

Origin of giant stellar clumps in high-redshift galaxies

Xi Meng¹★ and Oleg Y. Gnedin¹

Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

Accepted 2020 March 16. Received 2020 February 11; in original form 2019 September 23

ABSTRACT

We examine the nature of kpc-scale clumps seen in high-redshift galaxies using a suite of cosmological simulations of galaxy formation. We identify rest-frame UV clumps in mock *HST* images smoothed to 500 pc resolution, and compare them with the intrinsic 3D clumps of young stars identified in the simulations with 100 pc resolution. According to this comparison for the progenitors of Milky Way-sized galaxies probed by our simulations, we expect that the stellar masses of the observed clumps are overestimated by as much as an order of magnitude, and that the sizes of these clumps are also overestimated by factor of several, due to a combination of spatial resolution and projection. The masses of young stars contributing most of the UV emission can also be overestimated by factor of a few. We find that most clumps of young stars present in a simulation at one time dissolve on a timescale shorter than ~ 150 Myr. Some clumps with dense cores can last longer but eventually disperse. Most of the clumps are not bound structures, with virial parameter $\alpha_{\text{vir}} > 1$. We find similar results for clumps identified in mock maps of $\text{H}\alpha$ emission measure. We examine the predictions for effective clump sizes from the linear theory of gravitational perturbations and conclude that they are inconsistent with being formed by global disc instabilities. Instead, the observed clumps represent random projections of multiple compact star-forming regions.

Key words: galaxies: formation – galaxies: high redshift – galaxies: star formation – galaxies: structure.

1 INTRODUCTION

While most nearby L^* galaxies present disc-like morphologies, deep observations with the *Hubble Space Telescope* (*HST*) reveal irregular and clumpy shapes of high-redshift galaxies at rest-frame ultraviolet (UV) and optical wavelengths (Elmegreen et al. 2007; Overzier et al. 2010; Swinbank et al. 2010a). Integral field spectroscopic surveys of these galaxies show both rotation and turbulent motions (Förster Schreiber et al. 2006; Shapiro et al. 2008; Wisnioski et al. 2015). The fraction of galaxies that are clumpy at rest-frame UV evolves with time and varies with galaxy mass. The clumpy fraction for star-forming galaxies increases from $z \simeq 8$ to $z \simeq 1$ –3, reaches a peak, and subsequently decreases until $z \simeq 0$ (Murata et al. 2014; Guo et al. 2015; Shibuya et al. 2016). This trend is similar to the evolution of the cosmic star formation rate (SFR) density (Madau & Dickinson 2014). Moreover, the clumpy fraction tends to increase with SFR at $z \simeq 0$ –2 (Shibuya et al. 2016), suggesting a correlation with star formation activity. The clumpy fraction decreases with stellar mass in galaxies at $z \simeq 0.8$ –2 (Tadaki et al. 2014; Guo et al. 2015).

Some of the earliest found clumpy galaxies were characterized as ‘chain’, ‘tadpole’, and ‘clump cluster’ galaxies (Elmegreen,

Elmegreen & Sheets 2004; Elmegreen et al. 2005). There are typically 2–8 clumps per galaxy, with estimated stellar mass $\sim 10^7$ – $10^9 M_\odot$ (Förster Schreiber et al. 2011; Guo et al. 2012; Soto et al. 2017). The clumps typically have high SFR, resembling mini-starbursts in their galaxies (Bournaud et al. 2015; Zanella et al. 2015). While the clumps contribute only a few per cent individually and ~ 20 per cent altogether to their host galaxy’s total luminosity, their contribution to SFR is larger, ~ 10 –50 per cent (Förster Schreiber et al. 2011; Guo et al. 2012; Wuyts et al. 2012). The SFR of individual clumps varies from $10^{-1} M_\odot \text{ yr}^{-1}$ (Soto et al. 2017) to 1 – $10 M_\odot \text{ yr}^{-1}$ (Guo et al. 2012). The inferred ages of the clump stars range from 10^6 to 10^{10} yr (Soto et al. 2017). There is also wavelength dependence: clumps identified at different wavebands do not fully overlap (Förster Schreiber et al. 2011).

The clumpy structure is also observed in $\text{H}\alpha$ (e.g. Livermore et al. 2012; Mieda et al. 2016) and CO maps (e.g. Swinbank et al. 2010b; Dessauges-Zavadsky et al. 2017a). Kinematic studies of high-redshift massive clumpy galaxies ($M \sim 10^{10.6} M_\odot$) in $\text{H}\alpha$ emission using SINFONI/VLT show that these galaxies are turbulent and rotation dominated. Some galaxies have a massive stellar bulge (Genzel et al. 2008, 2011).

Typical sizes of clumps in *HST* images ~ 1 kpc are at the limit of angular resolution at high redshift. Gravitational lensing has afforded us a magnified view of these galaxies. Adamo et al. (2013) identified 30 clumps of 10^6 – $10^9 M_\odot$ in a lensed spiral galaxy Sp

* E-mail: xim@umich.edu

1149 at redshift 1.5 with spatial resolution ~ 100 pc. Girard et al. (2018) observed a lensed rotating galaxy at $z = 1.59$ and identified three H α clumps, which together contribute ~ 40 per cent of total SFR inferred from the H α flux. The SFR density in these clumps is ~ 100 times higher than in nearby H II regions. Livermore et al. (2012) obtained the luminosity function of clumps with median source plane spatial resolution ~ 360 pc and compared it with the luminosity function of H II regions in galaxies at $z \approx 1-1.5$. They conclude that high-redshift clumps are H II regions that are larger and brighter than local H II regions. The clump sizes in lensed galaxies are smaller than those in unlensed galaxies. Jones et al. (2010) found clumps with diameter 300–1 kpc in lensed galaxies at $z = 1.7-3.1$ with spatial resolution achieving ~ 100 pc. Livermore et al. (2015) extracted 50 star-forming H α and H β clumps with sizes in the range 60–1 kpc in 17 lensed galaxies at $1 < z < 4$. Wuyts et al. (2014) found clumps of diameter $\sim 300-600$ pc in a highly magnified lensed galaxy at $z = 1.70$ and found a radial gradient of their rest-frame UV colour. Johnson et al. (2017) found star-forming clumps of radius smaller than 100 pc in a lensed galaxy at $z = 2.5$. Olmstead et al. (2014) quantified relative stellar-to-nebular extinction in two $z = 0.91$ galaxies with ~ 0.3 kpc resolution. They found that the integrated extinction measurements agree with other studies in that the ionized gas is more obscured than stars. However, when examining on a clump-by-clump basis, they show that the hypothesis that stars and ionized gas experience identical extinction cannot be ruled out.

Unlike the clumpy structure observed in rest-frame UV, ALMA observations of cold dust of massive ($M_* \sim 10^{11} M_\odot$) star forming galaxies at $z \sim 3$ with 200 pc resolution show smooth, disc-like morphology (Rujopakarn et al. 2019).

The puzzling appearance of giant clumps has inspired many theoretical studies that have investigated the formation and evolution of these high-redshift clumps. They include isolated disc simulations (e.g. Tamburello et al. 2015; Inoue & Yoshida 2018) and cosmological zoom-in simulations (e.g. Ceverino, Dekel & Bournaud 2010; Oklopčić et al. 2017). The clumpy fraction in simulations decreases from high redshift ($z \approx 2$) to low redshift (e.g. Buck et al. 2017; Mandelker et al. 2017). These studies identify clumps using a variety of methods: in projected gas density maps (e.g. Oklopčić et al. 2017; Benincasa et al. 2019), projected stellar density maps (e.g. Mayer et al. 2016), mock observational maps (e.g. Hopkins et al. 2012; Moody et al. 2014; Buck et al. 2017), in 3D gas or stellar density distributions (e.g. Mandelker et al. 2014, 2017), or as gravitationally bound objects (e.g. Tamburello et al. 2015; Benincasa et al. 2019). Clumps identified in one type of maps do not necessarily correspond to clumps found in other maps (Moody et al. 2014). In stellar maps, galaxies are only clumpy in UV light, but not in projected stellar mass density (Buck et al. 2017).

The origin of these high-redshift clumps is yet unclear. There are two general scenarios for clump formation: one in which clumps grow through gravitational instability within galactic discs, the other in which clumps are caused by external perturbations, such as mergers. The internal scenario, including violent disc instability and spiral arm instability, is supported by many simulations (e.g. Ceverino et al. 2010; Genel et al. 2012; Inoue & Yoshida 2018) and observations (e.g. Elmegreen et al. 2007; Genzel et al. 2008; Guo et al. 2012; Zanella et al. 2015). Studies of nearby turbulent disc galaxies that resemble the high-redshift clumpy galaxies (Fisher et al. 2017a) show that the clump sizes are consistent with the results of instabilities in self-gravitating gas-rich discs (Fisher et al. 2017b). If clumps form from fragmentation driven by turbulence,

the clump stellar mass function should follow a power law of slope close to -2 . Dessauges-Zavadsky & Adamo (2018) derived the mass function of star-forming clumps at $z \sim 1-3.5$ and found the power-law slope ≈ -1.7 at $M_* > 2 \times 10^7 M_\odot$, in agreement with the turbulence-driven scenario.

On the other hand, some simulation results indicate that ex situ mergers contribute to at least a portion of the clumps (Mandelker et al. 2017), and that massive clumps could form from minor galactic mergers (Mandelker et al. 2014) and clump-clump mergers (Tamburello et al. 2015). Observations of merging galaxies appear to support such merger-driven clump formation (Puech et al. 2009; Puech 2010; Guo et al. 2015; Ribeiro et al. 2017).

The final fate of the clumps is also currently under debate. One alternative is that these clumps would migrate to the galactic centre and potentially contribute to the galactic bulge (Ceverino et al. 2010; Inoue & Saitoh 2012; Perez et al. 2013; Bournaud et al. 2014). Evidence for this scenario is built on the observed radial gradient of clump's colour or age, such that clumps closer to the galaxy centre are older (Noguchi 1999; Genzel et al. 2008; Adamo et al. 2013). This scenario requires clump lifetime to be longer than a few orbital times. Another alternative is that these clumps dissolve in a relatively short time and may contribute to the thick disc, whereas the colour gradient could instead be a result of the inside-out disc growth (Murray, Quataert & Thompson 2010; Genel et al. 2012; Hopkins et al. 2012; Buck et al. 2017; Oklopčić et al. 2017). Some studies advocate both scenarios: low-mass clumps get disrupted in a short time, while more massive clumps survive and migrate to the centre (Genzel et al. 2011; Mandelker et al. 2017).

The interpretation of clump origins is complicated by possible overestimation of the clump masses and sizes due to limited angular resolution and sensitivity (e.g. Tamburello et al. 2015, 2017; Cava et al. 2018). The observed kpc-scale clumps may also be clusters of clumps or blending of smaller structures (Behrendt, Burkert & Schartmann 2016). Due to limited sensitivity, the observed clumps may be biased against low-mass structures (Dessauges-Zavadsky et al. 2017b). These effects need to be taken into consideration when measuring clump properties.

In this work, we revisit the nature of giant clumps using our state-of-the-art simulations of galaxy formation. We explore how resolution and sensitivity affect the inferred clump mass and size, and investigate the final state of these clumps. Our high-resolution cosmological simulations include novel and most realistic modelling of star formation and stellar feedback, which allows us to produce galaxy structures that closely resemble observed high-redshift galaxies. In Section 2 we describe our simulations and clump identification method, both in 3D and in 2D projection. We present the properties of real 3D clumps and mock 2D clumps in our simulations in Section 3. In Section 4 we discuss how simulated clumps compare with the observations and other simulations. We present our conclusions in Section 5.

2 IDENTIFICATION OF CLUMPS

2.1 Simulation suite

We use a suite of cosmological simulations performed with the Adaptive Refinement Tree (ART) code (Kravtsov, Klypin & Khokhlov 1997; Kravtsov 1999, 2003; Rudd, Zentner & Kravtsov 2008) and described in Li, Gnedin & Gnedin (2018) and Meng, Gnedin & Li (2019). All runs start with the same initial conditions in a periodic box of 4 comoving Mpc, so that the main halo has total mass $M_{200} \approx 10^{12} M_\odot$ at $z = 0$, similar to that of the Milky Way.

The ART code uses adaptive mesh refinement to increase spatial resolution in dense regions. There are 128^3 root grid cells, setting the dark matter particle mass $m_{\text{DM}} \approx 10^6 M_{\odot}$. The finest refinement level is adjusted in runtime to keep the physical size of gas cells at that level between 3 and 6 pc. Because of strong stellar feedback, few cells reach the finest refinement level and the typical spatial resolution of molecular gas is 36–63 pc. We calculated the halo spin parameter $\lambda = J/(\sqrt{2}MVR)$ (Bullock, Kravtsov & Weinberg 2001) for our dark matter haloes, where J is the total angular momentum, M is the halo mass, V and R are virial velocity and virial radius of the dark matter halo. The halo spin parameter varies in the range of ~ 0.01 – 0.05 , depending on the redshift. Here we list the halo spin parameters of our dark matter haloes at the analysed snapshots: $\lambda \sim 0.05$ for SFE50 and SFE100 runs at $z = 1.50$, $\lambda \sim 0.01$ for SFE200 and SFE10 at $z = 1.78$, and $\lambda \sim 0.04$ for SFEturb run at $z = 1.98$.

The simulations include 3D radiative transfer (Gnedin & Abel 2001) of ionizing and UV radiation from stars (Gnedin 2014) and extragalactic UV background (Haardt & Madau 2001), non-equilibrium chemical network of ionization states of hydrogen and helium, and phenomenological molecular hydrogen formation and destruction (Gnedin & Kravtsov 2011). The simulations incorporate a sub-grid-scale (SGS) model for unresolved gas turbulence (Schmidt et al. 2014; Semenov, Kravtsov & Gnedin 2016). Star formation is implemented with the continuous cluster formation (CCF) algorithm (Li et al. 2017) where each star particle represents a star cluster that forms at a local density peak and grows mass via accretion of gas until the feedback from its own young stars terminates the accretion. The feedback recipe includes early radiative and stellar wind feedback, as well as a supernova (SN) remnant feedback model (Martizzi, Faucher-Giguère & Quataert 2015; Semenov, Kravtsov & Gnedin 2016). The momentum feedback of the SN remnant model is boosted by a factor $f_{\text{boost}} = 5$ to compensate for numerical underestimation and to match the star formation history (SFH) expected from abundance matching. To explore the variation of results with the speed of star formation, we ran a suite of simulations with different value of the local star formation efficiency (SFE) per free-fall time, ϵ_{ff} . For full description of star formation and feedback recipe, see Li et al. (2017, 2018).

Our simulation suite with $f_{\text{boost}} = 5$ produces the SFH expected for $10^{12} M_{\odot}$ haloes from abundance matching. The initial distribution of star clusters also matches observations of young clusters in nearby galaxies. The cluster initial mass function can be described by a Schechter function, the slope of which is close to the observed value -2 for $\epsilon_{\text{ff}} = 0.5$ – 1.0 . The fraction of clustered star formation correlates with the SFR density, consistent with observations for $\epsilon_{\text{ff}} = 0.5$ – 2.0 . The formation time-scales of clusters in runs with $\epsilon_{\text{ff}} \geq 0.1$ are shorter than 3 Myr, within the range of the observed age spread of young star clusters.

In this paper we analyse several runs with different value of ϵ_{ff} . The number after ‘SFE’ in the run names corresponds to the percentage of local ϵ_{ff} . In SFEturb run ϵ_{ff} is variable and turbulence-dependent (as implemented by Semenov et al. 2016). The typical values are 3 percent, with a lognormal scatter of about 0.3 dex (Li et al. 2018). For each run we focus on the main galaxy in the simulation box. In all the runs the galaxies have a similar SFH, reproducing the expectation of abundance matching. The galaxies also have similar axisymmetric stellar surface brightness profile (Meng et al. 2019). The SFE has systematic effects on small spatial and temporal scales: higher SFE leads to shorter formation time-scales for star clusters and more concentrated stellar feedback. This results in more bursty SFR and lower fraction of star-forming gas

in galaxies with higher SFE. Our simulated galaxies have stellar masses $(2.4 - 8.2) \times 10^9 M_{\odot}$ and neutral gas mass $(2.7 - 8.4) \times 10^9 M_{\odot}$ at the analysed snapshots ($z \approx 1.5$ – 3 , depending on the run). Other global properties of the simulated galaxies are listed in our previous paper (Meng et al. 2019).

2.2 Identification criteria

To match the procedures for finding clumps in rest-frame UV images of high-redshift galaxies, we made mock observations of our simulated galaxies and identified clumps in the mock images. We use the *Flexible Stellar Population Synthesis* (FSPS) model (Conroy, Gunn & White 2009; Conroy & Gunn 2010) to generate spectral energy distributions of all star particles and shift them to their corresponding redshifts in the simulation outputs.

To focus on rest-frame near-UV ($\sim 2800 \text{ \AA}$) light, Guo et al. (2015) detected clumps in *HST*/ACS F435W band for the galaxy redshift range $0.5 \leq z < 1.0$, F606W band for $1.0 \leq z < 2.0$, and F775W band for $2.0 \leq z < 3.0$. Note that the CANDELS galaxies used in Guo et al. (2018) have stellar masses $10^9 - 10^{11} M_{\odot}$ at $z = 1.5$ – 3 , covering the range of our galaxy masses but extending to higher values. We follow their choice and calculate the simulated galaxy fluxes in the latter two bands, because our outputs lie in the range $1 < z < 3$. First we bin the mock observational images in 2D patches of 100 pc side (physical, not co-moving) to avoid uneven resolution of cells at different refinement level. Then we degrade the images to patches of 500 pc to approximate the *HST* resolution at these redshifts. Galaxy orientation is chosen to be face-on, where the galaxy plane is determined using the shape tensor of neutral gas (Meng et al. 2019).

We do not take into account the effect of dust because dust is not modelled explicitly in our simulations. Including dust attenuation could change the degree of clumpiness of the mock images, suppressing intrinsically bright clumps and enhancing less luminous ones (Buck et al. 2017).

We use the PYTHON package ASTRODENDRO¹ to identify clumps in our mock observations. This package computes dendrograms, which are tree diagrams particularly useful for identifying hierarchical structures. A dendrogram contains *leaves*, which have the highest values and no substructure, and combines them into *branches*, which merge hierarchically into the largest branch, the tree *trunk*. We consider only leaves in the dendrogram of mock observational maps, since structures larger than leaves contain more than one clump. We focus our analysis on the brightest clumps, by setting a lower limit on the fractional contribution of each clump to the galaxy UV luminosity: $f_{\text{LUV}} \equiv L_{\text{clump}}^{\text{UV}}/L_{\text{galaxy}}^{\text{UV}} > 3$ per cent, following Guo et al. (2018).

There are three parameters in the ASTRODENDRO setup: *min-value*, which is the minimum value of the surface brightness for a structure to be identified; *min-delta*, which is the minimum difference between adjacent structures; and *min-npix*, which is the minimum number of pixels required to form a structure. Experimentation showed that the results depend largely on *min-npix*, but not the other two parameters. We set the value of *min-npix* according to the expected observational resolution, but then vary this parameter to find the sensitivity of the results.

We set the parameter *min-value* to be the typical surface brightness in the outer parts of the galaxy, about $26 \text{ mag arcsec}^{-2}$, to include most of it in the dendrogram tree. Raising *min-value*

¹<http://dendrograms.org/>

would make the smallest clumps disappear, most of which are excluded from our sample anyway due to the f_{LUV} cut. It would also make some of the largest clumps to be restricted only to their brightest parts.

The parameter `min-delta` determines whether a peak is considered a clump or noise. We found `min-delta` not as important as `min-npix` in clump identification, as long as it is smaller than the mean difference among pixels that are in the region of clumps. We set the value of `min-delta` at $0.5 \text{ mag arcsec}^{-2}$.

Since high-redshift galaxies do not have well-defined thin discs, intrinsic clumps could be overlapping with each other, even when the galaxies are viewed face-on. Identifying clumps in 3D uncovers this projection effect. We also identify clumps in 3D, using dendrograms of stellar 3D density. We calculate the mass density of all stars and of stars younger than 100 Myr on a uniform 3D grid of cubes of 100 pc side. The Z-direction of this uniform 3D grid is perpendicular to the galaxy plane, so that a column of 3D cubes coincides with one 2D patch. This way we can associate each 2D clump with one or several 3D clumps in the same projected area. The 3D dendrogram builds a similar tree-like structure starting from regions of highest density. We identify 3D clumps using the density of young stars because young stars are the main contributor to UV luminosity, and the distribution of old stars does not appear clumpy.

The choice of parameter `min-value` and `min-delta` for 3D clump identification is similar to that in 2D. We choose the parameter `min-value` by examining the 3D light and young stellar density profile of the faintest clumps and take the value at the boundary where the profile dissolves into the background. `min-delta` is set to be small enough that we do not artificially combine peaks due to this parameter. We take `min-value` = $10^{-1.6} M_{\odot} \text{pc}^{-3}$ and `min-delta` = 0.2 dex for 3D clump identification.

3 RESULTS

3.1 Clump properties in simulations

Using the clump finding method described above, we identified clumps in 2D surface brightness maps and in 3D density grid of young stars. In a given snapshot there are several 2D clumps in the main galaxy, and 10–40 3D clumps, depending on the parameter `min-npix`.

To explore the effects of parameter choice, we tried different values of `min-npix` for the identification of 3D clumps. We aim to find separate 3D clumps of young stars that are large enough to account for most of the UV luminosity from the corresponding projected 2D clump. Our 3D patch size is fixed at 100 pc, so that larger `min-npix` means larger minimum required clump volume. We concluded that `min-npix` = 14 is a good choice for 3D identification: lower value of `min-npix` would only produce more small 3D clumps, while higher value of `min-npix` would artificially lump together clumps that are already large enough. Also, if we use larger value of `min-npix`, we would miss one very small and compact clump in SFE50 that accounts for 48 per cent of the rest-frame UV luminosity of its 2D counterpart. Based on these tests, we think `min-npix` = 14 gives the most stable result for 3D clump finding.

For 2D clump identification in degraded mock observations with patches of 500 pc, we take `min-npix` = 2, since it best matches the 2D clumps identified by eye. This choice also corresponds to the roughly kpc scale of the observed clumps.

Therefore, we adopt the values `min-npix` = 2 for 2D and `min-npix` = 14 for 3D clump finding. This choice implies a

minimum size of the identified clumps. We define the effective clump size as the radius of a circle (in 2D) or a sphere (in 3D) with the same area/volume as the sum of patches in the clump, via $\text{Clump Area} \equiv \pi R_{2D}^2$ and $\text{Clump Volume} \equiv \frac{4\pi}{3} R_{3D}^3$. Our choice of `min-npix` corresponds to the minimum effective radius of 400 pc in 2D and 150 pc in 3D. Note that Oklopčić et al. (2017) uses similar minimum effective radius (125 pc) for 2D clump identification in FIRE simulations.

Fig. 1 shows the identified 2D clumps and 3D clumps in five simulation runs with different SFE. Colour maps show the surface brightness of each galaxy in *HST*/ACS F606W (for $1 \leq z < 2$) and F775W (for $2 \leq z < 3$) bands. The spatial resolution for the left-hand panels is a typical simulation patch of 100 pc, while the right-hand panels are degraded to 500 pc to approximate the *HST* resolution at $z > 1.5$. Contours in the right-hand panels show 2D clumps identified in the degraded maps. For each of these clumps, in the corresponding left-hand panel we show the projected contours of the 3D clump that contributes the most to the rest-frame UV luminosity within the area of that 2D clump. Reduced resolution tends to mix smaller clumps into one big peak, resulting in the kpc-scale clumps seen in the observations.

The SFE50 run has five 2D clumps. The top one is the brightest while the middle one is the largest. The large clumps are massive and extended and most of them include multiple 3D clumps in projection. This leads to significant overestimation of both mass and size. The most luminous corresponding 3D clump is bright and compact, but there is also another 3D clump that accounts for 15 per cent of the 2D luminosity, and four other clumps, each contributing a few per cent of the luminosity. At the top edge of the galaxy there is another structure which produces a single 2D pixel, but since we require at least two pixels for 2D clumps, this single pixel is not identified as a clump.

The SFE10 run has a big 2D clump in the centre, which is also a complex of multiple 3D clumps. Some of its smaller 3D clumps show elongated structure. The last snapshot of SFEturb run is at $z = 2.85$, much earlier than the other runs, so the main galaxy in SFEturb run is smaller. It has only three 2D clumps; one of the corresponding 3D clumps is compact, and the other two are more extended. The SFE100 run has three 2D clumps, two of which are large and extended, and so are the corresponding 3D clumps. The SFE200 run has four 2D clumps, all of which look elongated to some extent. The corresponding 3D clumps are small and compact, with many young stars located between the clumps, further leading to overestimation of mass and size of the 2D clumps.

Table 1 lists the properties of 2D and 3D clumps. The masses of the 2D clumps are in the range $10^7 - 10^9 M_{\odot}$, in agreement with clump masses in Guo et al. (2018). While the effective sizes of 2D clumps are of the order of kpc, as in the observations (Förster Schreiber et al. 2011; Guo et al. 2012; Soto et al. 2017), the sizes of 3D clumps are significantly smaller. The projected area of some 2D clumps may contain multiple 3D clumps. To demonstrate the effects of mixing and projection, we include two estimates of 3D mass: the total mass of all 3D clumps that contribute more than 32 magnitude in the corresponding *HST* band (but not necessarily centred) within the 2D clump column, and the mass of only one clump that contributes the highest rest-frame UV luminosity. Although mixing of multiple 3D clumps increases the mass by a factor of a few in some cases, inclusion of stars that are not in any identified 3D clump contributes much more to the overestimation of 2D clump masses.

The flux of the 2D clumps in their corresponding wavebands is similar to that in the observations of Guo et al. (2018). The flux of their clumps at $1.5 < z < 3$ in F606W and F775W bands is mostly

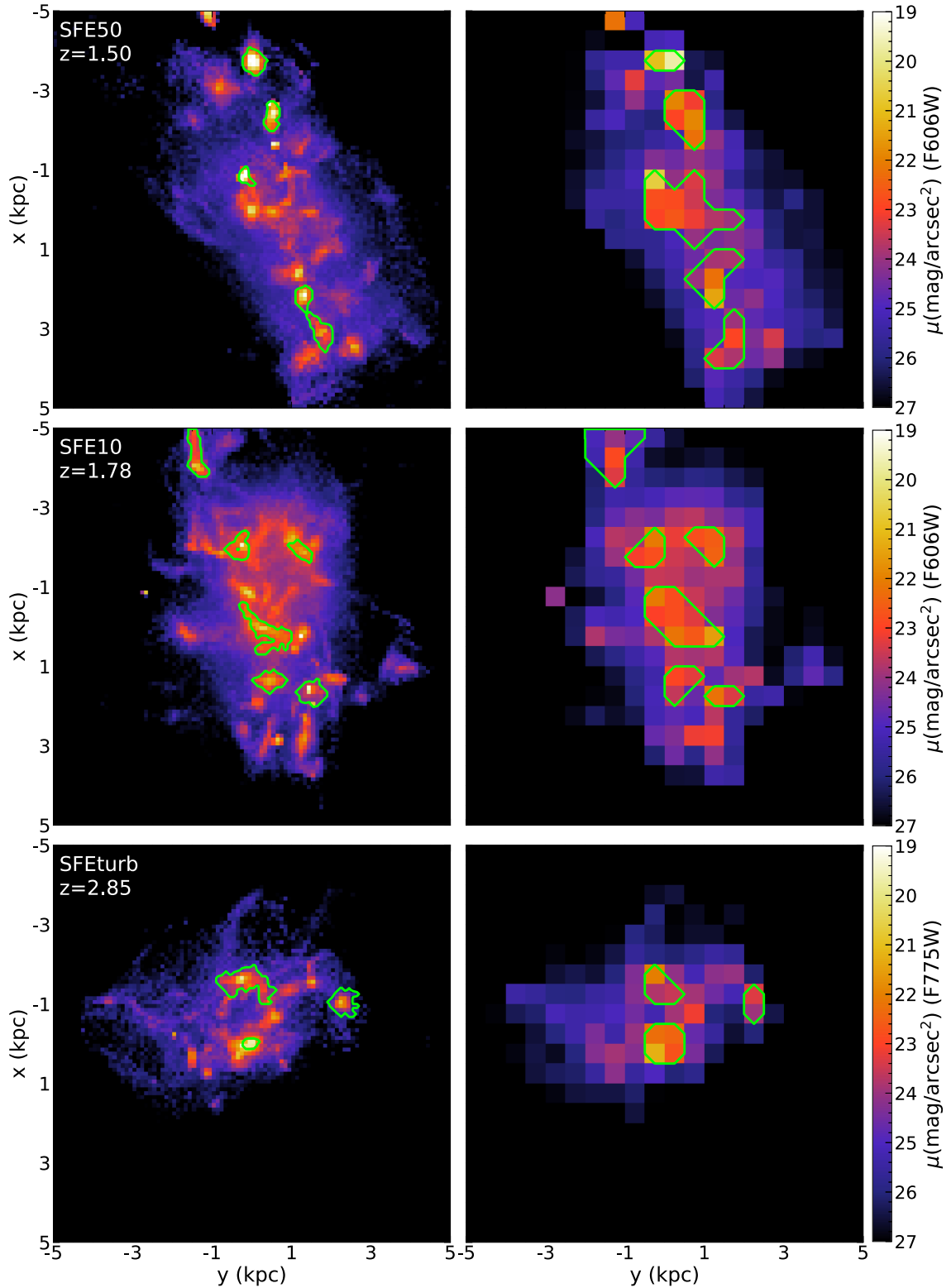
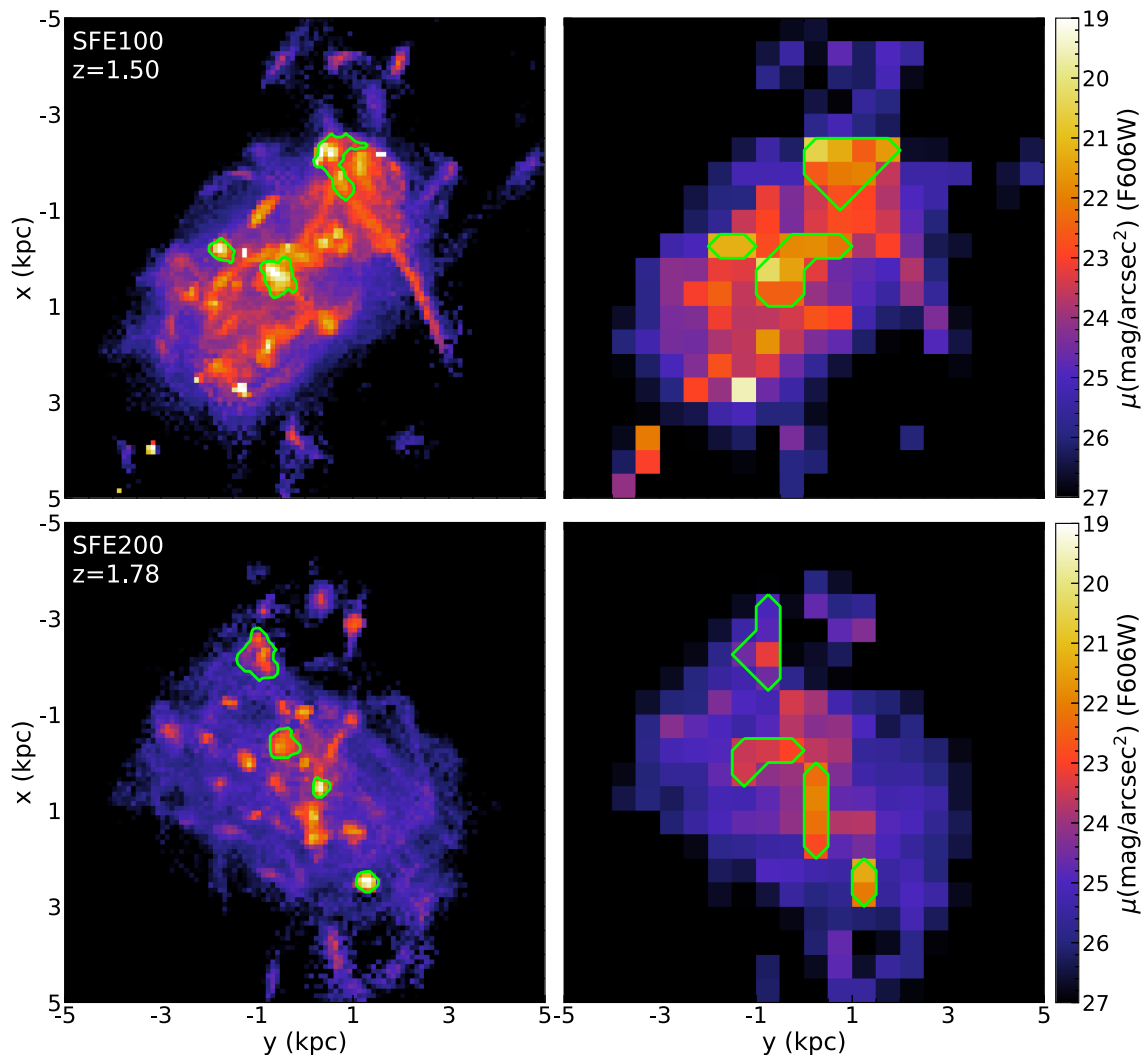


Figure 1. Comparison of 2D and 3D clumps in mock images of simulated galaxies. Colour maps show the surface brightness in *HST*/ACS F606W band (except lower panels for SFEturb which are in the F775W band because of higher redshift of the simulation output) and do not include internal extinction. Left-hand panels have pixels of 100 pc, right-hand panels are degraded to *HST* resolution (500 pc) at these redshifts. Contours in the right-hand panels show 2D clumps identified in the degraded maps using ASTRODENDRO with `min-npix` = 2. Only clumps with fractional luminosity > 3 per cent are shown. Contours in the left-hand panels show the most luminous 3D clump of young stars within the projected area of the corresponding 2D clump, with `min-npix` = 14.

Figure 1 – *continued*

between 0.02 and 0.3 μJy . Similarly, the largest of our simulated 2D clumps can contain between 20 per cent and 37 per cent of their host galaxy UV luminosity.

In the remaining columns we show the ratio of luminosities and SFRs of the most luminous 3D clump and the corresponding 2D clump. Although the masses of 2D clumps are significantly overestimated, the difference in luminosity is smaller. This is because we identify 3D clumps using young stars and they are the main contributor to the rest-frame UV luminosity. Similar to the luminosity ratio, the SFR in 2D clumps is also less overestimated than mass and size, because UV luminosity traces young stars.

The total contribution of 3D clumps to SFR in 100 Myr is 50–80 per cent. Each individual 3D clump typically contributes a few per cent to total SFR, up to ~ 30 per cent. The contribution of individual 2D clumps to the total SFR typically varies from a few per cent to ~ 15 per cent, and can be up to 60 per cent. This is generally consistent with the observational results of Guo et al. (2015) who find that the clumps typically contain a few per cent of the total SFR, and of Soto et al. (2017) who find median clump contribution of 5 per cent to the total SFR.

3.2 Comparison of 2D and 3D masses and sizes

The first obvious result of our comparison of clumps identified in 2D and 3D is the differences in mass and size. We compare the properties of 2D and 3D clumps in Fig. 2. Each point corresponds to the 3D clump that contributes the most to the luminosity of the 2D clump when projected on to the galaxy plane. The stellar mass of the identified 2D clumps range from several times $10^7 M_\odot$ to $10^9 M_\odot$, while few of the corresponding 3D clump masses exceed $10^8 M_\odot$. All the points are below the 1:1 line, which means that 2D clump masses and sizes are overestimated.

The overestimation of the 2D clump mass can be as large as an order of magnitude. Fractional luminosity of 2D clumps is not predictive of the extent of mass overestimation. The most overestimated 2D clumps are also more likely to contain multiple 3D clumps of comparable mass. Combining multiple snapshots, the most massive clump mass in both 2D and 3D seems to increase with decreasing SFE. The most overestimated clumps tend to have large fractional luminosity, since they are more likely to be combination of multiple clumps.

Similarly, the effective sizes of 2D clumps are also overestimated, especially for those that are made up of multiple 3D clumps. The

Table 1. Clump masses and sizes identified in 2D and 3D.

Run	L_{2D}/L_{gal}^a	Flux ^b (μ Jy)	M_{2D}^c ($10^7 M_\odot$)	R_{2D}^d (pc)	$M_{3D,tot}^e$ ($10^7 M_\odot$)	$M_{3D,max}^f$ ($10^7 M_\odot$)	R_{3D}^g (pc)	$L_{3D,max}/L_{2D}^h$	$SFR_{3D,max}/SFR_{2D}^i$	$L_{3D,H\alpha}/L_{2D,H\alpha}^j$
SFE200	0.195	0.060	22.9	560	5.6	0.7	160	0.42	0.10	0.11
	0.176	0.054	4.9	400	1.1	1.1	200	0.95	0.97	–
	0.074	0.023	15.4	560	2.1	0.8	230	0.37	0.56	0.88
	0.041	0.013	6.1	630	2.1	2.1	330	0.74	0.90	–
SFE100	0.198	0.213	34.9	800	9.8	6.4	300	0.57	0.52	0.43
	0.198	0.213	68.7	750	13.6	6.5	300	0.50	0.48	0.14
	0.073	0.078	12.4	400	1.3	1.3	220	0.46	0.65	–
SFE50	0.320	0.211	6.3	400	4.7	4.7	260	0.99	0.97	0.73
	0.201	0.132	58.2	940	6.8	0.7	150	0.48	0.09	0.82
	0.111	0.073	10.3	630	0.7	0.7	190	0.69	0.58	0.70
	0.090	0.059	24.3	630	3.9	0.8	210	0.51	0.17	0.19
	0.032	0.021	9.6	560	1.8	1.5	270	0.38	0.61	–
SFE10	0.236	0.108	82.9	800	8.3	3.7	280	0.21	0.41	0.26
	0.076	0.035	21.6	490	3.9	3.1	220	0.52	0.57	–
	0.074	0.034	26.0	490	4.2	1.1	180	0.23	0.16	–
	0.047	0.022	5.0	400	1.0	1.0	220	0.89	0.81	–
	0.046	0.021	4.5	690	0.8	0.8	260	0.65	0.60	0.49
	0.042	0.019	14.6	490	1.2	1.2	230	0.56	0.83	0.35
SFEturb	0.372	0.091	35.6	560	20.3	8.0	160	0.31	0.29	0.28
	0.132	0.032	15.6	490	11.7	9.8	350	0.82	0.84	–
	0.043	0.011	4.9	400	3.1	3.1	290	0.91	0.94	–

Notes. ^aFraction of the galaxy's luminosity contributed by the 2D clump. Only clumps with fractional luminosity > 3 percent are listed.

^bClump's flux in *HST*/ACS F606W waveband, except for SFEturb run which is in F775W because of higher output redshift.

^cStellar mass of 2D clumps, defined using surface density as: $M_{2D} = \Sigma_* \times \text{ClumpArea}$.

^dEffective radius of 2D clumps, defined as: $\text{Clump Area} = \pi R_{2D}^2$.

^eSum of stellar mass of all 3D clumps within the projected area of the 2D clump.

^fStellar mass of the most luminous 3D clump within the projected area of the 2D clump.

^gEffective radius of 3D clumps, defined as: $\text{Clump Volume} = \frac{4\pi}{3} R_{3D}^3$.

^hRatio of luminosity of 3D clumps to luminosity of 2D clumps in each galaxy's corresponding waveband.

ⁱRatio of star formation rates in 3D clumps and 2D clumps, averaged over 100 Myr.

^jRatio of H α luminosity in 3D clumps and 2D clumps, for corresponding 2D UV clumps (matched by eye). '–' means there is no corresponding H α 2D clump for a 2D UV clump.

ratio of 2D to 3D effective radii can be a factor of several. We do not see any correlation between clump sizes and SFE.

All of the 3D clumps have intrinsic sizes below the *HST* resolution of ~ 500 pc, even after combining simulation grid cells to uniform patches of 100 pc. Therefore, the sizes of clumps are *always overestimated* in observations, unless they benefit from the magnification by gravitational lensing.

3.3 Clump longevity

Important questions for the interpretation of these giant clumps are: Are they gravitationally bound? How long do they remain identifiable as distinct clumps? One way to address this is to look at how the average distance between pairs of clump stars varies over time. If majority of stars are self-bound, the average distance should not change much. On the other hand, if the stars that appear to be in a clump are unbound after the initial gas clearance, then their average distance would increase steadily over time. The maximum rate of expansion is set by the velocity dispersion of all stars in that region of the galaxy. Below we conduct several tests of the boundedness and longevity of the clumps in our simulations.

Fig. 3 shows the evolution of the average distance between pairs of young stars in 3D clumps, over the last three available simulation snapshots. After identifying a clump in the first snapshot, we calculate the mass-weighted average pair distance of the same

stars in the following two snapshots, about 150 and 300 Myr later. This can be used as a proxy for characteristic size of the clump. We can see that the average sizes of most clumps increase significantly after only 150 Myr, which means they dissolve and spread out. There are some less massive clumps that spread out slowly after 150 Myr, but they eventually dissolve after 300 Myr. We note that this analysis is limited by the time between saved simulation snapshots, and that clumps could dissolve in shorter than 150 Myr.

To estimate whether the 3D clumps are bound, we calculate the virial parameter

$$\alpha_{\text{vir}} = a \frac{\sigma^2 R}{GM}, \quad (1)$$

where M is clump mass, R is the effective radius, and σ is the 3D velocity dispersion of young stars. We take the parameter $a = 5/3$ for a constant density sphere (Bertoldi & McKee 1992). Values of $\alpha_{\text{vir}} \lesssim 1$ would indicate the clump is bound, although the precise value depends on the structure of the stellar distribution.

To quantify the evolution of the mass and size of a clump of young stars identified in one simulation snapshot, we look for clumps in the same group of stars in the next two consecutive snapshots. We keep the same density threshold for clump identification, so that stellar particles that dissolve into the background are not counted in the successor clumps. Then we calculate the fraction of mass of the original stars that remain in the most massive successor clump. We also take the mass-weighted average distance between pairs of

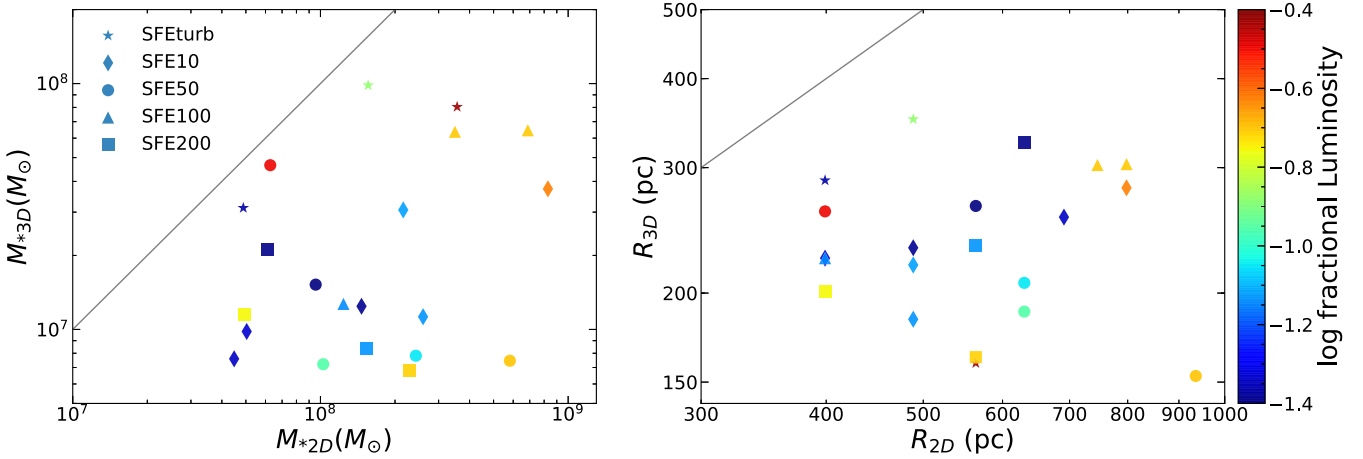


Figure 2. Stellar mass (left-hand panel) and effective radius (right-hand panel) of clumps identified in 2D degraded mock observation map (resolution 500 pc) with $\text{min-npix} = 2$ and 3D density of young stars (younger than 100 Myr) identified with $\text{min-npix} = 14$. Here stellar mass in 2D is total stellar mass within the projected column, and stellar mass in 3D is total stellar mass of the most luminous 3D clump within that 2D clump. Effective radius is the radius of a circle (in 2D) or a sphere (in 3D) that has the same area/volume as the corresponding clump. The colour of each point is the fractional luminosity of the 2D clump, L_{2D}/L_{gal} . The grey thick solid lines are 1:1.

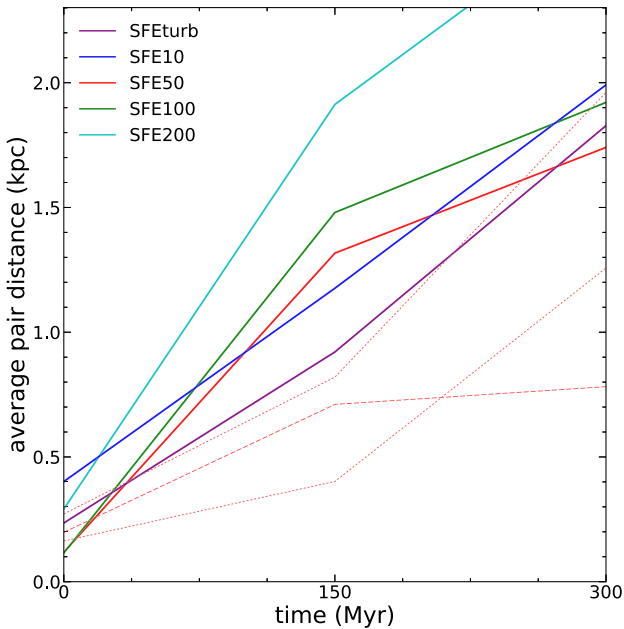


Figure 3. Evolution of mass-weighted average pair distance of young stars in 3D clumps over three consecutive snapshots. The time interval between snapshots is about 150 Myr. The solid lines show the most massive clump in each run. For SFE50 we show also the second most massive clump (dashed line) and the less massive ones (dotted lines). The upper limit of the y-axis corresponds to the distance travelled at 15 km s^{-1} over 150 Myr.

the original stars remaining in the successor clump as a measure of size of the successor clump.

We find a weak correlation between α_{vir} and the mass fraction of young stars that can still be found in the successor clump: clumps with high remaining mass fraction have small $\alpha_{\text{vir}} \lesssim 1$. For example, all (20) but two clumps with remaining mass fraction above 70 per cent have $\alpha_{\text{vir}} < 1$. Equivalently, clumps with small α_{vir} tend to retain more of their original mass: the median remaining mass fraction for clumps with $\alpha_{\text{vir}} < 1$ is 61 per cent, while the median

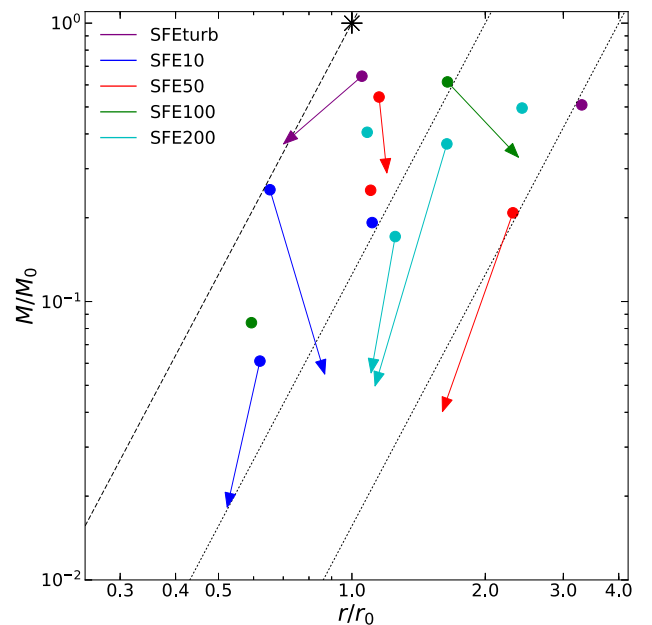


Figure 4. Evolution of the fraction of mass and mean separation of stars that remain in the most massive successor of clumps identified two snapshots before last available. The separation r is calculated as the mass-weighted average distance between all pairs of stars. By construction every clump starts at (1,1), marked by the star symbol. Circles show the ratios one snapshot later (~ 150 Myr), and arrows point to the values two snapshots later (~ 300 Myr). The dashed line marks the initial average density of the clumps, while dotted lines mark constant density 8 and 64 times lower. We note that, since r is a proxy for size, ‘density’ here is also a proxy for the actual density.

for clumps with $\alpha_{\text{vir}} > 1$ is only 7 per cent. Mandelker et al. (2017) similarly found that clumps with shorter lifetime have larger α_{vir} .

The overall evolution of clump mass and size, relative to the moment it was first identified, is shown in Fig. 4. Every clump begins at point (1,1) and then appears as a circle one snapshot later, with an

arrow pointing to the second snapshot later. If there is only a circle without an arrow, it means that the successor was not found after two snapshots. Since we use the average pair distance as a proxy for size of the surviving clump, we can illustrate decrease in the average density with lines of constant M/r^3 . Most circles (one snapshot later) lie below the dashed line, which means that the remaining clumps are less dense than the original clumps. Many arrows point in the direction of decreasing density, indicating that the remaining part of clumps keeps dissolving over time. After 300 Myr the density of most clumps decreases by an order of magnitude, after which they would no longer be interpreted as clumps.

There is one clump in SFETurb run whose average density increases between the second and third snapshots, as the mass continues to decrease. In this case the less bound outer part of the clump dissolves, leaving the remaining inner part more dense. At the time it was first identified, the clump was gravitationally bound, with the virial parameter $\alpha_{\text{vir}} = 0.37$. After 125 Myr, the inner part of this clump still remains bound with $\alpha_{\text{vir}} = 0.19$. However, after 250 Myr, half of this clump is dissolved and spread out all over the galaxy, while core now has $\alpha_{\text{vir}} = 4.9$, with a significant increase in the velocity dispersion. We can expect that this clump will dissolve soon afterwards.

To examine one possible fate of the clumps, that they migrate to the galaxy centre due to dynamical friction and contribute to the bulge (Ceverino et al. 2010), we calculated the evolution of the average galactocentric distance of the young stars in clumps. We found no net inward or outward migration of clump stars, similar to the conclusion of Buck et al. (2017). Instead of massive bulges, our galaxies contain nuclear star clusters that are consistent with being formed by in-situ star formation (Brown, Gnedin & Li 2018).

It has been argued that feedback recipe affects the longevity of clumps in simulations. Simulations with only thermal feedback from supernovae often produce long-lived clumps which then migrate to the galaxy centre (e.g. Ceverino et al. 2010, 2012; Mandelker et al. 2014). In simulations with strong momentum feedback, clumps usually dissolve in a short time $\lesssim 100$ Myr (e.g. Oklopčić et al. 2017). The results of Mandelker et al. (2017) with radiation pressure feedback fall into an intermediate category, where massive clumps survive and low-mass clumps disrupt. Our simulation results are consistent with the strong feedback regime: most clumps dissolve in a short time and do not migrate to the centre. As shown by Li et al. (2018) this strong feedback is required to match the SFH of Milky Way-sized galaxies inferred from abundance matching, and therefore, we favour the conclusions obtained with our simulations.

We have available one weaker feedback run, which produced too high SFR and metallicity. The distribution of young stars in that run is smoother and the disc appears regular even at high redshift. If we run 3D clump identification algorithm over that galaxy, we get 3D clumps that are similar to clumps in other galaxies at face value, except that they are slightly larger. However, these 3D clumps would be less prominent than clumps in other galaxies, given the smoother distribution of young stars. Thus, these clumps should not be viewed as the same ‘clumps’ expected in a clumpy galaxy.

Some observed clumps appear to be old: e.g. Guo et al. (2012) found clump ages in the range $10^8 - 10^9$ yr, Soto et al. (2017) in the range $10^6 - 10^{10}$ yr, Zanella et al. (2019) in the range $10^6 - 10^9$ yr. The oldest end of these intervals is used to argue the longevity of the clumps. Some clumps even appear older than the underlying disc stellar population. However, inferred age can be severely affected by contamination by disc stars, as well as measurement systematics.

When we calculate the mass-weighted average age of all stars in the clump regions in our simulations, we also find rather large ages 0.1–1 Gyr, because of the large amount of old stars that do not contribute much to the rest-frame UV light. The true age of stars producing most of the UV light is of course below 100 Myr and cannot be used to set constraints on the dynamical longevity of the clumps.

4 DISCUSSION

4.1 Effects of angular resolution

In our simulations, masses of the 2D clumps are overestimated by about an order of magnitude. Comparison of clump masses in high-redshift field galaxies with those magnified by gravitational lensing points to a similar overestimation due to insufficient resolution. The observed clump stellar mass increases from around $10^7 M_{\odot}$ for lensed galaxies to $\sim 10^9 M_{\odot}$ for field galaxies (Dessauges-Zavadsky et al. 2017b; Cava et al. 2018). Cava et al. (2018) observed the strongly lensed Cosmic Snake galaxy with the spatial resolution down to ~ 30 pc, as well as the less strongly lensed counterimage of the same galaxy with the resolution ~ 300 pc. They found rest-frame UV clumps in the counterimage with masses $\gtrsim 10^8 M_{\odot}$, while the corresponding clumps in the Cosmic Snake image have masses down to $\sim 10^7 M_{\odot}$. Our 3D clumps and 2D clumps show similar overestimation in mass. Thus angular resolution is critical for correct characterization of giant clumps in high-redshift galaxies.

Clump sizes in lensed galaxies range from several hundred pc to kpc (Adamo et al. 2013; Wuyts et al. 2014; Livermore et al. 2015), smaller than clumps found in unlensed galaxies. Tamburello et al. (2017) tested the resolution effect using the H α map smoothed with different Gaussian full width at half-maximum (FWHM) and found that clump sizes increase from ~ 120 pc to ~ 800 – 900 pc when FWHM increases from 100 to 1 kpc.

Two effects can increase the observed mass of 2D clumps: mixing of multiple 3D clumps within the projected column and inclusion of stars that are not in clumps. From Table 1 we can see that when clump masses are dramatically overestimated, the mass of intraclump stars usually contributes more to the total mass than additional smaller 3D clumps. Including the mass of all 3D clumps usually increases the mass only by a factor of few, while including intraclump stars can inflate the clump mass by over an order of magnitude.

The amount of discrepancy between the 2D and 3D masses is much larger when counting stars of all ages, relative to counting only young stars. Fig. 5 illustrates both ways of calculating the mass: lines connect the mass of young stars (lower left) to the mass of all stars (upper right) within the same clump. Counting all stars instead of only young stars is roughly equivalent to overestimating the SFR. From the plot we can see that SFR in 2D clumps could be overestimated by a factor of 3 or more. Similar result is found in Fisher et al. (2017a), where blurred H α clumps in the DYNAMO-HST sample of low-redshift turbulent disc galaxies have ~ 2 – 3 times higher SFR than the associated full-resolution clumps. The overestimation of SFR is larger for more massive 2D clumps, because they may contain multiple distinct 3D clumps.

However, we do not find systematic dependence of the mass ratios on the output redshift. We plot masses of clumps identified at several epochs in all five runs: 8 outputs for SFE50 and 5 each for SFE100, SFE200, and SFE10 at redshift between 1.5 and 2, and 2 outputs for SFETurb at $z \approx 3$. The distributions of points all these epochs are comparably broad.

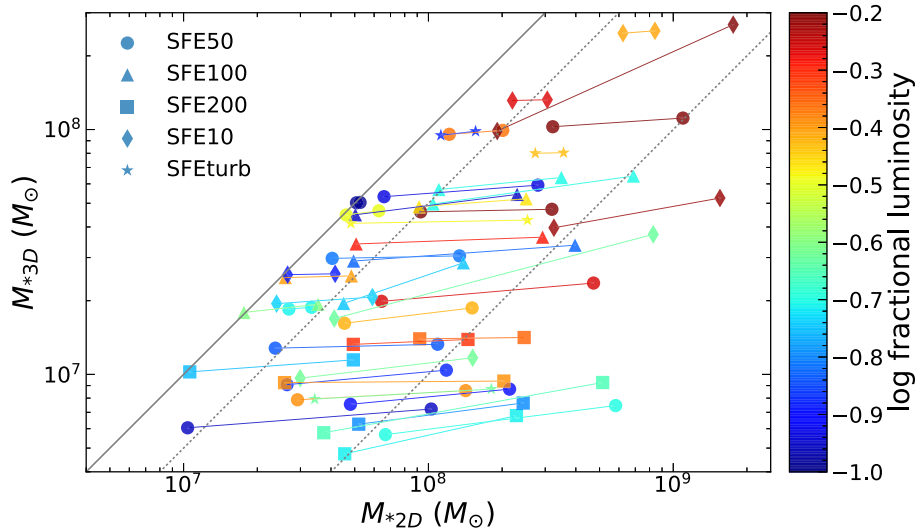


Figure 5. Masses of 2D clumps and the most luminous 3D clumps within them in different snapshots at redshifts from 1.5 to 3 (between 2 and 8 snapshots per run). The lines between symbols connect the mass of young stars (age less than 100 Myr) to the mass of all stars. To reduce the number of plotted points, we include only 2D clumps that contribute fractional luminosity > 10 per cent. The solid line shows 1:1 ratio, and the dotted lines show 1:2 and 1:10 ratios.

A commonly used observational proxy for SFR is the luminosity L_ν in the wavelength range 1500–2800 Å (Kennicutt 1998):

$$\text{SFR} (M_\odot \text{ yr}^{-1}) = 1.4 \times 10^{-28} L_\nu (\text{erg s}^{-1} \text{ Hz}^{-1}). \quad (2)$$

This relation holds for the Salpeter stellar initial mass function (IMF). In this study we adopt the Kroupa (2001) IMF, which requires a correction factor of 0.64 (Kennicutt & Evans 2012). Using this relation, we compared the SFR calculated from the near-UV L_ν in our mock images to the actual SFR in the simulation, averaged over 100 Myr. The SFR inferred from L_ν is systematically 0.13 dex higher than the actual SFR, but with a large scatter of about 0.4 dex. The scatter mainly comes from the difference in age and metallicity of young stars. Stars older than 100 Myr could also contribute to the UV luminosity but are not counted in the SFR, which adds even more scatter.

The ratio of the UV luminosities and SFRs of the 3D and 2D clumps are listed in Table 1. The difference in the two ratios is caused by the scatter in the SFR– L_ν relation discussed above. Most clumps have similar values of $L_{3D, \text{max}}/L_{2D}$ and $\text{SFR}_{3D, \text{max}}/\text{SFR}_{2D}$. Some clumps, however, show large discrepancies. For example, the first clump in SFE200 run and the second and fourth clumps in SFE50 run contain fairly large luminosity fractions, but smaller SFRs by a factor 3–5. This occurs because these 2D clumps combine multiple 3D clumps, and the selected most luminous 3D clump happens to be much younger than the other 3D clumps. For instance, in the second 2D clump in SFE50 run, the average age of the selected 3D clump is 5.9 Myr, while all other overlapping 3D clumps are about 50 Myr old.

4.2 Effects of numerical resolution

We find that clumps identified at one snapshot in our simulations disperse and disappear in subsequent snapshots. To check that the dissolution of clumps is not caused artificially by the numerical relaxation effects, we calculated the half-mass relaxation time

(Spitzer 1969) for stellar particles making up the identified clumps:

$$\begin{aligned} t_{\text{th}} &= \frac{0.17 N}{\ln(\lambda N)} \sqrt{\frac{r_h^3}{GM}} \\ &= \frac{0.78 \text{ Gyr}}{\ln(\lambda N)} \frac{1 M_\odot}{m} \left(\frac{M}{10^5 M_\odot} \right)^{1/2} \left(\frac{r_h}{1 \text{ pc}} \right)^{3/2}. \end{aligned} \quad (3)$$

Here m , M , and N are the average mass, total mass, and number of the stellar particles, respectively, and r_h is the half-mass radius. We take $\lambda = 0.2$ following Binney & Tremaine (2008). Using another suggested value $\lambda = 0.1$ does not change the results.

The relaxation times for our clumps are typically several hundred Myr, with the smallest ones being about 100 Myr. Since it usually takes several relaxation times for a stellar system to dissolve, which is much longer than the time between two snapshots (about 150 Myr), we can conclude that the dissolution of the 3D clumps in our simulations is not due to numerical relaxation.

4.3 Gaseous clumps in line emission

While we mainly focus on clumps of young stars in this paper, angular resolution also affects inferred sizes and masses of analogous clumps of ionized gas, found via their emission lines. For example, studies of H α emission from lensed high-redshift galaxies ($z = 1$ –4) find clumps with sizes ranging from ~ 100 pc to 1 kpc (Jones et al. 2010; Livermore et al. 2012, 2015). The clump luminosity and SFR density increase with redshift. Measuring shifts of the H α provides also kinematic information on the high-redshift clumps (e.g. Mieda et al. 2016).

H α emission traces even younger stars (age less than ~ 10 Myr) than the UV light, because youngest stars contribute most of the ionizing radiation. To make more direct comparison with H α observations, we created mock maps of the H α emission measure. Similar to the UV clumps, we identify H α clumps in both 2D and 3D.

The emission measure is defined as $\int n^2 ds$, where n is the density of ionized hydrogen (here we use ionized hydrogen in cells with $T < 20\,000$ K, which contribute most to the cross-section), and s

is the length in the line of sight. In a simulation output the ionized hydrogen density is given on the adaptive mesh, with some cells on low refinement levels that are larger than our preferred uniform grid of 100 pc. Therefore, we resample all cells in the galaxy region to level 9, which corresponds to physical size 15–25 pc, and then map the gas density on a uniform grid of 100 pc. We calculate the emission measure on this grid and then identify 3D $H\alpha$ clumps using *astrodendro*. The parameters we use here are: `min-pix` = 14 and `min-delta` = 0.4 dex. The parameter `min-value` is chosen using a similar approach of matching the background value as we did for 3D UV clumps. We choose different `min-value` for different runs because these galaxies are at different redshift and have different n_{HII} : `min-value` = $\log n^2/\text{cm}^{-6} = 1$ for SFE50 and SFE100 runs, $\log n^2/\text{cm}^{-6} = 0.2$ for SFE10 run, $\log n^2/\text{cm}^{-6} = 0.5$ for SFE200, and $\log n^2/\text{cm}^{-6} = 0$ for SFEturb run.

To obtain a 2D map of the emission measure, we integrate it in the vertical direction and degrade it to 500 pc patches. Then we identify 2D $H\alpha$ clumps using the following *astrodendro* parameters: `min-npix` = 2, `min-delta` = 0.1 dex, `min-value` = 200 pc cm^{-6} .

The $H\alpha$ clumps do not necessarily correspond to rest-frame UV clumps since they trace star formation on different time-scales. For those UV and $H\alpha$ clumps that match spatially, we show the ratios of $H\alpha$ luminosity in 3D and 2D clumps in the last column of Table 1. For the UV clumps that do not have a corresponding $H\alpha$ clump, we leave dashes in the table.

The first clump in SFE200 run has $L_{3\text{D}, H\alpha}/L_{2\text{D}, H\alpha}$ of only 11 per cent. The emission measure distribution in this galaxy is relatively smooth, thus producing this very small 3D $H\alpha$ clump.

We can compare the 3D/2D ratio of $H\alpha$ luminosities, which trace the SFR over ~ 10 Myr, with the corresponding ratio of SFR averaged over 100 Myr. For many clumps they are close in value, and in these cases the 3D and 2D $H\alpha$ and UV clumps correspond to each other fairly well. Some clumps have large discrepancies, generally because the UV clump and the $H\alpha$ clump differ significantly in geometry.

The $H\alpha$ light is more clumpy than UV light distribution, since it traces younger stellar population. The total fractional luminosity of 3D $H\alpha$ clumps ranges from ~ 10 per cent–95 per cent, while the contribution to $H\alpha$ luminosity from individual 3D clumps ranges from less than 1 per cent to ~ 90 per cent. This large range results from the short time-scale of star formation traced by $H\alpha$ and thus the clumpiness of $H\alpha$ emission. Most individual 3D clumps contribute less than 10 per cent to total $H\alpha$ luminosity, while only one or two largest clumps contribute most of the $H\alpha$ flux. The gas in $H\alpha$ clumps is not bound, with virial parameters (equation 1) larger than 1 by orders of magnitude.

4.4 Origins of giant stellar clumps

Gravitational instability can induce fragmentation of the galactic disc, thus leading to growth of clumps. Stability criterion to linear axisymmetric perturbations for gaseous discs can be described by the Toomre parameter (Safronov 1960; Toomre 1964):

$$Q = \frac{\sigma \kappa}{\pi G \Sigma}, \quad (4)$$

where κ is the epicycle frequency, σ is the velocity dispersion, and Σ is the mass surface density. A disc is unstable if $Q \lesssim 1$. Observations (e.g. Puech 2010; Girard et al. 2018) found that clumpy galaxies are marginally stable, with Toomre $Q \sim 1$. Most $H\alpha$ clumps in high-redshift galaxies are located in regions where Q is low (Genzel et al.

2011; Wisnioski et al. 2012; Mieda et al. 2016). In galaxy formation simulations (Inoue et al. 2016; Oklopčić et al. 2017) gas clumps also coincide with regions of $Q < 1$, albeit with large scatter of Q value. Inoue et al. (2016) also found that stellar clumps coincide with regions of low stellar Q .

In Meng et al. (2019) we studied the distribution of Q in our simulations, accounting for different phases of the gas and stars. Such multicomponent Q more accurately describes the linear stability criterion in realistic galaxies. We found that strong feedback from young stars disperses gas around them, leading to spatial anticorrelation of dense gas and stars up to 50 Myr old. Here we analogously find that locations of 2D UV clumps do not coincide with the regions of low Q . This is because low Q regions closely trace high gas density, while rest-frame UV clumps trace young stars. The average value of Q in 2D UV clump regions ranges from 2.1 to 3.6, which is larger than the median value of $Q = 0.5$ –1.0 weighted by molecular gas mass in Meng et al. (2019). For $H\alpha$ clumps $Q = 1.4$ –3.3 is closer to the gas values, as expected for very young stars. The discrepancy in location of gas and young stars is also seen in the simulations of Oklopčić et al. (2017), where they find that the gas clumps coincide with the location of instantaneous star formation, but not of the SFR averaged over more than 10 Myr.

If clumps form out of instability of a self-gravitating disc, the wavelength of the fastest growing perturbation is

$$\lambda_T = \frac{2\sigma^2}{G\Sigma}. \quad (5)$$

The effective radius of a fully formed clump may not exactly correspond to this scale, but it can still be used as a rough guide. If we take $\lambda_T/2$ for R_{clump} , and approximate the disc rotation velocity profile as flat, then $\kappa \approx \sqrt{2}V_{\text{rot}}/R_{\text{disc}}$, and we obtain

$$\frac{R_{\text{clump}}}{R_{\text{disc}}} \approx \frac{\pi Q}{\sqrt{2}} \left(\frac{\sigma}{V_{\text{rot}}} \right). \quad (6)$$

Here R_{disc} is a characteristic size of the gas distribution. When the gaseous disc is marginally stable, i.e. $Q \sim 1$, we expect a linear relation between $R_{\text{clump}}/R_{\text{disc}}$ and σ/V_{rot} .

Fisher et al. (2017b) found the relation $R_{\text{clump}}/R_{\text{disc}} = (0.38 \pm 0.02)\sigma/V_{\text{rot}}$ in nearby turbulent discs with properties closely resembling $z \sim 2$ star-forming galaxies. They take twice the half-light radius of $H\alpha$ light as R_{disc} and identify clumps in $H\alpha$ + [N II] map. They use flux-weighted $\sigma(H\alpha)$ and the modelled rotation velocity V_{rot} at 2.2 disc scale lengths. Note that the proportionality coefficient 0.38 is a factor of 6 smaller than that expected from the linear perturbation theory above. This already serves as a warning that the observed clump sizes may not be related to global disc instabilities.

We examined whether the 2D or 3D clumps in our simulations obey a similar relation. We take R_{disc} to be twice the half-light radius in our mock $H\alpha$ emission measure map. We use σ_{HII} in the brightest pixel and the circular velocity at galactocentric radius of 10 kpc as σ and V_{rot} . In the 25 snapshots of our five runs at different redshifts, most galaxies have the value of σ/V_{rot} between 0.2 and 1.2. The ratio of sizes $R_{\text{clump}}/R_{\text{disc}}$ for 2D clumps is in the range 0.1 to 0.4; for 3D clumps in the range 0.02 to 0.2. We do not find any correlation between the ratios $R_{\text{clump}}/R_{\text{disc}}$ and σ/V_{rot} , either for 2D or 3D clumps. We also examined this relation for the projected UV clumps and did not find any correlation. This shows that our clump sizes are not set by gravitational instability in thin axisymmetric discs. In fact, if the relation (6) were to hold, we would expect clump sizes of the order of λ_T , which is several kpc (Meng et al. 2019), and therefore significantly larger than the effective radii of our identified clumps. We also calculated the radius $(M/\pi\Sigma)^{1/2}$ corresponding to

the ‘fragmentation mass’ $M_f = 2\lambda_T \Sigma_{c,s}/(\kappa f_g)$ defined by Tamburello et al. (2015) through their equation (10), and found that it still overestimates the clump sizes, by factor of a few.

We also investigated the dependence of clump properties on the specific SFR (sSFR). Shibuya et al. (2016) found that the fraction of high-redshift clumpy galaxies increases with sSFR. Fisher et al. (2017a) found a positive relation between sSFR and the maximum fractional luminosity of low-redshift $H\alpha$ clumps, both blurred and with full resolution (~ 100 pc). Our $H\alpha$ clumps also show a positive correlation between the maximum fractional luminosity and sSFR. For most galaxies with $\text{sSFR} > 2 \times 10^{-10} \text{ yr}^{-1}$, the maximum fractional luminosity of 3D $H\alpha$ clumps is larger than 10 per cent, while the maximum fractional luminosity of 2D $H\alpha$ clumps (resembling their blurred clumps) is larger than 20 per cent. These results agree with those of Fisher et al. (2017a).

4.5 Comparison with other simulation studies

A number of studies have investigated the formation of giant clumps in galaxy formation simulations. Some simulation analyses identify 2D clumps in surface density maps. Oklopčić et al. (2017) identified clumps using gas surface density in a massive galaxy in the FIRE simulation at $z = 1-2$. They use a similar setup to our study: bin the surface density in patches of 50 pc and use *astrodendro* with a minimum of 20 patches. Thus their clump sizes are several hundred pc, with the baryon masses $10^{6.5-9.5} M_\odot$. Moody et al. (2014) identified clumps in projected stellar mass maps and obtained clump masses $\sim 10^{6.5} - 10^9 M_\odot$. Our 2D clump masses are similar to the high-mass end of these clumps, but we miss the low-mass end possibly because of the adopted cut on fractional luminosity.

Other simulation studies identify clumps using 3D density, which is similar to our clump identification in 3D. Mandelker et al. (2017) identified clumps in the 3D distribution of both the cold gas component and stellar component in the VELA simulation galaxies, with halo masses $10^{11} - 10^{12} M_\odot$ at $z = 2$. They identified clumps as connected regions, which is similar to our hierarchical tree construction, containing at least eight cells of $(70 \text{ pc})^3$ each and found clumps of baryon mass $\sim 10^7 - 10^9 M_\odot$. Tamburello et al. (2015) and Mayer et al. (2016) used the *SKID* algorithm to identify bound structures in their simulations. Their typical clump gas masses and stellar masses are $\sim 10^7 - 10^8 M_\odot$. Our 3D clumps are similar to these clumps in mass.

Some simulations investigate the effect of spatial resolution on identified clump masses using maps smoothed to different resolution. Tamburello et al. (2017) identified clumps in mock $H\alpha$ images of their simulations, convolved with the 2D Gaussian aperture with full width at half-maximum $\text{FWHM} = 1 \text{ kpc}$ and 100 pc. They found clump gas mass to be $\sim 10^8 - 10^9 M_\odot$ for $\text{FWHM} = 1 \text{ kpc}$, and $\sim 10^{6.5} - 10^{8.5} M_\odot$ for $\text{FWHM} = 100 \text{ pc}$. The clumps in the full-resolution $H\alpha$ maps are similar to the 3D bound structures found in Tamburello et al. (2015), but clumps found with $\text{FWHM} = 1 \text{ kpc}$ have masses overestimated by an order of magnitude. Our results support this order-of-magnitude overestimation. We also agree on the amount of overestimate of the clump size: Tamburello et al. (2017) find a median intrinsic radius $\sim 120 \text{ pc}$ versus $\sim 800 \text{ pc}$ in the blurred images, while our 3D and 2D median clump sizes are ~ 200 and $\sim 630 \text{ pc}$, respectively. How much clump sizes are overestimated depends largely on the spatial resolution, and this comparison shows that the clump sizes in observations of unlensed galaxies are likely severely overestimated.

Behrendt et al. (2016) also identified clumps in mock $H\alpha$ maps. They convolved the surface density map with a 2D Gaussian of $\text{FWHM} = 1.6 \text{ kpc}$ to mimic the instrumental response, and found clump baryon masses to be $(1.5 - 3) \times 10^9 M_\odot$. Buck et al. (2017) identified clumps in luminosity maps in the NIHAO galaxy sample, using both intrinsic clumps in non-dust-attenuated rest-frame U band and clumps in *HST* bands with dust taken into account. Their clump masses range from a few times $10^6 M_\odot$ to $10^9 M_\odot$, and sizes are $\sim 300-900 \text{ pc}$. The high-mass end is similar to our 2D rest-frame UV clumps in mock observation maps. Benincasa et al. (2019) identified clumps in isolated disc galaxy simulations both in 3D using *SKID* and in 2D gas surface density map using *astrodendro*. They used 2D resolution of 10 and 100 pc, and found only a factor of a few difference between the 2D and 3D clump masses. This result guarantees that our 3D clumps identified on grid of 100 pc are not much overestimated compared to the intrinsic clumps.

5 CONCLUSIONS

We investigated the nature of giant kpc-scale clumps in high-redshift progenitors of the Milky Way-sized galaxies. We identified both 2D clumps in rest-frame UV mock *HST* observation maps with 500 pc resolution and intrinsic clumps in 3D density of young stars with 100 pc resolution (Fig. 1). The 2D clumps are chosen to resemble observed giant clumps. The main results of our comparison are summarized below:

- (i) The masses and sizes of 2D clumps are overestimated due to limited angular resolution and projection of several 3D clumps along the line of sight. The overestimate of mass can be as large as an order of magnitude, and the overestimate of size can be a factor 2–3, compared to the most luminous corresponding 3D clump (Fig. 2 and Table 1). The intrinsic sizes of clumps (150–300 pc) are below the *HST* resolution at $z > 1$, unless the source galaxy is strongly lensed.
- (ii) Most clumps of young stars in our simulated galaxies dissolve on a time-scale shorter than $\sim 150 \text{ Myr}$. The average pair distance between young stars in a clump increases dramatically after 150 Myr (Fig. 3), and the density of the remaining parts of these clumps continues to decrease over time (Fig. 4).
- (iii) Most of the 3D clumps are not gravitationally bound structures, with the virial parameter $\alpha_{\text{vir}} > 1$. However, a few clumps with $\alpha_{\text{vir}} < 1$ are more likely to have a fraction of mass remain bound in the next simulation snapshot, after $\sim 150 \text{ Myr}$.
- (iv) Although total stellar masses of 2D clumps are significantly overestimated, the masses of young stars are only overestimated by factor of a few relative to the most luminous corresponding 3D clump (Fig. 5).
- (v) We created mock images of $H\alpha$ emission measure and compared the sizes of $H\alpha$ clumps with the expectation of the linear theory of global disc instabilities. We do not find the expected correlation between $R_{\text{clump}}/R_{\text{disc}}$ and σ/V_{rot} , primarily because the clump sizes are much smaller than expected. Therefore, we conclude that the observed clumps are not the result of gravitational instabilities in thin axisymmetric discs, at least in Milky Way-sized galaxies.

ACKNOWLEDGEMENTS

We are grateful to Hui Li for producing the simulation suite used in this analysis. We thank the referee for useful comments. This work

was supported in part by the National Science Foundation through grants 1412144 and 1909063.

REFERENCES

- Adamo A., Östlin G., Bastian N., Zackrisson E., Livermore R. C., Guaita L., 2013, *ApJ*, 766, 105
- Behrendt M., Burkert A., Scharmann M., 2016, *ApJ*, 819, L2
- Benincasa S. M., Wadsley J. W., Couchman H. M. P., Pettitt A. R., Tasker E. J., 2019, *MNRAS*, 486, 5022
- Bertoldi F., McKee C. F., 1992, *ApJ*, 395, 140
- Binney J., Tremaine S., 2008, *Galactic Dynamics*. Princeton Univ. Press, Princeton NJ
- Bournaud F. et al., 2014, *ApJ*, 780, 57
- Bournaud F., Daddi E., Weiß A., Renaud F., Mastropietro C., Teyssier R., 2015, *A&A*, 575, A56
- Brown G., Gnedin O. Y., Li H., 2018, *ApJ*, 864, 94
- Buck T., Macciò A. V., Obreja A., Dutton A. A., Domínguez-Tenreiro R., Granato G. L., 2017, *MNRAS*, 468, 3628
- Bullock J. S., Kravtsov A. V., Weinberg D. H., 2001, *ApJ*, 548, 33
- Cava A., Schaerer D., Richard J., Pérez-González P. G., Dessauges-Zavadsky M., Mayer L., Tamburello V., 2018, *Nat. Astron.*, 2, 76
- Ceverino D., Dekel A., Bournaud F., 2010, *MNRAS*, 404, 2151
- Ceverino D., Dekel A., Mandelker N., Bournaud F., Burkert A., Genzel R., Primack J., 2012, *MNRAS*, 420, 3490
- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Dessauges-Zavadsky M., Adamo A., 2018, *MNRAS*, 479, L118
- Dessauges-Zavadsky M. et al., 2017a, *A&A*, 605, A81
- Dessauges-Zavadsky M., Schaerer D., Cava A., Mayer L., Tamburello V., 2017b, *ApJ*, 836, L22
- Elmegreen D. M., Elmegreen B. G., Sheets C. M., 2004, *ApJ*, 603, 74
- Elmegreen D. M., Elmegreen B. G., Rubin D. S., Schaffer M. A., 2005, *ApJ*, 631, 85
- Elmegreen D. M., Elmegreen B. G., Ravindranath S., Coe D. A., 2007, *ApJ*, 658, 763
- Fisher D. B. et al., 2017a, *MNRAS*, 464, 491
- Fisher D. B. et al., 2017b, *ApJ*, 839, L5
- Förster Schreiber N. M. et al., 2006, *ApJ*, 645, 1062
- Förster Schreiber N. M. et al., 2011, *ApJ*, 739, 45
- Genel S. et al., 2012, *ApJ*, 745, 11
- Genzel R. et al., 2008, *ApJ*, 687, 59
- Genzel R. et al., 2011, *ApJ*, 733, 101
- Girard M., Dessauges-Zavadsky M., Schaerer D., Richard J., Nakajima K., Cava A., 2018, *A&A*, 619, A15
- Gnedin N. Y., 2014, *ApJ*, 793, 29
- Gnedin N. Y., Abel T., 2001, *New Astron.*, 6, 437
- Gnedin N. Y., Kravtsov A. V., 2011, *ApJ*, 728, 88
- Guo Y., Giallisco M., Ferguson H. C., Cassata P., Koekemoer A. M., 2012, *ApJ*, 757, 120
- Guo Y. et al., 2015, *ApJ*, 800, 39
- Guo Y. et al., 2018, *ApJ*, 853, 108
- Haardt F., Madau P., 2001, in Neumann D., Durret F., Tran Thanh Van J., eds, *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*. CEA, France, p. 64
- Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, *MNRAS*, 427, 968
- Inoue S., Saitoh T. R., 2012, *MNRAS*, 422, 1902
- Inoue S., Yoshida N., 2018, *MNRAS*, 474, 3466
- Inoue S., Dekel A., Mandelker N., Ceverino D., Bournaud F., Primack J., 2016, *MNRAS*, 456, 2052
- Johnson T. L. et al., 2017, *ApJ*, 843, L21
- Jones T. A., Swinbank A. M., Ellis R. S., Richard J., Stark D. P., 2010, *MNRAS*, 404, 1247
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kennicutt R. C. J., 1998, *ARA&A*, 36, 189
- Kravtsov A. V., 1999, PhD thesis, New Mexico State University
- Kravtsov A. V., 2003, *ApJ*, 590, L1
- Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, *ApJS*, 111, 73
- Kroupa P., 2001, *MNRAS*, 322, 231
- Li H., Gnedin O. Y., Gnedin N. Y., Meng X., Semenov V. A., Kravtsov A. V., 2017, *ApJ*, 834, 69
- Li H., Gnedin O. Y., Gnedin N. Y., 2018, *ApJ*, 861, 107
- Livermore R. C. et al., 2012, *MNRAS*, 427, 688
- Livermore R. C. et al., 2015, *MNRAS*, 450, 1812
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Mandelker N., Dekel A., Ceverino D., Tweed D., Moody C. E., Primack J., 2014, *MNRAS*, 443, 3675
- Mandelker N., Dekel A., Ceverino D., DeGraf C., Guo Y., Primack J., 2017, *MNRAS*, 464, 635
- Martizzi D., Faucher-Giguère C.-A., Quataert E., 2015, *MNRAS*, 450, 504
- Mayer L., Tamburello V., Lupi A., Keller B., Wadsley J., Madau P., 2016, *ApJ*, 830, L13
- Meng X., Gnedin O. Y., Li H., 2019, *MNRAS*, 486, 1574
- Mieda E., Wright S. A., Larkin J. E., Armus L., Juneau S., Salim S., Murray N., 2016, *ApJ*, 831, 78
- Moody C. E., Guo Y., Mandelker N., Ceverino D., Mozena M., Koo D. C., Dekel A., Primack J., 2014, *MNRAS*, 444, 1389
- Murata K. L. et al., 2014, *ApJ*, 786, 15
- Murray N., Quataert E., Thompson T. A., 2010, *ApJ*, 709, 191
- Noguchi M., 1999, *ApJ*, 514, 77
- Oklopčić A., Hopkins P. F., Feldmann R., Kereš D., Faucher-Giguère C.-A., Murray N., 2017, *MNRAS*, 465, 952
- Olmstead A., Rigby J. R., Swinbank M., Veilleux S., 2014, *AJ*, 148, 65
- Overzier R. A., Heckman T. M., Schiminovich D., Basu-Zych A., Gonçalves T., Martin D. C., Rich R. M., 2010, *ApJ*, 710, 979
- Perez J., Valenzuela O., Tissera P. B., Michel-Dansac L., 2013, *MNRAS*, 436, 259
- Puech M., 2010, *MNRAS*, 406, 535
- Puech M., Hammer F., Flores H., Neichel B., Yang Y., 2009, *A&A*, 493, 899
- Ribeiro B. et al., 2017, *A&A*, 608, A16
- Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, *ApJ*, 672, 19
- Rujopakarn W. et al., 2019, *ApJ*, 882, 107
- Safronov V. S., 1960, *Ann. Astrophys.*, 23, 979
- Schmidt W. et al., 2014, *MNRAS*, 440, 3051
- Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2016, *ApJ*, 826, 200
- Shapiro K. L. et al., 2008, *ApJ*, 682, 231
- Shibuya T., Ouchi M., Kubo M., Harikane Y., 2016, *ApJ*, 821, 72
- Soto E. et al., 2017, *ApJ*, 837, 6
- Spitzer L., Jr, 1969, *ApJ*, 158, L139
- Swinbank A. M. et al., 2010a, *MNRAS*, 405, 234
- Swinbank A. M. et al., 2010b, *Nature*, 464, 733
- Tadaki K.-I., Kodama T., Tanaka I., Hayashi M., Koyama Y., Shimakawa R., 2014, *ApJ*, 780, 77
- Tamburello V., Mayer L., Shen S., Wadsley J., 2015, *MNRAS*, 453, 2490
- Tamburello V., Rahmati A., Mayer L., Cava A., Dessauges-Zavadsky M., Schaerer D., 2017, *MNRAS*, 468, 4792
- Toomre A., 1964, *ApJ*, 139, 1217
- Wisnioski E., Glazebrook K., Blake C., Poole G. B., Green A. W., Wyder T., Martin C., 2012, *MNRAS*, 422, 3339
- Wisnioski E. et al., 2015, *ApJ*, 799, 209
- Wuyts S. et al., 2012, *ApJ*, 753, 114
- Wuyts E., Rigby J. R., Gladders M. D., Sharon K., 2014, *ApJ*, 781, 61
- Zanella A. et al., 2015, *Nature*, 521, 54
- Zanella A. et al., 2019, *MNRAS*, 489, 2792

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