Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **468,** 3628–3649 (2017)
Advance Access publication 2017 March 22

doi:10.1093/mnras/stx685

NIHAO XIII: Clumpy discs or clumpy light in high-redshift galaxies?

Tobias Buck,[1,2⋆] Andrea V. Macciò,[1,3] Aura Obreja,[3] Aaron A. Dutton,[3] Rosa Domínguez-Tenreiro[4] and Gian Luigi Granato[5]

[1]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*Member of the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg, IMPRS-HD, Germany*
[3]*New York University Abu Dhabi, PO Box 129188, Saadiyat Island, Abu Dhabi, UAE*
[4]*Departamento de Física Teórica, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[5]*INAF, Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34131 Trieste, Italy*

Accepted 2017 March 16. Received 2017 March 16; in original form 2016 December 15

ABSTRACT

Many massive star-forming disc galaxies in the redshift range 3–0.5 are observed to have a clumpy morphology showing giant clumps of size ∼1 kpc and masses of about 10^7–10^{10} $M_\odot$. The nature and fate of these giant clumps are still under debate. In this work, we use 19 high-resolution simulations of disc galaxies from the Numerical Investigations of Hundred Astrophysical Objects (NIHAO) sample to study the formation and evolution of clumps in the discs of high-redshift galaxies. We use mock *Hubble Space Telescope*-CANDELS observations created with the radiative transfer code GRASIL-3D to carry out, for the first time, a quantitative comparison of the observed fraction of clumpy galaxies and its evolution with redshift with simulations. We find a good agreement between the observed clumpy fraction and the one of the NIHAO galaxies. We find that dust attenuation can suppress intrinsically bright clumps and enhance less luminous ones. In our galaxy sample, we find clumps only in light (*U* band) from young stars but not in stellar mass surface density maps. This means that the NIHAO sample does not show clumpy stellar discs but rather a clumpy light distribution originating from clumpy star formation events. The clumps found in the NIHAO sample match observed age/colour gradients as a function of distance from the galaxy centre, but they show no sign of inward migration. Clumps in our simulations disperse on time-scales of about a hundred Myr and their contribution to bulge growth is negligible.

Key words: methods: numerical – Galaxy: formation – galaxies: bulges – galaxies: evolution – galaxies: high-redshift – galaxies: ISM.

1 INTRODUCTION

The presence of bright clumps in the light distribution of high-redshift ($z \sim 0.5$–3) star-forming galaxies has attracted a lot of attention both from an observational and from a theoretical point of view. Observations have shown that disc galaxies in the early universe have higher gas fractions (Daddi et al. 2010; Tacconi et al. 2010, 2013; Genzel et al. 2015), star formation rates (SFRs; Genzel et al. 2006; Förster Schreiber et al. 2006; Genzel et al. 2008) velocity dispersions (Elmegreen & Elmegreen 2005; Förster Schreiber et al. 2006) and a clumpy morphology (Genzel et al. 2008; Förster Schreiber et al. 2011; Guo et al. 2015) compared to their counterparts at $z = 0$.

Observationally, these clumps are mostly identified in the rest-frame ultraviolet (UV), optical or Hα maps (Elmegreen et al. 2007, 2009; Förster Schreiber et al. 2011; Genzel et al. 2011; Guo et al. 2012; Wuyts et al. 2012; Murata et al. 2014; Tadaki et al. 2014), although sometimes they are also found in the CO line emission of lensed galaxies (Jones et al. 2010; Swinbank et al. 2010). They have sizes of about ∼1 kpc with inferred masses between 10^7and10^9 $M_\odot$. The overall fraction of star-forming galaxies with observed clumps shows a significant evolution with redshift (Guo et al. 2015; Shibuya et al. 2016) and increases up to ∼50 per cent. However, the number of galaxies with observed clumps shows a large uncertainty among different measurements.

Wuyts et al. (2012) found that the clumpy fraction depends sensitively on the selected wavelength to define clumps. They find that the clumpy fraction decreases from ∼75 per cent for clumps selected in 2800 Å images to ∼40 per cent in the *V* band. This suggests a connection between the presence of luminous clumps and recent star formation episodes, as also indicated by the observation that some clumps are seen to be the launching spots of

⋆ E-mail: buck@mpia.de

outflows from the disc (Genzel et al. 2011). However, carefully constructed stellar mass maps of galaxies show no signs of prominent clumps (Wuyts et al. 2012).

Several theoretical studies have focused on an explanation for the formation of these giant star-forming clumps analytically (Dekel, Sari & Ceverino 2009b) either in isolated disc galaxy simulations (Bournaud, Elmegreen & Elmegreen 2007; Bournaud et al. 2008; Bournaud & Elmegreen 2009; Bournaud et al. 2014; Inoue & Saitoh 2014; Tamburello et al. 2015; Mayer et al. 2016) or in cosmological simulations of galaxy formation (Ceverino, Dekel & Bournaud 2010; Ceverino et al. 2012; Genel et al. 2012; Hopkins et al. 2012; Mandelker et al. 2014; Moody et al. 2014; Mandelker et al. 2017).

While some clumpy galaxies might be the result of an ongoing merger (Somerville, Primack & Faber 2001), their overall high fraction cannot be explained by the expected merger rate at these redshifts (Dekel et al. 2009a; Stewart et al. 2009; Hopkins et al. 2010, 2012). Another explanation for the origin of clumps, given the high gas fractions and high gas surface densities in $z > 0$ galaxies, is disc fragmentation via gravitational (disc) instabilities (Toomre 1964). Indeed, clumps are mainly observed to reside in gravitationally unstable regions (Genzel et al. 2011). The formation scenario invoked for clumps is the same as for local giant molecular clouds (MCs): Local collapse and fragmentation happens in regions where the self-gravity of gas and stars is stronger than the internal support by pressure and turbulent motions.

As already mentioned, several groups have studied the dynamics and stability of discs with the aid of high-resolution numerical simulations (Noguchi 1998, 1999; Agertz 2009; Dekel et al. 2009b; Cacciato, Dekel & Genel 2012; Genel et al. 2012; Inoue & Saitoh 2012; Perez et al. 2013; Ceverino et al. 2015; Tamburello et al. 2015). In most of these studies, discs do fragment and break up into large clumps; however, there are two different possible scenarios for the fate of the clumps and their subsequent impact on the global evolution of the host galaxy. If clump collapse is very efficient, it could lead to the formation of gravitationally self-bound, long-lived giant star clusters. These clusters will then lose angular momentum within the disc via gravitational torques and dynamical friction and migrate inwards on time-scales of about 10^8 yr to build up the central bulge (e.g. Bournaud et al. 2014). This picture seems to be supported by observed colour gradients of the clumps within the disc. Clumps closer to the centre of a galaxy show redder colours (Förster Schreiber et al. 2011; Guo et al. 2012; Shibuya et al. 2016), although this trend appears to be weak and might be caused by underlying evolved structures, for example, bulges (van Dokkum et al. 2010; Patel et al. 2013; Morishita et al. 2015; Nelson et al. 2016b).

Conversely, if gas cooling is suppressed, as, for example, in the presence of substantial stellar feedback, the resulting stellar clumps will not be self-bound and will then be quickly dispersed within the disc before experiencing any drag to the centre.

Recent numerical work (Mayer et al. 2016; Oklopčić et al. 2017) has somehow suggested that when robust feedback (needed to reproduce observed galaxy properties at redshift $z = 0$) is invoked, clumps are either short-lived transient features or do not form at all. Thus, a treatment of feedback that is able to prevent overcooling seems to be crucial to understand the origin of clumpy galaxies at high redshift. On the other hand, these studies were either based on a single simulated galaxy (Oklopčić et al. 2017) or were lacking cosmological gas inflow (Mayer et al. 2016), and hence more work on the simulation side is needed.

In this work, we revise the issue of the formation and evolution of luminous clumps in cosmological simulations of disc galaxies at $z = 1$–3 using the Numerical Investigations of Hundred Astrophysical Objects (NIHAO) simulation suite. The NIHAO[1] project is a suite of 100 high-resolution hydrodynamical cosmological simulations. We first compare the 'clumpiness' of our galaxies with real galaxies using an observational motivated clump selection procedure based on the UV luminosity of our objects. For this purpose, we have post-processed all our galaxies with the radiative transfer (RT) code GRASIL-3D (Domínguez-Tenreiro et al. 2014). Subsequently, we look at the evolution of these luminous clumps and their relation with the underlying stellar mass distribution to assess their final fate and their overall impact on galaxy evolution.

2 SIMULATIONS

This work is based on the NIHAO simulation suite (Wang et al. 2015), which contains nearly 100 zoomed cosmological simulations of galaxy formation across a wide range in stellar mass of $10^5 < M_\star/\mathrm{M}_\odot < 10^{11}$. One peculiarity of the NIHAO suite is that the numerical resolution, i.e. the number of elements (particles) representing each galaxy, is kept roughly constant over the whole mass range with of the order of one million particles (gas+stars+dm) in each simulated galaxy.

The NIHAO galaxies have been run using cosmological parameters from Planck Collaboration XVI (2014), namely $\Omega_\mathrm{m} = 0.3175$, $\Omega_\Lambda = 0.6825$, $\Omega_\mathrm{b} = 0.049$, $H_0 = 67.1$ km s^{-1} Mpc^{-1} and $\sigma_8 = 0.8344$. For the subset of simulations used here, the mass resolution is either $m_\mathrm{dark} = 1.735 \times 10^6$ or 2.169×10^5 $\mathrm{M}_\odot$ for dark matter particles and $m_\mathrm{gas} = 3.166 \times 10^5$ or 3.958×10^4 $\mathrm{M}_\odot$ for the gas particles. The corresponding force softenings are $\epsilon_\mathrm{dark} = 931.4$ and 465.7 pc for the dark matter particles and $\epsilon_\mathrm{gas} = 397.9$ and 199.0 pc for the gas and star particles, which is the spatial scale more relevant for this study. However, the smoothing length of the gas particles can be much smaller, for example, as low as $h_\mathrm{smooth} \sim 45$ pc. See Table 1 for a full list of the our galaxies parameters.

NIHAO galaxies have been proved to match remarkably well many of the properties of observed galaxies, for example, results from abundance matching (Wang et al. 2015), metals distribution in the circumgalactic medium (Gutcke et al. 2017), the local velocity function (Macciò et al. 2016) and the properties of stellar and gaseous discs (Dutton et al. 2016; Obreja et al. 2016). Overall, NIHAO simulated galaxies are among the most realistic simulations run in a cosmological context. This means that the NIHAO sample is a perfect testbed to study the occurrence and properties of clumps in high-redshift galaxies.

2.1 Hydrodynamics

The NIHAO simulations were run with a modified version of the smoothed particle hydrodynamics (SPH) solver GASOLINE (Wadsley, Stadel & Quinn 2004) with substantial updates made to the hydrodynamics, as described in Keller et al. (2014). We will refer to this version of GASOLINE as ESF-GASOLINE2.

This modified version of hydrodynamics removes spurious numerical surface tension and improves multiphase mixing by calculating P/ρ^2 as a geometrical average over the particles in the smoothing kernel, as proposed by Ritchie & Thomas (2001). The

[1] Nihao is the chinese word for *hello*.

Table 1. Properties of the high-mass end ($M_{\rm star} > 10^9$ M$_\odot$) of NIHAO galaxies at $z = 1.5$. The first column contains the name of the galaxy, columns 2 to 4 state the number of particles within R_{200} of all species, dark matter and stars. Column 5 shows the total mass within R_{200} of the galaxy, column 6 gives the stellar mass measure within $0.2R_{200}$, column 7 is the value of R_{200}, column 8 states the gas mass within $0.2R_{200}$, column 9 shows the fraction of cold gas (<30 000 K) and the last column shows the SFR measured within $0.2R_{200}$.

Name	N_{200}	$N_{\rm dark}$	$N_{\rm star}$	M_{200}	$M_{\rm star}$	R_{200}	$M_{\rm gas}$	$f_{\rm coldgas}$	SFR
g2.04e11	1122 554	635 540	110 861	1.53×10^{11}	9.01×10^{8}	62.74	5.03×10^{9}	0.77	0.61
g2.41e11	814 971	467 054	55 838	1.13×10^{11}	4.64×10^{8}	57.18	3.35×10^{9}	0.84	0.22
g2.42e11	1282 787	663 987	219 824	1.61×10^{11}	1.92×10^{9}	64.32	7.48×10^{9}	0.75	1.08
g2.57e11	1408 427	714 783	314 763	1.71×10^{11}	2.81×10^{9}	64.89	6.16×10^{9}	0.61	1.99
g3.06e11	1077 199	556 033	139 713	1.36×10^{11}	1.22×10^{9}	60.28	5.08×10^{9}	0.73	0.56
g3.49e11	1219 383	693 736	118 777	1.67×10^{11}	1.00×10^{9}	64.48	4.12×10^{9}	0.70	0.60
g3.71e11	591 465	334 290	58 670	8.06×10^{10}	4.99×10^{8}	51.19	2.95×10^{9}	0.80	0.21
g4.90e11	824 643	454 348	92 880	1.10×10^{11}	8.16×10^{8}	56.47	4.39×10^{9}	0.76	0.68
g5.02e11	1478 107	793 862	56 186	1.97×10^{11}	4.13×10^{8}	70.73	7.18×10^{9}	0.89	0.23
g5.38e11	1920 057	1123 330	121 085	2.75×10^{11}	3.25×10^{9}	76.13	0.98×10^{10}	0.72	2.24
g5.46e11	808 030	454 991	84 139	1.10×10^{11}	7.40×10^{8}	56.17	4.08×10^{9}	0.77	0.54
g5.55e11	814 019	446 562	81 263	1.08×10^{11}	6.81×10^{8}	55.99	3.82×10^{9}	0.74	0.74
g7.55e11	345 668	171 936	53 240	3.38×10^{11}	3.73×10^{9}	81.78	1.71×10^{10}	0.76	2.09
g7.66e11	357 279	180 174	31 052	3.60×10^{11}	1.35×10^{9}	83.61	1.17×10^{10}	0.78	1.93
g8.13e11	836 211	277 144	404 458	5.52×10^{11}	2.93×10^{10}	95.69	1.36×10^{10}	0.27	17.54
g8.26e11	662 237	226 691	307 793	4.49×10^{11}	2.20×10^{10}	89.60	1.47×10^{10}	0.35	10.22
g1.12e12	766 034	294 512	267 954	5.87×10^{11}	1.96×10^{10}	97.76	1.95×10^{10}	0.42	17.64
g1.92e12	2081 091	752 770	906 338	1.49×10^{12}	6.46×10^{10}	133.25	2.24×10^{10}	0.22	38.21
g2.79e12	2207 915	806 995	853 758	1.62×10^{12}	6.10×10^{10}	137.89	4.76×10^{10}	0.37	40.42

Saitoh & Makino (2009) time-step limiter was implemented so that cool particles behave correctly when a hot blastwave hits them, and ESF-GASOLINE2 uses the Wendland C2 smoothing kernel (Dehnen & Aly 2012) to avoid pairing instabilities. The treatment of artificial viscosity has been modified to use the signal velocity, as described in Price (2008), and the number of neighbour particles used in the calculation of the smoothed hydrodynamic properties was increased from 32 to 50.

Cooling via hydrogen, helium and various metal lines is included, as described in Shen, Wadsley & Stinson (2010), and was calculated using CLOUDY (version 07.02; Ferland et al. 1998). These calculations include photoionization and heating from the Haardt & Madau (2005) UV background and Compton cooling in a temperature range from 10 to 10^9 K. Finally, we adopted a metal diffusion algorithm between particles, as described in Wadsley, Veeravalli & Couchman (2008).

2.2 Star formation and feedback

In the fiducial NIHAO runs, gas is eligible to form stars according to the Kennicutt–Schmidt Law when it satisfies a temperature and density threshold. The simulations employ the star formation recipe, as described in Stinson et al. (2006), which is summarized below. Stars form from cool (T <15 000K), dense gas ($n_{\rm th} > 10.3$ cm^{-3}). The threshold number density $n_{\rm th}$ is set to the maximum density at which gravitational instabilities can be resolved in the simulation: $n_{\rm th} = 50 m_{\rm gas}/\epsilon^3_{\rm gas} = 10.3$ cm^{-3}, where $m_{\rm gas}$ denotes the gas particle mass and $\epsilon_{\rm gas}$ denotes the gravitational softening of the gas. The value of 50 denotes the number of neighbouring particles. The gas fulfilling these requirements will then be converted into stars according to the following equation:

$$\frac{\Delta M_{\rm star}}{\Delta t} = c_{\rm star} \frac{M_{\rm gas}}{t_{\rm dyn}}, \quad (1)$$

where $\Delta M_{\rm star}$ is the mass of the star particle formed, $M_{\rm gas}$ is the gas particle's mass, Δt is the time-step between star formation events (here 8×10^5 yr) and $t_{\rm dyn}$ is the gas particle dynamical time. $c_{\rm star}$ is the star formation efficiency, i.e. the fraction of gas that will be converted into stars during the time $t_{\rm dyn}$ and is taken to be $c_{\rm star} = 0.1$.

Stellar feedback is implemented in two modes, as described in Stinson et al. (2013). The first mode accounts for the energy input from stellar winds and photoionization from luminous young stars. This pre-supernova (pre-SN) feedback, which was dubbed early stellar feedback (ESF) in Stinson et al. (2013), happens before any SNe explode and consists of 10 per cent of the total stellar flux, 2×10^{50} erg of thermal energy per M$_\odot$ of the entire stellar population. The efficiency parameter for the coupling of the energy input is set to $\epsilon_{\rm ESF} = 13$ per cent. And, for the pre-SN feedback, unlike the SN feedback, the radiative cooling is left on.

The second mode accounts for the energy input via SNe and starts 4 Myr after the formation of the star particle, and it is implemented using the blastwave formalism described in Stinson et al. (2006). In this approach, SNe input energy (i.e. thermal feedback) into the interstellar gas surrounding the region where they formed, but since the gas receiving the energy is dense, it would quickly be radiated away due to its efficient cooling. For this reason, cooling is delayed for particles inside the blast region for ∼30 Myr. See Stinson et al. (2013) for an extended feedback parameter search.

3 THE GALAXY SAMPLE

In order to match the stellar range of the observed galaxies at high redshift, we have selected from the NIHAO sample all the galaxies with stellar masses larger than $M_{\rm star} > 10^9$ M$_\odot$ at redshift $z = 1.5$, obtaining a final sample of 19 galaxies. From these galaxies, we use all the snapshots in the redshift range $3 > z > 0.25$, where the stellar mass of the galaxy is larger than the above threshold mass. This selection criterion leaves us with 203 snapshots in the redshift range $3 > z > 1$, 155 in the redshift range $1 > z > 0.5$ and 130 in the redshift range $0.5 > z > 0.25$.

For this work, the virial mass, M_{200}, of each halo is defined as the mass of all particles within a sphere containing $\Delta = 200$ times

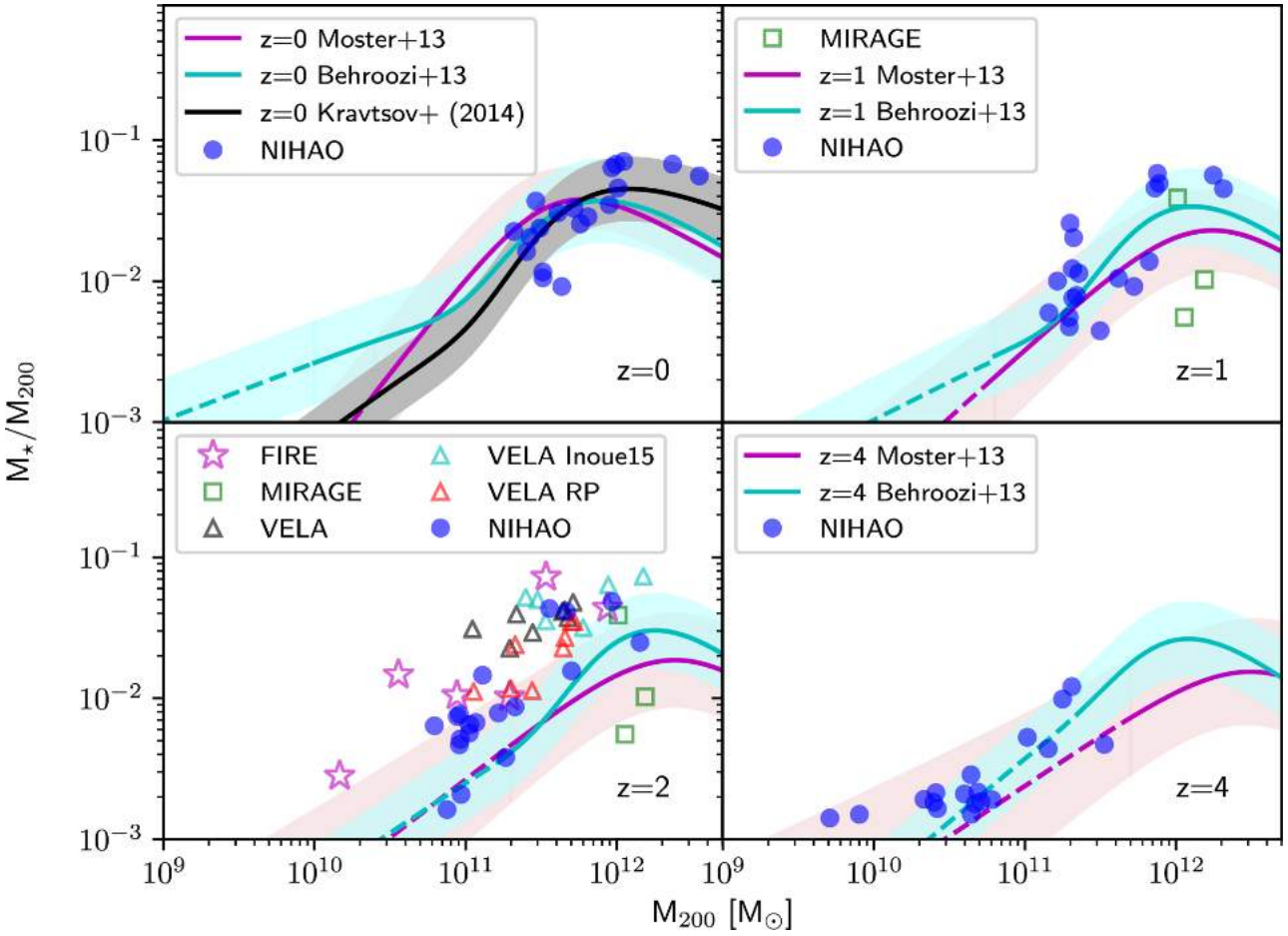

Figure 1. Evolution of the stellar mass versus halo mass relation for the selection of hosts. Redshift $z = 0$ (top left-hand panel), 1 (top right-hand panel), 2 (bottom left-hand panel) and 4 (bottom right-hand panel). For all redshifts shown, the simulations agree well with constraints from halo abundance matching (Behroozi, Wechsler & Conroy 2013; Moster, Naab & White 2013; Kravtsov, Vikhlinin & Meshscheryakov 2014). Where the relations are extrapolated into mass scales without observational constraints, the lines are shown in dashed line style.

the cosmic critical matter density, $\rho_{\rm crit}$. The virial radius, R_{200}, is defined accordingly as the radius of this sphere. The haloes in the zoom-in simulations were identified using the MPI+OpenMP hybrid halo finder AHF2 (Gill et al. 2004; Knollmann & Knebe 2009). The stellar mass $M_{\rm star}$ and the gas mass $M_{\rm gas}$ are measured within a sphere of radius, $r_{\rm gal} \equiv 0.2R_{200}$. The SFR is measured as the mass of stars formed inside $r_{\rm gal}$ over the preceding Gyr and the gas fraction $f_{\rm gas}$ is defined as the fraction of cold gas ($T < 3 \times 10^4$ K) over the total baryonic mass within $r_{\rm gal}$. The main parameters of the 19 NIHAO galaxies are listed in Table 1.

As already mentioned before, the strength of the (stellar) feedback has a strong impact in promoting or suppressing the formation of stellar clumps. The efficiency of the feedback can be tested by looking at how realistic the properties of simulated galaxies are with respect to the observed ones.

3.1 Stellar mass–halo mass relation

Fig. 1 shows the evolution of the stellar mass versus halo mass relation since redshift $z = 4$ (a lookback time of ∼12 Gyr); NIHAO simulations (blue circles) show a very good agreement with the abundance-matching relations from Behroozi et al. (2013) and Moster et al. (2013). At redshifts 2 and 1, some of the higher mass galaxies show about a factor of 2 to many stars, but this is consistent with, for example, the FIRE simulation (Hopkins et al. 2014). Some of this discrepancy with respect to abundance-matching results might be due to systematic uncertainties in the form of the stellar initial mass function (IMF; e.g. Conroy & van Dokkum 2012; Dutton et al. 2013a,b). Thus, for massive galaxies ($M_\star \gtrsim 10^{11}$ M$_\odot$), the stellar masses may be underestimated by a factor of 2 when assuming a Milky Way IMF.

Comparison to the VELA galaxies (Moody et al. 2014; Ceverino et al. 2015; Inoue et al. 2016) (coloured triangles in Fig. 1) shows that these galaxies tend to substantially overproduce stellar masses at redshift 2. Inclusion of radiation pressure as an additional source of feedback in these galaxies (VELARP) brings them in better agreement with the abundance-matching relation. This in comparison to the results from the FIRE simulation and from NIHAO indicates that inclusion of some sort of feedback prior to SN (Ceverino: radiation pressure; Hopkins: radiation pressure, stellar winds and photoionization; NIHAO: strong photoionization included as thermal energy) is needed to reproduce the stellar mass–halo mass relation at high redshifts.

It is worth noting that the VELA galaxies were not run down to redshift 0, so there is no information whether these simulations do or do not provide realistic present-day galaxies. For reference, we also include the isolated simulations by Perret et al. (2014) taken from the MIRAGE sample.

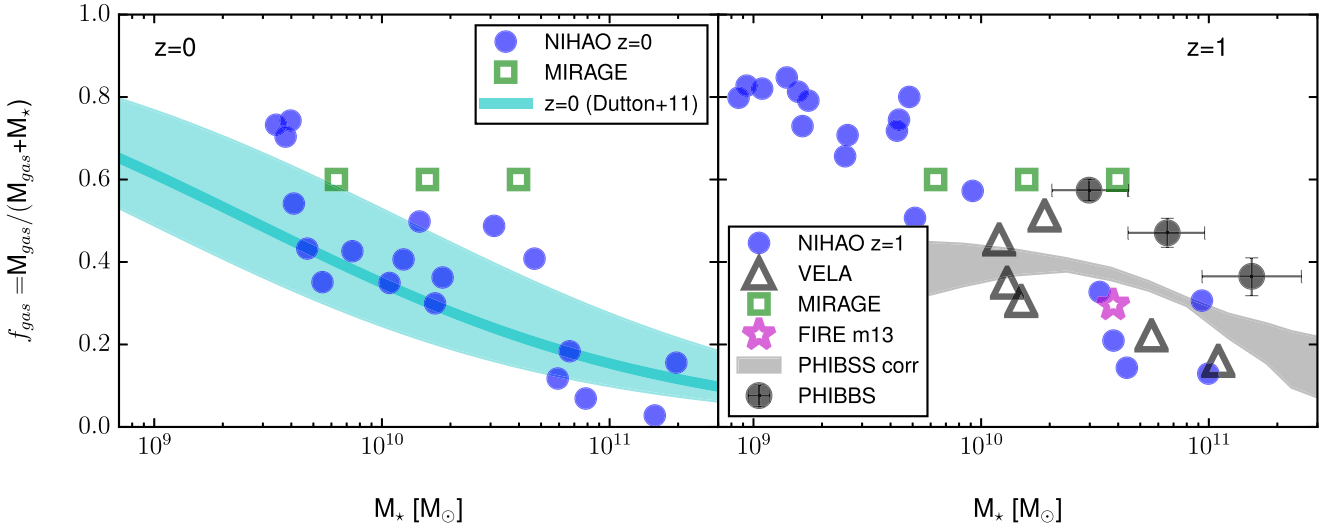

Figure 2. The galaxy gas fraction as a function of stellar mass at redshift $z = 0$ (left-hand panel) and 1 (right-hand panel). The selected NIHAO galaxies are shown with blue points and the relation from Dutton et al. (2011) for redshift 0 galaxies is shown with a cyan line with a shaded region indicating the scatter. At redshift ∼1.5, the PHIBBS survey data (Tacconi et al. 2013) are shown as black dots. The incompleteness corrected relation from the PHIBBS survey is shown as the grey shaded area.

3.2 Gas fractions

Together with the total stellar mass, another important quantity in determining the stability of a disc in a galaxy is the fraction of cold gas.

In Fig. 2, we show the gas fraction defined as $f_{\rm gas} = M_{\rm gas}/(M_\star + M_{\rm gas})$ as a function of stellar mass compared to observations at redshift 0 from Dutton et al. (2011) (left-hand panel) and observations from the PHIBBS survey (Tacconi et al. 2013) for galaxies at redshifts 1–1.5 (right-hand panel). The selected NIHAO galaxies nicely follow the trend observed by Dutton et al. (2011).

The discrepancy at these lower masses might be due to different measurement methods in the simulations and the observations. In the simulations, we use all the gas with a temperature smaller than 3×10^4 K in a sphere of radius $r_{\rm gal} = 0.2R_{200}$, while the observations measure atomic and molecular gas with a correction factor for helium (Dutton et al. 2011).

As was shown in Stinson et al. (2015), a simple temperature cut overestimates the amount of neutral gas, especially at lower masses. The galaxies are also broadly in agreement with the incompleteness-corrected measurements from the PHIBBS survey (grey band) for galaxies in the redshift range $z = 1$–1.5. For reference, we also show the gas fractions of the VELA galaxies (Moody et al. 2014; Ceverino et al. 2015; Inoue et al. 2016), the FIRE galaxy from Oklopčić et al. (2017) and the MIRAGE sample (Perret et al. 2014). The VELA sample and the FIRE galaxy are well in agreement with the NIHAO sample as well as the observed gas fractions. The two higher mass MIRAGE galaxies show very high gas fractions of 60 per cent in slight tension with the completeness-corrected PHIBBS observations. Although galaxies with such high gas fractions are observed, they are likely not typical.

4 RADIATIVE TRANSFER

In order to compare the fraction of galaxies with clumps in the simulations and in the CANDELS galaxy sample, we post-process our simulations with the RT code GRASIL-3D (Domínguez-Tenreiro et al. 2014). The post-processing step ensures that we account for the effects of dust attenuation and cosmological redshift, thus allowing for a meaningful comparison between simulations and observations.

GRASIL-3D is a three-dimensional (3D) RT code designed to be used with the outputs of hydrodynamical simulations. The code solves the RT equation for dusty media on a regular grid. The treatment of dust is based on the formalism of the GRASIL model (Silva et al. 1998; Granato et al. 2000), which has been successfully used with semi-analytical models of galaxy formation. The key feature of this dust model is that it does a detailed non-equilibrium calculation for polycyclic aromatic hydrocarbon molecules and dust grains smaller than 150 Å, thus allowing for a proper description of the cirrus emission in the mid-infrared (IR; Guhathakurta & Draine 1989).

An important point to remember is that any RT post-processing of simulated galaxies introduces a few more sub-grid parameters on top of those already included in the hydrodynamic codes. In the case of GRASIL-3D, these parameters are particularly related to the properties of the MCs. Since short-lived massive stars are spatially associated with MCs, it is well established that much of the dust reprocessing of stellar light occurs inside MCs. In the case of most cosmological simulations, these small scales are not resolved, and as such, dust reprocessing has to be modelled. In GRASIL-3D, the interstellar medium is split into MCs and diffuse cirrus, by assuming the densities of unit hydrodynamical gas masses to follow a lognormal probability distribution function with a mean equal to the local gas density. The dispersion of the distribution is a free parameter. In this manner, the fraction of the gas mass above a certain density threshold gives the MC contribution, while the rest is considered cirrus. Thus, the first two parameters GRASIL-3D needs are the MC density threshold and the dispersion of the density distribution function.

In GRASIL-3D, the dust reprocessing of stellar populations is age-dependent, similar to the implementation in GRASIL, which was the first model to take age into account. Practically, stars younger than a certain age, t_0, radiate all their energy inside MCs, while stars older than $2t_0$ have already dispersed their MC cocoons. In the intermediate-age regime, the fraction of energy dumped inside the MC is a decreasing function of stellar age. The last parameter the code needs is the spatial extent of MCs. Once these sub-grid

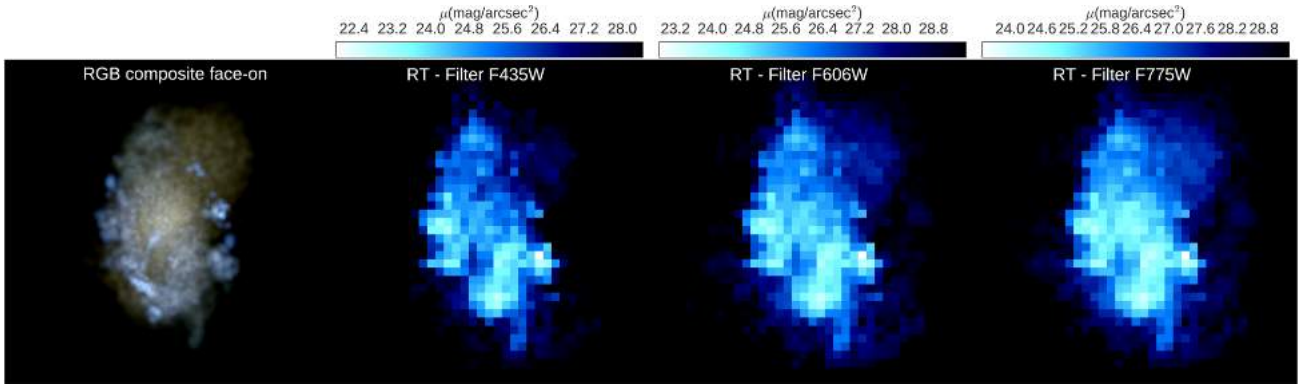

Figure 3. RGB image (left-hand panel) and RT maps in different filters. The second panel shows the map of the *F435W* filter, the third panel shows the *F606W* image and the rightmost panel shows the *F775W* filter. Image sizes are 20×20 kpc^2.

parameters are set, GRASIL-3D solves the RT equation by treating separately the light reprocessing in the dense and diffuse interstellar medium.

Finally, the stellar particle luminosities are computed according to the SSP models of Bruzual & Charlot (2003) (which is the only option available in the current version of GRASIL-3D) assuming a Chabrier IMF (Chabrier 2003).

The code has already been used to study the star formation main sequence of simulated galaxies (Obreja et al. 2014), the properties of high-redshift clusters and protoclusters in sub-mm and IR (Granato et al. 2015), and the correlations between IR fluxes of $z = 0$ simulated galaxies and their baryonic content (Goz et al. 2016). These studies have shown that GRASIL-3D reproduces observables (e.g. broad-band fluxes, spectral energy distributions) when used with realistic galaxy simulations.

In order to run GRASIL-3D on the simulated NIHAO galaxies, we need to specify the four parameters discussed above. We chose the fiducial values for disc galaxies of Domínguez-Tenreiro et al. (2014): 14 pc for the MC radius, 5 Myr for the MC destruction time-scale, 3.3×10^9 M$_\odot$ kpc^{-3} for the MC threshold density and a dispersion of the lognormal gas density probability distribution function of 3. We do not perform any GRASIL-3D parameter study in this work, given both the high computational cost implied and the fact that the above-mentioned values have been shown to reproduce normal star-forming galaxies (e.g. Silva et al. 1998; Goz et al. 2016). Finally, we chose pixel sizes corresponding to the resolution of the *Hubble Space Telescope* (*HST*) at the given redshift of the snapshots (0.06 arcsec).

After running GRASIL-3D on all the snapshots of our galaxy sample in the redshift range ∼0.25–3, we apply the same *HST* filter selection with the same magnitude cuts and surface brightness limits as chosen in Guo et al. (2015). For snapshots in the redshift range $3 > z > 2$, we select the *F775W* filter; for $2 > z > 1$, we select the *F606W* filter; and for $z < 1$, we select the *F435W* filter to detect clumps. The outcome of the RT calculations results in mock observations of our simulations closely matching the ones of CANDELS galaxies used by Guo et al. (2015). A selection of images for the galaxy g7.55e11 at redshift $z \sim 1.3$ in all three filters used are shown in Fig. 3 in comparison to a red giant branch (RGB) map of the stellar luminosity (left-hand panel). For the RGB map we calculate the stellar luminosity of a star particle, given its age and metallicity in three different bands (see the next section for a more detailed description). The RT images show the luminosity maps of the same galaxy in the three filters used for the analysis (from the left-hand to right-hand side: *F435W*, *F606W* and *F775W*). As described above, for the given snapshot time of $z \sim 1.3$, we would use for the clump analysis the *F606W* image, while for higher redshifts, we would use the *F435W* filter, and for lower redshifts, we would use the *F775W* one.

5 CLUMP DETECTION

In this study, we perform an observationally motivated clump selection, while other theoretical works on this subject focused on identifying clumps as regions of high surface density of gas or stars (Genel et al. 2012; Ceverino et al. 2015; Tamburello et al. 2015; Inoue et al. 2016; Mayer et al. 2016; Oklopčić et al. 2017)

Observed clumps are mostly detected as UV-bright or Hα-bright clumps, and we thus decided to select clumps in the luminosity maps of our galaxies with and without RT post-processing. We will refer to clumps detected in non-dust-attenuated images as *intrinsic* clumps, while we will call clumps in the RT-processed images, i.e. *observed* clumps. This allows us to do a proper comparison with both observations and previous theoretical studies.

5.1 Intrinsic clump selection

For every galaxy and every snapshot in the redshift range $z = 0.25$–3, we create UV-light images by calculating the UV luminosity of every star particle, given its age, metallicity and its IMF under the assumption that these particles represent simple stellar populations (SSPs). We use the PYNBODY package[2] to perform these calculations. This package includes a grid of SSP luminosities for different stellar ages and metallicities in several bandpasses. The grids are calculated using Padova SSPs from Girardi[3] (Marigo et al. 2008; Girardi et al. 2010).

Our clump-finding procedure is similar to the one used by Oklopčić et al. (2017) and we will briefly describe it here. For all snapshots, we first rotate the galaxies face-on, using the total angular momentum of the stars within a sphere of 10 kpc around the centre of the halo. We then select all star particles in a cylinder of radius 10 kpc and height of 6 kpc centred on the galaxy and then construct the luminosity maps by binning the particle positions on to a 2D grid of bin size 100 pc.

We sum up all luminosities of the star particles in one cell to get its total luminosity. After that, we smooth the luminosity map by

[2] https://pynbody.github.io/pynbody/
[3] http://stev.oapd.inaf.it/cgi-bin/cmd

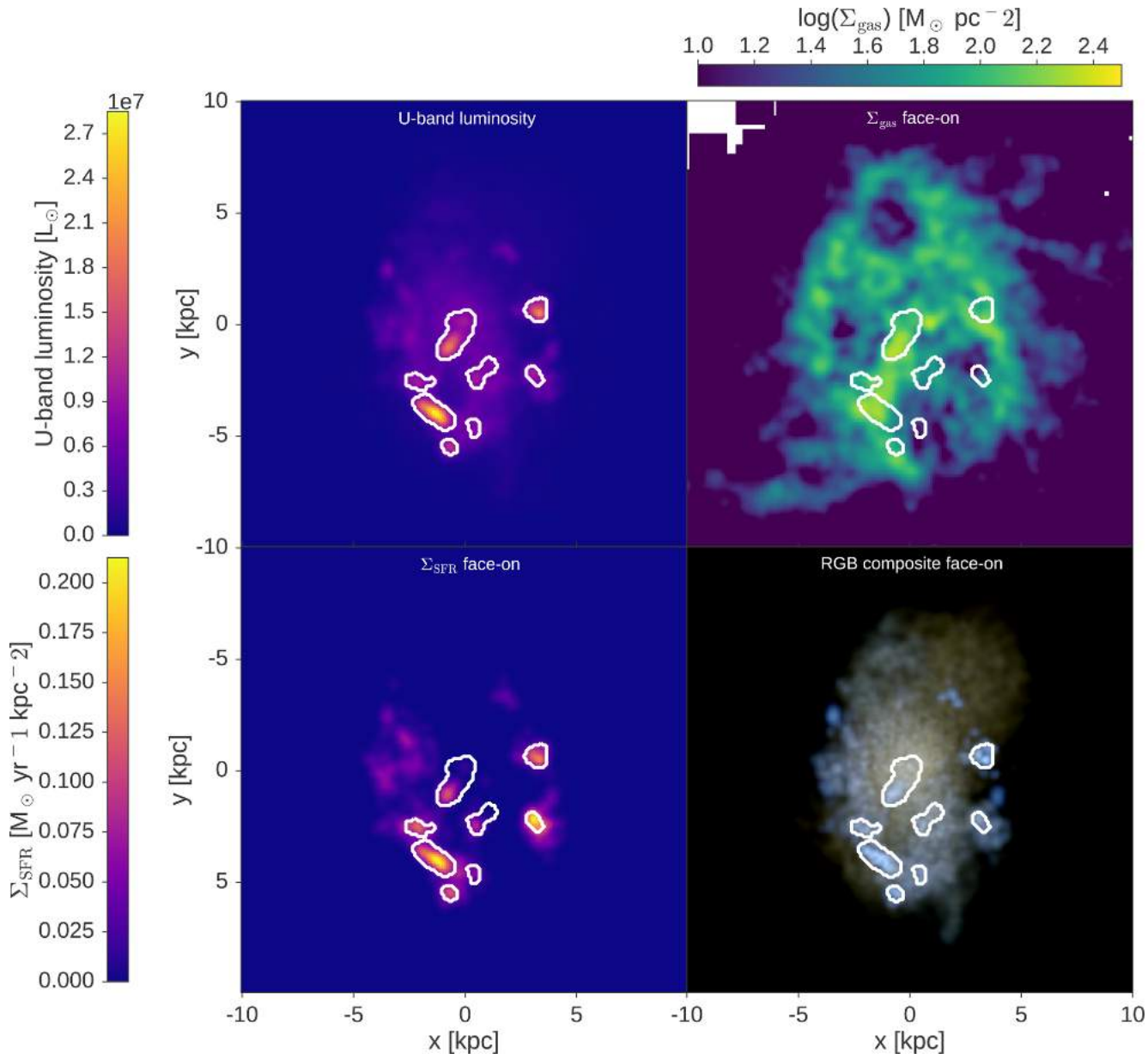

Figure 4. Face-on maps of intrinsic *U*-band luminosity (upper left-hand panel), SFR surface density (lower left-hand panel), cold gas surface density (upper right-hand panel) and RGB stellar composite (lower right-hand panel). The white contours show the clumps selected in the intrinsic *U*-band luminosity map.

convolving it with a Gaussian filter of 200 pc standard deviation (FWHM $\sim$ 470 pc, where FWHM stands for full width at half-maximum) to account for the particle's softening. We checked that taking the softening into account before binning the data on to the grid does not make a significant difference in terms of clump detection.

Using the procedure described above, we can also create maps of different quantities, for example, the surface density maps of gas can be calculated by summing up all the mass in one bin and then dividing the result by the surface area of the bin. Fig. 4 shows the *U*-band luminosity map (upper left-hand panel), the SFR surface density map (lower left-hand panel) with the SFR calculated as the stellar mass formed over the last 500 Myr, surface density of cold gas (upper right-hand panel) and an RGB stellar composite image of the stars (lower right-hand panel), where RGB stands for the colours used to render the stars (old stars are rendered with a red colour, intermediate-aged stars with green and young stars with blue). As expected, intrinsic clumps found in the *U*-band luminosity agree very well with clumps in the SFR surface density. Once we have created a map of a given quantity, we further calculate the mean value of all non-empty bins and the corresponding standard deviation. We then use the PYTHON package ASTRODENDRO[4] to find overdensities in the maps. The ASTRODENDRO package calculates hierarchical trees of structures, so-called dendrograms, in the maps. There are three parameters needed by the package to calculate the tree: (i) a threshold value to define the clump, which we set to three standard deviations above the mean value (we checked that our results are robust if we altered this value, see Fig. A1); (ii) the minimum difference between two close structures to count them as

[4] http://www.dendrograms.org/

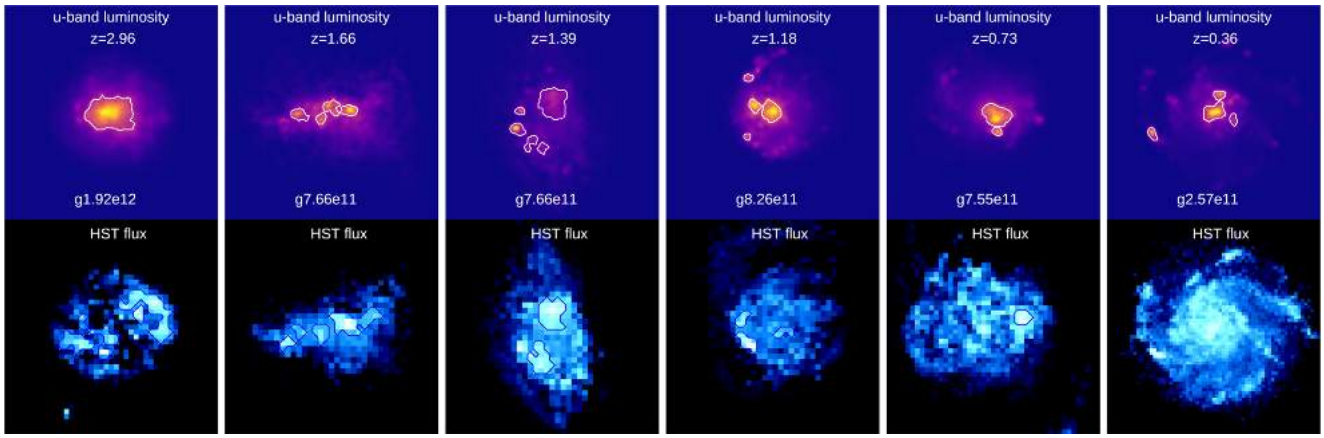

Figure 5. Comparison between clumps in the rest-frame *U*-band luminosity maps (upper panels) and clumps in the RT images (lower panels). For the RT images, the galaxies were redshifted to the proper redshift of the snapshot and observed in the same filter bands as the observations by Guo et al. (2015), accounting for dust absorption, scattering and cosmological redshift. For galaxies in the redshift range $3 < z < 2$, we selected the *F775W* filter; for $2 < z < 1$, we selected *F606W*; and for $z < 1$, we selected *F435W*, where the redshift of each panel is indicated in the *U*-band maps.

separate clumps, set to 10 per cent following Oklopčić et al. (2017); and (iii) the minimum number of pixels within a clump, which is set to be 30 pixels, resulting in an effective radius of ∼300 pc, which is consistent with the gravitational softening of the gas and star particles of our simulations.

An example of the outcome of this clump-finding algorithm is shown in Fig. 4, where *U*-band-selected clumps are overplotted on all four face-on images with white contours.

The ASTRODENDRO package already comes with tools to measure the size (surface area) of each clump in the plane of the disc (*x*–*y* plane). Given the surface area, A, of each clump, we can calculate an effective radius of the clump, $R^2 = A/\pi$. We follow the assumption by Oklopčić et al. (2017) and take the extent of each clump perpendicular to the plane of the disc as equal to $2R$ centred on the densest part of the clump. All stellar and gas particles falling into this volume of space are counted as belonging to the clump, and clump properties such as luminosity, mass, etc., can be calculated from these particles. We checked that the enclosed mass does not depend strongly on the exact choice of the clump's vertical extent, as long as it is of the order of the disc scaleheight. Most of the snapshots show one large clump in the centre of the galaxy, which can be matched to the bulge component. Thus, we excluded from the search area the innermost 1 kpc around the centre of the galaxy.

5.2 Clump selection in RT images

While the above clump selection is useful to study physical properties of clumps, a comparison to observed galaxies is difficult due to missing dust attenuation and cosmological redshift. Therefore, we look for clumps directly in the dust-attenuated RT maps computed as explained in Section 4.

In the RT images, we adjust our clump finder to better match the observational clump selection, as described in Guo et al. (2015, section 3). We first calculate the background mean after applying a 3σ clipping and then we select clumps as local maxima that are at least 3σ above the mean. Again, we checked how results change if this threshold is changed to 2σ or 4σ (see Fig. A2). Because we adjusted the pixel size of the RT images to match the resolution of *HST*, we now set the minimum number of pixels to 5 per clump following Guo et al. (2015), and we neglect clumps within the inner 1 kpc around the galaxy centre.

Fig. 5 shows that the RT images recover some of the clumps found in the *U*-band images but not all. Sometimes, the dust attenuation and the cosmological redshift dim down clumps such that they fall below the detection threshold of 3σ above the mean (e.g. fourth panel of Fig. 5). On the other hand, a non-clumpy galaxy in the *U* band shows clumps in the RT image due to a non-uniform gas distribution (e.g. first and fifth panels of Fig. 5). This underlines the importance of using RT images to compare simulation and observations quantitatively.

6 RESULTS

6.1 The *observed* NIHAO clumpy fraction

As a first test, we want to check how the clumpiness of the NIHAO galaxies compares with observed galaxies, to assess if our galaxies have a realistic number of structures in their light distribution.

In order to compare our simulations with the observations used by Guo et al. (2015), we constructed *HST* mock images with GRASIL-3D (Domínguez-Tenreiro et al. 2014) for all snapshots in the redshift range $0.3 < z < 3$ of our 19 NIHAO galaxies. We then apply the same filters (*F435W*, *F606W* and *F775W*) and surface brightness cuts as used by Guo et al. (2015) to our mock images and look for clumps as described in Section 5.2.

In Fig. 6, we compare the fraction of simulated galaxies with at least one off-centre clump with that of observations. The left-hand panel shows the comparison between the complete simulation sample (blue) and different observations. Overall, the clumpy fraction in NIHAO galaxies follows quite well the clumpy fraction derived by Shibuya et al. (2016) (purple dotted line) and agrees within the error bars with the results of Guo et al. (2015) (black short–long dashed line). The right-hand panel shows the results when galaxies are separated into three stellar mass bins $\log(M_*/\mathrm{M}_\odot) < 9.8$, blue; $9.8 < \log(M_*/\mathrm{M}_\odot) < 10.6$, green; and $\log(M_*/\mathrm{M}_\odot) > 10.6$, red.

There is a peak in the clumpy fraction of about 60 per cent–70 per cent at $z = 1.5$–2 in both simulations and observations. At higher and lower redshifts, the clumpy fraction goes down to 50 per cent at $z \sim 3$ and to 40 per cent at $z \sim 0.5$, dropping even further in the simulations to 20 per cent at $z \sim 0.25$, thus matching observations by Murata et al. (2014).

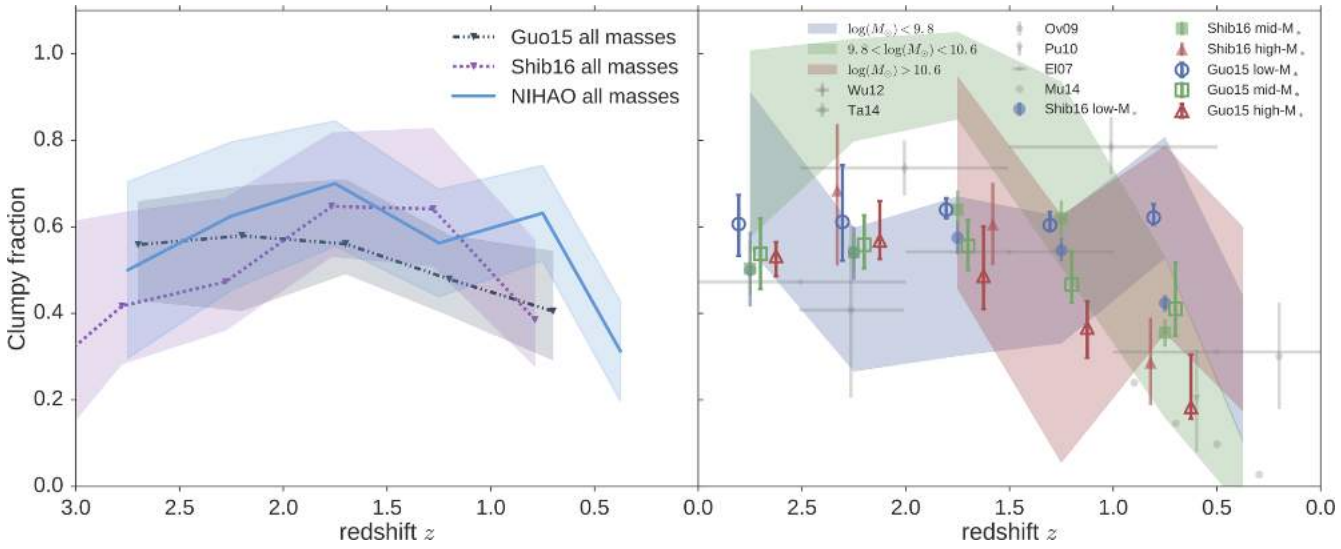

Figure 6. Evolution of the fraction of galaxies with at least one *observed* off-centre clump. The left-hand panel shows the evolution of the clumpy fraction for the whole NIHAO sample (blue line) compared to observations from Guo et al. (2015) (black dash–dotted line) and Shibuya et al. (2016) (purple dotted line). The shaded band shows the 1σ scatter. The right-hand panel shows the evolution of the clumpy fraction split into different mass bins. Coloured bands show the result for our simulations, coloured open symbols show the corresponding values from Guo et al. (2015) and coloured filled symbols show the results from Shibuya et al. (2016). Grey open symbols show results from other observational studies: diamonds are from Wuyts et al. (2012), squares are from Tadaki et al. (2014), pentagons are from Overzier et al. (2009), inverse triangles are from Puech (2010), stars are from Elmegreen et al. (2007) and dots are from Murata et al. (2014).

When inspecting the different mass bins separately, we still find a quite good agreement between simulations and observations. Specifically, we see that at low redshifts, the two less massive bins (blue and green bands and points) agree very well with observed clumpy fractions for the same mass ranges. However, the highest mass galaxies (red band) in NIHAO show a clumpy fraction slightly too high but still in agreement with the observations within their error bars. At intermediate redshifts, the high-mass and low-mass bins agree again well with the observations, but the intermediate-mass bin (green band) shows a too high clumpy fraction, although in agreement with measurements from Wuyts et al. (2012). Finally, due to the selection function of the NIHAO galaxies, we do not have any data for the highest mass bin; however, the other two mass bins agree well with the observed clumpy fraction.

Another observable to which we can compare our RT calculations is the fractional contribution of the clump's UV flux to the total galaxy UV flux. Following Guo et al. (2015), we call this quantity $C_{\rm UV}$. In Figs 7 and 8, we show $C_{\rm UV}$ as a function of the clump size and of the galaxy mass, respectively.

Clump sizes found in the RT calculations of the NIHAO sample are in agreement with observed clump sizes from Förster Schreiber et al. (2011) and with intrinsic clump sizes if the pixel scale is matched to the *HST* pixel scale (see Appendix B).

Fig. 8 shows the comparison between the observed $C_{\rm UV}$ as a function of stellar mass and the NIHAO one in three different redshift bins: $0.5 < z < 1$, $1 < z < 2$ and $2 < z < 3$. As in Guo et al. (2015), the clump contribution to the total UV flux of the galaxy is calculated for all galaxies, not only the clumpy ones. In the two lower redshift bins, simulations (orange points) agree well with observations (black squares) within their error bars, while the contribution of UV light from clumps of intermediate-mass galaxies in the highest redshift bin is too high in NIHAO. A partial explanation of this excess could be related to the difference in the sample size in this redshift bin: ∼70 simulation snapshots versus several thousand galaxies in CANDELS. Furthermore, we did not add noise to our images, which might lead to recovering more clumps than the observers would do. Despite this discrepancy, NIHAO recovers quite well the observed UV light fraction of clumps.

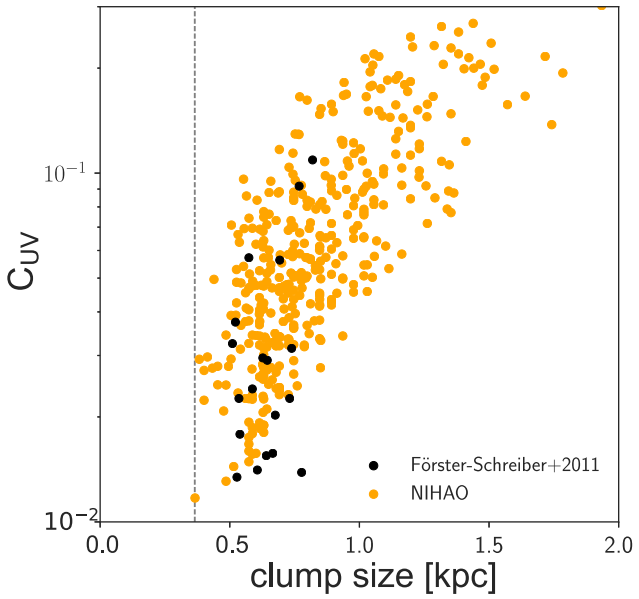

Figure 7. Clump contribution to the rest-frame dust-attenuated UV light of galaxies as a function of clump size for the NIHAO sample (orange dots) and observations from Förster Schreiber et al. (2011) (black dots). The grey dashed line indicates the lower limit on clump sizes set by our clump-finding procedure.

Mock observations of simulated clumpy galaxies have already been studied before using different methods (see e.g. Genel et al. 2012; Moody et al. 2014; Tamburello et al. 2016), and similar results to ours were found. Genel et al. (2012) converted the SFR of one of their zoom-in cosmological galaxies (s224) into Hα flux,

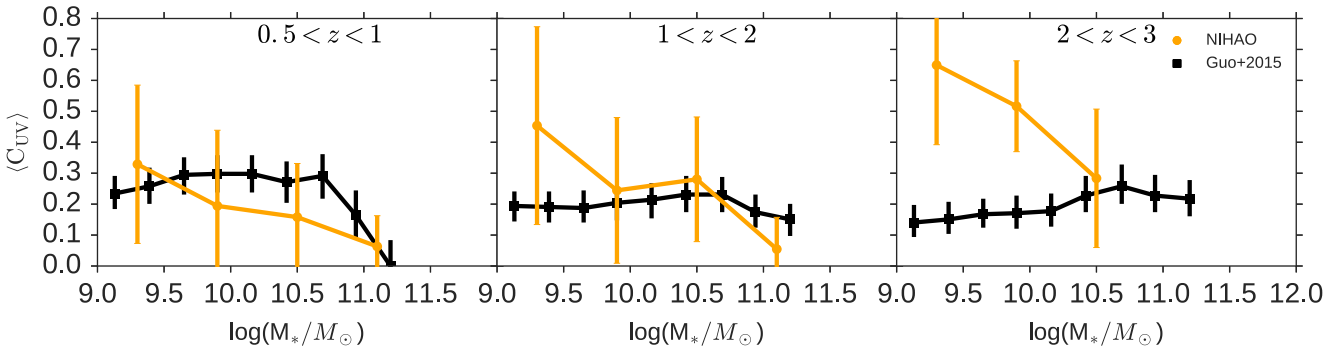

Figure 8. Average intrinsic clump contribution to the rest-frame dust- attenuated UV light of galaxies as a function of stellar mass in three different mass bins. The observations from Guo et al. (2015) are shown as black squares and the results from NIHAO are shown as orange dots.

convolving it with a Gaussian of FWHM = 0.17 arcsec and putting it at redshift $z = 2.2$ to mimick SINFONI observations. They found that their short-lived clumps are consistent with the SINFONI observations of galaxies at redshift $z = 2.2$. Moody et al. (2014) run the RT code SUNRISE (Jonsson 2006; Jonsson & Primack 2010; Jonsson, Groves & Cox 2010) on their sample of eight cosmological zoom-in simulated clumpy galaxies producing mock *HST* observations in four different filters. Unlike this work, these authors additionally add noise to their images. They find qualitatively very similar results to the ones found here; a clump selection in longer wavelength maps results in lower clump counts compared to shorter wavelength maps. Furthermore, these authors find that if clumps are selected in mock observations, stellar mass maps or gas surface density maps, then the outcome shows vastly different clumps (see e.g. their fig. 5). There is only a minority of clumps showing up simultaneously in two or more maps. This is also confirmed by Tamburello et al. (2016), who used the RT code TRAPHIC (Pawlik & Schaye 2008, 2011) to create Hα maps of ionizing radiation of their non-cosmological simulations of galaxies. These authors add different levels of noise to their mock images and convolve the images with Gaussians with different values of FWHM to mimick different spatial resolutions. They find that the recovered properties of clumps strongly depend on the noise level and the spatial resolution. Clump sizes and masses can change by more than a factor of 2 depending on the sensitivity and the spatial resolution.

However, in this study, we focus on the evolution of the clumpy fraction. This is the first time that such an observationally motivated comparison of the clumpy fraction in the light distribution between observations and simulations has been carried out. Our results show that the NIHAO galaxies have a realistic light distribution, suggesting that they are a good testbeds for a better understanding of the origin and fate of these luminous clumps. Therefore, in the next sections, we analyse the physical properties of the intrinsic clumps.

6.2 Properties of the intrinsic clumps in NIHAO

In order to better understand the physical properties of the clumps, we analyse in detail the *intrinsic* clumps.

Selecting *U*-band luminous clumps results in a variety of clumps with different properties for every galaxy and over a vast redshift range ($0 < z < 3$). In total, we find 682 *U*-band clumps in 488 snapshots in the selected redshift range (203 snapshots within $3 > z > 1$, 155 within $1 > z > 0.5$ and 130 within $0.5 > z > 0.25$).

In Fig. 9, we show some of the properties of NIHAO UV *intrinsic* clumps. The mass (gas+stars) distributions are shown in the top left-hand panel of the figure. Independently of redshift, the mass of clumps in *U*-band maps always peaks at $\sim 10^8$ M$_\odot$ with a maximum clump mass around $\sim 10^{9.5}$ M$_\odot$ and a minimum around $\sim 10^{6.5}$ M$_\odot$. This minimum value is mainly due to the resolution of our simulation, which fixes the minimum pixel size of our maps (see Section 5.1). These results do not change if we select only the six most massive galaxies that most closely resemble the observed galaxies, as can be seen by comparing the thick dashed lines with the thin dotted lines in Fig. 9.

The upper right-hand panel shows the distribution of the clump gas fraction defined as $f_g = M_{gas}/(M_{gas} + M_{star})$. Our clump-finding algorithm has no specific lower mass limit. The lower limit on the clump mass is set by the size limit (300 pc) and the surface density of the selected clump. The lowest clump masses we find in the redshift bins are as follows: $0.25 < z < 0.5 : 2.7 \times 10^4$ M$_\odot$, $0.5 < z < 1 : 7 \times 10^3$ M$_\odot$, $1 < z < 3 : 1.3 \times 10^5$ M$_\odot$ The clump gas fraction shows a slight trend to higher gas fractions at higher redshifts ($1 < z < 3$) and a more flat distribution at lower redshifts ($0.5 < z$). For intermediate redshifts, we see a slight bimodality with clumps showing either very low gas fractions or very high ones close to $f_g = 1$. This is in contrast with Oklopčić et al. (2017), who find a peaked distribution of clump gas fraction with the maximum around gas fractions of $f_g = 0.3$. This discrepancy comes from the difference in the selection method. We select *U*-band bright clumps, while Oklopčić et al. (2017) select clumps in gas surface density.

Despite the fact that we find giant clumps of masses greater than 10^8 M$_\odot$, the contribution of these clumps to the total baryonic mass of the galaxy disc is less than 1 per cent, as shown in the lower left-hand panel of Fig. 9. The sizes of the clumps found in the NIHAO galaxies range from ~300 to ~900 pc with a median value of 450 pc for every redshift bin (see the appendices for a comparison between observed and intrinsic clump sizes, as well as for the clump size dependence on pixel scale). The lower bound is given again by our resolution (we fixed the minimum effective radius to be 300 pc). Finally, most of the clumps we found are round(ish) in shape and sometimes slightly elongated (see e.g. Fig. 4).

The next question to address is what determines whether a galaxy has one or more luminous clumps. In Fig. 10, we show the correlation between stellar mass, SFR, cold gas fraction (f_g) and the mean surface density within the half-mass radius (Σ_{Re}) and colour code of each galaxy according to the 'morphology' of the light distribution (clumpy = orange, non-clumpy = green).

In the lower left-hand triangle of the plot, every galaxy is shown at every snapshot in the redshift range $0 < z < 3$ as a point, while in the

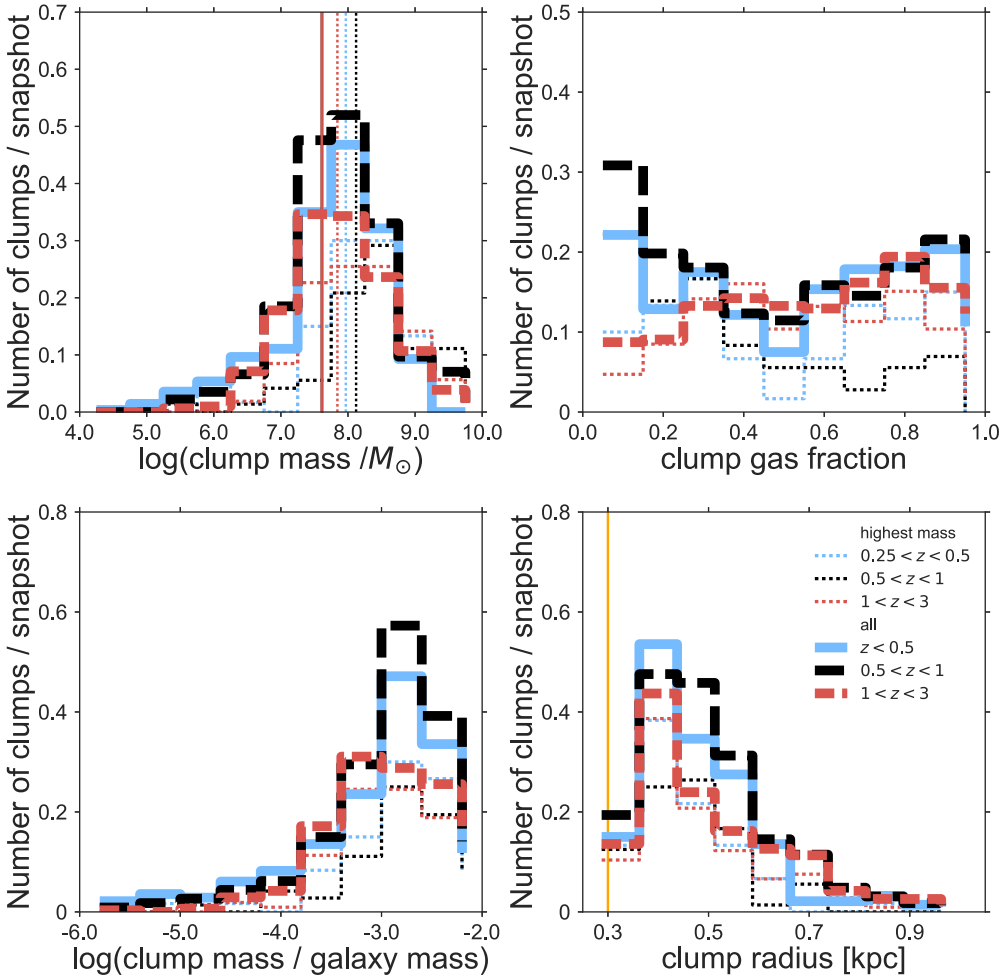

Figure 9. Clump properties for three different redshift bins ($z < 0.5$, $0.5 < z < 1$, $1 < z < 3$) for *U*-band selected clumps. The top left-hand panel shows the total clump mass, the top right-hand panel shows the clump gas fraction, the bottom left-hand panel shows the clump mass as a fraction of galaxy mass and the bottom right-hand panel shows the clump effective radii with the solid vertical orange line indicating the lower limit of clump sizes set by our selection criteria. Thick dashed lines show the results for the whole galaxy sample, while thin dotted lines show results for only the six most massive galaxies. The coloured solid lines in the top left-hand panel show the median mass of clumps in the whole sample, and the dotted lines show the median mass of clumps for the most massive galaxies.

upper right-hand triangle, we show a kernel density estimation of the point distribution. In the diagonal, we show the marginal histograms of the property in the corresponding column for clumpy/non-clumpy galaxies.

Clumpy and non-clumpy galaxies mostly separate in two distinct populations in these parameter spaces. We find that clumpy galaxies show high cold gas fractions, are less centrally concentrated (lower value of $\Sigma_{\rm Re}$), and show low and average SFRs and stellar masses. In contrast, non-clumpy galaxies are more centrally concentrated, have low gas fractions and are among the highest mass galaxies with high SFR. Qualitatively similar correlations are found by (Tadaki et al. 2014, e.g. their fig. 1) for galaxies from the SXDF-UDS-CANDELS field and by (Shibuya et al. 2016, e.g. their figs 7–9) for *HST* photo-*z* and Lyman break galaxies.

There is one notable difference between simulations and observations. From observations, one would expect the highest mass, highest star forming galaxies to be clumpy. Here we find that the lower mass, lower star forming galaxies are preferentially clumpy. This can be explained by the fact that our simulations are more centrally concentrated than observed galaxies, and as a consequence, the gas in the simulated discs is less prone to gravitational instabilities (Martig et al. 2009). Furthermore, we use non-dust-attenuated properties for this plot. As we showed in Fig. 5, the degree of clumpiness in non-dust-attenuated images can be different from that of dust-attenuated ones. Some of the most massive galaxies, in particular, are not clumpy in the non-dust-attenuated maps, but clumpy in the mock RT images. However, other galaxies show the opposite behaviour; they are clumpy in the non-dust-attenuated maps and not clumpy in the mock images. Therefore, in Fig. 11, we give the clumpy fraction as a function of stellar mass, SFR, cold gas fraction and mean surface density within the half-mass radius for both *intrinsic* and *observed* clumps, as well as the observational data from Tadaki et al. (2014). For our *intrinsic* clumps, we find an anticorrelation between clumpy fraction and stellar mass, SFR and mean central surface density, and a correlation of clumpy fraction with the cold gas fraction of the galaxies. While *intrinsic* clumps show a strong evolution with all four galaxy parameters (anticorrelation with the galaxy stellar mass, SFR and central

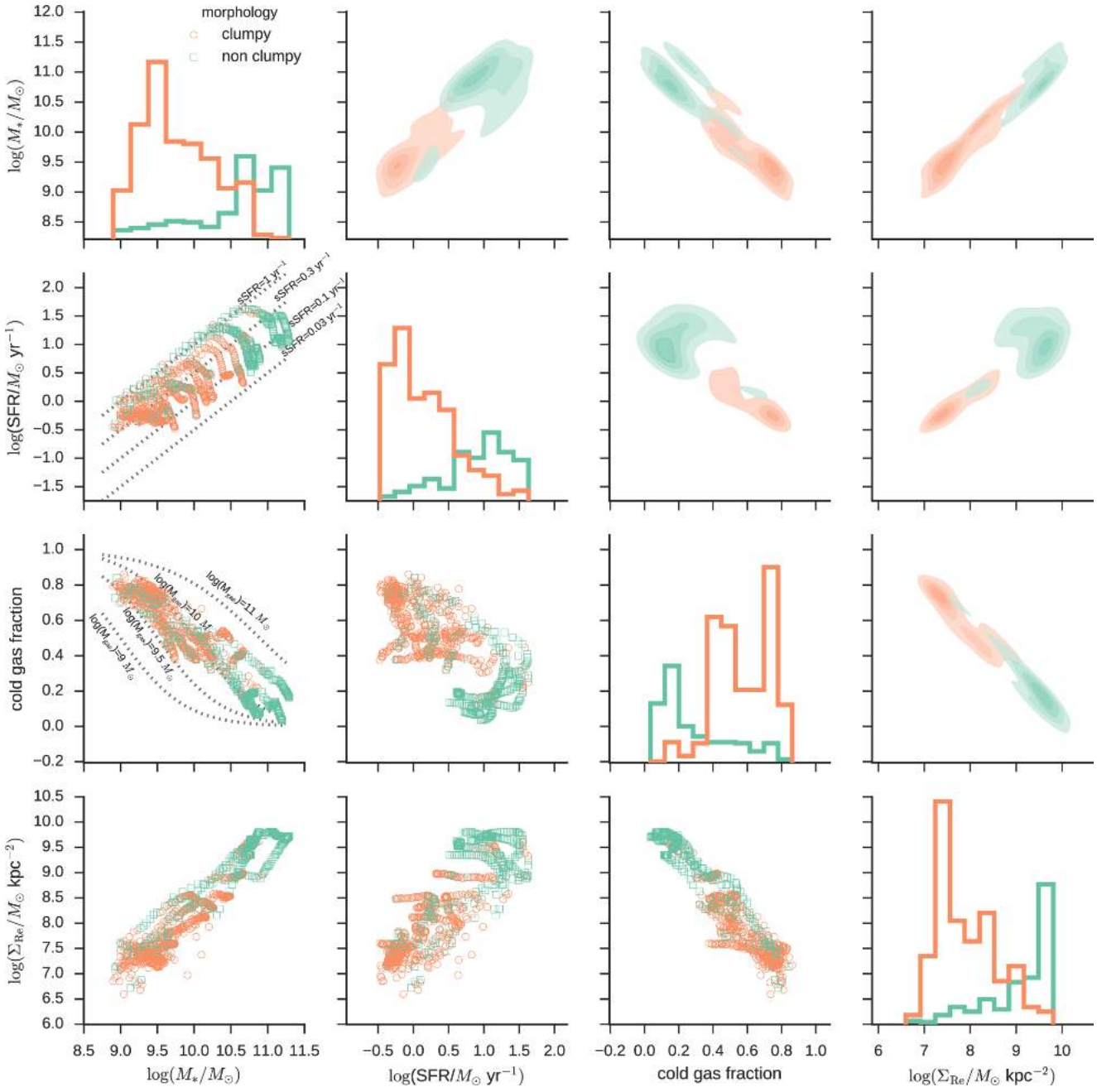

Figure 10. Properties of host galaxies and their correlation with clumpy morphology for all galaxies and all redshifts in our sample. We show correlations of stellar mass (left-hand column), SFR (second column), cold gas fraction (third column) and the mean surface density within the half-mass radius (right-hand column) with each of the other quantities and distinguish clumpy from non-clumpy galaxies. Red coloured points and lines show galaxies with a clumpy morphology (at least one off-centre clump) and green squares/lines show non-clumpy galaxies. The lower left-hand triangle shows all snapshots for all galaxies with single dots, while the upper right-hand triangle shows a kernel-density estimation of the point distribution. The diagonal shows the marginal histogram of the property of the given column.

surface density, and correlation with the cold gas fraction), we do not find such an evolution for our *observed* clumps. This is due to the beforehand-mentioned dust attenuation. Our *observed* clumps, however, show very good agreement with the observational data, which also shows no strong correlation between the clumpy fraction and galaxy parameters. This result further supports the need for a careful modelling of dust obscuration when comparing galaxy morphologies between simulations and observations.

6.3 Clumps in light or clumps in mass?

A lot of discussion is going on whether the observed clumps in high-redshift galaxies represent self-bound clumps of stars orbiting within the disc or whether they are simply a concentration of luminous young stars with practically non-dynamical influence on the disc.

To answer this question, in the upper panel of Fig. 12, we looked at the clumpiness of stars in different wavelength bands

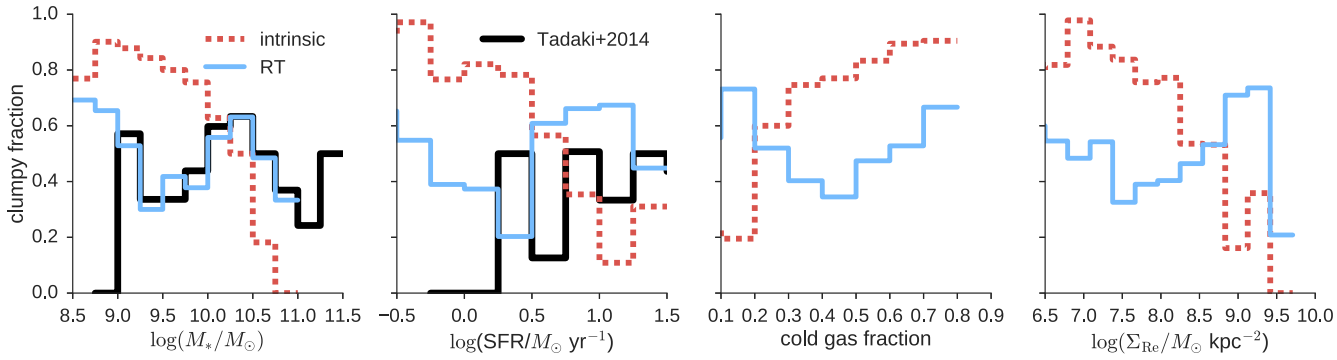

Figure 11. The fraction of clumpy galaxies as a function of host galaxy parameters for *intrinsic* and *observed* clumps in comparison to observations from Tadaki et al. (2014). From the left-hand to right-hand side, we show the clumpy fraction as a function of stellar mass M_{star}, SFR, cold gas fraction and the mean surface density within the half-mass radius.

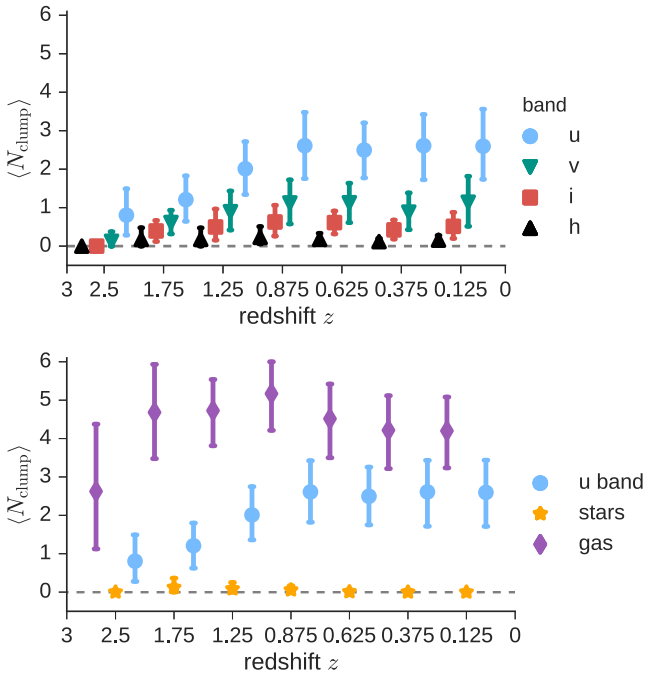

Figure 12. Mean number of clumps per galaxy for clumps selected in different wavelength bands. The upper panel shows the mean number of clumps for clumps selected in the *U*-band (blue dots), the *v* band (green lower triangles), the *i* band (red squares) and the *h* band (black upper triangles). The lower panel shows a comparison of the mean number of clumps selected in the *U*-band (blue dots) with clumps selected in the gas surface density maps (purple diamonds) and the stellar surface density maps (orange stars). The points are slightly offset to avoid overlap.

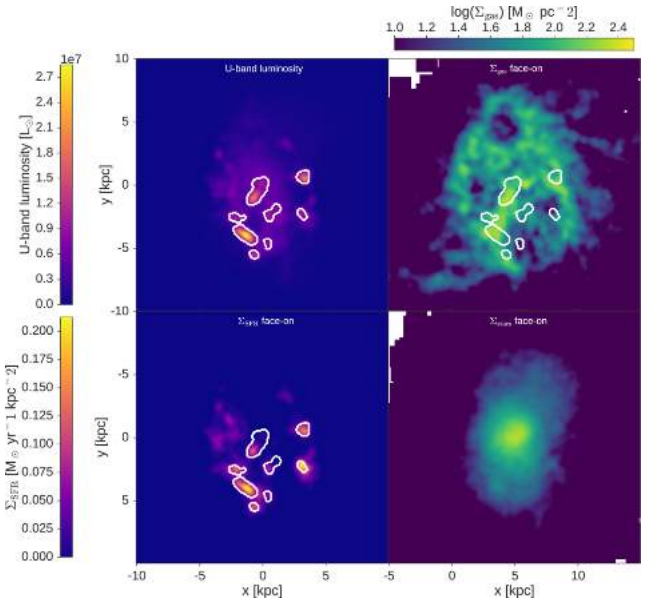

Figure 13. Face-on maps of *U*-band luminosity (upper left-hand panel), SFR surface density (lower left-hand panel), cold gas surface density (upper right-hand panel) and stellar mass surface density maps (lower right-hand panel). The stellar mass surface density map is extremely smooth. White contours show again the clumps selected in the *U*-band luminosity maps.

(*u*, *v*, *i*, *h*). It is clear from the plot that moving to longer wavelength, the disc clumpiness tends to disappear, and practically no clumps are detected in the *h* band.

This seems to suggest that observed and simulated clumps in the non-dust-attenuated *U*-band are not actually bound structures of stars. This result is even more clear in the lower panel of Fig. 12, where we look at the presence (or lack thereof) of clumps in maps constructed for different quantities: *U*-band luminosity, gas and stellar masses. As already noted in Fig. 4, clumps in the *U*-band match clumps in the gas surface density, but when we look at the stellar mass maps, they are extremely smooth and no clumps are detected by our algorithm in practically any of our galaxies.

This is confirmed by a visual inspection of one of the NIHAO galaxies (as already shown in Fig. 4) in Fig. 13, where there is clearly a lack of any substructure in the stellar mass map. This is maybe the most important result of our study. Despite having galaxies as clumpy as the observed ones in *light* maps, we found no evidence of any self-bound stellar structure in the *mass* maps of the same galaxies. We are facing here luminous clumps and not dynamical ones. The good match between clumps in the *U*-band and the gas surface density maps suggests that observed clumps are simply a manifestation of localized (clumpy) star formation regions, as also observed in redshift zero galaxies.

This is also confirmed by the lifetime of the these luminous clumps: Most of the clumps disappear between two consecutive simulation snapshots in NIHAO, setting an upper limit to their dissolution time of about 200 Myr, less than one dynamical time of the galaxy. Fig. 14 shows the fraction of clump stars still close together in consecutive snapshots ($\sim$ 200 Myr). We track the star

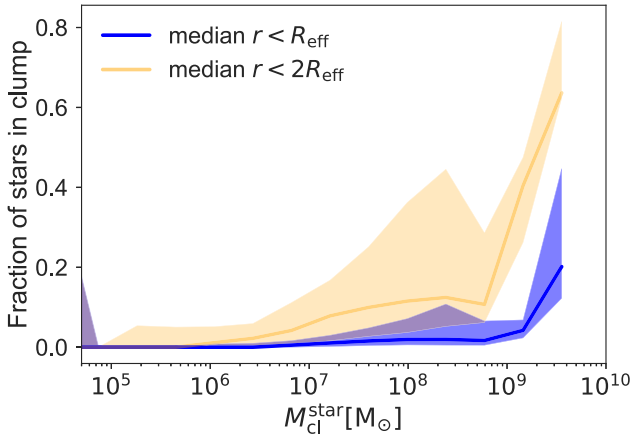

Figure 14. Fraction of clump stars still close together after 200 Myr. For every identified clump, we show the fraction of clump stars still close together in the subsequent snapshot. The blue line shows the median fraction of clump stars still within a region of $1R_{\rm eff}$ around the clump's centre of mass, while the yellow line shows the median fraction of clumps stars still within $2R_{\rm eff}$.

particles of every clump identified in a given snapshot via their particle IDs. In this manner, we evaluate how many stellar particles are still within a region of $1R_{\rm eff}$ or $2R_{\rm eff}$ around the clump centre of mass in the next snapshots. We show the median fraction of clump stars still within a region of $1R_{\rm eff}$ with a blue line and the corresponding 16th and 84th percentile as a blue shaded region. The yellow line and yellow shaded region give the median fraction of clump stars still within a region of $2R_{\rm eff}$. With the exception of a few of the most massive clumps, we find that all clumps lose more than 90 per cent of their mass between two consecutive snapshots.

These results are at odds with previous works (Bournaud & Elmegreen 2009; Bournaud et al. 2014; Dekel et al. 2009b; Ceverino et al. 2010; Mandelker et al. 2014; Ceverino et al. 2015), which suggested that clumps are gravitationally bound structures that migrate towards the centre under the influence of dynamical friction. In support of this picture, there is the observation of a colour gradient of clumps as a function of distance from the galaxy centre (Förster Schreiber et al. 2011; Shibuya et al. 2016), with clumps closer to the centre being redder than clumps at the outskirts. The colour gradient is supposed to be a proxy for the clump age. On the other hand, it has already been suggested that such a gradient could simply be due to the underlying stellar disc population (Genel et al. 2012; Nelson et al. 2016a; Oklopčić et al. 2017).

In Fig. 15, we show the average age of simulated clumps as a function of their distance from the centre. The stellar age of a clump is the (mass-weighted) average age of *all* stars within the clump, which, of course, is a mix of stars actually formed in the clump and the underlying stellar population. As reference, we show in the same figure the mean stellar age of the global stellar disc population as a function of distance from the Galactic Centre. The gradient is quite weak with a slight increase of stellar ages in the outskirts of the disc, consistent with expectations from stellar migration (El-Badry et al. 2016).

We find that most clumps are in the central parts of the galaxy within 5 kpc from the centre, (but keep in mind that we excluded the innermost 1 kpc from our analysis). In NIHAO, we find a wide range of clump stellar ages from ∼200 Myr up to about 2 Gyr in the lowest redshift snapshots in agreement with observed ages of stellar populations in high-redshift clumps (Elmegreen & Elmegreen 2005; Elmegreen et al. 2009; Förster Schreiber et al. 2011; Genzel et al. 2011; Guo et al. 2012; Wuyts et al. 2012). Stellar ages of clumps are generally slightly younger than the underlying mean age of the disc stars, consistent with the picture of clumps being sites of intense star formation.

Interestingly, we recover the observed trend of clumps in the outskirts being younger than clumps in the central parts of the galaxies. In contrast to what was suggested by some authors, who ascribe this gradient to a migration of clumps, we find that this effect is strongly due to a selection bias. Clumps in the central parts of the galaxy include more underlying disc stars, which increases the mean stellar age, while in the outskirts the density of the stellar disc decreases and clumps are less polluted by disc stars and thus appear younger.

Although we recover the observed trend of clump ages as a function of radius, this is not a signature of clumps spiralling inwards and moving to the centre, as we will show in the next section.

6.4 The evolution and final fate of light clumps

If light clumps were self-bound structures of stars, it would be natural to expect them to spiral in under the influence of dynamical friction. Previous studies have indeed argued that high-redshift

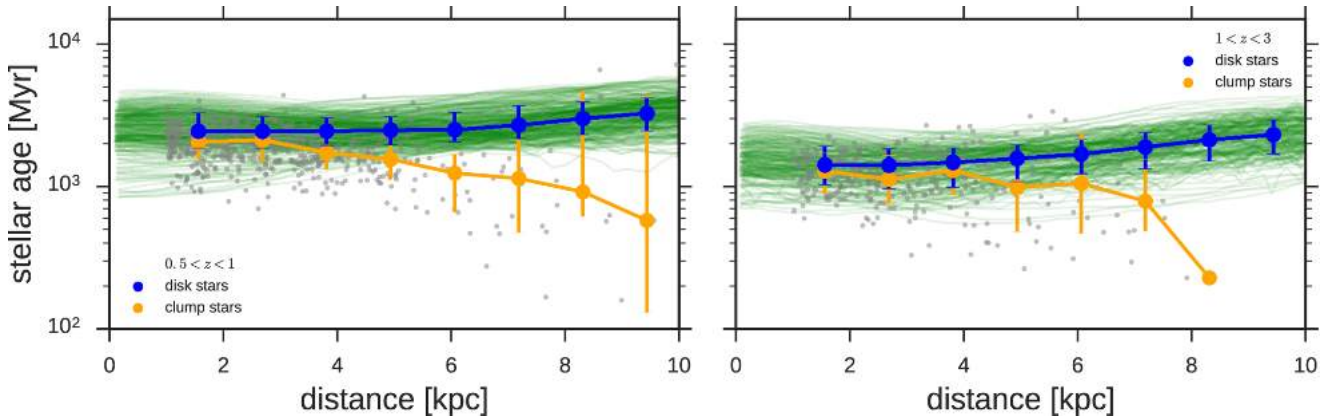

Figure 15. Mean mass-weighted stellar ages of individual clumps (grey dots) and the underlying disc stars (green lines) as a function of radius in two different redshift bins. The left-hand panel shows the mean mass-weighted stellar ages in the redshift range $0.5 < z < 1$ and the right-hand panel shows the stellar ages for the redshift range $1 < z < 3$. The yellow dots show the median stellar age of individual clumps as a function of distance from the galaxy centre and the blue dots show the median age of stars in the disc. The error bars show the 16th and 84th percentile.

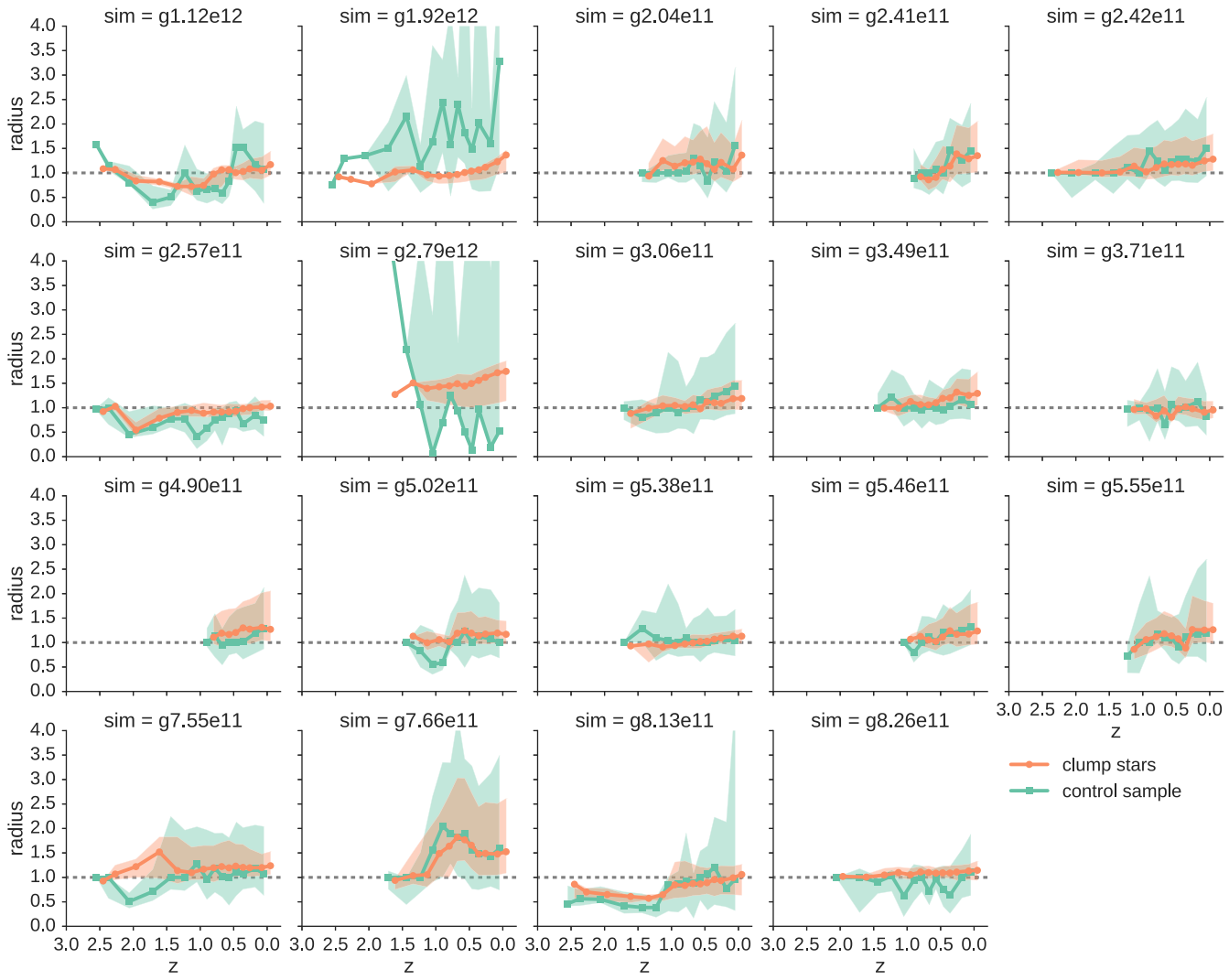

Figure 16. Median distance of clump stars (red dots) and control stars (green squares) as a function of redshift. For every galaxy and for every clump, we track the stars found to be in one clump at a given time and calculate their mean galactocentric distance at later times. We show the median value of all clump radii normalized to the initial clump radius. The same procedure is done for a control sample of stars of the same number and same average radius as the clump stars. The shaded region shows the 16th and 84th percentile.

clumps can be responsible for the building of bulges at high redshift (Ceverino et al. 2010; Bournaud et al. 2014).

As discussed in the previous section, our simulations suggest a quite different scenario, and hence it is interesting to ask what is the evolution of our luminous (but not bound) clumps.

For every clump in every galaxy, we track its evolution by calculating the mean galactocentric distance of all its stars at later times [$r(z)$], and we normalize this number by the distance at the time the clump was first detected [$r(z_{\rm form})$]. For stars moving inwards, we expect the ratio $r(z)/r(z_{\rm form})$ to be below 1 and the opposite for particles moving outwards.

Fig. 16 shows the evolution of the ratio $r(z)/r(z_{\rm form})$ as a function of redshift for all our NIHAO galaxies. In the plot, we also show the same quantity for a 'control sample' of disc stars; for each clump with a given amount of stars, we select an identical number of disc stars at the same galactocentric distance as the clump of interest at the time of its first detection. The shaded region shows the 16th and 84th percentile. The orange dots in Fig. 16 show that there is no preferential inward migration for clump stars. If anything, there is a slight trend for clump stars to migrate outwards. When compared to the control sample of disc stars, clump stars do not show any particular difference; they behave in a very similar manner. Similar results for the lack of inward migration of clumps are found by Oklopčić et al. (2017), who looked at the angular momentum change between the final and initial snapshots in which a clump is visible. These authors do find a roughly equal likelihood for clumps to lose and to gain angular momentum. Thus, these results indicate that clump migration is not governed by dynamical friction but rather by gravitational torquing or tidal forces.

Since we plot the average distance, it could still be possible that a substantial migration of stars inwards is compensated by a same amount of stars moving outwards. Therefore, we look in Fig. 17 at the mass contribution from clump stars to the galaxy bulge at $z = 0$. We first decompose each galaxy at redshift zero into a disc and a spheroidal component using the same procedure as described in Obreja et al. (2016). All stars marked as belonging to the spheroid and having a galactocentric distance smaller than 1 kpc are regarded as bulge stars.

We then calculate the fraction of disc and bulge mass due to stars that have been found in clumps at some earlier point in time.

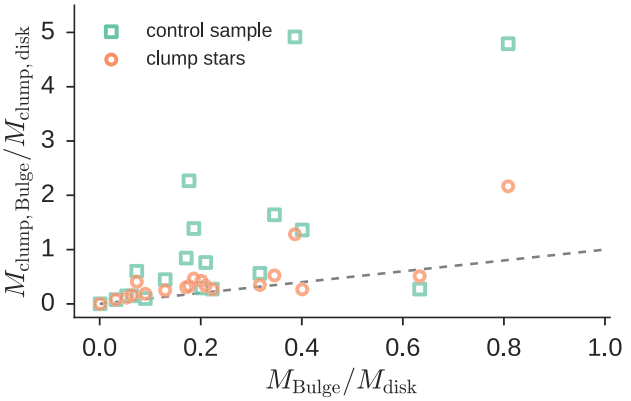

Figure 17. Contribution of clump stars to the bulge and to the disc at redshift zero. Shown is the ratio of stellar mass from clumps in the bulge to stellar mass from clumps in the disc as a function of bulge-to-disc ratio. The red dots show the result for stellar mass from clumps, while the green squares show a control sample of disc stars chosen to have the same size and radius as the clump stars.

In Fig. 17, we plot the ratio of mass from clumps in the bulge $M_{\rm clump,\,Bulge}$ to mass from clumps in the disc $M_{\rm clump,\,disc}$ as a function of bulge-to-disc ratio $M_{\rm Bulge}/M_{\rm disc}$ of the galaxy. Similar to the previous plot, we show clump stars as orange dots and our control sample as green squares. The dashed grey line shows the 1:1 relation. Overall, clump stars seems to be equally distributed between the bulge and the disc at $z = 0$. Furthermore, when compared to the control sample, clump stars seem to have the same final fate as any other stars in the disc.

We can then conclude that, consistently with not being gravitational bound, light clumps in high-redshift galaxies do not preferentially move inwards as time goes by, and do not preferentially contribute to the bulge growth.

7 SUMMARY

We use 19 galaxies from the high-mass end ($M_\star > 10^9$ M$_\odot$ at $z \sim 1.5$) of the NIHAO sample to analyse them in detail for their clumpy morphology and quantify the clumpy fraction of this simulation suite in the redshift range $0 < z < 3$. The NIHAO sample is a suite of high-resolution cosmological hydrodynamical simulations of galaxies in the mass range $10^9 < M_{200} < 4 \times 10^{12}$ M$_\odot$, which reproduces realistic galaxy properties over this huge range of galaxy masses. Unlike most other theoretical studies that looked for giant clumps in galaxies, we do not select clumps in gas or stellar mass maps, but we select clumps in luminosity maps. We apply two different selection methods: *Intrinsic clumps* are selected in the non-dust-attenuated rest-frame *U*-band images closely matching the observational selection method, while for *observed clumps*, we run the RT code GRASIL-3D (Domínguez-Tenreiro et al. 2014) on the 488 snapshots of our sample to create realistic mock *HST* observations of the galaxies. In this way, we can directly compare the clumpy fraction of the NIHAO suite to the observed clumpy fraction from Guo et al. (2015) and Shibuya et al. (2016).

Our main findings can be summarized as follows:

(i) Comparing the observed clumpy fraction, the number of galaxies/snapshots with at least one off-centre clump, of our RT images to the observed clumpy fraction of Guo et al. (2015) and Shibuya et al. (2016), we find very good agreement between our simulations and the observations (see Fig. 6). The NIHAO sample can well reproduce the observed number of clumpy galaxies and their evolution with redshift and perfectly matches the correlation of the clumpy fraction of galaxies with stellar mass (compare Fig. 11). Sizes of clumps found in the NIHAO sample agree well with observed clump sizes (Fig. 7). Furthermore, we recover the observed UV-light contribution of clumps to the total UV light of the galaxy with ∼30 per cent of the light coming from clumps.

(ii) Selecting *intrinsic* clumps in the rest-frame *U*-band images results in clump masses of a few times 10^6–10^9 M$_\odot$, sizes of about 300–900 pc and gas fractions spanning a wide range from 0.1 to 0.9. These findings agree well with observed sizes and masses of giant clumps. However, although clumps are prominent in *U*-band luminosity maps, they contribute only a small fraction of less than 1 per cent to the disc mass (compare with Fig. 9).

(iii) For this work, *intrinsic* clumps are selected in stellar light and can only be found in young stars showing up in the *U* band. Selecting clumps in longer wavelength bands like the *v* band, *h* band or *i* band, the number of clumps found drops to zero (see Fig. 12), and we cannot find clumps in stellar mass at all. This recovers the findings of Wuyts et al. (2012) that clumps are present only in short-wavelength images but not in the inferred stellar mass maps (compare also with Fig. 13). Thus, we find clumpy star formation but no clumpy stellar discs in the NIHAO galaxies.

(iv) Comparing the properties of *intrinsic* clumpy and non-clumpy galaxies in Fig. 10, we find a bimodality between these two. Clumpy galaxies show high cold gas fractions, are less centrally concentrated, and show low and average SFRs and stellar masses. In contrast, non-clumpy galaxies are more centrally concentrated, have low gas fractions and are among the highest mass galaxies with high SFR. These correlations for *intrinsic* clumps are strongly altered if we use *observed* clumps to divide our galaxy sample into clumpy and non-clumpy galaxies, as shown in Fig. 11. Especially, the clumpy fraction as a function of stellar mass for our *observed* clumps is in very good agreement with the observed relation. Thus, we conclude that a careful modelling of dust obscuration has to be taken into account for a direct comparison of galaxy morphologies between simulations and observations.

(v) The mean mass-weighted stellar ages of clumps in the simulations show the same trend as observed colour gradients of clumps (e.g. Shibuya et al. 2016). Clumps in the outskirts of the galaxies are younger and clumps in the centre are almost as old as the underlying mean mass-weighted stellar age of the disc stars (see Fig. 15). This trend can be attributed to the fact that the stellar density of disc stars is lower in the outskirts and higher in the centre; thus, when selecting clumps in images, the contribution/pollution of disc stars to the clump is higher in the centre than in the outskirts. That is why the mean mass-weighted stellar ages of clumps and stellar discs are more similar in the galaxy centre than in the outskirts.

(vi) We find that clumps in the NIHAO sample do not spiral inwards and do not contribute much mass to the bulge. Clumps get quickly disrupted and disperse. They lose about 90 per cent of their mass in less than 200 Myr (see Fig. 14). We trace the stars of the clumps through time down to redshift zero and follow their mean radius and do not find any net inward migration of clump stars. Indeed, we find that clump stars and disc stars behave the same way: Both do not show signs of a net inward migration, as we have shown in Figs 16 and 17. Furthermore, we quantify how much mass these stars contribute to the bulge and to the disc of the galaxies. Clump stars and randomly selected disc stars contribute the same mass to the bulge. Thus, we do not see any indications that clumps in NIHAO would contribute to preferentially build up the bulge.

ACKNOWLEDGEMENTS

We thank the referee for valuable comments that improved the readability of this paper. The authors like to thank Hans-Walter Rix for fruitful discussions and very helpful comments on this work. This research made use of the PYNBODY package Pontzen et al. (2013) to analyse the simulations and used ASTRODENDRO, a PYTHON package to compute dendrograms of astronomical data (http://www.dendrograms.org/), to find clumps. The authors acknowledge support from the Sonderforschungsbereich SFB 881 "The Milky Way System" (subproject A2) of the German Research Foundation (DFG). Simulations have been performed on the THEO clusters of the Max-Planck-Institut für Astronomie at the Rechenzentrum in Garching and the HYDRA and DRACO clusters at the Rechenzentrum in Garching. Further computations used the High Performance Computing resources at the New York University Abu Dhabi. We highly appreciate the contributions of all these computing allocations.

REFERENCES

Agertz O., Teyssier R., Moore B., 2009, MNRAS, 397, L64
Behrendt M., Burkert A., Schartmann M., 2016, ApJ, 819, L2
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Bournaud F., Elmegreen B. G., 2009, ApJ, 694, L158
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237
Bournaud F. et al., 2008, A&A, 486, 741
Bournaud F. et al., 2014, ApJ, 780, 57
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cacciato M., Dekel A., Genel S., 2012, MNRAS, 421, 818
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
Ceverino D., Dekel A., Tweed D., Primack J., 2015, MNRAS, 447, 3291
Chabrier G., 2003, PASP, 115, 763
Conroy C., van Dokkum P. G., 2012, ApJ, 760, 71
Daddi E. et al., 2010, ApJ, 713, 686
Dehnen W., Aly H., 2012, MNRAS, 425, 1068
Dekel A. et al., 2009a, Nature, 457, 451
Dekel A., Sari R., Ceverino D., 2009b, ApJ, 703, 785
Domínguez-Tenreiro R., Obreja A., Granato G. L., Schurer A., Alpresa P., Silva L., Brook C. B., Serna A., 2014, MNRAS, 439, 3868
Dutton A. A. et al., 2011, MNRAS, 416, 322
Dutton A. A. et al., 2013a, MNRAS, 428, 3183
Dutton A. A., Macciò A. V., Mendel J. T., Simard L., 2013b, MNRAS, 432, 2496
Dutton A. A. et al., 2016, MNRAS, preprint (arXiv:1610.06375)
El-Badry K., Wetzel A., Geha M., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2016, ApJ, 820, 131
Elmegreen B. G., Elmegreen D. M., 2005, ApJ, 627, 632
Elmegreen D. M., Elmegreen B. G., Ravindranath S., Coe D. A., 2007, ApJ, 658, 763
Elmegreen B. G., Elmegreen D. M., Fernandez M. X., Lemonias J. J., 2009, ApJ, 692, 12
Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761
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., 2006, Nature, 442, 786
Genzel R. et al., 2008, ApJ, 687, 59
Genzel R. et al., 2011, ApJ, 733, 101
Genzel R. et al., 2015, ApJ, 800, 20
Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
Girardi L. et al., 2010, ApJ, 724, 1030
Goz D., Monaco P., Granato G. L., Murante G., Domínguez-Tenreiro R., Obreja A., Annunziatella M., Tescari E., 2016, MNRAS, preprint (arXiv:1610.09843)
Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, ApJ, 542, 710
Granato G. L., Ragone-Figueroa C., Domínguez-Tenreiro R., Obreja A., Borgani S., De Lucia G., Murante G., 2015, MNRAS, 450, 1320
Guhathakurta P., Draine B. T., 1989, ApJ, 345, 230
Guo Y., Giavalisco M., Ferguson H. C., Cassata P., Koekemoer A. M., 2012, ApJ, 757, 120
Guo Y. et al., 2015, ApJ, 800, 39
Gutcke T. A., Stinson G. S., Macciò A. V., Wang L., Dutton A. A., 2017, MNRAS, 464, 2796
Haardt F., Madau P., 2005, preprint (arXiv:e-prints)
Hopkins P. F. et al., 2010, ApJ, 724, 915
Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, MNRAS, 427, 968
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Inoue S., Saitoh T. R., 2012, MNRAS, 422, 1902
Inoue S., Saitoh T. R., 2014, MNRAS, 441, 243
Inoue S., Dekel A., Mandelker N., Ceverino D., Bournaud F., Primack J., 2016, MNRAS, 456, 2052
Jones T. A., Swinbank A. M., Ellis R. S., Richard J., Stark D. P., 2010, MNRAS, 404, 1247
Jonsson P., 2006, MNRAS, 372, 2
Jonsson P., Primack J. R., 2010, New Astron., 15, 509
Jonsson P., Groves B. A., Cox T. J., 2010, MNRAS, 403, 17
Keller B. W., Wadsley J., Benincasa S. M., Couchman H. M. P., 2014, MNRAS, 442, 3013
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kravtsov A., Vikhlinin A., Meshscheryakov A., 2014, ApJ, preprint (arXiv:1401.7329)
Macciò A. V., Udrescu S. M., Dutton A. A., Obreja A., Wang L., Stinson G. R., Kang X., 2016, MNRAS, 463, L69
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
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Martig M., Bournaud F., Teyssier R., Dekel A., 2009, ApJ, 707, 250
Mayer L., Tamburello V., Lupi A., Keller B., Wadsley J., Madau P., 2016, ApJ, 830, L13
Moody C. E., Guo Y., Mandelker N., Ceverino D., Mozena M., Koo D. C., Dekel A., Primack J., 2014, MNRAS, 444, 1389
Morishita T., Ichikawa T., Noguchi M., Akiyama M., Patel S. G., Kajisawa M., Obata T., 2015, ApJ, 805, 34
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Murata K. L. et al., 2014, ApJ, 786, 15
Nelson E. J. et al., 2016a, ApJ, 817, L9
Nelson E. J. et al., 2016b, ApJ, 828, 27
Noguchi M., 1998, Nature, 392, 253
Noguchi M., 1999, ApJ, 514, 77
Obreja A., Brook C. B., Stinson G., Domínguez-Tenreiro R., Gibson B. K., Silva L., Granato G. L., 2014, MNRAS, 442, 1794
Obreja A., Stinson G. S., Dutton A. A., Macciò A. V., Wang L., Kang X., 2016, MNRAS, 459, 467
Oklopčić A., Hopkins P. F., Feldmann R., Kereš D., Faucher-Giguère C.-A., Murray N., 2017, MNRAS, 465, 952
Overzier R. A. et al., 2009, ApJ, 706, 203
Patel S. G. et al., 2013, ApJ, 766, 15
Pawlik A. H., Schaye J., 2008, MNRAS, 389, 651
Pawlik A. H., Schaye J., 2011, MNRAS, 412, 1943
Perez J., Valenzuela O., Tissera P. B., Michel-Dansac L., 2013, MNRAS, 436, 259
Perret V., Renaud F., Epinat B., Amram P., Bournaud F., Contini T., Teyssier R., Lambert J.-C., 2014, A&A, 562, A1

Planck Collaboration XVI, 2014, A&A, 571, A16
Pontzen A., Roškar R., Stinson G. S., Woods R., Reed D. M., Coles J., Quinn T. R., 2013, Pynbody: Astrophysics Simulation Analysis for Python
Price D. J., 2008, J. Comput. Phys., 227, 10040
Puech M., 2010, MNRAS, 406, 535
Ritchie B. W., Thomas P. A., 2001, MNRAS, 323, 743
Saitoh T. R., Makino J., 2009, ApJ, 697, L99
Shen S., Wadsley J., Stinson G., 2010, MNRAS, 407, 1581
Shibuya T., Ouchi M., Kubo M., Harikane Y., 2016, ApJ, 821, 72
Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., 2009, ApJ, 702, 307
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, MNRAS, 428, 129
Stinson G. S. et al., 2015, MNRAS, 454, 1105
Swinbank A. M. et al., 2010, Nature, 464, 733
Tacconi L. J. et al., 2010, Nature, 463, 781
Tacconi L. J. et al., 2013, ApJ, 768, 74
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., 2016, preprint (arXiv:1610.05304)
Toomre A., 1964, ApJ, 139, 1217
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
Wadsley J. W., Stadel J., Quinn T., 2004, New Astron., 9, 137
Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83
Wuyts S. et al., 2012, ApJ, 753, 114

APPENDIX A: SENSITIVITY TO DETECTION THRESHOLD

In this section, we analyse how the properties of intrinsic clumps change if we change the sensitivity of our clump-finding algorithm. Our fiducial clump selection picks up clumps in the *U*-band images that stick out at least 3σ above the mean *U*-band luminosity of the 10 kpc × 10 kpc images. In order to quantify the robustness of our results, we rerun our clump-finding algorithm on all our snapshots with a reduced threshold of 2σ and with an increased threshold of 4σ. In Fig. A1, we compare the clump properties found with all three thresholds. The upper panel shows the clump masses found, the middle panel shows clump gas fractions and the lower panel shows clump sizes. We do not see any differences in the properties of the clumps identified using different thresholds. Clump masses of all three selection thresholds peak at around 10^8 $M_\odot$. However, we see, as would have been expected, that a clump selection with a threshold of 2σ picks up much more clumps compared to the 3σ and 4σ selection. The same is true for the clump selection in the RT images. Fig. A2 shows the evolution of the clumpy fraction for a selection threshold of 2σ above the mean flux in the given band (upper panel) and for a selection threshold of 4σ (lower panel) as opposed to our fiducial selection threshold of 3σ (compare with the left-hand panel of Fig. 6). For a selection threshold of 2σ above the mean flux, we find a clumpy fraction that is higher than observed, and for a selection threshold of 4σ, we find a clumpy fraction that is well in agreement with observations and the one obtained with a threshold of 3σ. By visual inspection of the clumps found, we verify that a selection threshold of 2σ picks up a significant amount of spurious, low-surface brightness clumps in galaxies that, by eye, would not have been classified as clumpy. This is the reason for the

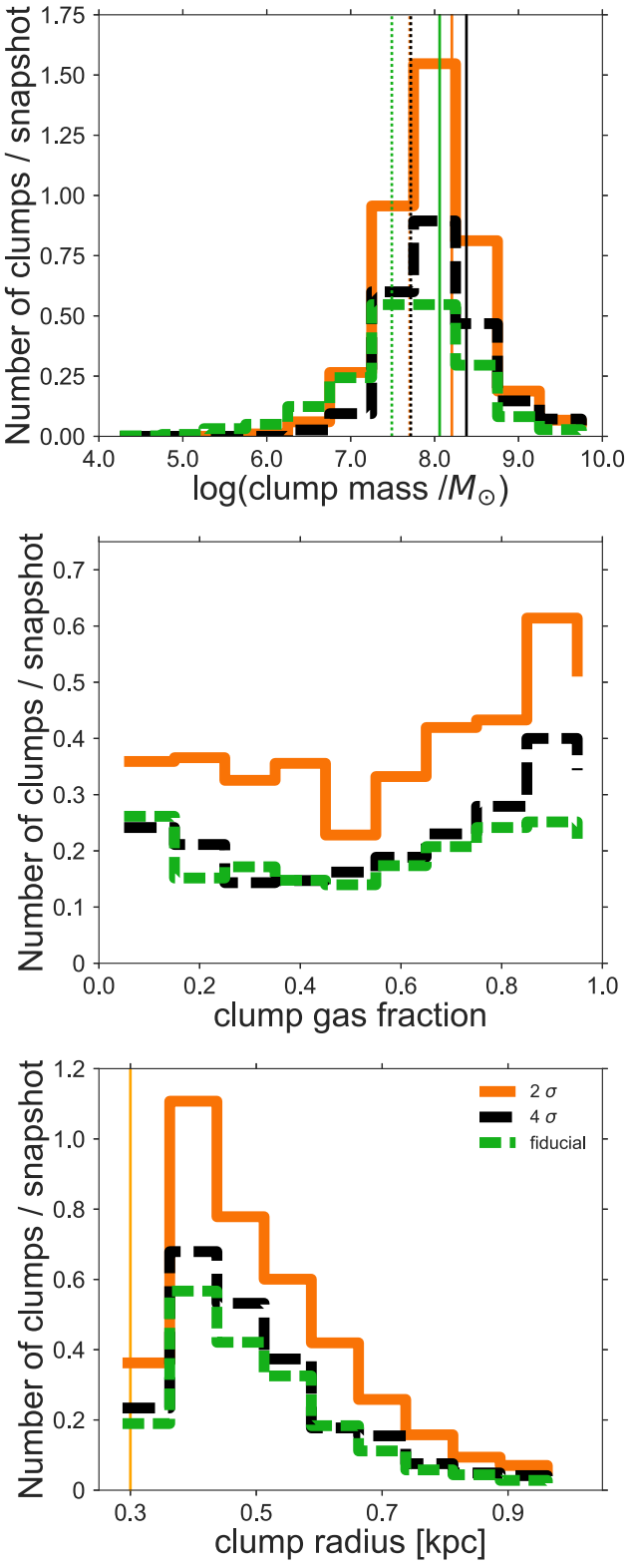

Figure A1. Clump properties for three different clump selection thresholds (2σ, 3σ, 4σ) for *U*-band-selected clumps. The top left-hand panel shows the total clump mass, the middle panel shows the clump gas fraction and the bottom left-hand panel shows the clump effective radii with the solid vertical orange line indicating the lower limit of clump sizes set by our selection criteria. The coloured solid lines in the top left-hand panel show the median mass of clumps and the dotted lines show the mean mass of clumps.

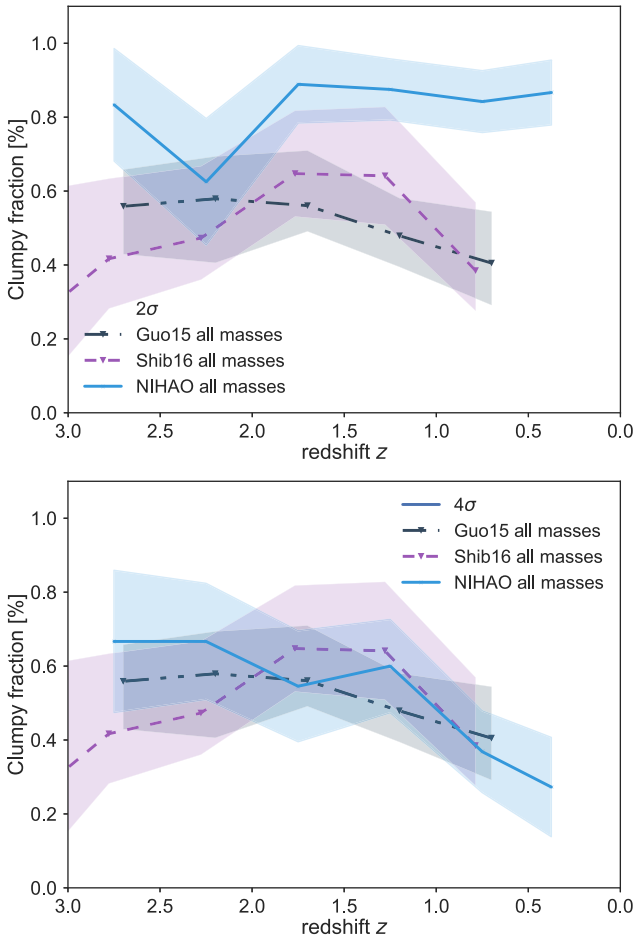

Figure A2. Evolution of the fraction of galaxies with at least one *observed* off-centre clump for a selection threshold of 2σ (upper panel) and 4σ (lower panel). Both panels show the evolution of the clumpy fraction for the whole NIHAO sample (blue line) compared to observations from Guo et al. (2015) (black dash–dotted line) and Shibuya et al. (2016) (purple dotted line). The shaded band shows the 1σ scatter.

increased clumpy fraction for this threshold value. The agreement between the threshold values of 3σ and 4σ makes us conclude that our fiducial selection threshold of 3σ works properly.

APPENDIX B: DEPENDENCE OF CLUMP SIZE ON PIXEL SCALE

Recent work by Behrendt, Burkert & Schartmann (2016) and by Tamburello et al. (2016) found that clump sizes depend on the chosen spatial resolution and giant clumps would break up into smaller clumps if the spatial resolution is increased. In this section, we test what happens to the clump sizes found in this study if the spatial resolution is changed. For our intrinsic clump selection, we have chosen a pixel scale that is comparable to the physical resolution of our simulation. Therefore, we are already using the highest spatial resolution our simulations allow for. However, we change the pixel scale of our intrinsic clump selection to match the one of our RT runs (*HST* spatial resolution). Accordingly, we changed the minimum pixel per clump of our clump finder from 30

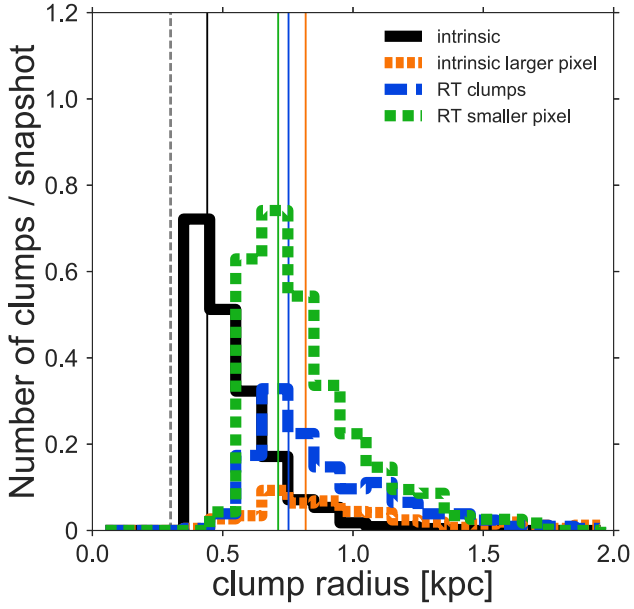

Figure B1. Comparison of clump sizes for intrinsic clumps and clumps in the RT runs for different pixel scales. The black line shows our fiducial intrinsic clump selection, the orange line shows clump sizes for intrinsic clump selection with the *HST* pixel scale, the blue line shows clump sizes from our fiducial clump selection in the RT run and the green line shows clump sizes for RT runs with doubled resolution (half the pixel scale of *HST*). The coloured vertical lines show the corresponding median clump sizes and the vertical grey dashed line shows our resolution limit on clump sizes.

to 5 to keep the minimum effective radius of the clumps to about 300 pc. In Fig. B1, we show the comparison of clumps sizes between the fiducial and the lower spatial resolution runs. We confirm that our clump finder picks up larger clump sizes in the lower spatial resolution runs, the median clump size increasing from ∼0.5 to ∼0.8 kpc. From a visual inspection of the galaxies, we find that this is due to two reasons: (1) Some larger clumps in the lower spatial resolution run indeed break up into several clumps if the spatial resolution is increased, and (2) most small clumps in the higher spatial resolution run are not found in the lower resolution run.

For the RT clumps, we selected the six highest mass galaxies and rerun GRASIL-3D in the redshift range $z \sim 3$–1, with a spatial resolution twice as large as that of the fiducial runs. Again, we changed the minimum pixel per clump to match the physical resolution of the simulation. For these runs, we adopted a value of at least 10 pixel per clump, which again results in a minimum effective radius of ∼300 pc for the clumps. As shown in Fig. B1, there is almost no change in the clump sizes between the two different RT runs, and clump sizes seem to be stable if the pixel scale is increased. It is unclear if doubling the spatial resolution is not enough to see a break-up of clumps in the RT runs or if clumps do not break up in the RT run. We did not test the effect of increasing the resolution of the RT calculations because a resolution twice as good as the fiducial one already requires four times more memory per core. In our case, this would exceed the typical 2 Gb available on HPC machines. Finally, clump sizes of intrinsic clumps found in the images with the *HST* pixel scale agree well with clump sizes found in the RT run of the same pixel scale.

APPENDIX C: IMAGES

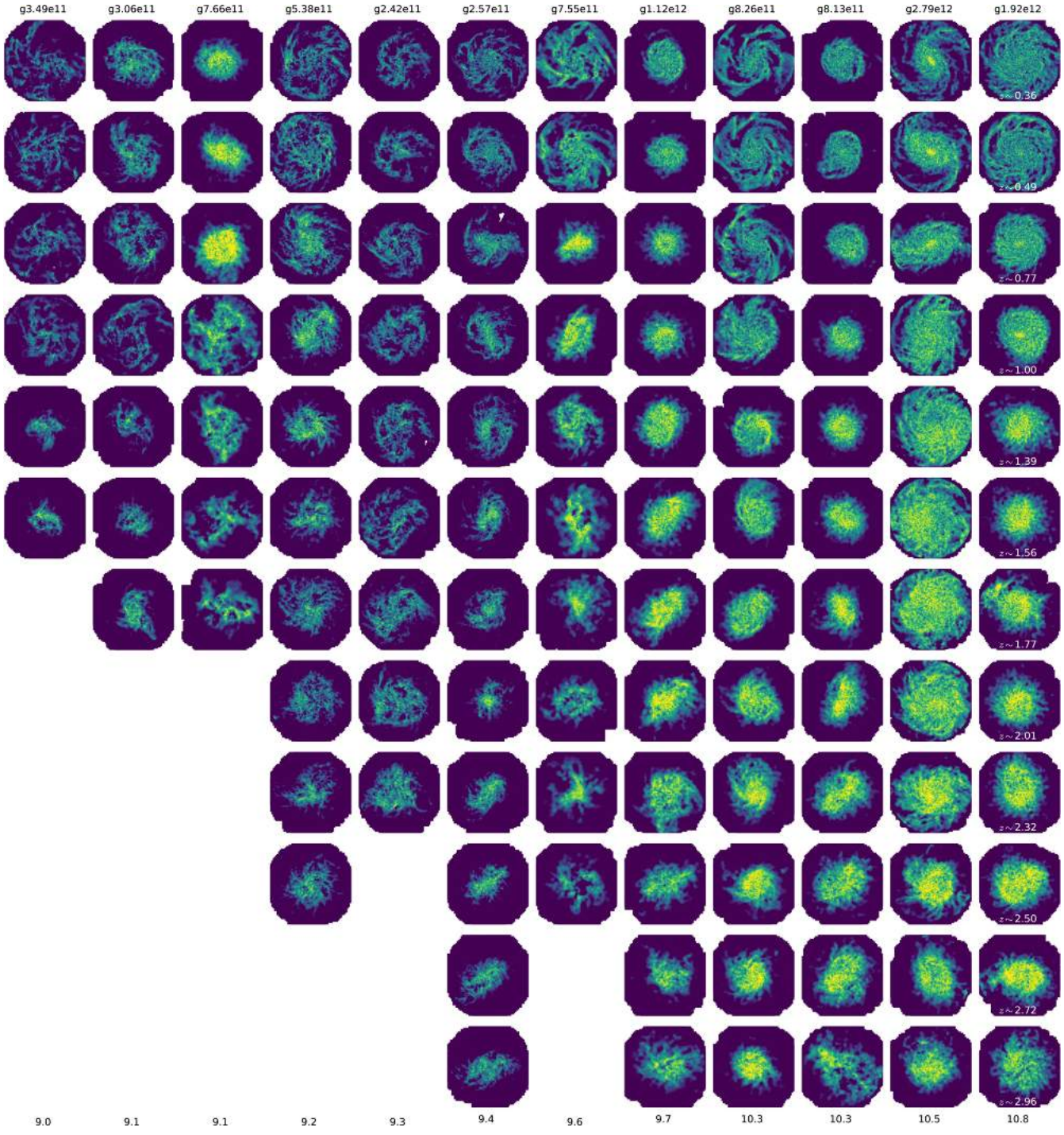

Figure C1. Impression of the cold gas surface density. From the left-hand to right-hand side: increasing stellar mass at redshift zero; from bottom to top: increasing time. Lower bounds of gas surface densities are set to $10\,M_{\odot}\ pc^{-2}$ and maximum values are set to $300\,M_{\odot}\ pc^{-2}$

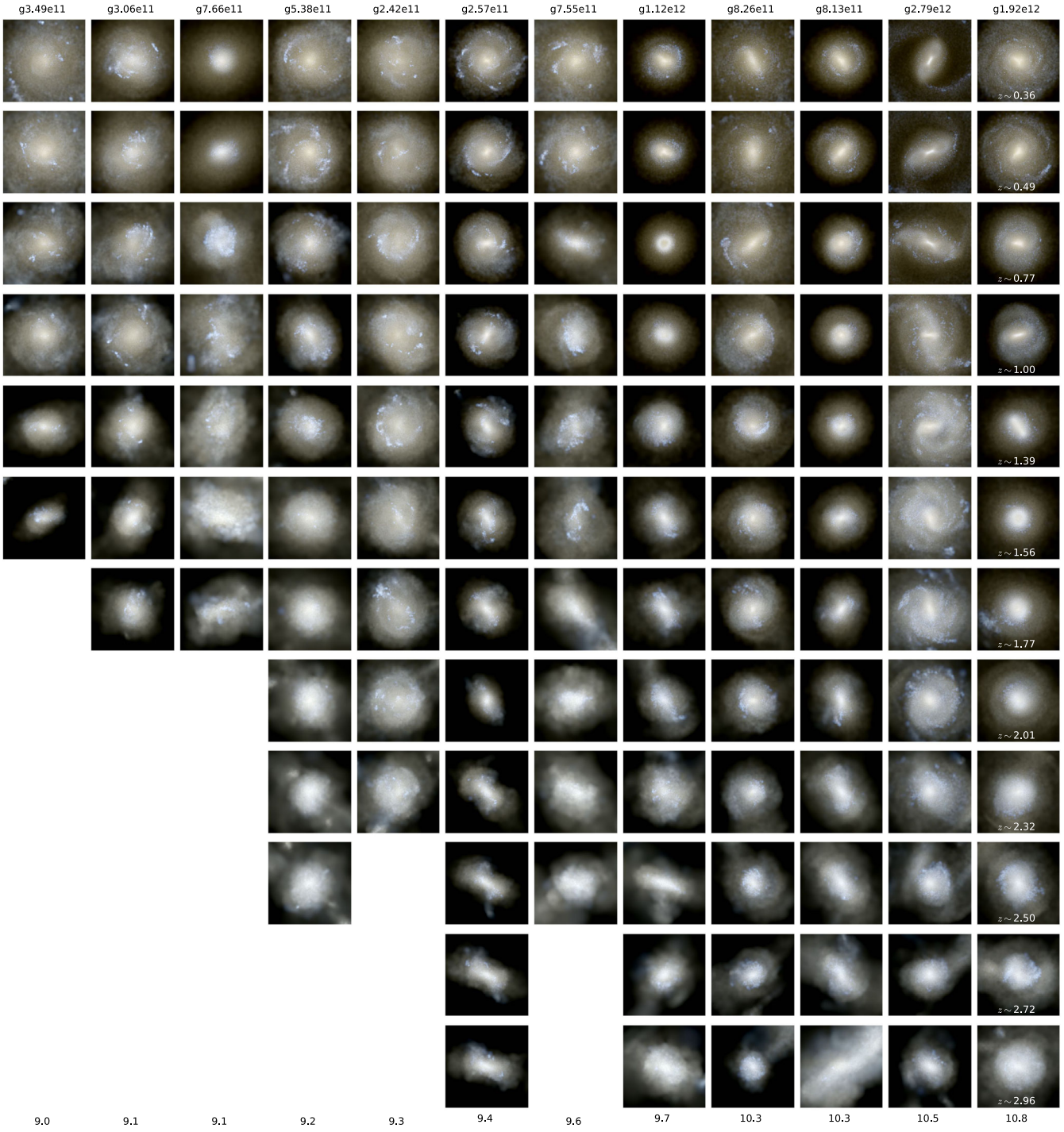

Figure C2. Impression of RGB composite images. From the left- to right-hand side: increasing stellar mass at redshift zero; from bottom to top: increasing time.

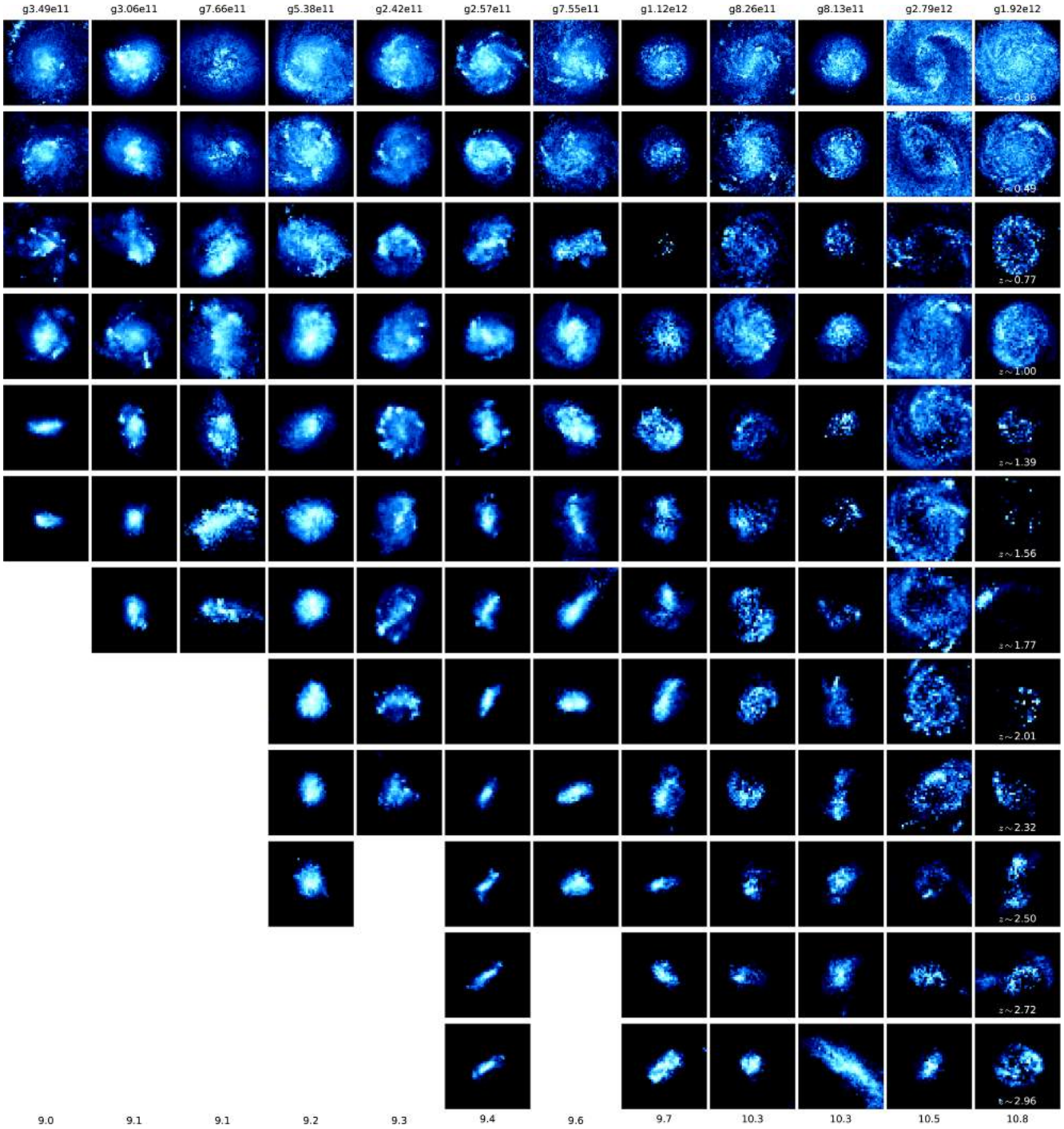

Figure C3. Impression of the outcome of the RT runs. From the left-hand to right-hand side: increasing stellar mass at redshift zero; from bottom to top: increasing time. For snapshots in the redshift range $3 > z > 2$, we select the *F775W* filter; for $2 > z > 1$, we select the *F606W* filter; and for $z < 1$, we select the *F435W* filter to detect clumps. This filter selection roughly corresponds to the rest-frame UV at these redshifts (Guo et al. 2015).

This paper has been typeset from a TEX/LATEX file prepared by the author.