First light and reionisation epoch simulations (FLARES) – IV. The size evolution of galaxies at z ≥ 5

Roper, William J, Lovell, Christopher C, Vijayan, Aswin P, Marshall, Madeline A, Irodotou, Dimitrios, Kuusisto, Jussi K, Thomas, Peter A and Wilkins, Stephen M (2022) First light and reionisation epoch simulations (FLARES) – IV. The size evolution of galaxies at z ≥ 5. Monthly Notices of the Royal Astronomical Society, 514 (2). pp. 1921-1939. ISSN 0035-8711

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/111951/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
First Light And Reionisation Epoch Simulations (FLARES) – IV. The size evolution of galaxies at $z \geq 5$

William J. Roper, Christopher C. Lovell, Aswin P. Vijayan, Madeline A. Marshall, Dimitrios Irodotou, Jussi K. Kuusisto, Peter A. Thomas, and Stephen M. Wilkins

1 Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK
2 Centre for Astrophysics Research, School of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield AL10 9AB, UK
3 Cosmic Dawn Center (DAWN), Rindammsgade 62, DK-2200 København N, Denmark
4 DTU Space, Technical University of Denmark, Elektrovej 327, DK-2800 Kongens Lyngby, Denmark
5 National Research Council of Canada, Herzberg Astronomy and Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
6 ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Stromlo, ACT 2611, Australia
7 Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014 Helsinki, Finland

ABSTRACT

We present the intrinsic and observed sizes of galaxies at $z \geq 5$ in the First Light And Reionisation Epoch Simulations (FLARES). We employ the large effective volume of FLARES to produce a sizeable sample of high-redshift galaxies with intrinsic and observed luminosities and half-light radii in a range of rest-frame ultraviolet (UV) and visual photometric bands. This sample contains a significant number of intrinsically ultracompact galaxies in the far-UV (1500 Å), leading to a negative intrinsic far-UV size–luminosity relation. However, after the inclusion of the effects of dust these same compact galaxies exhibit observed sizes that are as much as 50 times larger than those measured from the intrinsic emission, and broadly agree with a range of observational samples. This increase in size is driven by the concentration of dust in the core of galaxies, heavily attenuating the intrinsically brightest regions. At fixed luminosity we find a galaxy size redshift evolution with a slope of $m = 1.21\pm0.87$ depending on the luminosity sample in question, and we demonstrate the wavelength dependence of the size–luminosity relation that will soon be probed by the James Webb Space Telescope.

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

1 INTRODUCTION

Galaxy sizes are governed by a range of processes including galaxy mergers, instabilities, gas accretion, gas transport, star formation, and feedback (Conselice 2014). Studying galaxy sizes helps us to understand the interplay between these key astrophysical processes and galactic structure. By extension, understanding how galaxy sizes evolve tells us how these fundamental physical mechanisms, and the interplay between them, change over time.

At fixed redshift, the size–luminosity relation can be expressed as a power law of the form:

$$R = R_0 \left( \frac{L}{L_{\nu=3}} \right)^\beta,$$

where $R_0$ is a normalization factor, $\beta$ is the slope of the size–luminosity relation, and $L_{\nu=3}$ is the characteristic ultraviolet (UV) luminosity for $z \sim 3$ Lyman-break galaxies (with value $L_{\nu=3} = 10^{20.03}$ erg s$^{-1}$ Hz$^{-1}$), which corresponds to $M_{1609} = -21.0$ (Steidel et al. 1999). As a function of redshift the size evolution can be expressed as

$$R(z) = R_0 |z=0| (1 + z)^{-m},$$

where $R_0 |z=0|$ is another normalization factor corresponding to the size of a galaxy at $z = 0$ and $m$ is the slope of the redshift evolution. In addition to its importance to understanding physical processes, probes of the size–luminosity relation and its evolution are indispensable to our understanding of survey completeness and by extension the luminosity function (Kawamata et al. 2018; Bouwens et al. 2022).

In observations at low redshifts ($z < 3$), galaxies have sizes of the order $1\text{–}30$ proper kpc (pkpc), with actively star-forming galaxies typically larger than their quiescent counterparts (Zhang & Yang 2019; Kawinwanichakij et al. 2021). These galaxies exhibit a positive size–luminosity relation (van der Wel et al. 2014; Suess et al. 2019; Kawinwanichakij et al. 2021), although van der Wel et al. (2014) find a significant number density of compact and massive ($R < 2$ pkpc, $M/M_\odot > 10^{11}$) galaxies at $z = 1.5\text{–}3$, whose number density drops drastically by the current day.

The landscape is different at high redshift where we are primarily probing star-forming galaxies. A number of studies using deep Hubble Space Telescope (HST) fields have measured the sizes of galaxies ($z = 6\text{–}12$ Lyman-break galaxies (Oesch et al. 2010; Grazian et al. 2012; Mosleh et al. 2012; Huang et al. 2013; Ono et al. 2013; Holwerda et al. 2015, 2020; Kawamata et al. 2015, 2018; Shibuya, Ouchi & Harikane 2015). In contrast to the low-redshift size regime, these studies found bright star-forming galaxies with compact half-light radii of $0.5\text{–}1.0$ pkpc.

* E-mail: w.roper@sussex.ac.uk

© 2022 The Author(s). Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
There is a growing consensus that the high-redshift size–luminosity relation is positively sloped ($\beta > 0$), as it is at low redshift, with a range of reported slopes and differing reports of $\beta$’s redshift evolution.

(i) Grazian et al. (2012) find $\beta = 0.3\pm0.5$ at $z \approx 7$.
(ii) Huang et al. (2013) find $\beta = [0.22, 0.25]$ for $z = 4$ and $z = 5$, respectively.
(iii) Holwerda et al. (2015) find $\beta = 0.24 \pm 0.06$ at $z \approx 7$ and $\beta = 0.12 \pm 0.09$ at $z \approx 9$–10.
(iv) Shibuya et al. (2015) find a redshift-independent slope of $\beta = 0.27 \pm 0.01$ in the range $z = 0$–8.
(v) Kawamata et al. (2018) find steeply sloped relations with $\beta = [0.46, 0.46, 0.38, 0.56]$ at $z = [6, 7, 8, 9]$.

Recent lensing studies agree with the steeper slope of Kawamata et al. (2018), itself using a sample including lensed sources. Bouwens et al. (2022) find $\beta = 0.40 \pm 0.04$ for a galaxy sample in the redshift range $z \approx 6$–8, while Yang et al. (2022) find $\beta = 0.48 \pm 0.08$ for $z \approx 6$–7 and $\beta = 0.68 \pm 0.14$ for $z \approx 8.5$ (assuming the Braudac lens model; Bradač et al. 2005). This steeper slope is driven by compact dim galaxies that are better sampled in lensing studies. Neufeld et al. (2021) also find extremely compact lensed galaxies with $R_{1/2} < 200$ ppc in the Reionization Lensing Cluster Survey (REILCS) data set; they identify these galaxies as potential Lyman continuum (LyC) leakers, possibly representing a large contribution to reionization.

A similar range of results exists within measurements of the redshift dependence of galaxy size at fixed luminosity with slopes in the range $1 < m < 1.5$ (Bouwens et al. 2004; Oesch et al. 2010; Ono et al. 2013; Kawamata et al. 2015, 2018; Shibuya et al. 2015; Laporte et al. 2016). This is consistent with two theoretical scenarios: $m = 1$, the expected scaling for systems of fixed mass (e.g. Bouwens et al. 2004), and $m = 1.5$, the expected evolution for systems with fixed circular velocity (e.g. Ferguson et al. 2004; Hathi, Malhotra & Rhoads 2008). However, galaxy sizes are not wholly dependent on these theoretical scalings with significant contributions from baryonic processes such as stellar and active galactic nuclei (AGN) feedback (Wyithe & Loeb 2011).

Simulations provide detailed information on the properties of the underlying components that make up galaxies. From this information we can probe large samples of galaxies with knowledge of the intrinsic physical processes governing their evolution, albeit processes that are themselves dictated by subgrid models that are sensitive to their physical model and parameter assumptions. The intrinsic properties of particles and their spatial distribution can be utilized to measure galaxy properties such as their half-mass/light radii at the mass resolution of the simulation without the associated uncertainties inherent in measurements of this kind in observations. Using this fidelity, the size–mass and size–luminosity relations have been probed by many simulations. However, much of this analysis still focuses on comparatively low redshifts. Furlong et al. (2017) analysed the EAGLE simulation and found a good agreement with observed trends using intrinsic particle measurements to find a positive ($\beta > 0$) size–mass relation that flattens at $z = 2$, and an increase in size with decreasing redshift over the range $z = 0$–2.

At higher redshift ($z = 6$), the SIMBA simulations (Davé et al. 2019) find a positive far-UV attenuated size–luminosity relation while showing the dust attenuated size is significantly larger than the intrinsic size, with the magnitude of this increase a function of stellar mass (Wu et al. 2020). This implies a flatter intrinsic size–luminosity relation at high redshift. This flattened intrinsic size–luminosity relation is particularly evident in the BLUETIDES simulation (Feng et al. 2016; Marshall et al. 2022), which has been used to probe the UV and visual size–luminosity relations with synthetic observations at $z \geq 7$. In doing so they find a negative intrinsic size–luminosity relation ($\beta < 0$) in the far-UV that flips to positive after the inclusion of dust attenuation ($\beta > 0$). They also probe the redshift evolution of size, finding a shallow redshift evolution of $m = 0.662 \pm 0.008$ in agreement with the redshift evolution of Holwerda et al. (2015).

In addition to the higher redshift results derived from BLUETIDES, the ILLUSTRIS-TNG simulations have also exhibited a negative size–luminosity relation at $z = 5$ (Popping et al. 2022).

The FIRE-2 simulations (Ma et al. 2018) present a sample of compact galaxies with sizes of $0.05$–1 pkpc, in the range $-22 < M_{UV} < -7$ at $z = [6, 8, 10]$. The sizes in this sample are measured from synthetic galaxy images of the intrinsic stellar emission using a non-parametric pixel method, which converts the pixel area containing half the total luminosity to a half-light radius. Unlike Marshall et al. (2022) this sample exhibits a size–mass relation and $B$-band size–luminosity relation with $\beta > 0$. The FIRE-2 galaxy sample extends to galaxies far fainter than those present in other simulated samples, which could explain the differences in size–mass and size–luminosity relations. They also present redshift evolution slopes derived in fixed stellar mass regimes that produce values of $1 < m < 2$, encompassing many of the observational measurements but extending to more extreme values for the brightest and most massive galaxies.

Clearly there is much work to be done in understanding galaxy size at this epoch, especially with the impending first light of the James Webb Space Telescope (JWST) and other next–generation observatories. In this paper, we analyse the large sample of galaxies produced by the First Light And Reionisation Epoch Simulations (FLARES) simulations (Lovell et al. 2021; Vijayan et al. 2021). FLARES is uniquely placed to complement previous studies of high-redshift galaxy size due to its enormous effective volume, coverage a wide array of environments during the Epoch of Reionization (EoR), and sufficient mass resolution, producing a large and robust galaxy sample. In previous work, we have shown that FLARES reproduces the distributions of stellar mass, star formation rate (SFR), and UV luminosity up to $z = 10$.

The rest of this paper is structured as follows. In Section 2, we detail the simulations themselves. In Section 3, we detail the methods used to make synthetic photometry and observations. In Section 4, we detail the galaxy sample and size measurement methods. In Section 5, we present the results of this analysis of the size–luminosity relation. We present our conclusions in Section 6. Throughout this work we assume a Planck Year 1 cosmology ($\Omega_0 = 0.307, \Omega_{\Lambda} = 0.693, h = 0.6777$; Planck Collaboration I 2014) and a Chabrier stellar initial mass function (IMF; Chabrier 2003).

### 2 FIRST LIGHT AND REIONISATION EPOCH SIMULATIONS (FLARES)

FLARES is a simulation program targeting the EoR. It consists of 40 zoom simulations, targeting regions with a range of overdensities drawn from an enormous (3.2 comoving Gpc – cGpc$^3$) dark matter only simulation (Barnes et al. 2017a), which we will refer to as the ‘parent’. The regions are selected at $z = 4.67$, which ensures that extreme overdensities are only mildly non-linear, and thus the rank ordering of overdensities at higher redshifts is approximately preserved. Regions are defined as spheres with radius 14 cMpc $h^{-1}$, and their overdensities are selected to span a wide range ($\delta = 0.479$–0.970; see table A1 of Lovell et al. 2021) in order to sample the most under- and overdense environments at this cosmic time, the latter containing a large sample of the most massive galaxies, thought to be biased to such regions (Chiang, Overzier & Gebhardt 2013; Lovell,
These regions are then resimulated with full hydrodynamics using the EAGLE model (Crain et al. 2015; Schaye et al. 2015).

The EAGLE project consists of a series of hydrodynamic cosmological simulations, with varying resolutions and box sizes. The code is based on a heavily modified version of P-GADGET-3, a smooth particle hydrodynamics (SPH) code last described in Springel et al. (2005b). The hydrodynamic solver is collectively known as ANARCHY (described in Schaller et al. 2015; Schaye et al. 2015), and adopts the pressure–entropy formulation described by Hopkins (2013), an artificial viscosity switch (Cullen & Dehnen 2010), and an artificial conduction switch (e.g. Price 2008). The model includes prescriptions for radiative cooling and photoheating (Wiersma, Schaye & Smith 2009a), star formation (Schaye & Dalla Vecchia 2008), stellar evolution and mass loss (Wiersma et al. 2009b), feedback from star formation (Dalla Vecchia & Schaye 2012), and black hole growth and AGN feedback (Springel, Di Matteo & Hernquist 2005a; Booth & Schaye 2009; Rosas-Guevara et al. 2015). The $z=0$ galaxy mass function, the mass–size relation for discs, and the gas mass–halo mass relation were used to calibrate the free parameters of the subgrid model. The model is in good agreement with a number of observables at low redshift not considered in the calibration (e.g. Furlong et al. 2015; Lagos et al. 2015; Trayford et al. 2015).

FLARES uses the AGN iT9 configuration of the model, which produces similar mass functions to the fiducial reference model, but better reproduces the hot gas properties of groups and clusters (Barnes et al. 2017b). It uses a higher value for $C_{\text{visc}}$, a parameter for the effective viscosity of the subgrid accretion, and a higher gas temperature increase from AGN feedback, $\Delta T$. These modifications give less frequent, more energetic AGN outbursts.

The FLARES simulations have an identical resolution to the 100 cMpc EAGLE reference simulation box, with a dark matter and an initial gas particle mass of $m_{\text{dm}} = 9.7 \times 10^9 M_\odot$ and $m_g = 1.8 \times 10^9 M_\odot$, respectively, and has a gravitational softening length of 2.66 ckpc at $z \geq 2.8$.

In order to obtain a representative sample of the Universe, by combining these regions using appropriate weightings corresponding to their relative overdensity, we are able to create composite distribution functions that represent much larger volumes than those explicitly simulated. For a more detailed description of the simulation and weighting method we refer the reader to Lovell et al. (2021).

### 2.1 Galaxy extraction

We follow the same structure extraction method as the EAGLE project: this is explained in detail in McAlpine et al. (2016). In brief, dark matter overdensities are identified using a Friends-of-Friends (FoF) approach (Davis et al. 1985) with the usual linking length of $\ell = 0.2\bar{x}$, where $\bar{x}$ is the mean interparticle separation. All other particle types are then assigned to the halo containing their nearest dark matter neighbour. These FoF haloes are then refined to produced self-bound ‘subgroups’ (galaxies) containing both dark matter and baryonic particles using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009).

The SUBFIND method involves finding saddle points in the density field in a FoF halo to identify self-bound substructures. This can lead to spurious oversplitting of extremely dense galaxies where saddle points are misidentified near density peaks. These objects often contain mainly a single particle type and have anomalous integrated properties. Although they make up $<0.1$ per cent of all galaxies $>10^9 M_\odot$ at $z=5$, we identify and recombine them into their parent structure in post-processing. To do this we label a ‘galaxy’ as spurious if it has any zero mass contributions in the stellar, gas, or dark matter components. We remove the spurious galaxies from the SUBFIND catalogue and add their particle properties to the parent ‘central’ subhalo, including the reassigned particles in any integrated quantities.

In a minority of pathological cases tidal stripping can cause galaxies to exhibit diffuse populations of particles at large radii. Although identified by SUBFIND as belonging to a galaxy, these distributions can have a large effect on integrated quantities such as the total luminosity and the half-light radius. For this reason we adopt a 30 pkpc aperture in line with all EAGLE and FLARES papers and calculate all integrated properties using only particles associated with each subgroup within this aperture. This aperture ensures the majority of galaxies have mass distributions that are wholly within this aperture and any erroneous distributions at large radii are omitted.

### 3 MODELLING PHOTOMETRY

We use the approach presented in Vijayan et al. (2021), (henceforth FLARES II) to produce resolved galaxy images, both including and excluding the effects of dust. We first produce spectral energy distribution (SEDs) and then apply top-hat rest-frame UV and visual band filters to extract photometry. As in FLARES II we focus on the stellar emission, deferring the treatment of accretion on to the supermassive black holes to a future work. However, as will be shown in the coming sections this simplification does not pose a significant challenge to the results of this work. This approach broadly follows Wilkins et al. (2016, 2017, 2018, 2020), with modifications to the dust treatment. For a full description of this method and discussion of the free parameters see FLARESII. What follows is a brief summary of the approach to compute galaxy images.

#### 3.1 Spectral energy distribution modelling

In this work, we use the SYNTHOBS module\(^1\) to produce synthetic rest-frame photometry primarily focusing on a top-hat far-UV (1500 Å) filter with a wavelength range of $1300 \leq \lambda \leq 1700$ Å. We do however calculate results for a range of different filters all shown in the example SED in Fig. 1. Each component of the stellar luminosity can be included independently enabling the probing of both the intrinsic luminosity and the effects of dust extinction. In this section, we briefly detail each component.

##### 3.1.1 Stellar emission

For the pure stellar emission we start with a simple stellar population model (SSP) by associating each stellar particle with a stellar SED based on the particle’s age and metallicity. As with FLARES II we use v2.2.1 of the Binary Population and Spectral Synthesis (BPASS) stellar population synthesis (SPS) models (Stanway & Eldridge 2018) and assume a Chabrier (2003) IMF. As shown in Wilkins et al. (2016, 2017, 2018) the resulting luminosities are sensitive to the choice of SPS and IMF used in their derivation.

##### 3.1.2 Nebular emission

To account for the LyC emission of young stellar populations we associate young stellar particles ($t < 10$ Myr, following the

\(^1\)github.com/stephenmwilkins/SynthObs
assumption from Charlot & Fall 2000 that birth clouds dissipate on these scales) to a H II region (or birth cloud). To include the LyC emission for each stellar particle we follow the approach detailed in Wilkins et al. (2020), in which the pure stellar spectrum is processed with the CLOUDY photoionization code (Ferland et al. 2017) assuming:

(i) the H II region’s metallicity is identical to the stellar particle’s;
(ii) dust depletion and relative abundances from Gutkin, Charlot & Bruzual (2016);
(iii) a reference ionization parameter (defined at $t = 1$ Myr and $Z = 0.02$) of $\log_{10}(U_{\text{LyC}}) = -2$;
(iv) a hydrogen density of $\log_{10}(n_H/cm^{-3}) = 2.5$;
(v) CLOUDY’s default Orion-type graphite and silicate grains.

### 3.1.3 Dust attenuation

To include the effects of dust attenuation from the interstellar medium (ISM) we adopt a line-of-sight (LoS) attenuation model. In this model we treat stellar particles as emitters along an LoS (in this paper, we select the z-axis of the simulation) and account for the attenuation due to gas particles that intersect this LoS. Using an LoS approach means stellar emission undergoes spatially resolved attenuation rather than the uniform attenuation of a simple screen model, enabling considerably more robust photometry.

To do this we find all gas particle SPH kernels that intersect the stellar particle’s LoS and integrate along it to get the metal column density, $\Sigma(x, y)$. We then link this metal column density to the ISM dust optical depth in the V band (550 nm), $\tau_{\text{ISM}}(x, y)$, with a similar approach as in Wilkins et al. (2017). This gives the expression

$$\tau_{\text{SMY}}(x, y) = D_{\text{ISM}} \Sigma(x, y),$$

where $D_{\text{ISM}}$ is the galaxy specific dust-to-metal ratio from the fitting function presented in Vijayan et al. (2019). This is a function of the mass-weighted stellar age ($t$) and the gas-phase metallicity of a galaxy ($Z$),

$$D_{\text{ISM}} = D_0 + (D_1 - D_0) \left[ 1 - \exp \left( -\alpha Z^\beta (t/\gamma)^\gamma \right) \right],$$

where $D_0$ and $D_1$ represent the initial Type II supernovae (SNe) dust injection and saturation, respectively, and $\tau$ is an estimate of the initial dust growth time-scale after dust injection from Type II SNe but prior to the initiation of dust growth on grains. The normalization factor $k_{\text{ISM}}$ was chosen to match the rest-frame ultraviolet luminosity function (UVLF) from Bouwens et al. (2015) and acts as a proxy for dust properties such as average grain size, shape, and composition ($k_{\text{ISM}} = 0.0795$). The FLARES simulations do not inherently model dust production and destruction, thus we have to resort to these data-driven proxies.

In addition to attenuation due to the ISM, young stellar populations ($t < 10$ Myr) are still embedded in their birth clouds and thus need to take into account attenuation due to this cloud. For these young stellar particles we include the additional attenuation expression:

$$\tau_{\text{BCY}}(x, y) = k_{\text{BC}}(Z/0.01),$$

where $Z$ is the metallicity of the young stellar particle and $k_{\text{BC}}$ is another normalization factor encapsulating the dust properties of the birth cloud, for this we assume a constant value of $k_{\text{BC}} = 1$. For stellar particles older than 10 Myr, $\tau_{\text{BCY}}(x, y) = 0$ and there is no contribution. In Appendix D, we present sizes omitting this birth cloud contribution to quantify its effect on galaxy size.

We then combine these optical depths in the $V$ band,

$$\tau_s = (\tau_{\text{BCY}} + \tau_{\text{SMY}}) \left( \frac{\lambda}{550 \text{ nm}} \right)^{-1},$$

yielding an expression for the optical depth at other wavelengths that can be applied to the stellar particle SEDs to account for dust attenuation.

### 3.2 Image creation

We then apply top-hat photometric band filters to the SEDs producing photometry for each stellar particle. Using this photometry we produce synthetic observations with a field of view (FoV) of $60 \times 60$ pkpc$^2$ encompassing the entire 30 pkpc aperture in which a galaxy’s integrated quantities are measured (corresponding to 9.34, 12.20, and 14.13 arcsec at $z = 5, 8$, and $10$, respectively); see Section 2.1. We adopt a resolution equal to the redshift-dependent softening length of the simulation ($s = 2.66/(1 + z)$ pkpc). In Appendix C, we present a comparison of image resolutions.

Synthetic images are often created by treating each stellar particle as a two-dimensional Gaussian kernel. The standard deviation of this kernel can either be defined by the softening length ($\sigma = s$ producing minimal smoothing), the stellar particle’s smoothing length ($\sigma = h_{\text{nml}}$, accounting for the local density), or, most often, the proximity to the Nth neighbouring stellar particle ($\sigma = r_N$) (e.g. Torrey et al. 2015; Ma et al. 2018; Marshall et al. 2022). The full image is then

\[\text{Figure 1. The median rest-frame SEDs for all galaxies in all FLARES regions at } z = 5 \text{ with } 10^{10} \leq M_*/M_\odot \leq 10^{11.3} \text{ produced by SYNTHOBS. The top panel shows the intrinsic stellar SED in green and the dust attenuated SED (including LoS effects) in red. The lower panel shows the rest-frame top-hat photometric filters used throughout this analysis, plotted with an arbitrary y-axis to aid interpretation. The black lines correspond to the location and bandwidth of the James Webb Space Telescope (JWST)’s Near-Infrared Camera (NIRCam)’s reddest wide-band filter (F444W) at the indicated redshifts. This indicates the reddest rest-frame bands accessible by the JWST at high enough resolution to measure robust sizes with NIRCam (0.062 arcsec) at } z > 5.\]
a sum over these contributions. In this method an image \( I \) can therefore be expressed mathematically as
\[
I_i = \exp \left( -\frac{(X - x_i)^2 + (Y - y_i)^2}{2\sigma_i^2} \right),
\]
\[
I = \sum_{j=1}^{N_i} \frac{I_{ij}}{\sum_{i=1}^{N_{pix}} I_i},
\]
where \( I_i \) is the smoothed image (kernel) produced for the \( i \)th stellar particle, \( \sigma_i \) is the standard deviation of the \( i \)th stellar particle’s kernel, \( X \) and \( Y \) are a grid of pixel positions, \( x_i \) and \( y_i \) are the \( i \)th stellar particle’s \( x \)-axis and \( y \)-axis positions in the desired projection, \( L_i \) is the luminosity of the \( i \)th particle, and the sum in the denominator is a sum over all pixels for the \( i \)th stellar particle to normalize the kernel.

However, this approach not only differs from the SPH treatment of a stellar particle but is also extremely computationally expensive. Unless artificially truncated a Gaussian kernel encompasses the whole image, leading to insignificant but time-consuming calculations. In fact, in SPH simulations a stellar particle is treated as a representation of a fluid with the full extent of the stellar population described by a spline kernel with a definitive cut-off where the kernel falls to 0 (Borrow et al. 2022). Using a spline kernel-based approach is not only a better representation of the underlying simulation’s treatment of stellar particles but also greatly reduces the size of the computation by limiting the number of pixels computed per stellar particle.

For these reasons we implement a method of smoothing employing the SPH kernel used in the simulation to describe a stellar particle’s ‘extent’. In the ANARCHY SPH scheme, used in the EAGLE model (Schaye et al. 2015), this kernel is the C2 Wendland kernel (Wendland 1995; Dehnen & Aly 2012). We therefore adopt this kernel in this work, but note that for other simulations the kernel corresponding to that particular simulation should be used to maximize the fidelity of this method.

As with the Gaussian approach, an image can be described as a sum over kernels; unlike the Gaussian approach however, the spline kernels are necessarily three-dimensional and need projecting into the \( x-y \) plane. To achieve this we calculate the spline kernels on a voxel grid and sum over the \( z \)-axis,
\[
I = \sum_{i=0}^{N_{vox}} \sum_{j=1}^{N_i} \frac{K_{ij}}{K_i} L_i,
\]
where each stellar particle’s kernel \( K_i \) is now
\[
K_i = \frac{21}{2r^3 h^3_{\text{gal}}},
\]
with the kernel \( w_i \) given by
\[
w_i(q_i = r/h_i) = \begin{cases} (1 - q_i)^3 (1 + 4q_i), & q_i \leq 1, \\ 0, & q_i > 1, \end{cases}
\]
where \( r \) is the distance between the particle and any given voxel within the kernel.

To compute this kernel efficiently we employ a KD-Tree algorithm, building a tree based on voxel coordinates. We query the tree for all non-zero pixels where the distance between the pixel and the stellar particle \( r \) is less than the limits of the smoothing kernel (here \( r < h \)), greatly reducing the computation from \( O(N_i N_{vox}) \) in the Gaussian case to \( O(N_i N_{vox} r < h) \) using the more representative spline approach.

In Fig. 2, we present a grid of randomly selected galaxies in the far-UV filter along with their stellar mass (derived by summing the underlying particle distribution), luminosities, central surface densities, and half-light radii measured including the effects of dust.

It should be noted that throughout this analysis we do not rotate galaxies, instead adopting their existing orientation in the box to emulate the stochastic viewing angles of galaxies in the real Universe. Henceforth, all analysis derived from images will use this method of stellar particle smoothing (implemented from Section 4.2.2 onwards), unless explicitly stated otherwise. In Appendix A, we present comparisons between the Gaussian and spline approach for this simulation.

4 GALAXY SELECTION AND SIZE MEASUREMENT

In this section, we describe our galaxy sample, and describe the two measurement methods used to derive sizes.

4.1 Extracting the galaxy sample

To ensure all galaxies in the sample have enough particles to be considered morphologically resolved, we omit all subgroups with fewer than 100 stellar particles \( N_i < 100 \). We apply a 95 per cent completeness criterion, dividing the sample of galaxies into those above and below the completeness limits in mass and luminosity. These completeness limits are given by the mass and luminosity at which the galaxy sample is missing 5 per cent due to galaxies having \( N_i < 100 \). We adopt 95 per cent complete rather than 100 per cent complete to avoid the luminosity threshold being defined by anomalously bright galaxies with \( N_i < 100 \). These limits are presented in Table 1 at each redshift for the far-UV band. This ensures we present results motivated by a complete galaxy sample. We note the less present the incomplete sample at low opacity in all scatter plots for context.

We further distinguish between two morphological populations by applying a threshold derived from the intrinsic size–luminosity relation of \( 5 \geq 10^{29} \text{erg s}^{-1} \text{Hz}^{-1} \text{pkpc}^{-2} \) to their central surface flux density (i.e. the surface flux density within the half-light radius). This threshold splits the sample into a population of centrally compact galaxies and a population of diffuse galaxies; in subsequent plots we will denote the compact population by coloured hexbins and the diffuse population by greyscale hexbins.

This division of the galaxy sample is shown in the mass–luminosity relation in Fig. 3 at \( z = 5 \); here we have adopted the previously described colouring and have used opacity to distinguish the complete and incomplete populations. The dashed lines denote the completeness limits in mass and luminosity. The histograms on the axes show the galaxy distribution along each axis with the full galaxy population in grey and galaxies with \( N_i \geq 100 \) shown in black. All following plots will follow these plotting conventions, with greyscale colours denoting the diffuse galaxy distribution and coloured hexbins denoting the compact population (as defined by their central surface density). The hexbins themselves indicate the weighted number density of galaxies, using the weights derived in Lovell et al. (2021). All fits are performed on the complete sample. This division of the galaxy sample leads to:

(i) 50,238 galaxies in the sample with more than 100 stellar particles (25,556, 2863, and 492 at \( z = 5, 8, \) and 10, respectively);
(ii) 7172 in the compact population with more than 100 stellar particles (2701, 696, and 240 at \( z = 5, 8, \) and 10, respectively);
(iii) 43,066 in the diffuse population with more than 100 stellar particles (22,855, 2167, and 252 at \( z = 5, 8, \) and 10, respectively);
(iv) 31,697 galaxies in total above the completeness limit (16,238, 1700, and 273 at \( z = 5, 8, \) and 10, respectively).
Figure 2. A subset of $z = 5$ synthetic far-UV galaxy images computed using the method outlined in Section 3.2. Each panel is the full $60 \times 60$ pkpc$^2$ FoV for each galaxy. Galaxies increase in mass left to right and increase in central surface density top to bottom. The pixel values of these images are linearly normalized across all panels with their mass, luminosity, central surface density, and half-light radius included in each panel. The galaxies included in this subset were randomly selected from each mass and central surface density bin, even so they display the variety of morphologies already present by $z = 5$ in FLARES.

4.2 Size measurement methods

There are a myriad of methods used to define the sizes of galaxies present in the literature including Sérsic profile fitting (Sérsic 1963, 1968), curves of growth (e.g. Bouwens et al. 2004; Ferguson et al. 2004; Oesch et al. 2010), Petrosian radius (Petrosian 1976), and simulation specific methods that use the particle distribution to find the radius enclosing a percentage of the total mass/luminosity.

Each measurement method introduces its own dependencies and challenges. In this section, we detail and compare the two methods utilized in this analysis: a particle-based method, and a non-parametric pixel-based method (e.g. Ribeiro et al. 2016; Ma et al. 2018; Marshall et al. 2022). We neglect curves of growth, Sérsic profiles entirely; at these redshifts the clumpy nature of galaxies, particularly at lower masses (Jiang et al. 2013; Bowler et al. 2017), makes these methods unreliable. Throughout this work we use $R$ to refer to the half-light radius (size) of a galaxy.

4.2.1 Particle-based method

We take the underlying particle distribution within a 30 pkpc aperture and find the radius of the particle bounding half the total luminosity inside this aperture. We then interpolate around this initial measurement to better sample the radial density profile, mitigating its discretization into individual, comparatively low resolution, particles.
Table 1. The mass and luminosity 95 per cent completeness limits for the galaxy sample in each redshift bin. The mass limits are consistent across all bands, but the luminosity limits are band specific. Here we present the far-UV (1500 Å) limits focused on for the majority of the analysis presented in this paper.

| Redshift (z) | \( \log_{10}(M/M_\odot) \) | \( \log_{10}(L_{\text{bol}}/\text{erg s}^{-1} \text{ Hz}^{-1}) \) | \( \log_{10}(L_{\alpha}/\text{erg s}^{-1} \text{ Hz}^{-1}) \) |
|-------------|-------------------------------|-----------------------------------------------------------------|----------------|
| 12          | 8.16                          | 28.60                                                           | 28.43          |
| 11          | 8.15                          | 28.55                                                           | 28.42          |
| 10          | 8.15                          | 28.52                                                           | 28.39          |
| 9           | 8.14                          | 28.46                                                           | 28.34          |
| 8           | 8.13                          | 28.40                                                           | 28.28          |
| 7           | 8.13                          | 28.31                                                           | 28.19          |
| 6           | 8.12                          | 28.24                                                           | 28.12          |
| 5           | 8.11                          | 28.16                                                           | 28.03          |

It should be noted that this measurement method is sensitive to the chosen galactic centre; in this work we use the centre of potential calculated by SUBFIND. Other choices, such as the centroid, can give different results for diffuse and irregular structures since the centre of potential may be located within one of the clumps, which may not necessarily lie in the centre of the galaxy. This offset centre leads to larger size measurements, as the majority of the stellar material of the galaxy is offset from the centre from which the radius is measured.

In all plots including this measurement we take the luminosity to be the sum of each individual particle’s luminosity within the aperture, neglecting any smoothing over the SPH kernel.

4.2.2 Pixel-based method

In the non-parametric pixel approach, the pixels of the image are ordered from most luminous to least luminous. We then find the pixel area containing half the total luminosity before converting to a radius assuming a circular area, \( R = \sqrt{A/\pi} \), and then interpolating around this radius as in the particle method. Unlike the particle method this method of measurement has a minimum possible size where half the total luminosity falls within a single pixel, resulting in a radius of \( R_{\text{min}} = \sqrt{A_{\text{pix}}/\pi} \) before interpolation between 0 and \( R_{\text{min}} \). The interpolation here allows for the measurement of half-light radii smaller than a single pixel, however this does not remove the limitation caused by the finite pixel resolution.

This method is particularly robust at high redshifts, where the independence from a centre definition and non-contiguous size definition better encapsulate the morphology of clumpy structures.

In all plots using this measurement we present the luminosities as detected from the image, i.e. the sum of all pixels within the FoV. This can subtly differ from the particle luminosities where a particle’s kernel extends beyond the bounds of the FoV, spreading the particles light outside the image in contrast to the particle-based method.

4.2.3 Comparing particle and pixel methods

In Fig. 4, we present a comparison of these methods for the sizes of all galaxies at \( z = 5.0 \) using their intrinsic luminosities. For the compact galaxies (colour) we see a reasonable correspondence between the two methods with a scatter around the 1:1 relation. However, as the size of a galaxy increases the particle method begins to produce larger sizes than the pixel method due to a combination of centring effects and luminous structures within the outskirts of galaxies, such as those shown in a number of panels in Fig. 2. Conversely, for the smallest galaxies, the pixel size is larger than the particle size; this is a manifestation of the stellar particle smoothing used in the creation of the images, where light concentrated in densely packed particles is smoothed over a larger pixel area.

For the diffuse (greyscale) population the scatter is more pronounced and extends towards larger particle values across the full range of sizes. This is because of the aforementioned strength of the pixel method when it comes to clumpy diffuse structures and the issue of defining a centre for these structures in the particle method. The size floor is also evident in the smallest galaxies in the diffuse (and incomplete) sample where a single pixel contains half the total luminosity of the dim galaxy.

5 SIZE–LUMINOSITY RELATIONS

Here we present results for the sizes of galaxies in the EoR. All plots that compare to observational quantities are derived from the pixel measurement method (Section 4.2.2) measured from the synthetic images detailed in Section 3.2. Intrinsic properties such as the intrinsic size–luminosity relation (Section 5.1) and half-dust
mass

MNRAS

compact

more

mass

with

(Popping

in

distinct

particle

Although

5.1

galaxies,

relationship.

the

diffuse

galaxy

population,

while

the

lower

panel

of

coloured

points

shows

the

compact
galaxy

population.

The

dashed

black

line

corresponds
to

a

1:1

relationship.

Each

hexbin

is

coloured

by

the

weighted

number

density

of
galaxies,

using

the

FLARES

region

weighting

scheme.

radius

(Section

5.2.1)

are

measured

using

the

particle

method
to

focus

on

the

intrinsic

nature

of

these

properties.

5.1 Intrinsic UV size–luminosity relation

Although impossible to probe in observations, we can use the intrinsic UV size–luminosity relation to trace the underlying stellar population in galaxies. Fig. 5 shows this relation at $z = 5$ for the particle measurements. This shows two surprising features: two distinct populations, and a clear negative slope to the intrinsic size–luminosity relation.

Although the negative slope of the intrinsic size–luminosity relation is somewhat counter-intuitive, it has been seen at these redshifts in other recent simulations, particularly in BLUE/TIDES (Marshall et al. 2022) with a negative size–mass relation at $z = 7$ and ILLU/TIS-TNG (Popping et al. 2022) with a negative observed-frame 850 µm size–mass relation at $z = 5$. Indeed, there are also hints in observations with evidence for a constant dependence between galaxy size and mass (Lang et al. 2014; Mosleh et al. 2020).

Here the division in central surface density is particularly evident. In terms of luminosity we have one dim ($L \lesssim 10^{39}$ erg s$^{-1}$ Hz$^{-1}$) and more diffuse population, and one bright ($L \gtrsim 10^{39}$ erg s$^{-1}$ Hz$^{-1}$) and compact ($R_{1/2} \lesssim 1$ pkpc) population.

As shown in Furlong et al. (2017), the EAGLE low-redshift intrinsic size–mass relation, a good tracer for the intrinsic size–luminosity relation, is positively sloped with a good agreement with observational results. Below we briefly outline the physical mechanisms in FLARES and the EAGLE model that cause the bimodality and negative intrinsic size–luminosity relation at high redshift, and how they evolve leading to the results in the low-redshift regime. We will present our investigation into the physical mechanisms governing the cause and evolution of the two populations in detail in an upcoming paper.

(i) At $z \gtrsim 5$, galaxies that reach $M/M_\odot \gtrsim 10^9$ develop extremely dense cores and begin a spike in core star formation at high stellar birth densities.

(ii) This begins to seed the gas in the galaxy’s core with metals, increasing the effectiveness of metal line cooling, inhibiting stellar and AGN feedback, and further driving star formation.

(iii) This overcooling causes a feedback loop of star formation in the galaxy’s core, allowing the galaxy to become massive and ultracompact during this early epoch.

(iv) While this process takes place in the galaxy’s core the galaxy acquires an extended gas distribution up to 100 times larger than the stellar distribution. Because of the high densities in the core, stellar feedback is unable to mix the core’s metals into this surrounding gas distribution. This lack of metals inhibits cooling and leaves the extended gas distribution unable to efficiently form stars.

(v) At $z \lesssim 4$ the extended gas distribution reaches the density and metallicity necessary for efficient star formation. This is facilitated partly by their own collapse and partly due to the growing efficiency of stellar and AGN feedback (Crain et al. 2015), mixing metals from the core into the surroundings. This extended star formation manifests as an increase in intrinsic galaxy size at late times, yielding a positively sloped intrinsic size–luminosity relation.

In the upper panel of Fig. 6, we present a stack of the central intrinsic emission of all galaxies at $z = 5$ in FLARES (irrespective of completeness) split into mass bins of $\log_{10}(M/M_\odot) = [8–9, 9–9.5,$
9.5–10, >10]. This qualitatively shows how the negative gradient in the size–luminosity relation translates to the compactification of a galaxy’s intrinsic emission in relation to a galaxy’s mass. In the lower panel of Fig. 6, we plot one-dimensional profiles of the stacked mass bin images to explicitly show the compactification. As with the stacked images, the profiles exhibit a narrowing and increasing central concentration with increasing mass. The overcooling begins to take effect between the leftmost mass bin (10^6 < M/M⊙ < 10^7) and the next mass bin of 10^7 < M/M⊙ < 10^8. At this crossover between regimes there is a narrowing of the profile and steeper concentrated peak, which becomes more peaked as the mass increases. The growth of this central peak then drops off in the final mass bin due to an increased contribution by the wings of the profile; galaxies in this mass bin exist in the most dense environments and thus include more luminous substructure at large radii.

In Fig. 7, we show the mass dependence of the half-light radii shown in Fig. 5 and exponential fits to profiles derived from intrinsic stacked images, such as those shown in Fig. 6, split into mass bins of ΔM = 10^{8.4} M⊙. Stacking the images mitigates the issue raised by clumpy structures at high redshift, as detailed in Section 4.2, enabling exponential profile fitting. The scale lengths follow the same negative trend as the intrinsic half-light radii with smaller sizes at low masses relative to the half-light radii. The exponential profiles fit the central regions of the stacks well but fail to fit the profile in the wings, hence the smaller scale lengths for the most diffuse galaxies.

5.2 The effects of dust

We now move on from the intrinsic size–luminosity relation to discuss the effects of dust on the observed UV size–luminosity relation. All plots from this point on will present the pixel measured sizes unless explicitly stated otherwise.
concentrated the dust, the larger the increase in observed size due to the attenuation of the galaxy’s bright core. With the central regions strongly attenuated, the more extended regions are able to contribute more to the total luminosity of the galaxy, increasing the perceived size. In the most extreme cases, galaxies can appear ~50 times larger when including dust attenuation. There are hints of this behaviour in observations at high redshift (z ~ 7) with Bowler et al. (2022) demonstrating that between 35 and 75 per cent of a galaxy’s SFR is obscured by dust and is often highly compact, pinpointing the most star-forming regions in a galaxy.

The vast majority of the diffuse galaxy population (grayscale) also have diffuse dust distributions (\(R_{1/2,\text{dust}} > 1 \, \text{pkpc}\)) and exhibit a more conservative increase in size between intrinsic and attenuated size. Compared to the compact population, the more diffuse dust distributions (and galaxies) have a flatter relation between the ratio of sizes and half-dust radius. Both the smaller increase in size and the flattening of this relation can be explained by a more uniform distribution of dust in these diffuse clumpy structures.

Galaxies that fall below the dashed line, indicating a ratio of 1, represent a decrease in size with the inclusion of dust effects. These are instances where the dust is more uniformly distributed, and results in greater attenuation of their extremities, driving down the apparent size.

### 5.2.2 The observed UV size–luminosity distribution

The negative gradient in the intrinsic size–luminosity relation presented in Fig. 5 is in direct conflict with observational results that necessarily include the effects of dust attenuation (e.g. Hathi et al. 2008; Grazian et al. 2011, 2012; Kawamata et al. 2015, 2018; Shibuya et al. 2015; Calvi et al. 2016; Morishita et al. 2018; Bridge et al. 2019; Bouwens et al. 2022; Yang et al. 2022). However, in Section 5.2.1, we have shown that the inclusion of dust attenuation can result in large increases in size for the most intrinsically compact galaxies. Ascertain if this effect is enough to yield sizes in line with observations is imperative to probe the validity of the negative intrinsic size–luminosity relation, and thus the physical models used in FLARES.

To compare to the observed results we use the method detailed in Section 3.2 for synthetic image creation and the pixel measurement method (Section 4.2.2) to produce the observed size–luminosity relation and compare to a wide array of observations in integer redshift bins from \(z = 5\) to 9. This observed size–luminosity relation is shown in Fig. 9.

Evidently, the concentration of dust in compact cores and increase in size between intrinsic and attenuated sizes, detailed in Section 5.2.1, has completely reversed the slope of the size–luminosity relation relative to the intrinsic relation.

Focusing on the high central surface density distribution (coloured hexbins), beyond the positive relation between size and luminosity, we can already see a power-law relation with minimal scatter. This

---

3 This strong attenuation of the core justifies the omission of the AGN contribution to the UV luminosity. We have confirmed the AGN contribution is heavily attenuated at these wavelengths, in fact only a handful of galaxies in the sample have AGN that are comparable to their host galaxy in the UV luminosity.

4 Those galaxies in the diffuse population that do not follow this trend (i.e. exhibit large increases in size with the inclusion of dust and have compact dust distributions) are galaxies very close to the central surface flux density threshold used to split the populations.
scatter is increased for the diffuse, low central surface density population (greyscale hexbins), particularly for low-luminosity galaxies that exhibit a large range of sizes at fixed luminosity. We can also see that the FLARES galaxy sample extends to larger sizes and higher luminosities than the observed results, this is because of FLARES's focus on rare and extreme environments where the most luminous galaxies reside.

There is a fair agreement between the scatter of observational measurements and the FLARES distribution with the exception of galaxies in the Kawamata et al. (2018) (lensed) sample that have sizes smaller than the resolution of FLARES. Particularly evident when comparing the FLARES and observational scatter are the Grazian et al. (2011) and Hathi et al. (2008) (dropout selected) points at \( z = 7 \) and 6, respectively, with similar normalization to the low central surface density galaxies that scatter further from the power-law relation evident in the compact population. This could be a tautological observational evidence for the galaxies that populate the diffuse population.

To quantify the agreement between the observational scatter and the FLARES sample we use \textsc{curvefit} (non-linear least-squares fitting), from \textsc{scipy} (Virtanen et al. 2020), to produce fits of the form of equation (1). The results of this fitting are shown in Table 2.

Fig. 10 shows a comparison of these fits (solid red lines) to fits from observed samples: Huang et al. (2013) at \( z = 5 \); Holwerda et al. (2015) at \( z = 7 \) and 9; Kawamata et al. (2018) at \( z = 6–9 \); Bouwens et al. (2022) at \( z = 6–8 \); and Yang et al. (2022) at \( z = 6–7 \), the latter three of these including lensed sources. We also compare to two simulations: the \textsc{meraxes} semi-analytic model (SAM; Liu et al. 2016; Marshall et al. 2019) at \( z = 5–9 \); and the \textsc{bluetides} simulation (Marshall et al. 2022) at \( z = 7–9 \). We denote observations by dashed lines and simulations (other than FLARES) by dotted lines. Each fit is plotted using their published fitting parameters.

At \( z > 7 \) the FLARES fits exhibit a good agreement in slope with the observational studies including lensed samples. These fits are significantly steeper than the observational samples that do not have a contribution of lensed galaxies, as demonstrated in Bouwens et al. (2022). At \( z \leq 7 \) the FLARES fits begin to flatten relative to the studies including lensed sources as galaxies in the dim and diffuse size–luminosity regime become more numerous.

Compared to \textsc{bluetides}, we find FLARES has a steeper size–luminosity relation at \( z = 8–9 \) and a stronger redshift evolution in

### Table 2

| Redshift (\( z \)) | \( R_o/(\text{pkpc}) \) | \( \beta \) |
|-------------------|----------------|-----|
| 9                 | 0.793 ± 0.019 | 0.519 ± 0.026 |
| 8                 | 0.842 ± 0.012 | 0.319 ± 0.013 |
| 7                 | 1.126 ± 0.011 | 0.290 ± 0.008 |
| 6                 | 1.370 ± 0.007 | 0.279 ± 0.004 |
| 5                 | 1.692 ± 0.006 | 0.300 ± 0.003 |

Figure 9. The attenuated far-UV (1500 Å) size–luminosity relation measured using the pixel method. The hexbins are again coloured by the weighted number density. The galaxy sample is divided into the compact galaxy population (top row, colour) and the diffuse galaxy population (middle row, greyscale). The dashed line shows the pixel resolution of the images used to make the FLARES measurements. Galaxies can fall below this line due to the interpolation used in the calculation of the pixel half-light radius. The bottom row contains both galaxy populations with a comparison to high-redshifts observations using the \textit{Hubble Space Telescope} (Hathi et al. 2008; Grazian et al. 2011, 2012; Calvi et al. 2016; Kawamata et al. 2018; Morishita et al. 2018; Bridge et al. 2019).
the normalization over the redshift range \(7 \leq z \leq 9\). With respect to MERAxes we find a good agreement in slopes at \(z < 9\) with a consistently higher normalization at all redshifts.

Each work predicts a different normalization of the size–luminosity relation. This is particularly evident at \(z < 8\) where FLARES has consistently higher normalization than all other studies. One explanation for this difference is the resolution and measurement methods in each study. The pixel method used in this work is sensitive to the resolution of the image (for which we adopt the softening length of the simulation), observational studies on the other hand use images with a higher resolution than the softening length of FLARES and use an array of measurement techniques that are less sensitive to the pixel resolution. BLUETIDES uses the pixel method but adopts a higher pixel resolution below the softening length of the simulation, and MERAxes derive their sizes (scale radius of the disc) from the SAM galaxy properties. In addition to methodological differences, there is likely a significant contribution to the normalization by the diffuse galaxies, which at fixed luminosity extend to larger sizes in the FLARES sample.

The slopes reported in Table 2 for the attenuated size–luminosity relation are in broad agreement with the results of Grazian et al. (2012), Huang et al. (2013), Holwerda et al. (2015), Shibuya et al. (2015), Kawamata et al. (2018), Bouwens et al. (2022), and Yang et al. (2022) in various different redshift regimes. At \(z > 7\) the FLARES results exhibit the steeper slopes present in Kawamata et al. (2018), Bouwens et al. (2022), and Yang et al. (2022) before flattening into closer agreement with Grazian et al. (2012), Huang et al. (2013), Holwerda et al. (2015), and Shibuya et al. (2015) at \(z \leq 7\). Again, this is due to the aforementioned compact low-luminosity galaxies present in the lensed samples, which are absent from the other studies, and the diffuse low-luminosity galaxies in the FLARES sample that become more numerous with decreasing redshift.

Many of the compact galaxies that strongly affect the slope of the size–luminosity relation in lensing studies fall below the resolution limit of FLARES (indicated by the dashed line in Fig. 9) and BLUETIDES. Higher resolution simulations are necessary to ascertain if these galaxies are present in the simulated sample and produce the same steepening behaviour. All observational samples also lack the most diffuse galaxies in the simulated samples due to their low surface densities. These would act to flatten the size–luminosity relation if present. Future works will aim to address both these issues with higher resolution simulations and fully synthetic observations including survey limits, instrument noise, point spread functions, and observational methods of structure detection; the former addressing the missing dim and compact galaxies in the simulated sample and the latter addressing the diffuse galaxies that are likely undetected in the observational sample.

5.3 The size–luminosity relation as a function of wavelength

In Fig. 11, we present the size–luminosity relation across a range of rest-frame filters (shown in Fig. 1), and compare to the corresponding fits from Marshall et al. (2022) at \(z = [8, 7]\). We present the fitting parameters in Appendix B.

As the probed wavelength regime reddens, the slope of the size–luminosity relation decreases, becoming increasingly negative for the reddest filters. These red filters probe the underlying stellar distribution with the least attenuation. The increasing representation of the underlying intrinsic distribution is clearly shown in the bottom row of panels as the slope of the ratio between attenuated and intrinsic size flattens with increasing wavelength. The slope of the size–luminosity relation for the reddest filters increases with decreasing redshift, implying that the intrinsic stellar population is becoming more diffuse as galaxies evolve.

This variation with wavelength is also predicted by BLUETIDES (Marshall et al. 2022) at \(z = [7, 8]\), although they predict a shallower size–luminosity relation for the reddest filters relative to those produced in this work. It is also consistent with observations at low redshift (e.g. La Barbera et al. 2010; Kelvin et al. 2012; Vulcani et al. 2014; Kennedy et al. 2015; Tacchella et al. 2015).

None the less, these results present a tantalizing prediction that will allow the JWST to ascertain the validity of the negative intrinsic size–luminosity relation. The JWST’s reddest broad-band Near-Infrared Camera (NIRCam) filter (F444W) will probe as blue as the \(B\) band at \(z = 9\) and \(I\) band at \(z = 5\) (as shown in Fig. 1) allowing for high-resolution measurements of galaxy sizes in this regime.

5.4 Redshift evolution

In the literature there has been a wide range of presented methods for measuring the redshift evolution of galaxy sizes, with various approaches and galaxy sample definitions used for the computation. To produce a comprehensive comparison with FLARES we employ non-linear least-squares fitting (again using \textsc{scipy}.$\text{curvefit}$) to produce fits to equation (2) from various sample definitions pulled from the complete galaxy sample, all weighted with the FLARES.
weighting scheme. The results of this fitting are presented in Table 3.

In Fig. 12, we present these fits for a number of different sample definitions found in the literature. Fig. 13 shows a comparison of the slope (m) from various studies, left to right: FLARES, Marshall et al. (2022), Oesch et al. (2010), Holwerda et al. (2015), Kawamata et al. (2018), and Ono et al. (2013), with a shaded region representing the range of slopes from Ma et al. (2018). We present the fitting parameters for these fits in Table 3.

For the low-luminosity sample we see a good agreement in slope between FLARES and Oesch et al. (2010) and Ono et al. (2013). For the other FLARES samples we find comparatively high slopes compared to the other works. However, these values are in agreement with Ma et al. (2018) who predict values in the range $m = 1–2$ depending on the fixed mass or luminosity regime (shown by the shaded region). All but the low-luminosity sample’s slopes are larger than the evolution of systems at fixed circular velocity, implying an increasing feedback contribution to the evolution with decreasing redshift. Conversely, the low-luminosity sample’s evolution is closer to that of a system at fixed mass with the same additional feedback contribution. As feedback becomes more efficient with decreasing redshift the star-forming gas will be given more thermal energy and thus change the dynamics of the star-forming gas, increasing the radii at which stars can form and thus the half-light radii.

Limited the included redshifts in the FLARES sample cannot only be used to compare to the more limited samples of BLUETIDES, with no galaxies at $z < 7$, and observations, where $z \geq 10$ galaxies are exceedingly rare, but can also probe the evolution of size during particular epochs. To do this we limited the sample to a high-$z$ sample limited to $z \geq 7$ and a low-$z$ sample with $z \leq 10$, the results of which are also included in Table 3. Limiting to $z \geq 7$ resulted in a larger increase in the slope of the redshift evolution alongside unrealistically high normalizations, predicting $z = 0$ sizes of the order $\sim 300$ pkpc for the low-luminosity sample and over double the $z = 0$ size in the limited and high-luminosity samples in the other redshift selections. Conversely, limiting to $z \leq 10$ instead results in fitting results consistent with those produced by the full redshift range. This casts doubt on the sparse $z > 10$ measurements in observations causing the differences in slope between the FLARES measurement and observational measurements. More interestingly the differences in fits between redshift regimes imply a significantly faster evolution of galaxy size at the earliest times, even for the most dim and diffuse galaxies in the low-luminosity sample. It is clear from Fig. 12 that a piecewise fit produces a considerably better fit to the data than fitting across the entire redshift range.

Tensions between FLARES and the observations are far less stark than those between FLARES and BLUETIDES samples but are none the less evident for the capped and high-luminosity samples, we do however see a good agreement in the low-luminosity sample. The tensions here could be explained by how sparse observations are at the highest redshifts due to the small area covered at the required depth; given that the low-luminosity sample in FLARES is also sparse.
at the highest redshifts, the agreement between observations and FLARES here could be due to this luminosity regime being where the simulation and observations have the largest overlap in sampling strength. Additional observations from upcoming observatories populating the highest redshifts will increase the area and depth sampled in at this epoch and could rectify this tension. It should also be noted however that subgrid models require intensive investigation at this epoch, with comparison to robust observations to ascertain the validity of their behaviour. Future work will be able to converge the results of both simulations and observations to a consistent story of galaxy size evolution.

6 CONCLUSIONS

In this paper, we have presented an analysis of galaxy sizes at $z \geq 5$ in the FLARES simulations across a wide array of environments. To do this we produced synthetic galaxy images using photometry in rest-frame UV and visual bands derived using the LoS attenuation method presented in Vijayan et al. (2021). We presented an efficient method of image computation by utilizing a KD-Tree of pixel coordinates and smoothing stellar particles over their SPH kernels. We employed this imaging method to produce synthetic galaxy images, from which the size of galaxies were measured using a non-parametric pixel-based method to account for the clumpy nature of galaxies at high redshift.

Using these measurements we probed both the intrinsic and observed size–luminosity relation in the rest-frame far-UV (1500 Å), findings are as follows.

(i) The intrinsic size–luminosity relation is bimodal, with one intrinsically compact and bright population and one intrinsically diffuse and dim population.

(ii) These two populations result in a negative slope to the rest-frame far-UV intrinsic size–luminosity distribution.

(iii) Including the effects of dust attenuation results in the perceived size of galaxies to increase, with the most intrinsically compact galaxies increase in size by as much as 50 times.

(iv) The increase in size due to dust attenuation inverts the slope of the size–luminosity relation, resulting in a fair agreement between observations and in this work. However, the FLARES sample lacks low-luminosity compact galaxies that have been shown to steepen the size–luminosity relation in lensing studies. Conversely, the observational samples lack the diffuse and dim galaxies that are present in this work, these act to flatten the size–luminosity relation. The effects of these missing galaxies highlight the need for high-resolution simulations in the future and observationally motivated measurement methods.

(v) Dust distributions in these compact galaxies are highly concentrated with half-metal radii of <1 pkpc, heavily attenuating the intrinsically bright cores and increasing the observed half-light radius. This may be observable as strong dust gradients.

We performed size measurements for a range of rest-frame UV and visual bands, finding an anticorrelation between the slope of the size–luminosity relation and wavelength. This anticorrelation becomes weaker with decreasing redshift as the intrinsic stellar distribution.
increases in size. This represents a falsifiable prediction that the JWST will be able to probe at high resolution with NICCam.

We then investigated the evolution of size with redshift in the far-UV, finding slopes for multiple sample definitions in the range $m = 1.21$–1.87. These values are consistent with theoretical predictions modified by additional contributions to the evolution by feedback mechanisms. At low luminosity the evolution is consistent with an evolution at fixed mass ($m = 1$) with additional evolution due to feedback, while high-luminosity galaxies are consistent with a fixed circular velocity evolution ($m = 1.5$), again with an additional contribution from feedback. With the exception of the low-luminosity sample giving a good agreement, these results are in tension with observations. They do however broadly agree with the range found in the FIRE-2 simulations. The limited observational galaxy sample at extremely high redshifts could contribute to this tension. Limiting the galaxy sample to both a low ($5 \leq z \leq 10$) and high ($7 \leq z \leq 12$) redshift sample yielded little change in the results for the low-redshift sample but resulted in significantly higher slopes for the high-redshift sample. This implies a non-constant size evolution with faster evolution in the highest redshift bins. Further observations from future high-redshift surveys are needed to probe the differences highlighted here in addition to future simulations adding to the theory.

With the launch of the JWST we will soon be able to probe these high-redshift regimes with far greater fidelity and further strengthen our understanding of the earliest epochs of galaxy evolution. The JWST will allow us to probe higher redshifts at high resolution with NICCam. Not only will this further populate galaxy samples at $z > 8$, it will also increase the completeness of the high-redshift observational surveys at low luminosity.

Future work will include the next generation of FLaRES simulating a wider range of environments, probing more regions, and simulating a significant amount at high mass resolution. Including higher resolution simulations will enable comparison to the dim and compact galaxies found in lensing studies, while increasing the effective volume with more resimulated regions will allow FLaRES to reach a volume comparable to the largest upcoming observational surveys from Euclid.

In addition to the next generation of FLaRES, the underlying physical processes governing the size evolution in the subgrid model will be probed. This will include stellar and AGN feedback, star formation conditions, and chemical enrichment. The effects of simulation and observational structure detection methods will be investigated to quantify the effect of survey depth and the segmentation of substructures. In particular this will aim to probe the effects of structure detection methods on the diffuse galaxy population and the effect this has on the size–luminosity relation.

ACKNOWLEDGEMENTS

We thank the EAGLE team for their efforts in developing the EAGLE simulation code.

We acknowledge the indispensable contribution from the publicly available programming language PYTHON (van Rossum 1995), including the NUMPY (Harris et al. 2020), ASTROPY (Astropy Collaboration et al. 2013), MATPLOTLIB (Hunter 2007), SCIPY (Virtanen et al. 2020), and hSPY (Collette 2013) packages.

This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002571/1, and ST/S002502/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

CCL acknowledges support from the Royal Society under grant RG/E/181016. DI acknowledges support by the European Research Council via ERC Consolidator Grant KETJU (no. 818930). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant no. 140. MAM acknowledges the support of a National Research Council Canada Plaskett Fellowship, and the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

DATA AVAILABILITY

The integrated galaxy properties used to generate the plots in this paper are available at flaresimulations.github.io/data. More detailed data, including particle data, can be provided upon request. All the code used to produce the analysis in this paper is public and available at github.com/WillJRoper/flares-sizes-obs.

REFERENCES

Astropy Collaboration et al., 2013, A&A, 558, A33
Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 213
Barnes D. J. et al., 2017b, MNRAS, 471, 1088
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borrow J., Schaller M., Bower R. G., Schaye J., 2022, MNRAS, 511, 2367
Bouwens R. J., Illingworth G. D., Blakeslee J. P., Broadhurst T. J., Franx M., 2004, ApJ, 611, L1
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bouwens R. J., Illingworth G. D., van Dokkum P. G., Oesch P. A., Stefanon M., Ribeiro B., 2022, ApJ, 927, 81
Bowler R. A. A., Dunlop J. S., McIver R. J., McLeod D. J., 2017, MNRAS, 466, 3612
Bowler R. A. A., Cullen F., McIver R. J., Dunlop J. S., Avison A., 2022, MNRAS, 510, 5088
Bradač M., Schneider P., Lombardi M., Erben T., 2005, A&A, 437, 39
Bridge J. S. et al., 2019, ApJ, 882, 42
Calvi V. et al., 2016, ApJ, 817, 120
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Chiang Y.-K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127
Collette A., 2013, Python and HDF5: Unlocking Scientific Data. O’Reilly Media, Sebastopol, CA
Conselice C. J., 2014, ARA&A, 52, 291
Crain R. A. et al., 2015, MNRAS, 450, 1937
Cullen L., Dehnen W., 2010, MNRAS, 408, 669
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Dehnen W., Aly H., 2012, MNRAS, 425, 1068
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, MNRAS, 455, 2778
Ferguson H. C. et al., 2004, ApJ, 600, L107
Ferland G. J. et al., 2017, Rev. Mex. Astron. Astrofís., 53, 385
Furlong M. et al., 2015, MNRAS, 450, 4486
Furlong M. et al., 2017, MNRAS, 465, 722
Grauzan A. et al., 2011, A&A, 532, A35
Grauzan A. et al., 2012, A&A, 547, A51
Gutkin J., Charlot S., Bruzual G., 2016, MNRAS, 462, 1757

MNRAS 514, 1921–1939 (2022)
Here we present comparisons between smoothing methods used in image creation first comparing Gaussian and spline kernel smoothing and then the differences between smoothing and ignoring smoothing.

**A1 Comparing kernel averaging to Gaussian smoothing**

Fig. A1 shows a comparison between the Gaussian and spline smoothing methods. Qualitatively it can be seen the Gaussian method results in a smoother light distribution due to the indefinite boundaries of the Gaussian smoothing kernel, this spreads light beyond the ‘extent’ given by the SPH kernel. The spline method produces a more granular image with clearer small structures at the outskirts of the FoV. The residual image shows that the Gaussian method’s spreading of light leads to differences at large radii where the Gaussian image is brighter due to the spreading of light. However, this does not mean the Gaussian image is consistently more luminous at large radii, compact structures at large radii in the spline image have more concentrated emission causing these regions to outshine the Gaussian image. This effect is also noticeable in the centre of the image where there is a ring of spline dominated pixels due to this concentration of light. These effects are however minimal with each image differing at most by 0.1 dex.

We further show the effects of smoothing method in Fig. A2 where we compare the measured sizes of galaxies in each method. In the vast majority of cases the Gaussian smoothing results in a larger perceived size due to the increased spread of a single stellar particle’s luminosity. The instances where the spline method yields larger sizes are dominated by smaller galaxies where the dilution of the Gaussian method causes structures to occupy more pixels relative to the more concentrated spline method and thus a larger area is used in the pixel-driven size calculation. It should be noted here that the spline method produces a better agreement with observations with the Gaussian method producing size–luminosity relations that overestimate galaxy sizes relative to observations.

**APPENDIX A: THE EFFECTS OF SMOOTHING**
**Figure A1.** A comparison of log-scaled stacked images produced using the Gaussian smoothing method (left), spline kernel method (middle), and a residual image showing the difference between the log of the two methods images. The images themselves are stacks in the far-UV of all galaxies in the FLARES sample (irrespective of completeness).

**Figure A2.** A comparison between the sizes of galaxies measured using the pixel method from the spline (y-axis) and Gaussian (x-axis) smoothing methods. The dashed line represents a 1:1 relation. In this plot we do not differentiate between the compact and diffuse galaxy populations and only present the full complete sample.

**A2 Smoothing versus no smoothing**

In Fig. A3, we compare the spline smoothing method to galaxy sizes measured from images where no smoothing has been performed on the stellar particles. In some cases there is minimal difference between the smoothed and unsmoothed measurements, particularly for compact galaxies where the stellar kernels themselves are very small resulting in minimal smoothing. In the vast majority of cases the smoothing increases the measured size, with the most diffuse incomplete galaxies (transparent distribution) extending to much larger sizes when smoothed.

**APPENDIX B: SIZE–LUMINOSITY RELATION WAVELENGTH VARIATION**

In this appendix, we present the fitting parameters for the wavelength evolution of the size–luminosity relation shown in Fig. 11 (Tables B1 and B2).
Table B1. The fitting results for equation (1) for \( z = 7–9 \) and all rest-frame bands in Fig. 11. \( R_0 \) is a normalization factor, \( \beta \) is the slope of the size–luminosity relation, and \( N \) is the number of galaxies used in each fit.

| Redshift (z) | Band | \( R_0 \)       | \( \beta \)   | \( R_0 \)       | \( \beta \)   | \( R_0 \)       | \( \beta \)   |
|-------------|------|----------------|--------------|----------------|--------------|----------------|--------------|
|             |      | 9              | 8            | 7              |              | 6              | 5            |
| FUV         | 0.793 ± 0.019 | 0.519 ± 0.026 | 0.842 ± 0.012 | 0.319 ± 0.013 | 1.126 ± 0.011 | 0.290 ± 0.008 |
| MUV         | 0.773 ± 0.020 | 0.493 ± 0.026 | 0.821 ± 0.012 | 0.313 ± 0.013 | 1.070 ± 0.011 | 0.263 ± 0.008 |
| NUV         | 0.777 ± 0.021 | 0.485 ± 0.026 | 0.813 ± 0.013 | 0.296 ± 0.014 | 1.020 ± 0.013 | 0.211 ± 0.009 |
| U           | 0.687 ± 0.017 | 0.434 ± 0.026 | 0.743 ± 0.011 | 0.262 ± 0.014 | 0.878 ± 0.011 | 0.092 ± 0.010 |
| B           | 0.660 ± 0.014 | 0.428 ± 0.025 | 0.704 ± 0.010 | 0.133 ± 0.014 | 0.854 ± 0.010 | 0.017 ± 0.010 |
| V           | 0.702 ± 0.018 | 0.375 ± 0.024 | 0.689 ± 0.013 | 0.022 ± 0.014 | 0.823 ± 0.011 | 0.114 ± 0.009 |
| R           | 0.573 ± 0.011 | 0.397 ± 0.027 | 0.638 ± 0.008 | 0.154 ± 0.014 | 0.765 ± 0.009 | 0.030 ± 0.011 |
| I           | 0.598 ± 0.019 | 0.178 ± 0.026 | 0.601 ± 0.013 | 0.110 ± 0.015 | 0.763 ± 0.011 | 0.167 ± 0.009 |
| Z           | 0.558 ± 0.015 | 0.229 ± 0.027 | 0.583 ± 0.011 | 0.077 ± 0.015 | 0.715 ± 0.010 | 0.186 ± 0.010 |
| Y           | 0.595 ± 0.014 | 0.312 ± 0.026 | 0.616 ± 0.010 | 0.000 ± 0.014 | 0.715 ± 0.010 | 0.183 ± 0.010 |
| J           | 0.532 ± 0.017 | 0.090 ± 0.027 | 0.525 ± 0.011 | 0.220 ± 0.015 | 0.698 ± 0.010 | 0.228 ± 0.009 |
| H           | 0.476 ± 0.017 | −0.035 ± 0.027 | 0.503 ± 0.011 | −0.268 ± 0.014 | 0.688 ± 0.010 | −0.250 ± 0.008 |

Table B2. The fitting results for equation (1) for \( z = 6–5 \) and all rest-frame bands in Fig. 11. \( R_0 \) is a normalization factor, \( \beta \) is the slope of the size–luminosity relation, and \( N \) is the number of galaxies used in each fit.

| Redshift (z) | Band | \( R_0 \)       | \( \beta \)   | \( R_0 \)       | \( \beta \)   |
|-------------|------|----------------|--------------|----------------|--------------|
|             |      | 6              | 5            | 6              | 5            |
| FUV         | 1.370 ± 0.007 | 0.279 ± 0.004 | 1.692 ± 0.006 | 0.300 ± 0.003 |
| MUV         | 1.326 ± 0.007 | 0.256 ± 0.004 | 1.639 ± 0.006 | 0.280 ± 0.003 |
| NUV         | 1.315 ± 0.008 | 0.238 ± 0.004 | 1.627 ± 0.006 | 0.261 ± 0.003 |
| U           | 1.218 ± 0.007 | 0.184 ± 0.004 | 1.514 ± 0.006 | 0.215 ± 0.003 |
| B           | 1.227 ± 0.007 | 0.111 ± 0.004 | 1.526 ± 0.005 | 0.149 ± 0.003 |
| V           | 1.285 ± 0.008 | 0.060 ± 0.004 | 1.604 ± 0.006 | 0.104 ± 0.002 |
| R           | 1.106 ± 0.005 | 0.124 ± 0.005 | 1.383 ± 0.004 | 0.156 ± 0.003 |
| I           | 1.238 ± 0.008 | 0.021 ± 0.004 | 1.554 ± 0.006 | 0.069 ± 0.002 |
| Z           | 1.155 ± 0.007 | 0.013 ± 0.004 | 1.455 ± 0.005 | 0.064 ± 0.002 |
| Y           | 1.143 ± 0.007 | 0.019 ± 0.004 | 1.439 ± 0.005 | 0.061 ± 0.002 |
| J           | 1.161 ± 0.007 | −0.023 ± 0.004 | 1.455 ± 0.005 | 0.024 ± 0.002 |
| H           | 1.146 ± 0.007 | −0.053 ± 0.004 | 1.430 ± 0.005 | −0.004 ± 0.002 |

APPENDIX C: THE EFFECTS OF IMAGE RESOLUTION

Image resolution can have a significant effect on measured quantities. In addition to this fact, the resolution one should use to produce synthetic observations from simulations can be unclear, with the softening length often used as a representation of the simulations spatial resolution. We note that the gravity calculations are in fact softened by a factor of \( \sim 2 \), this value indicating a possible larger spatial resolution definition. In Fig. C1, we present a comparison of sizes derived using the softening length resolution used throughout this paper and a resolution two times worse (\( 2 \times s \)). Our sizes are insensitive to this change producing a clear 1:1 relation with rare pathological instances of large scatter in cases that have little effect on the overall results. This insensitivity is largely due to the fact that many of the structures in galaxies in FLARES already fall below the softening length scale at these redshifts and the measurement methods employ robust interpolation methods to combat resolution effects.
APPENDIX D: REMOVING BIRTH CLOUD ATTENUATION

Of the parameters used in our synthetic photometry model the birth cloud attenuation is the most uncertain at the epoch investigated in this work. In FLARES II (Vijayan et al. 2021) the values of the photometry model parameters were derived from fits to observational studies and the effects of different values for birth cloud attenuation were probed in Appendix A. For completeness, in Fig. D1, we present a comparison of sizes with and without the birth cloud contribution to attenuation to show its effect on galaxy size.

We can see for the largest (most luminous) galaxies the omission of birth cloud attenuation has little effect on the measured size. These large galaxies have extremely concentrated dust distributions, as shown in Fig. 8, where the attenuation is dominated by the ISM contribution. For the smaller (less luminous) galaxies we see a larger scatter in size with a maximum increase of ∼0.5 dex. These galaxies have a more diffuse dust distribution limiting the contribution to attenuation from the ISM. For galaxies in the diffuse population with young stellar populations the birth cloud attenuation can represent a considerably larger contribution to the overall attenuation. We stress that this comparison is presented to quantify the effect of birth cloud attenuation in the extreme. The model parameters were derived using fits to observational data and thus provide robust results that give a good agreement with observational high-redshift studies not used in the fitting of model parameters.

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