

Simulating galaxies in the reionization era with FIRE-2: morphologies and sizes

Xiangcheng Ma,¹★ Philip F. Hopkins,¹ Michael Boylan-Kolchin,² Claude-André Faucher-Giguère,³ Eliot Quataert,⁴ Robert Feldmann,⁵ Shea Garrison-Kimmel,¹ Christopher C. Hayward,⁶ Dušan Kereš⁷ and Andrew Wetzel⁸

¹ TAPIR, MC 350-17, California Institute of Technology, Pasadena, CA 91125, USA

² Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712-1205, USA

³ Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

⁴ Department of Astronomy and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA

⁵ Institute for Computational Science, University of Zurich, Zurich CH-8057, Switzerland

⁶ Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

⁷ Department of Physics, Center for Astrophysics and Space Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

⁸ Department of Physics, University of California, Davis, CA 95616, USA

Accepted 2018 March 12. Received 2018 March 12; in original form 2017 September 30

ABSTRACT

We study the morphologies and sizes of galaxies at $z \geq 5$ using high-resolution cosmological zoom-in simulations from the Feedback In Realistic Environments project. The galaxies show a variety of morphologies, from compact to clumpy to irregular. The simulated galaxies have more extended morphologies and larger sizes when measured using rest-frame optical B -band light than rest-frame UV light; sizes measured from stellar mass surface density are even larger. The UV morphologies are usually dominated by several small, bright young stellar clumps that are not always associated with significant stellar mass. The B -band light traces stellar mass better than the UV, but it can also be biased by the bright clumps. At all redshifts, galaxy size correlates with stellar mass/luminosity with large scatter. The half-light radii range from 0.01 to 0.2 arcsec (0.05–1 kpc physical) at fixed magnitude. At $z \geq 5$, the size of galaxies at fixed stellar mass/luminosity evolves as $(1 + z)^{-m}$, with $m \sim 1\text{--}2$. For galaxies less massive than $M_* \sim 10^8 M_\odot$, the ratio of the half-mass radius to the halo virial radius is ~ 10 per cent and does not evolve significantly at $z = 5\text{--}10$; this ratio is typically 1–5 per cent for more massive galaxies. A galaxy’s ‘observed’ size decreases dramatically at shallower surface brightness limits. This effect may account for the extremely small sizes of $z \geq 5$ galaxies measured in the Hubble Frontier Fields. We provide predictions for the cumulative light distribution as a function of surface brightness for typical galaxies at $z = 6$.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift – cosmology: theory.

1 INTRODUCTION

High-redshift galaxies are thought to be the dominant source of cosmic reionization (e.g. Kuhlen & Faucher-Giguère 2012; Robertson et al. 2013, 2015). The number density of these galaxies, as described by the ultraviolet (UV) luminosity function, is well constrained for galaxies brighter than $M_{\text{UV}} = -17$ at $z \sim 6$ (e.g.

McLure et al. 2013; Schenker et al. 2013; Bouwens et al. 2015; Finkelstein et al. 2015). Recently, the Hubble Frontier Fields (HFF) program (Lotz et al. 2017), which takes advantage of strong gravitational lensing by foreground galaxy clusters, has made it possible to estimate UV luminosity functions down to $M_{\text{UV}} \sim -13$ (e.g. Bouwens et al. 2017b; Livermore, Finkelstein & Lotz 2017). But one of the dominant outstanding uncertainties is the intrinsic size distribution of faint galaxies, which is necessary in order to determine the completeness of the observed sample due to surface brightness limits (Bouwens et al. 2017a).

* E-mail: xchma@caltech.edu

There are only a few galaxies at these redshifts that have robust size measurements. Oesch et al. (2010) measured the sizes of galaxies brighter than $M_{\text{UV}} = -19$ at $z = 4\text{--}8$ in the Hubble Ultra-Deep Field (HUDF) and found that the half-light radii of galaxies evolve according to $(1 + z)^{-m}$, with $m \sim 1\text{--}1.5$ (see also, e.g. Bouwens et al. 2004; Ferguson et al. 2004; Ono et al. 2013; Kawamata et al. 2015). It is also expected from analytic models that galaxy size decreases with increasing redshift (Mo, Mao & White 1998, 1999).

High-redshift galaxies tend to be intrinsically small. The half-light radii of bright galaxies ($M_{\text{UV}} < -19$) at $z \sim 6\text{--}8$ range from 0.5 to 1 kpc (e.g. Oesch et al. 2010). More recently, Kawamata et al. (2015) and Bouwens et al. (2017a) measured the half-light radii for a sample of fainter galaxies ($-19 < M_{\text{UV}} < -12$) from the HFF. They found that the size–luminosity relation has large scatter, with half-light radii spanning more than an order of magnitude (0.1–1 kpc) at fixed UV magnitude. A fraction of these faint galaxies have extremely small sizes from a few pc to 100 pc, although these results are very uncertain because they are far below the resolution of the *Hubble Space Telescope* (*HST*).

Morphological studies have revealed that galaxies at intermediate redshifts ($z \sim 0.5\text{--}3$) typically contain a number of star-forming clumps (e.g. Guo et al. 2015). These prominent clumps only contribute a small fraction of the total mass, however (e.g. Wuyts et al. 2012). So far, the sizes of $z \gtrsim 6$ galaxies are measured using noise-weighted stacked images over all available bands (dominated by rest-frame UV), so it is likely that the extremely small galaxy sizes in the HFF are biased by such clumps (e.g. Vanzella et al. 2017). With the *James Webb Space Telescope* (*JWST*, scheduled to launch in 2020), one will be able to probe these faint, high-redshift galaxies with deeper imaging, higher spatial resolution, and at longer wavelengths. This makes it possible to compare galaxy morphology and sizes in different bands and to recover the stellar mass distribution using multiband images via pixel-by-pixel spectral energy distribution (SED) modelling (e.g. Smith & Hayward 2015).

The goal of this paper is to make predictions of morphologies and sizes for $z \geq 5$ galaxies, which can be used to plan and interpret future observations. We use a suite of high-resolution cosmological zoom-in simulations from the Feedback In Realistic Environments (FIRE) project¹ (Hopkins et al. 2014). The FIRE simulations include explicit treatments of the multiphase interstellar medium (ISM), star formation, and stellar feedback. The simulations are evolved using the FIRE-2 code (Hopkins et al. 2017), which is an update of the original FIRE code with several improvements to the numerics. These simulations predict stellar mass functions and luminosity functions in broad agreement with observations at these redshifts. When evolved to $z = 0$, simulations with the same physics have been shown to also reproduce many other observed galaxy properties (Hopkins et al. 2017, and references therein).

The paper is organized as follows. In Section 2, we describe the simulations briefly and the methods we use to measure galaxy sizes. We present the results in Section 3 and discuss their implications to observations in Section 4. We conclude in Section 5. We adopt a standard flat Λ CDM cosmology with *Planck* 2015 cosmological parameters $H_0 = 68 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.69$, $\Omega_m = 1 - \Omega_\Lambda = 0.31$, $\Omega_b = 0.048$, $\sigma_8 = 0.82$, and $n = 0.97$ (Planck Collaboration XIII 2016). We use a Kroupa (2002) initial mass function (IMF) from 0.1 to 100 M_\odot , with IMF slopes of -1.30 from 0.1 to 0.5 M_\odot and -2.35 from 0.5 to 100 M_\odot . All magnitudes are in the AB system (Oke & Gunn 1983).

¹ <http://fire.northwestern.edu>

2 METHODS

2.1 The simulations

We use a suite of 15 high-resolution cosmological zoom-in simulations at $z \geq 5$ from the FIRE project, spanning a halo mass range $M_{\text{halo}} = 10^8\text{--}10^{12} \text{ M}_\odot$ at $z = 5$. These simulations are first presented in Ma et al. (2017). The mass resolution for baryonic particles (gas and stars) is $m_b = 100\text{--}7000 \text{ M}_\odot$ (more massive galaxies having larger particle mass). The minimum Plummer-equivalent force softening lengths for gas and star particles are $\epsilon_b = 0.14\text{--}0.42 \text{ pc}$ and $\epsilon_* = 0.7\text{--}2.1 \text{ pc}$ (see table 1 in Ma et al. 2017 for details). The softening lengths are in comoving units above $z = 9$ but switch to physical units thereafter. All of the simulations are run using exactly identical code GIZMO² (Hopkins 2015), in the mesh-less finite-mass (MFM) mode with the identical FIRE-2 implementation of star formation and stellar feedback.

The baryonic physics included in FIRE-2 simulations are described in Hopkins et al. (2017), but we briefly review them here. Gas follows an ionized-atomic-molecular cooling curve from 10 to 10^{10} K , including metallicity-dependent fine-structure and molecular cooling at low temperatures and high-temperature metal-line cooling for 11 separately tracked species (H, He, C, N, O, Ne, Mg, Si, S, Ca, and Fe). At each time-step, the ionization states and cooling rates H and He are calculated following Katz, Weinberg & Hernquist (1996), with the updated recombination rates from Verner & Ferland (1996), and cooling rates from heavier elements are computed from a compilation of CLOUDY runs (Ferland et al. 2013), applying a uniform but redshift-dependent photoionizing background from Faucher-Giguère et al. (2009), and an approximate model for H II regions generated by local sources. Gas self-shielding is accounted for with a local Jeans-length approximation. We do not include a primordial chemistry network nor Pop III star formation, but apply a metallicity floor of $Z = 10^{-4} \text{ Z}_\odot$.

We follow the star formation criteria in Hopkins, Narayanan & Murray (2013) and allow star formation to take place only in dense, molecular, and locally self-gravitating regions with hydrogen number density above a threshold $n_{\text{th}} = 1000 \text{ cm}^{-3}$. Stars form at 100 per cent efficiency per local free-fall time when the gas meets these criteria, and there is no star formation elsewhere. The galactic-scale star formation efficiency is regulated by feedback to ~ 1 per cent (e.g. Orr et al. 2017). The simulations include the following stellar feedback mechanisms: (1) local and long-range momentum flux from radiation pressure, (2) SNe, (3) stellar winds, and (4) photoionization and photoelectric heating. Every star particle is treated as a single stellar population with known mass, age, and metallicity. The energy, momentum, mass, and metal returns from each stellar feedback processes are directly calculated from STARBURST99 (Leitherer et al. 1999).

2.2 Post-processing and size definition

We use Amiga Halo Finder (AHF; Knollmann & Knebe 2009) to identify haloes in our simulations. The halo mass (M_{halo}) and virial radius (R_{vir}) are computed by AHF, applying the redshift-dependent virial overdensity criterion from Bryan & Norman (1998). Each zoom-in simulation contains one central halo, which is the most massive halo of the zoom-in region. In this work, we also consider other less massive haloes in the zoom-in region. We restrict our

² <http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html>

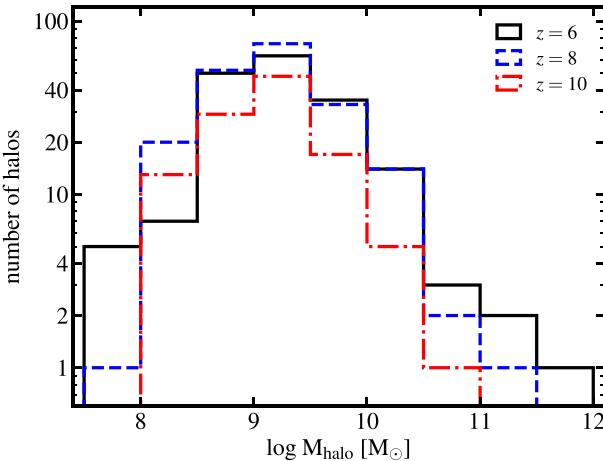


Figure 1. Number of haloes in our sample at $z = 6, 8$, and 10 .

analysis to haloes with zero contamination from low-resolution particles, which also having more than 100 star particles and 10^4 total particles within the virial radius. In Fig. 1, we show the number of haloes in our simulated sample at $z = 6, 8$, and 10 . At a given mass, our sample includes both central haloes and less massive haloes in the zoom-in regions, and include simulations run with different mass resolutions. In Appendix B, we show that our results converge reasonably well with resolution.

At a given redshift, we project all star particles inside the halo along a random direction on to a two-dimensional uniform grid. The pixel size of the grid is 0.0032 arcsec (3.2 mas), which equals to $1/10$ of the pixel size of *JWST*'s Near-Infrared Camera (NIRCam) and corresponds to 10 – 20 pc in the redshift range of our interest. Each star particle is smoothed over a cubic spline kernel with a smoothing length $h = 1.5 h_n$, where h_n is its distance to the n th nearest neighbour star particle. We adopt $n = 5$ as our default value, but varying $n = 5$ – 10 only makes small difference for a small fraction of our galaxies. We only consider intrinsic morphologies and sizes and do not include dust extinction in this work. The majority of galaxies (over 95 per cent) in our sample have intrinsic UV magnitude fainter than $M_{\text{UV}} = -18$ (stellar mass $M_* < 10^8 M_\odot$). We find that dust attenuation has little effect on these low-mass, faint galaxies (also see Ma et al. 2017), so most results in this paper are not affected by dust extinction.

We make projected images for stellar surface density and rest-frame UV (1500 Å) and rest-frame B -band (4300 Å) surface brightness. The rest-frame UV of galaxies at these redshifts falls in the short-wavelength channel of NIRCam (observed wavelengths 0.6 – 2.3 μm, spatial resolution 0.032 arcsec per pixel), in $F090W$ band for $z = 5$ galaxies and $F150W$ band for $z = 10$ galaxies. The rest-frame B -band falls in NIRCam's long-wavelength channel (2.4 – 5 μm, spatial resolution 0.065 arcsec per pixel), in $F277W$ band for $z = 5$ galaxies and $F444W$ band for $z = 10$ galaxies. The SED of each star particle is computed using the synthesis spectra predicted by the Binary Population and Spectral Synthesis (BPASS) models (version 2.0; Eldridge, Izzard & Tout 2008).³ We use the binary stellar population models in BPASS by default.⁴ In Fig. 2, we

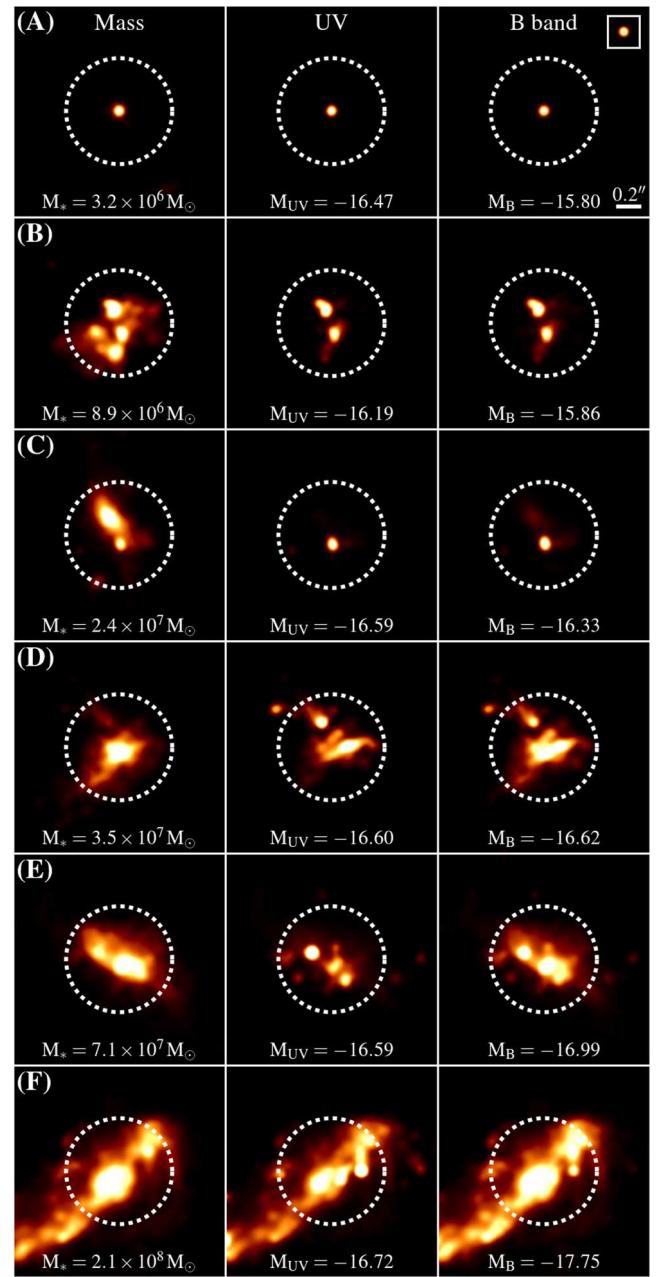


Figure 2. Example images for six galaxies (A–F) at $z = 6$ with rest-frame UV magnitude $M_{\text{UV}} \sim -16.5$. The left column shows stellar surface density. The middle and right columns show unattenuated, noise-free rest-frame UV and rest-frame B -band surface brightness, respectively. All images are 2 arcsec (11.6 kpc) along each side. Colours are in linear scale. These images are smoothed over a Gaussian kernel with 0.032 arcsec dispersion (10 pixels on the image) only for visualization purposes (otherwise the structures are too small to visualize on these images). The square in the top-right corner shows the appearance of a point source on these images for reference. The white dotted circles show the 1 arcsec-diameter aperture in which the sizes are measured. These galaxies span a wide range of stellar mass and B -band magnitude, and show a variety of morphologies. More massive galaxies appear to be larger than galaxies at lower masses. The UV images are largely dominated by bright, young stellar clumps, which do not necessarily trace the bulk of stars. The B -band light traces stellar mass better than the UV, but it can also be biased by the UV-bright clumps. Galaxies tend to be more clumpy, more concentrated, and smaller in size from stellar mass to rest-frame optical to the UV.

³ <http://bpass.auckland.ac.nz>

⁴ We prefer binary models to single-star models when calculating SEDs because they are able to reproduce nebular emission line features observed in high-redshift galaxies (e.g. Steidel et al. 2016; Strom et al. 2017). While we use STARBURST99 single-star models for stellar feedback calculations in

show example images for six galaxies labelled by A–F. Each panel is 2 arcsec (11.6 kpc) on each side. Some galaxies and structures are so small that they only occupy very few pixels, so we further smooth the images using a two-dimensional Gaussian kernel with a dispersion equal to the size of 10 pixels (0.032 arcsec) only for easier visualization here.

Most galaxies in our sample show clumpy, irregular morphologies that cannot be well described by a simple profile (see also figs 2 and 3 in Ma et al. 2017; cf. Jiang et al. 2013; Bowler et al. 2017). Therefore, we adopt a non-parametric approach to define galaxy sizes, in a way similar to Curtis-Lake et al. (2016) and Ribeiro et al. (2016). For every galaxy, we place a circular aperture of 1 arcsec in diameter, whose centre is located by iteratively finding the *B*-band surface brightness-weighted centre of all pixels within the 1 arcsec aperture, as illustrated by the white dotted circles in Fig. 2. We visually inspect all galaxies to ensure the apertures are reasonably located. The same aperture is used for the size measurement in stellar mass, UV, and *B*-band luminosity for the same galaxy as follows. We sort the pixels enclosed in the 1 arcsec-diameter aperture in descending order of surface density or surface brightness, and find the number of ‘brightest’ pixels that contribute 50 per cent of the total mass or luminosity within the 1 arcsec aperture. We calculate the area spanned by these pixels S_{50} and define the ‘half-mass’ or ‘half-light’ radius as $R_{50} = \sqrt{S_{50}/\pi}$. We quote the galaxy stellar mass and luminosity as the total amount enclosed in the 1 arcsec-diameter aperture.⁵

One may also measure the half-mass or half-light radius alternatively by finding the radius that encloses half of the mass or light within some larger aperture. This is close to the commonly used algorithms in observations for size measurements, such as SExtractor and GALFIT (e.g. Oesch et al. 2010), where concentric circular or ellipsoid apertures are usually assumed. However, this approach suffers from two main issues when applying to clumpy, irregular galaxies in our simulations. First, these galaxies do not have a well-defined centre: one may use the position of intensity peak on the image or intensity-weighted centre and get very different results. Secondly, for multicloud systems (e.g. galaxies B, D, E, and F in the rest-frame UV, see Fig. 2), the size defined in this way in fact represents the spatial separation between the bright clumps. The non-parametric definition we use better reflects the physical size of individual clumps. For single-component objects, such as galaxy A in Fig. 2 and well-ordered massive galaxies at intermediate and low redshifts, both definitions should give consistent results.

None the less, we note that our size measurement depends on how we smooth the star particles. In general, using a larger smoothing length results in slightly larger galaxy sizes, but the difference is usually small for most of the galaxies. We refer the readers to Appendix A for details.

our simulations, we expect BPASS binary models to give similar results in terms of feedback strengths, as bolometric luminosities and supernova rates are similar between these two models (see section 2.2 in Ma et al. 2017, for a more detailed discussion).

⁵ We have checked that a 1 arcsec aperture is sufficiently large for most galaxies in our sample. The exceptions are a small number of galaxies that are in late stages of merging. In our sample, galaxy F (shown in Fig. 2) is the object that is most strongly affected: its stellar mass and half-mass radius are underestimated by about 50 per cent because of an ongoing merger.

3 RESULTS

3.1 Qualitative behaviours: an overview

In Fig. 2, we show projected images of stellar mass (left), and noise-free rest-frame UV (middle) and rest-frame *B*-band (right) luminosity for six galaxies at $z = 6$. These galaxies are selected to have similar UV magnitudes around $M_{\text{UV}} \sim -16.5$, in increasing order of stellar mass from the top to the bottom. The images are smoothed for visualization purposes in Fig. 2. Our default size measurements are performed using the original images. The colours represent surface density or surface brightness in linear scale and they saturate at a level such that pixels above it contribute 10 per cent of the total intensity on the image. The square in the top-right corner shows the appearance of a point source on these images.

Despite all galaxies having $M_{\text{UV}} \sim -16.5$, they span two orders of magnitude in stellar mass from $M_* = 3 \times 10^6$ – $2 \times 10^8 M_{\odot}$. They show a wide range of morphologies in their surface density maps: galaxy A is compact; galaxies C, E, and F all have a small companion that is close (within 0.2 arcsec) to the main galaxy; galaxy B is made of several clumps of comparable sizes. High-mass galaxies are generally larger than low-mass galaxies in all bands.

The UV images of these galaxies are largely dominated by one or several bright clumps, where the stellar populations are relatively young (10 Myr on average). Most of the UV clumps are intrinsically small and appear like point sources on these images. More importantly, these bright clumps do not necessarily trace the bulk of stars. In galaxy C, for example, the UV-bright clump is associated with the small companion, while the larger, more massive main galaxy is very faint in the UV. In galaxy D, there are two dominant clumps in the UV image: the smaller one to the upper-left to the galaxy is not associated with any prominent stellar structure; the larger one also has a small spatial offset to the right of the stellar surface density peak. We visually inspected every galaxy in our sample and found this phenomenon to be very common in our simulated galaxies (e.g. galaxies E and F). This is consistent with the off-centre star-forming UV clumps observed in intermediate-redshift galaxies ($z \sim 0.5$ – 2.5 , e.g. Wuyts et al. 2012), which can contribute a large fraction of star formation but only a small fraction of stellar mass. These clumps are either small satellite galaxies or stars formed in individual star-forming regions (see Section 4.4 for more discussion).⁶

In contrast, the *B*-band light traces stellar mass better than the UV, although it can also be biased by the young, bright stellar clumps in some circumstances. In galaxy E, the UV-bright clump associated with the companion upper-left to the central galaxy is also bright in *B* band. The central galaxy also appears bright in the *B* band, but it is much fainter in the UV, due to a relatively older stellar population than the UV-bright stellar clump. In galaxy C, the *B*-band image is dominated by the only bright clump; the main galaxy, however, is faint in the *B* band, because its stellar population is much older.

In general, galaxy size increases with increasing stellar mass or luminosity, following the size–mass or size–luminosity relation. From stellar mass to rest-frame optical to the UV, galaxies tend to be more clumpy, more concentrated, and smaller in size. Moreover, there is a broad distribution of galaxy UV size at fixed UV magnitude. Galaxies A–C have intrinsically small UV sizes, because

⁶ Each system enclosed by a 1 arcsec aperture is counted as one galaxy in this paper even if it contains multiple clumps, as we deem such clumps to be sufficiently close that they can be classified as one galaxy.

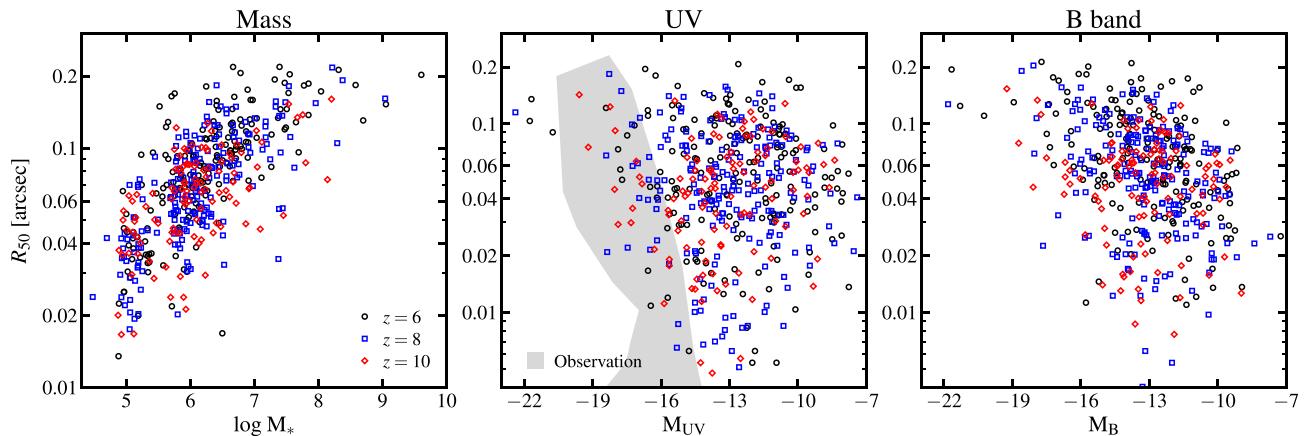


Figure 3. Size–stellar mass relation (left) and size–luminosity relation in the rest-frame UV (middle) and rest-frame B band (right) at $z = 6$ (black circles), 8 (blue squares), and 10 (red diamonds). At any redshift, the sizes of galaxies increase with stellar mass and/or luminosity, but all relations have considerable scatter. There is also a weak redshift evolution of galaxy sizes: high-redshift galaxies tend to have smaller sizes than low-redshift galaxies. The grey shaded region shows observational measurements from Ono et al. (2013), Kawamata et al. (2015), and Bouwens et al. (2017a), adopted from the compilation in Bouwens et al. 2017a (see Section 3.2 for a more detailed discussion).

nearly all of the UV light is emitted by the bright clumps. In contrast, in galaxies D–F, the more extended, low surface brightness pixels contribute a large fraction of the total UV luminosity, so they have larger half-light radius in the UV. However, we caution that low surface brightness regions may fall below the detection limit of a given observational campaign, so the observed half-light radius is consequently smaller if the imaging decreases in depth (Section 4.1). A better way to compare our simulations with observations is to add the background noise of an observing campaign and process the simulations with an identical pipeline for size measurement on the mock images. This is beyond the scope of the current paper, but it is worth exploring in the future.

3.2 Size–mass and size–luminosity relations

In Fig. 3, we show the galaxy size–mass relation (left) and size–luminosity relation in the UV (middle) and B band (right) for our simulated sample. The points represent galaxies at $z = 6$ (black circles), 8 (blue squares), and 10 (red diamonds). We follow Bouwens et al. (2017a) and express the sizes in arcsec in this section. At any redshift, there is a correlation between galaxy size and stellar mass and/or luminosity with considerable scatter. At $M_* < 10^8 M_\odot$, the half-mass radius spans a factor of 3 (0.5 dex) at fixed stellar mass. The scatter is likely driven by several different effects, including halo-to-halo variance, and dynamical effects connected to mergers and strong stellar feedback (El-Badry et al. 2016), and mergers. The size–luminosity relations show larger scatter: at $M_{\text{UV}} > -18$ and $M_B > -18$, the half-light radii spans nearly one dex at fixed magnitude. Most simulated galaxies have half-light radii within 0.01–0.2 arcsec, while some galaxies have even smaller half-light radii down to 0.005 arcsec. In contrast, very few galaxies have half-mass radii smaller than 0.02 arcsec, suggesting that galaxies with extremely small UV sizes should be larger in terms of stellar mass. This is because the bright clumps that dominate the light in these galaxies are very concentrated. At the more massive/brighter end, our simulations do not contain sufficient number of galaxies for a robust estimate of the scatter. In addition, there is a weak redshift evolution of galaxy sizes: the median angular size of galaxies decreases by 25 per cent (physical size by a factor of 2) from $z = 6$ to $z = 10$ at a fixed stellar mass and/or magnitude. This is much smaller

than the intrinsic scatter of the size–mass and size–luminosity relations (see Section 3.3 for quantitative results).

The grey shaded region in Fig. 3 shows the observational data taken from the compilation in Bouwens et al. (2017a) (also including the data from Ono et al. 2013 and Kawamata et al. 2015). Kawamata et al. (2015) measured the sizes of 31 lensed galaxies at $z = 6$ –8 in the Abell 2744 cluster field from the HFF. The half-light radii of galaxies at $M_{\text{UV}} \sim -19.5$ in their $z \sim 6$ –7 sample range from 0.1 to 1 kpc, corresponding to 0.02–0.2 arcsec at these redshifts. Similarly, Bouwens et al. (2017a) also found a similar range of half-light radius from 0.02 to 0.2 arcsec for galaxies with $M_{\text{UV}} \sim -18.5$ at $z \sim 6$ in a more complete HFF sample. Ono et al. (2013) found that $z \sim 7$ –8 galaxies of $M_{\text{UV}} \sim -19$ in the HUDF also have half-light radii from 0.02 to 0.2 arcsec with a median of about 0.06 arcsec. Brighter galaxies at $M_{\text{UV}} \sim -21$ in the HUDF have half-light radii about 0.15 arcsec at $z \sim 5$ –8 (e.g. Bouwens et al. 2004; Oesch et al. 2010). The most massive galaxies in our sample broadly agree with these measurements.

So far, only a small sample of galaxies fainter than $M_{\text{UV}} \sim -18$ from the HFF have size measurements (Kawamata et al. 2015; Bouwens et al. 2017a). These galaxies are reported to have very small intrinsic sizes from 0.01 to 0.06 arcsec, and a small fraction of them even have half-light radii down to 0.001 arcsec. Some of our simulated galaxies fall in the observed range, but our sample also contains a large number of galaxies that have much larger sizes (they tend to lie above the grey shaded region at a given magnitude in Fig. 3). We speculate two possible observational biases/uncertainties that may lead to such discrepancies. First, at fixed magnitude, larger galaxies tend to have lower surface brightness, so they are more likely to be missed in the observed sample. Secondly, for clumpy galaxies, one may only pick up the brightest clumps and thus the sizes are underestimated. Therefore, observations are likely biased towards intrinsically small galaxies and/or star-forming clumps (e.g. Vanzella et al. 2017). In Section 4.1, we will explicitly show the effect of limited surface brightness sensitivity on the observed galaxy sizes. Future deep observations on a few candidate clumpy galaxies with *JWST* can test our predictions. On the other hand, we note that our size measurements are different from those commonly used in observations (see Section 2.2 for a detailed discussion), which further complicates the comparison. It

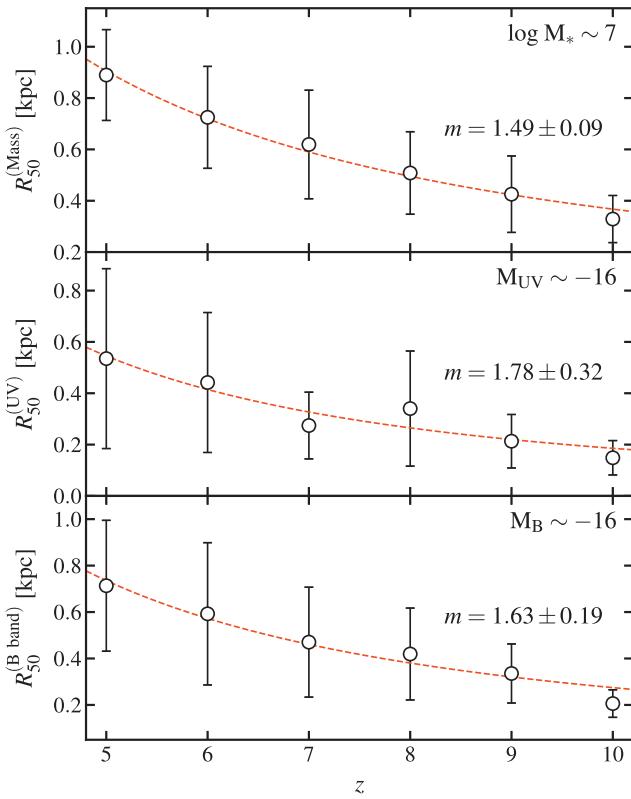


Figure 4. Size evolution of the simulated galaxies at $M_* \sim 10^7 M_\odot$ (top), $M_{\text{UV}} \sim -16$ (middle), and $M_B \sim -16$ (bottom). Points with errorbars show the mean and 1σ (16–84 per cent) distribution of physical half-light radii (in kpc) at $z = 5$ –10. The red lines show the best-fitting evolution following $R_{50} \sim (1+z)^{-m}$. The best-fitting parameters are also listed in Table 1.

would be interesting to carry out more detailed comparisons with specific observational campaigns to understand the discrepancies.

3.3 Size evolution

In this section, we quantify the redshift evolution of galaxy sizes using the simulated sample at $z = 5$ –10. At each redshift, we bin our data in stellar mass every $\Delta \log M_* = 1$ and/or in magnitude every 2 mag. In each bin, we calculate the mean and 1σ distribution (14–86 percentile) of galaxy half-mass and/or half-light radii. In Fig. 4, we show the results at $M_* \sim 10^7 M_\odot$ (top), $M_{\text{UV}} \sim -16$ (middle), and $M_B \sim -16$ (bottom) (same stellar mass/magnitude at all redshifts). Note that we show the physical sizes (in kpc) instead of angular sizes (in arcsec). We fit the evolution trend with a functional form $R_{50} \sim (1+z)^{-m}$. The red dashed lines in Fig. 4 show the best-fitting results at these bins. In Table 1, we list all the best-fitting parameters

Table 1. Best-fitting parameters of galaxy size evolution.

$\log M_*$ (M_\odot)	Stellar mass		Rest-frame UV				Rest-frame B band	
	$R_{50}(z=5)$ (kpc)	m	M_{UV} (mag)	$R_{50}(z=5)$ (kpc)	m	M_B (mag)	$R_{50}(z=5)$ (kpc)	m
8	1.09 ± 0.04	1.15 ± 0.12	–18	0.76 ± 0.08	1.87 ± 0.42	–18	0.99 ± 0.07	1.53 ± 0.25
7	0.91 ± 0.02	1.49 ± 0.09	–16	0.55 ± 0.05	1.78 ± 0.32	–16	0.74 ± 0.04	1.63 ± 0.19
6	0.52 ± 0.02	0.87 ± 0.09	–14	0.49 ± 0.01	1.78 ± 0.06	–14	0.66 ± 0.02	1.92 ± 0.11
5	0.29 ± 0.01	0.84 ± 0.11	–12	0.45 ± 0.02	1.00 ± 0.15	–12	0.46 ± 0.01	1.05 ± 0.09

Note. The size evolution of galaxies at a given stellar mass or magnitude is described as $R_{50} \sim (1+z)^{-m}$ (see Section 3.3 for details). $R_{50}(z=5)$ gives the best-fitting normalization at $z = 5$.

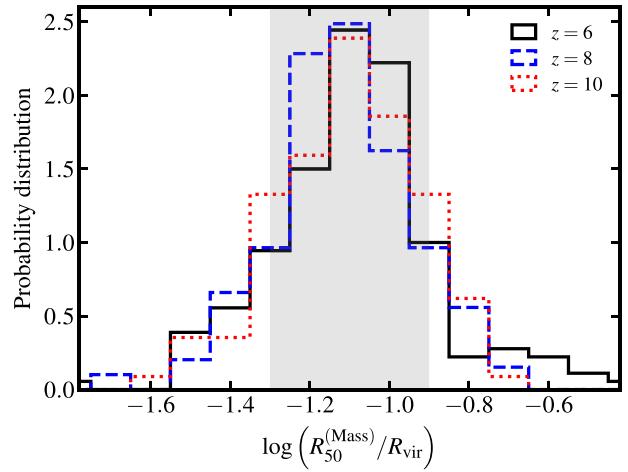


Figure 5. Probability distribution (normalized) of the ratio of galaxy half-mass radius to halo virial radius for the simulated sample at $z = 6$, 8, and 10. $R_{50}^{(\text{Mass})}/R_{\text{vir}}$ has a median value of 0.08 and 1σ range of 0.05–0.12 (the shaded region). The distribution does not evolve significantly over the redshift range considered in this figure.

from $M_* \sim 10^5$ to $10^8 M_\odot$, $-18 < M_{\text{UV}} < -12$, and $-18 < M_B < -12$. It is worth noting that the evolution of the physical sizes has power-law index $m \sim 1$ –2, which is steeper than the redshift dependence of the angular diameter distance [$D_A \sim (1+z)^{2/3}$]. This indicates that the angular sizes of galaxies also decrease with redshift, as shown in Fig. 3.

There are some observational constraints on the size evolution. Various authors have reported $m \sim 1$ –1.5 for galaxies brighter than $M_{\text{UV}} < -19$ across $z \sim 0$ –8 (e.g. Bouwens et al. 2004; Oesch et al. 2010; Kawamata et al. 2015; Shibuya, Ouchi & Harikane 2015). Our results show broad agreement with these constraints (within 1σ for most mass and magnitude bins), although we mainly study galaxies at lower masses and luminosities than the observed sample, and only focus on $z \geq 5$. We note that the best-fitting value of m is sensitive to the data and their uncertainties. For several stellar mass and magnitude bins, our sample only contains a small number of galaxies at some redshift. Stochastic effects in bins with small numbers of objects can strongly affect the results of fitting in those bins.

In Fig. 5, we show the distribution of $R_{50}^{(\text{Mass})}/R_{\text{vir}}$ for our simulated sample at $z = 6$, 8, and 10. We find that for the entire sample, this ratio has a median of 8 per cent and 1σ range from 5 to 12 per cent (the shaded region in Fig. 5). The median and dispersion do not strongly evolve with redshift at $z = 5$ –10. This is consistent with the fact that $R_{50}^{(\text{Mass})} \sim (1+z)^{-m}$ with $m \sim 1$ at these masses (see Table 1), given a non-evolving stellar mass–halo mass relation at these redshifts (Ma et al. 2017) and $R_{\text{vir}} \sim (1+z)^{-1}$ at a fixed halo

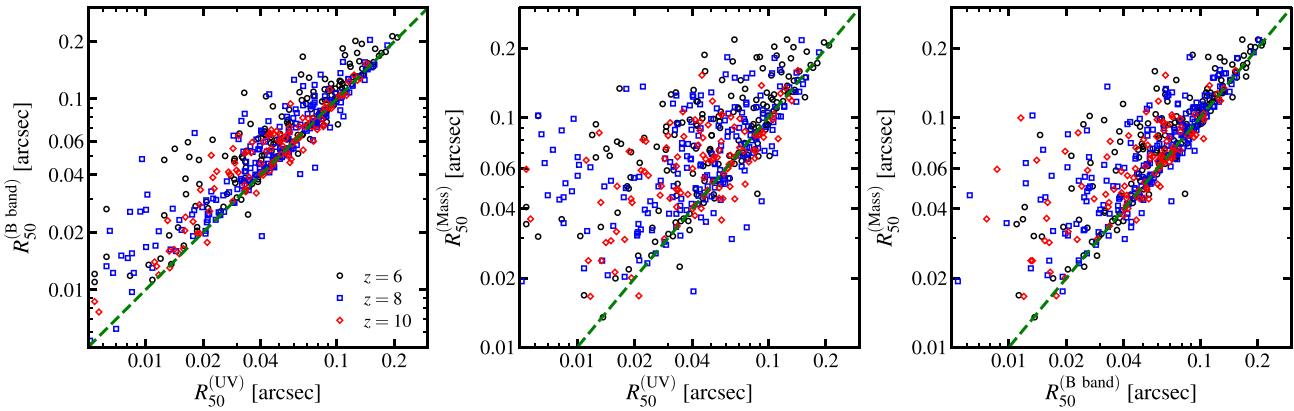


Figure 6. The comparison of galaxy sizes measured in stellar mass, rest-frame UV, and rest-frame B band. The half-mass (light) radii systematically decrease from stellar mass to B band to the UV, in line with a more concentrated morphology in this sequence. The $R_{50}^{(\text{Mass})}-R_{50}^{(\text{UV})}$ relation shows larger scatter than the $R_{50}^{(\text{Mass})}-R_{50}^{(\text{B band})}$ relation, suggesting the UV light is a relatively worse tracer of the stellar mass distribution. Galaxies with small UV sizes are also small in B band, but they usually have much larger half-mass radii, because the B -band light is biased by the bright clumps in these galaxies.

mass (the virial overdensity is nearly a constant at these redshift; see Bryan & Norman 1998). For the few more massive galaxies in our sample ($M_* > 10^8 M_\odot$), this ratio is smaller, mostly at 1–5 per cent: this is comparable to observational measurements for galaxies at similar masses (~ 3 per cent, e.g. Kawamata et al. 2015; Shibuya et al. 2015). Our simulations thus predict that at these redshifts, the stellar-to-halo size ratio (as defined above) is larger for low-mass galaxies, where there are no observational constraints so far.

Our results at $z \geq 5$ should not be extrapolated to lower redshifts. Recently, Fitts et al. (2017) presented a suite of cosmological zoom-in simulations of isolated dwarf galaxies run to $z = 0$ using the same FIRE-2 code. All of their galaxies are hosted in haloes of $M_{\text{halo}} \sim 10^{10} M_\odot$ at $z = 0$. Several galaxies in their sample have stellar mass $M_* \sim 10^7 M_\odot$: these are all early-forming galaxies with half-mass radii around 1 kpc. This is very close to our $z = 5$ galaxy sizes at the same stellar mass, likely due to the fact that the early-forming galaxies in Fitts et al. (2017) do not grow significantly at later times. Although this is a biased sample, and $M_* \sim 10^7 M_\odot$ galaxies may have a broad distribution of half-mass radius at $z = 0$, this suggests that the stellar-to-halo size ratio may be much at lower redshifts for low-mass galaxies (since the virial radius increases with decreasing redshift at fixed mass). This could be due, for example, to less efficient halo gas accretion at later times. For more massive galaxies, our simulations show that the stellar-to-halo size ratio at $z \geq 5$ is already comparable to that at $z \sim 0$ (~ 2 per cent), so it may not evolve strongly at later times (e.g. Shibuya et al. 2015). A mass-dependent evolution of the stellar-to-halo size ratio is consistent with recent analysis for $z \lesssim 3$ galaxies (e.g. Somerville et al. 2018). The galaxy size and morphology evolution down to $z \sim 0$ will be studied in details in a separate paper (Schmitz et al., in preparation).

3.4 Galaxy sizes at different bands

In Section 3.1, we show examples of our simulated galaxies to illustrate that galaxies tend to be more concentrated and smaller from stellar mass to B band to the UV. The UV light is dominated by small, bright, young stellar clumps, while the B -band morphology is determined by both bright clumps and more broadly distributed stars. In Fig. 6, we compare the half-mass (light) radii measured in one quantity against another. The green dashed lines show the $y = x$ relation. All three sizes correlate with each other, but $R_{50}^{(\text{B band})}$ is

systematically larger than $R_{50}^{(\text{UV})}$, and the $R_{50}^{(\text{Mass})}$ is larger than both $R_{50}^{(\text{UV})}$ and $R_{50}^{(\text{B band})}$. We also check the Gini coefficient (e.g. Lotz, Primack & Madau 2004), which is a parameter between 0 and 1 that describes the concentration of galaxy morphology (1 being the most concentrated). We find that the Gini coefficient increases from stellar mass to B band to the UV, in line with the decreasing galaxy sizes in the sequence.

The correlation between $R_{50}^{(\text{Mass})}$ and $R_{50}^{(\text{UV})}$ has a larger scatter than that between $R_{50}^{(\text{Mass})}$ and $R_{50}^{(\text{B band})}$, indicating that rest-frame UV light is a relatively worse tracer of stellar mass than the B band. Furthermore, galaxies with intrinsically small UV sizes (below 0.01 arcsec) mostly have small B -band sizes (below 0.02 arcsec) as well, although these galaxies usually have relatively large half-mass radii (0.04–0.1 arcsec). This is because the B -band light is also biased by the small, bright clumps with high light-to-mass ratios in these galaxies (e.g. galaxy C in Fig. 2).

4 DISCUSSION

4.1 How do surface brightness limits affect observed galaxy sizes?

In some galaxies, a large fraction of the total UV luminosity is contributed in low surface brightness, diffuse light (pixels). These regions are dominated by relatively older stars (10–100 Myr) with lower light-to-mass ratios than those in the young, bright clumps (e.g. galaxies D–F in Fig. 2). For a specific observing campaign, there is a surface brightness limit below which the signal-to-noise ratio is too low to be detectable: this can have a significant effect on the observed morphologies and size measurements of clump-dominated galaxies. In the top panel of Fig. 7, we illustrate this effect using example galaxies D and F from Fig. 2 at $z = 6$. We show the rest-frame UV half-light radii measured for the same galaxies as a function of surface brightness limit (assuming pixels below such limit have zero flux). The effect can be dramatic in some circumstances: for galaxy F, the ‘observed’ half-light radius decreases by over an order of magnitude (from 0.1 to 0.01 arcsec) if the surface brightness depth drops from 29 to 28 mag arcsec $^{-2}$. In the bottom panel of Fig. 7, we further show the rest-frame UV images of galaxies D and F at three surface brightness limits. From $\mu_{\text{min}} = 31.5$ to 28.5 mag arcsec $^{-2}$, most of the low surface brightness regions become ‘invisible’, and the

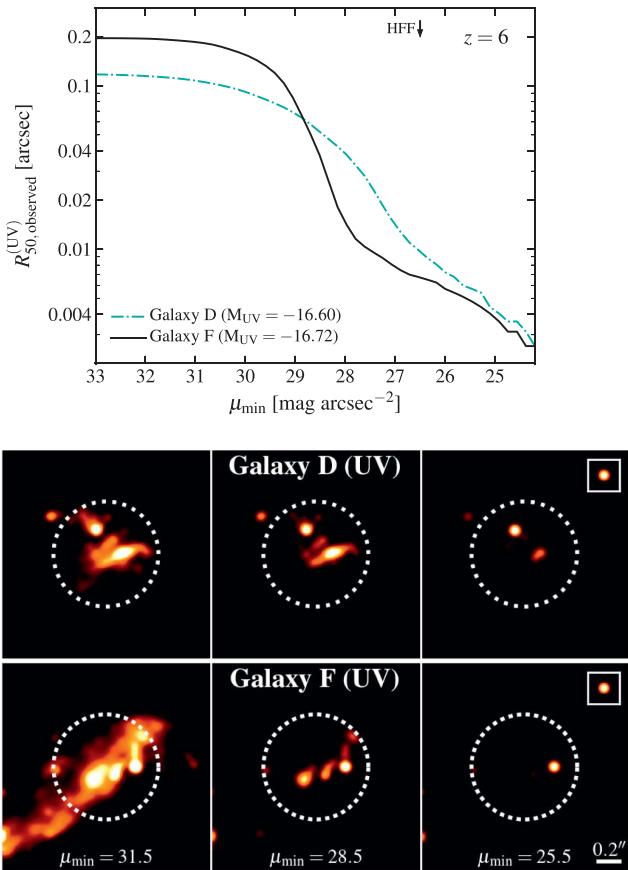


Figure 7. *Top:* Galaxy UV half-light radii measured assuming different surface brightness detection limits for galaxies D and F from Fig. 2. The ‘observed’ size increases with the depth of imaging. *Bottom:* Appearance of galaxies D and F (Fig. 2) at different rest-frame UV surface brightness limits. At a detection limit of $\mu_{\min} = 25.5 \text{ mag arcsec}^{-2}$, the galaxies appear as point sources, and only the brightest clump is dominant.

galaxy is dominated by a few clumps in the UV. Once the detection limit further drops to $\mu_{\min} = 25.5 \text{ mag arcsec}^{-2}$, only the intrinsically small, brightest clump is dominant in these galaxies as a point source.

Now we discuss the implications of this effect on the size-luminosity relation and extremely small sizes measured for galaxies in the HFF. The typical 5σ point-source detection limit in the rest-frame UV of $z = 5\text{--}10$ galaxies is $\sim 28.7\text{--}29.1$ mag within a 0.4 arcsec-diameter aperture (Coe, Bradley & Zitrin 2015). This corresponds to a surface brightness limit about $\mu_{\min} \sim 26.5 \text{ mag arcsec}^{-2}$ for extended sources if we demand the same signal-to-noise ratio within the same aperture. As a proof of concept, we perform a simple experiment on our simulated galaxies to mimic the HFF detection limit: we zero out all pixels below $26.5 \text{ mag arcsec}^{-2}$ and re-measure the luminosities and sizes. We find that *all* galaxies intrinsically brighter than $M_{\text{UV}} < -13$ are still detectable, but their ‘observed’ luminosities and sizes become smaller. *No* galaxies intrinsically fainter than $M_{\text{UV}} > -12$ are detectable. Approximately, the fraction of light *lost* due to such surface brightness cut is a linear function of intrinsic UV magnitude, from zero at $M_{\text{UV}} = -22$ to unity at $M_{\text{UV}} = -12$. In Fig. 8, we show the ‘observed’ size-luminosity relation in the rest-frame UV for our simulated sample. The intrinsic size-luminosity relation for the same galaxies is shown by grey points for reference (non-detectable galaxies are not shown). Most galaxies appear fainter and have much

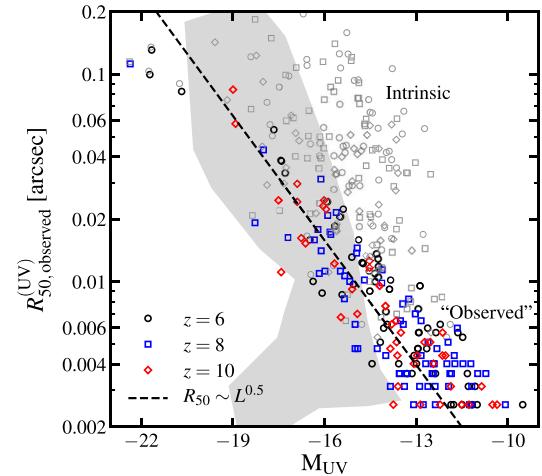


Figure 8. The ‘observed’ rest-frame UV size-luminosity relation for our simulated sample after mimicking the effect of the HFF surface brightness detection limit at $\mu_{\min} \sim 26.5 \text{ mag arcsec}^{-2}$. The grey shaded region represents the observational data as in Fig. 3. The grey points show the intrinsic size-luminosity relation for the same galaxies (non-detectable galaxies are not shown). Most galaxies appear fainter and show much smaller ‘observed’ sizes. The black dashed line shows the $R_{50} \sim L^{0.5}$ scaling as suggested in Bouwens et al. (2017a) for HFF galaxies; however, such scaling is expected due to the selection effect of a surface brightness-limited sample (L/R^2 is constant).

smaller ‘observed’ sizes. When taking into account surface brightness limits, our simulations broadly follow a $R_{50} \sim L^{0.5}$ relation (the black dashed line), as suggested in Bouwens et al. (2017a) for HFF galaxies, but this trend is affected by the $L/R^2 \sim \text{constant}$ selection for a given surface brightness limit. None the less, our simple experiment is by no means a one-to-one comparison with HFF observations. Ideally, one should post-process the high-resolution images of simulated galaxies with gravitational lensing, convolve them with *HST* PSF, add comparable background noise, run identical source finder, and measure the luminosities and sizes using the same method (e.g. Price et al. 2017). This is beyond the scope of this paper, but is worth future exploration in parallel with *JWST* deep surveys.

4.2 Implications for the observed (faint-end) galaxy UV luminosity functions

Current observational constraints on the $z \gtrsim 6$ galaxy UV luminosity functions fainter than $M_{\text{UV}} \sim -17$ come from the HFF program, which takes advantages of foreground galaxy clusters to detect strongly gravitationally lensed high-redshift galaxies. Our results in this paper have two important implications for these observations. First, our simulations show a broad distribution of galaxy sizes at fixed UV magnitude. This affects the estimated completeness correction for the observed sample: if there are more galaxies that have large sizes than expected (they cannot be detected due to low surface brightness), their number densities may be underestimated (Bouwens et al. 2017a). Secondly, some galaxies are dominated by a few small, bright clumps in the rest-frame UV, so they can be mis-identified as several fainter galaxies. If this is the case, the UV luminosity function can be underestimated at intermediate magnitudes, but overestimated at fainter magnitudes. It is interesting that some faint-end UV luminosity functions derived from HFF samples show a small discontinuity at the magnitude where this

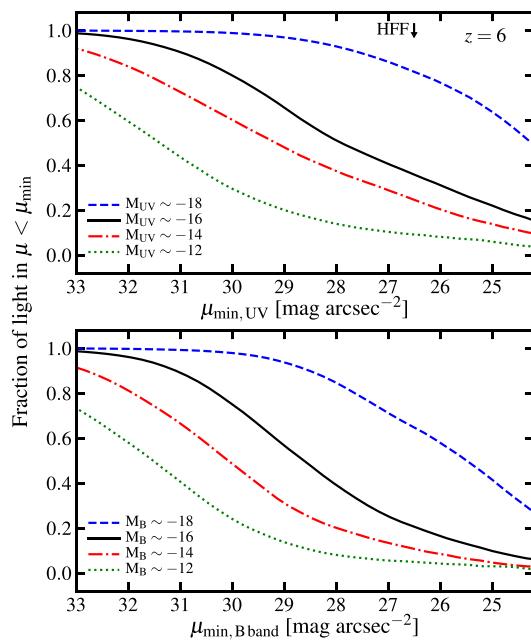


Figure 9. The fraction of light in pixels brighter than $\mu < \mu_{\min}$ as a function of μ_{\min} , averaged over the simulated galaxies at a given intrinsic magnitude (total luminosity) in the rest-frame UV (top) and B band (bottom) at $z = 6$. These results provide predictions on what depths one needs to reach, to target galaxies at a certain magnitude.

effect is likely to become important (although it may also be caused by other effects, e.g. Bouwens et al. 2017b; Livermore et al. 2017). A more quantitative analysis of the observational biases is worth future investigation.

4.3 What fraction of light come from low surface brightness regions?

In this section, we attempt to address the following question: for a given surface brightness limit, what fraction of a galaxy's light will be detected or missed? This is useful for planning future *JWST* deep surveys or follow-up deep imaging and understanding the completeness of an observed sample. In Fig. 9, we show the fraction of light in pixels brighter than $\mu < \mu_{\min}$ as a function of μ_{\min} for our simulated galaxies in the rest-frame UV (top) and B band (bottom) at $z = 6$. We show the results for galaxies at several intrinsic magnitudes in $-18 < M_{\text{UV}} < -12$ and $-18 < M_B < -12$ (averaged over all simulated galaxies at a given magnitude in our sample). Our calculation indicates that at the limits of HFF (26.5 mag arcsec $^{-2}$) and HUFD (~ 1 mag deeper than HFF), more than 80–90 per cent of the rest-frame UV light from galaxies brighter than $M_{\text{UV}} < -18$ should be detected, but this fraction is much smaller for fainter galaxies. In the rest-frame B band, a larger fraction of the light is in low surface brightness regions, as expected from the fact that B -band light is more spatially extended than the UV. Fig. 9 provides information on what depths the observations need to reach for certain targets, although in practice one also needs to account for the PSF of the observational facilities for quantitative comparison.

4.4 The nature of UV-bright clumps

Our simulations suggest that $z \geq 5$ galaxies are mostly irregular, with rest-frame UV images dominated by a few bright clumps.

These clumps mainly have two different origins: some of them are satellite galaxies falling on to their host (e.g. galaxy C in Fig. 2), while others are groups of young stars formed collectively from a parent cloud, i.e. massive giant molecular cloud-like complexes (galaxies D and F). The latter is similar to the clumps formed in gas-rich discs via disc instabilities in intermediate-redshift massive galaxies in simulations (e.g. Genel et al. 2012; Hopkins et al. 2012; Moody et al. 2014; Mandelker et al. 2017; Oklopčić et al. 2017) and observations (e.g. Guo et al. 2015). These high-redshift galaxies are gas-rich and highly turbulent, in part due to rapid accretion from the intergalactic medium. The high degree of turbulent support causes the gas to fragment into large clumps, which subsequently form stars. These early galaxies often do not have well-defined, rotationally supported discs.

The two formation channels mentioned above are essentially the same as the *ex situ* and *in situ* clumps defined in Mandelker et al. (2017). Many clumps formed ‘*in situ*’ are dynamically short-lived (as seen at intermediate-redshift galaxies in Oklopčić et al. 2017). For example, the brightest clump in galaxy F (also see the top-right panel in Fig. 7) contains a mass of $2 \times 10^5 M_{\odot}$ in stars within 100 pc (central surface density $\sim 50 M_{\odot} \text{ pc}^{-2}$) that are formed simultaneously 6 Myr ago; the clump is unbound with a virial parameter $\alpha_{\text{vir}} \sim 2E_{\text{k}}/|E_{\text{p}}| \sim 10$, and it will be dispersed to ~ 500 pc in size within ~ 30 Myr. However, these simulations also form long-lived bound stellar clumps that survive more than 400 Myr, after which the present simulations end. Some of these stellar clumps might survive and evolve into present-day globular clusters (Kim et al. 2018). Bound cluster are more likely to form once the initial gas surface density exceeds $\sim 500 M_{\odot} \text{ pc}^{-2}$ (also see Grudić et al. 2018).

Finally, we caution that these UV-bright clumps are observationally ‘short-lived’: they become much fainter after 30 Myr as the light-to-mass ratio decreases by more than a factor of 10 following stellar evolution and the loss of massive stars. Even the dynamically long-lived clumps are difficult to identify at later times if they only contribute a small fraction of the total stellar mass. Consequently, the rest-frame UV morphology and size of a galaxy can vary greatly on ~ 30 Myr time-scale due to stellar evolution, even if the stellar mass morphology and size do not change dramatically.

5 CONCLUSIONS

In this paper, we use high-resolution FIRE-2 cosmological zoom-in simulations to predict galaxy morphologies and sizes during the epoch of reionization. We project the star particles on to a two-dimensional grid to make stellar surface density and UV and B -band surface brightness images, and measure the half-mass and half-light radii in UV and B band for our simulated galaxies at $z = 5$ –10. Our main findings are as follows:

(i) The simulated galaxies show a variety of morphologies at similar magnitude and/or stellar mass, from compact galaxies to clumpy, multicomponent galaxies to irregular galaxies. The rest-frame UV images are dominated by a few bright, small young stellar clumps that are often not always associated with a large stellar mass. The rest-frame B -band images are determined both by the bulk of stars and by the bright clumps (Section 3.1 and Fig. 2).

(ii) At any redshift, there is a correlation between galaxy size and stellar mass/luminosity with large scatter. At fixed stellar mass, the half-mass radius spans over a factor of 5, while at fixed magnitude, the half-light radius (both UV and B band) spans over a factor of 20 from less than 0.01 arcsec up to 0.2 arcsec (Fig. 3).

(iii) Galaxy morphologies and sizes in our simulations depend on the band in which they are observed. Going from the intrinsic stellar mass distribution to rest-frame *B* band to rest-frame UV, galaxies appear smaller and more concentrated. (Fig. 6). The half-mass radii correlate with *B*-band half-light radii better than those in the UV, suggesting that *B*-band light is a better tracer of stellar mass than the UV light, but it can also be strongly biased by the UV-bright clumps.

(iv) At $z \geq 5$, the physical sizes of galaxies at fixed stellar mass and/or magnitude decrease with increasing redshift as $(1+z)^{-m}$ with $m \sim 1-2$ (Fig. 4). For galaxies below $M_* \sim 10^8 M_\odot$, the ratio of the half-mass radius to the halo virial radius is ~ 10 per cent and does not evolve at $z = 5-10$. More massive galaxies have smaller stellar-to-halo size ratios, typically 1–5 per cent (Section 3.3).

(v) The observed half-light radius of a galaxy strongly depends on the surface brightness limit of the observational campaign (Fig. 7). This effect may account for the extremely small galaxy sizes and size–luminosity relation measured in the HFF observations (Fig. 8), as shallower observations can be dominated by single young stellar ‘clumps’. We provide the cumulative light distribution of surface brightness for typical $z = 6$ galaxies (Fig. 9).

In this paper, we make predictions to help understand current and plan future observations of faint galaxies at $z \geq 5$. Our prediction that these galaxies have small, bright clumps on top of more extended, low surface brightness regions can be tested in the near future by high-resolution deep imaging with *JWST* on a typical sample of galaxies. In future work, we intend to make more realistic comparisons with specific observational campaigns to understand the sample completeness and their implications for the faint-end UV luminosity functions. We will also study the size evolution for a broad range of galaxies from $z = 0$ to 10 (Schmitz et al., in preparation). Moreover, it is also worth quantifying the statistical and physical properties of the UV-bright clumps in $z \geq 5$ galaxies.

ACKNOWLEDGEMENTS

We acknowledge the anonymous referee for useful comments that help improve this manuscript. We thank Rychard Bouwens, Brian Siana, James Bullock, and Eros Vanzella for helpful discussion. The simulations used in this paper were run on XSEDE computational resources (allocations TG-AST120025, TG-AST130039, TG-AST140023, and TG-AST140064). The analysis was performed on the Caltech compute cluster ‘Zwicky’ (NSF MRI award #PHY-0960291). Support for PFH was provided by an Alfred P. Sloan Research Fellowship, NASA ATP Grant NNX14AH35G, and NSF Collaborative Research Grant #1411920 and CAREER grant #1455342. MBK acknowledges support from NSF grant AST-1517226 and from NASA grants NNX17AG29G and HST-AR-12836, HST-AR-13888, HST-AR-13896, and HST-AR-14282 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. CAFG was supported by NSF through grants AST-1412836, AST-1517491, AST-1715216 and CAREER award AST-1652522, by NASA through grant NNX15AB22G, and by STScI through grant HST-AR-14562.001. EQ was supported by NASA ATP grant 12-APT12-0183, a Simons Investigator award from the Simons Foundation, and the David and Lucile Packard Foundation. RF acknowledges financial support from the Swiss National Science Foundation (grant no. 157591). SGK was supported by NASA through Einstein Postdoctoral Fellowship grant number PF5-160136 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. DK was sup-

ported by NSF grant AST-1412153, funds from the University of California, San Diego, and a Cottrell Scholar Award from the Research Corporation for Science Advancement. AW was supported by NASA through grants HST-GO-14734 and HST-AR-15057 from STScI.

REFERENCES

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., Oesch P. A., Atek H., Lam D., Stefanon M., 2017a, *ApJ*, 843, 41
 Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017b, *ApJ*, 843, 129
 Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2017, *MNRAS*, 466, 3612
 Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
 Coe D., Bradley L., Zitrin A., 2015, *ApJ*, 800, 84
 Curtis-Lake E. et al., 2016, *MNRAS*, 457, 440
 El-Badry K., Wetzel A., Geha M., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2016, *ApJ*, 820, 131
 Eldridge J. J., Izzard R. G., Tout C. A., 2008, *MNRAS*, 384, 1109
 Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
 Ferguson H. C. et al., 2004, *ApJ*, 600, L107
 Ferland G. J. et al., 2013, *Rev. Mex. Astron. Astrofis.*, 49, 137
 Finkelstein S. L. et al., 2015, *ApJ*, 810, 71
 Fitts A. et al., 2017, *MNRAS*, 471, 3547
 Genel S. et al., 2012, *ApJ*, 745, 11
 Grudić M. Y., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Murray N., Kereš D., 2018, *MNRAS*, 475, 3511
 Guo Y. et al., 2015, *ApJ*, 800, 39
 Hopkins P. F., 2015, *MNRAS*, 450, 53
 Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, *MNRAS*, 427, 968
 Hopkins P. F., Narayanan D., Murray N., 2013, *MNRAS*, 432, 2647
 Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
 Hopkins P. F. et al., 2017, preprint ([arXiv:1702.06148](https://arxiv.org/abs/1702.06148))
 Jiang L. et al., 2013, *ApJ*, 773, 153
 Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
 Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., 2015, *ApJ*, 804, 103
 Kim J.-h. et al., 2018, *MNRAS*, 474, 4232
 Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
 Kroupa P., 2002, *Science*, 295, 82
 Kuhlen M., Faucher-Giguère C.-A., 2012, *MNRAS*, 423, 862
 Leitherer C. et al., 1999, *ApJS*, 123, 3
 Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, *ApJ*, 835, 113
 Lotz J. M., Primack J., Madau P., 2004, *AJ*, 128, 163
 Lotz J. M. et al., 2017, *ApJ*, 837, 97
 Ma X. et al., 2017, preprint ([arXiv:1706.06605](https://arxiv.org/abs/1706.06605))
 Mandelker N., Dekel A., Ceverino D., DeGraf C., Guo Y., Primack J., 2017, *MNRAS*, 464, 635
 McLure R. J. et al., 2013, *MNRAS*, 432, 2696
 Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, 295, 319
 Mo H. J., Mao S., White S. D. M., 1999, *MNRAS*, 304, 175
 Moody C. E., Guo Y., Mandelker N., Ceverino D., Mozena M., Koo D. C., Dekel A., Primack J., 2014, *MNRAS*, 444, 1389
 Oesch P. A. et al., 2010, *ApJ*, 709, L21
 Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
 Oklopčić A., Hopkins P. F., Feldmann R., Kereš D., Faucher-Giguère C.-A., Murray N., 2017, *MNRAS*, 465, 952
 Ono Y. et al., 2013, *ApJ*, 777, 155
 Orr M. et al., 2017, preprint ([arXiv:1701.01788](https://arxiv.org/abs/1701.01788))
 Planck Collaboration XIII, 2016, *A&A*, 594, A13

Price S. H., Kriek M., Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., Barro G., 2017, ApJ, 844, L6

Ribeiro B. et al., 2016, A&A, 593, A22

Robertson B. E. et al., 2013, ApJ, 768, 71

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19

Schenker M. A. et al., 2013, ApJ, 768, 196

Shibuya T., Ouchi M., Harikane Y., 2015, ApJS, 219, 15

Smith D. J. B., Hayward C. C., 2015, MNRAS, 453, 1597

Somerville R. S. et al., 2018, MNRAS, 473, 2714

Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, ApJ, 826, 159

Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, ApJ, 836, 164

Vanzella E. et al., 2017, MNRAS, 467, 4304

Verner D. A., Ferland G. J., 1996, ApJS, 103, 467

Wuyts S. et al., 2012, ApJ, 753, 114

APPENDIX A: PARTICLE SMOOTHING AND SIZE MEASUREMENTS

In this work, we adopt a non-parametric approach to define galaxy half-mass (light) radii by measuring the area spanned by the brightest pixels that contribute 50 per cent of the total intensity within an 1 arcsec-diameter, S_{50} , and taking $R_{50} = \sqrt{S_{50}/\pi}$ (Section 2.2). We note that the results weakly depend on how we smooth the star particles on the projected images. By default, each star particle is smoothed over a cubic spline kernel with a smoothing length h_5 equal to its distance to the 5th nearest particle. Here, we discuss two alternative smoothing approaches. First, we adopt a smoothing length h_{10} (the distance to the 10th nearest particle) instead of h_5 (but still use the cubic spline kernel) and repeat the size measurement. In the left-hand panel of Fig. A1, we compare the new half-mass radii with our default results for our simulated galaxies. The green dashed line shows the $y = x$ relation. By using h_{10} , the half-mass radii only increase by less than 10 per cent for most of the galaxies (5 per cent difference on average), and only a small fraction (1 per cent) of our galaxies are affected by 50 per cent or more.

Alternatively, we further smooth our default images using a two-dimensional Gaussian function with a dispersion corresponding to the size of 10 pixels (0.032 arcsec). This equals to the pixel size of the NIRCam on *JWST*. This is to mimic the observed galaxy images after convolving with the PSF. A point source thus has a

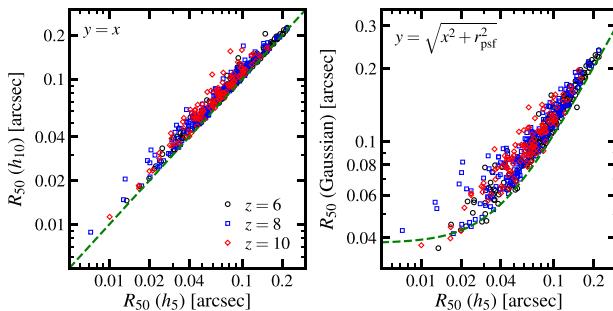


Figure A1. Galaxy sizes measured using different smoothing approaches (see the text for details). Our default size measurements are reasonably numerically robust to this choice.

half-mass (light) radius of $r_{\text{psf}} = 0.038$ arcsec. Note that the example images in Fig. 2 are generated in this way for better visualization. We repeat the non-parametric size measurement on the Gaussian-smoothed images and compare the results in the right-hand panel of Fig. A1. The green dashed line shows the $y = \sqrt{x^2 + r_{\text{psf}}^2}$ relation for reference: we note that this relation is also used in observations to convert apparent sizes to intrinsic sizes (e.g. Oesch et al. 2010). Nearly all of our simulated galaxies lie close to this curve (less than 20 per cent deviation) as expected.

These experiments suggest that we obtain numerically stable galaxy half-mass radii by using a cubic spline kernel with smoothing length h_5 . We find similar results for half-light radii in UV and B band: using h_{10} instead, the B -band sizes are not affected by more than 10 per cent for the vast majority of our galaxies. The differences are slightly larger for UV sizes. Five per cent of our galaxies have UV half-light radii increased by a factor of 1.5–2 when using h_{10} . This is because the UV light in these galaxies is dominated by few diffuse star particles that have large inter-particle distance, so the sizes we obtain can only be treated as upper limits. None the less, most galaxies in our sample are only affected by less than 20 per cent. We conclude that our non-parametric size measurement is robust to our particle smoothing method.

APPENDIX B: RESOLUTION CONVERGENCE

We note that our sample includes simulations using three different mass resolutions for baryonic particles ($m_b \sim 100, 900$, and $7000 M_\odot$). We showed in previous papers that galactic-scale quantities, such stellar mass, star formation rates, etc., converge reasonably well at these mass resolutions (e.g. Hopkins et al. 2017; Ma et al. 2017). Here in Fig. B1, we show the $z = 6$ galaxy size–mass relation for our simulated sample, where the colours represent simulations run with different mass resolution. There is no significant difference between different resolution levels in the size–mass relation, so we conclude that galaxy sizes are robust with respect to resolution in our simulations.

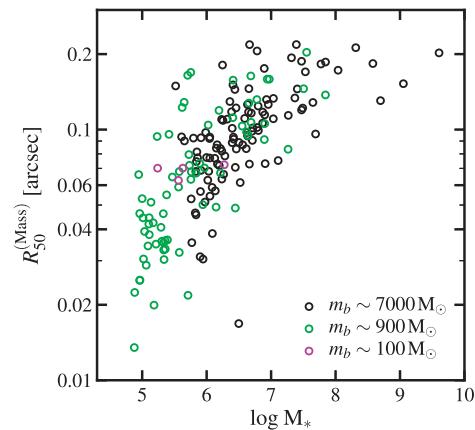


Figure B1. The $z = 6$ galaxy size–mass relation. Colours show simulations run with different mass resolution. Galaxy sizes converge reasonably well with resolution in our simulations.

This paper has been typeset from a $\text{TeX}/\text{\LaTeX}$ file prepared by the author.