in redder filters than their lower mass counterparts.

(iv) We find that visually classified spheroid galaxies are smaller than other galaxy types at all redshifts, and have a higher Sérsic index at all redshifts than other galaxy types, showing their small, compact nature.

(v) We verify that our results are due to real evolutionary effects only, shown by fitting simulated high-redshift galaxies with the intrinsic properties preserved, and we recover results that do not vary based on redshift effects.

Overall, this paper, we show that the evolution of galaxy size and structure continues to the highest redshifts, with disc-like galaxies forming most stars within the universe at all epochs. High-redshift morphology studies are revealing a new picture of the structural evolution of galaxies, which will continue further with increasing numbers of high-redshift galaxies being discovered with *JWST*.

This study is just the start of this type of analysis. The benefit of the CEERS fields is that we have very accurate photometric redshifts at lower redshifts, due to overlap with existing *HST* observations. In the near future this will be available for many other and larger fields where more subtle changes in the size and structural features of galaxies will be studied, and this will lead to a more complete understanding of galaxy evolution over nearly the universe's entire history. Future studies with a larger sample of high-redshift galaxies are needed to determine precisely when these aspects first formed, and future observations will provide these increased samples of high-redshift galaxies, through upcoming wide and deep surveys. An alternative method of probing only the highest redshift galaxies could also be used, removing the reliance on *HST* detections.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their comments, which improved this paper. QL, CJC, JT, and NJA acknowledge support from the ERC Advanced Investigator Grant EPOCHS (788113). DA and TH acknowledge support from STFC in the form of PhD studentships. LF acknowledges financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) in the form of a PhD studentship.

This work is based on observations made with the NASA/ESA *Hubble Space Telescope* (*HST*) and NASA/ESA/CSA *James Webb Space Telescope* (*JWST*) obtained from the Mikulski Archive for Space Telescopes (MAST) at the *Space Telescope Science Institute* (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for *JWST*, and NAS 5-26555 for *HST*. The authors

thank all involved in the construction and operations of the telescope as well as those who designed and executed these observations.

The authors thank Anthony Holloway and Sotirios Sanidas for providing their expertise in high-performance computing and other IT support throughout this work. This research made use of the following Python libraries: *ASTROPY* (Astropy Collaboration et al. 2022), *MORFOMETRYKA* (Ferrari, Carvalho & Trevisan 2015), *PANDAS* (Reback et al. 2022), *MATPLOTLIB* (Hunter 2007), and *PHOTUTILS* (Bradley et al. 2020).

DATA AVAILABILITY

The *JWST*/NIRCam images used in this work are available through the Mikulski Archive for Space Telescopes (<https://mast.stsci.edu/>) under Proposal ID 1345. Additional data products will be made available upon reasonable request to the corresponding author.

REFERENCES

- Adams N. J. et al., 2023a, preprint ([arXiv:2304.13721](https://arxiv.org/abs/2304.13721))
 Adams N. J. et al., 2023b, *MNRAS*, 518, 4755
 Almosallam I. A., Jarvis M. J., Roberts S. J., 2016, *MNRAS*, 462, 726
 Ashby M. L. N. et al., 2015, *ApJS*, 218, 33
 Astropy Collaboration, 2022, *ApJ*, 935, 167
 Atek H. et al., 2022, *MNRAS*, 519, 1201
 Austin D. et al., 2023, *ApJ*, 952, L7
 Bagley M. B. et al., 2023, *ApJ*, 946, L12
 Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
 Bluck A. F. L., Conselice C. J., Buitrago F., Grützbauch R., Hoyos C., Mortlock A., Bauer A. E., 2012, *ApJ*, 747, 34
 AstropyCollaboration, The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package, 2022, *ApJ*,
 Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503
 Bridge J. S. et al., 2019, *ApJ*, 882, 42
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Buitrago F., Trujillo I., Conselice C. J., Bouwens R. J., Dickinson M., Yan H., 2008, *ApJ*, 687, L61
 Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, *MNRAS*, 428, 1460
 Buitrago F., Conselice C. J., Epinat B., Bedregal A. G., Grützbauch R., Weiner B. J., 2014, *MNRAS*, 439, 1494
 Caon N., Capaccioli M., D'Onofrio M., 1993, *MNRAS*, 265, 1013
 Castellano M. et al., 2022, *ApJ*, 938, L15
 Chabrier G., 2003, *Publ. Astron. Soc. Pac.*, 115, 763
 Ciotti L., 1991, *A&A*, 249, 99
 Ciotti L., Bertin G., 1999, *A&A*, 352, 447
 Conselice C. J., 2003, *ApJS*, 147, 1
 Conselice C. J., 2014, *ARA&A*, 52, 291
 Costantin L. et al., 2023, *ApJ*, 946, 71
 Curtis-Lake E. et al., 2023, *Nat. Astron.*, 7, 622
 Delgado-Serrano R., Hammer F., Yang Y. B., Puech M., Flores H., Rodrigues M., 2010, *A&A*, 509, A78
 Donnan C. T. et al., 2022, *MNRAS*, 518, 6011
 Duncan K. et al., 2014, *MNRAS*, 444, 2960
 Duncan K. J. et al., 2018a, *MNRAS*, 473, 2655
 Duncan K. J., Jarvis M. J., Brown M. J. I., Röttgering H. J. A., 2018b, *MNRAS*, 477, 5177
 Duncan K. et al., 2019, *ApJ*, 876, 110
 Erwin P., 2015, *ApJ*, 799, 226
 Ferrari F., Carvalho R. R. d., Trevisan M., 2015, *ApJ*, 814, 55
 Ferreira L., Conselice C. J., Duncan K., Cheng T.-Y., Griffiths A., Whitney A., 2020, *ApJ*, 895, 115
 Ferreira L. et al., 2022a, *ApJ*, 955, 15
 Ferreira L. et al., 2022b, *ApJ*, 938, L2
 Finkelstein S. L. et al., 2017, The Cosmic Evolution Early Release Science (CEERS) Survey, *JWST* Proposal ID 1345. Cycle 0 EarlyRelease Science

- Finkelstein S. L. et al., 2022, *ApJ*, 946, L35
- Finkelstein S. L. et al., 2023, *ApJ*, 946, L13
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *Publ. Astron. Soc. Pac.*, 125, 306
- Furlong M. et al., 2017, *MNRAS*, 465, 722
- Gaia Collaboration et al., 2021, *A&A*, 649, A1
- Grogin N. A. et al., 2011, *ApJS*, 197, 35
- Harikane Y. et al., 2023, *ApJS*, 265, 5
- Herschel W., 1786, *Phil. Trans. R. Soc.*, 76, 457
- Hoaglin D. C., Mosteller F., Tukey J. W., 1983, *Understanding Robust and Exploratory Data Analysis*, WileyNew York
- Hoyos C. et al., 2011, *MNRAS*, 411, 2439
- Hoyos C. et al., 2012, *MNRAS*, 419, 2703
- Hubble E. P., 1926, *ApJ*, 64, 321
- Huertas-Company M. et al., 2023, preprint (arXiv:2305.02478)
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ito K. et al., 2023, preprint (arXiv:2307.06994)
- Jacobs C. et al., 2023, *ApJ*, 948, L13
- Kartaltepe J. S. et al., 2022, *ApJ*, 946, L17
- Koekemoer A. M. et al., 2011, *ApJS*, 197, 36
- Kron R. G., 1980, *ApJS*, 43, 305
- Kubo M., Yamada T., Ichikawa T., Kajisawa M., Matsuda Y., Tanaka I., Umehata H., 2017, *MNRAS*, 469, 2235
- Lotz J. M., Primack J., Madau P., 2004, *AJ*, 128, 163
- Lovell C. C. et al., 2023, *MNRAS*
- Ma X. et al., 2018, *MNRAS*, 477, 219
- Mager V. A., Conselice C. J., Seibert M., Gusbar C., Katona A. P., Villari J. M., Madore B. F., Windhorst R. A., 2018, *ApJ*, 864, 123
- Margalef-Bentabol B., Conselice C. J., Mortlock A., Hartley W., Duncan K., Ferguson H. C., Dekel A., Primack J. R., 2016, *MNRAS*, 461, 2728
- Marshall M. A., Wilkins S., Di Matteo T., Roper W. J., Vijayan A. P., Ni Y., Feng Y., Croft R. A. C., 2022, *MNRAS*, 511, 5475
- Matharu J. et al., 2022, *The Astrophysical Journal*, 937, 16
- Morishita T. et al., 2023, preprint (arXiv:2308.05018)
- Mortlock A. et al., 2013, *MNRAS*, 433, 1185
- Naidu R. P. et al., 2022, *ApJ*, 940, L14
- Nelson E. J. et al., 2016, *ApJ*, 828, 27
- Nelson E. J. et al., 2019, *ApJ*, 870, 130
- Oke J. B., 1974, *ApJS*, 27, 21
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Ono Y. et al., 2023, *ApJ*, 951, 72
- Papaderos P., Östlin G., Breda I., 2023, *A&A*, 673, A30
- Paulino-Afonso A., Sobral D., Buitrago F., Afonso J., 2016, *MNRAS*, 465, 2717
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, *AJ*, 139, 2097
- Perrin M. D., Sivaramakrishnan A., Lajoie C.-P., Elliott E., Pueyo L., Ravindranath S., Albert L., 2014, in Oschmann Jacobus M. J., Clampin M., Fazio G. G., MacEwen H. A. eds, *SPIE Conf. Ser. Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. SPIE, Bellingham. p. 91433X
- Popping G. et al., 2022, *MNRAS*, 510, 3321
- Pandas The , 2022, pandas-dev/pandas: Pandas 1.4.2, Zenodo
- Rieke M. J. et al., 2022, *Publ. Astron. Soc. Pac.*, 135, 028001
- Robertson B. E. et al., 2023, *ApJ*, 942, L42
- Roper W. J., Lovell C. C., Vijayan A. P., Marshall M. A., Irodotou D., Kuusisto J. K., Thomas P. A., Wilkins S. M., 2022, *MNRAS*, 514, 1921
- Roper W. J. et al., 2023, *MNRAS*, 526, 6128
- Rosse T. E. o., 1850, *Phil. Trans. R. Soc. London*, 140, 499
- Santini P. et al., 2015, *ApJ*, 801, 97
- Schawinski K. et al., 2014, *MNRAS*, 440, 889
- Sérsic J. L., 1963, *BAAA*, 6, 41
- Stefanon M. et al., 2017, *ApJS*, 229, 32
- Suess K. A. et al., 2022, *ApJ*, 937, L33
- Tacchella S. et al., 2015a, *Science*, 348, 314
- Tacchella S. et al., 2015b, *ApJ*, 802, 101
- Tacchella S. et al., 2018, *ApJ*, 859, 56
- Tacchella S. et al., 2023, *ApJ*, 952, 74
- Taylor-Mager V. A., Conselice C. J., Windhorst R. A., Jansen R. A., 2007, *ApJ*, 659, 162
- Tohill C., Ferreira L., Conselice C. J., Bamford S. P., Ferrari F., 2021, *ApJ*, 916, 4
- Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, *MNRAS*, 382, 109
- van Dokkum P. G. et al., 2010, *ApJ*, 709, 1018
- van der Wel A. et al., 2012, *ApJS*, 203, 24
- van der Wel A. et al., 2014, *ApJ*, 788, 28
- Whitney A., Conselice C. J., Duncan K., Spitler L. R., 2020, *ApJ*, 903, 14
- Whitney A., Ferreira L., Conselice C. J., Duncan K., 2021, *ApJ*, 919, 139
- Wilman D. J. et al., 2020, *ApJ*, 892, 1
- Windhorst R. A. et al., 2002, *ApJS*, 143, 113
- Yan H., Ma Z., Ling C., Cheng C., Huang J.-S., 2022, *ApJ*, 942, L9
- Yang L. et al., 2022, *ApJ*, 938, L17

This paper has been typeset from a \LaTeX file prepared by the author.