Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **493,** 4607–4621 (2020)
Advance Access publication 2020 February 27

doi:10.1093/mnras/staa541

Rapid early coeval star formation and assembly of the most-massive galaxies in the Universe

Douglas Rennehan [1]★ Arif Babul,[1] Christopher C. Hayward,[2] Connor Bottrell [1] Maan H. Hani [1]† and Scott C. Chapman[1,3,4,5]

[1]*Department of Physics and Astronomy, University of Victoria, BC V8X 4M6, Canada*
[2]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*
[3]*Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS B3H 4R2, Canada*
[4]*NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada*
[5]*Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada*

Accepted 2020 February 18. Received 2020 February 4; in original form 2019 July 1

ABSTRACT

The current consensus on the formation and evolution of the brightest cluster galaxies is that their stellar mass forms early ($z \gtrsim 4$) in separate galaxies that then eventually assemble the main structure at late times ($z \lesssim 1$). However, advances in observational techniques have led to the discovery of protoclusters out to $z \sim 7$. If these protoclusters assemble rapidly in the early Universe, they should form the brightest cluster galaxies much earlier than suspected by the late-assembly picture. Using a combination of observationally constrained hydrodynamical and dark-matter-only simulations, we show that the stellar assembly time of a sub-set of brightest cluster galaxies occurs at high redshifts ($z > 3$) rather than at low redshifts ($z < 1$), as is commonly thought. We find, using isolated non-cosmological hydrodynamical simulations, that highly overdense protoclusters assemble their stellar mass into brightest cluster galaxies within ∼1 Gyr of evolution – producing massive blue elliptical galaxies at high redshifts ($z \gtrsim 1.5$). We argue that there is a downsizing effect on the cluster scale wherein some of the brightest cluster galaxies in the cores of the most-massive clusters assemble earlier than those in lower mass clusters. In those clusters with $z = 0$ virial mass $\geq 5 \times 10^{14}$ $M_\odot$, we find that 9.8 per cent have their cores assembly early, and a higher fraction of 16.4 per cent in those clusters above 10^{15} $M_\odot$. The *James Webb Space Telescope* will be able to detect and confirm our prediction in the near future, and we discuss the implications to constraining the value of σ_8.

Key words: galaxies: clusters: general – galaxies: evolution – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

Galaxy clusters are at the peak of the mass-assembly hierarchy in the Lambda cold dark matter (ΛCDM) paradigm and share a common feature of hosting a distinct population of galaxies aptly referred to as the brightest cluster galaxies (BCGs: Peebles 1968; Sandage 1976; Tremaine & Richstone 1977). The BCGs are ultraluminous (with $\sim 10 L_{K,*}$ where $L_{K,*}$ is the characteristic luminosity of the general galaxy population), morphologically spheroidal, have a large spatial extent, and exhibit core stellar velocity dispersions of order 300–400 km s^{-1} (Lin & Mohr 2004; Pipino et al. 2011; Loubser et al. 2018). The BCGs are in fact the brightest and the most-massive galaxies in the present-day Universe. Many of the observed characteristics of the BCGs seem to scale with the properties of the hosting cluster halo (Brough et al. 2008; Lidman et al. 2012; Lavoie et al. 2016; Kravtsov, Vikhlinin & Meshcheryakov 2018) and they are frequently found close to, and typically have relatively small velocity offsets with respect to, the potential centres of the cluster halo (Lidman et al. 2013; Lauer et al. 2014). This is commonly interpreted as an indication that the formation and evolution of the BCGs and their host clusters are intimately linked, and that studying the former will provide clues about the formation and evolution of the latter. However, pinning down the assembly and growth histories of these gigantic galaxies is proving to be a challenge.

There are several pathways that could explain the origin of the BCGs: (i) extended *in-situ* star formation (e.g. via cooling flows); (ii) rapid star formation and early assembly; and (iii) early star formation in separate galaxies but relatively recent assembly via a sequence of late-time mergers. The key idea underlying the first proposal is that radiative cooling drives the hot intracluster medium

★ E-mail: douglas.rennehan@gmail.com
† Vanier Scholar.

to concentrate at the clusters' potential centre where it then forms stars at relatively high rates (Cowie & Binney 1977; Fabian & Nulsen 1977). However, observations show that not only are the bulk of the BCG stars old (Whiley et al. 2008), but heating from the central active galactic nucleus (AGN) also strongly suppresses the cooling of the intracluster medium (Tabor & Binney 1993; Ciotti & Ostriker 1997; Silk & Rees 1998). The second proposal suggests that massive elliptical galaxies, including the BCGs, form via essentially monolithic collapse of a mass density peak (Eggen, Lynden-Bell & Sandage 1962), with the galaxies' stellar mass building up rapidly in the process. One difficulty with this model is that BCGs show evidence of significant growth in their sizes over cosmic time (Daddi et al. 2005; van der Wel et al. 2008; Shankar et al. 2015). Once a leading theory, this scenario has fallen out of favour due to the emergence of the hierarchical assembly paradigm for cosmic structure formation. According to this paradigm, the third proposal, galaxies form via a series of mergers of lower mass systems – implying that the low-mass systems form first, and over time build-up the more-massive systems (Aragon-Salamanca, Baugh & Kauffmann 1998; Dubinski 1998).

Numerical studies investigating the formation and evolution of BCGs in the presently favoured hierarchical ΛCDM model find that the majority of stars that end up in the present-day BCGs typically form at $z \gtrsim 4$ in distinct progenitor galaxies (De Lucia et al. 2006). These galaxies then eventually merge to assemble the BCGs we observe today. As for the timing of this assembly, until recently, the theoretical consensus was that present-day BCGs are assembled through dissipationless mergers, with between 50–60 per cent of their present-day stellar mass being incorporated at relatively late times, from $z \sim 1.5$ to the present (Dubinski 1998; Conroy, Wechsler & Kravtsov 2007; De Lucia & Blaizot 2007; Ruszkowski & Springel 2009; Laporte et al. 2013; Contini et al. 2014; Laporte & White 2015). Recent cosmological hydrodynamical simulations suggest a slightly modified scenario (Martizzi et al. 2016; Ragone-Figueroa et al. 2018) and highlight the importance of distinguishing between the galaxy proper and its extended stellar envelope. Focusing on the galaxy proper (i.e. the mass within 50 kpc of the centre), Ragone-Figueroa et al. (2018) find that while half of the stars that end up in the BCG have formed by $z \sim 3.7$, the assembly of the BCG occurs over an extended time-span and half of the BCGs' stellar mass falls only into place typically by $z \sim 1.5$. In detail, the stellar mass of the BCGs grows on average by a factor of 2.5 between $z = 2$ and $z = 0$, with a median stellar mass growth factor over all BCGs of 3.5. Interestingly, Ragone-Figueroa et al. (2018) also find that *in-situ* star formation is not entirely negligible; it accounts for nearly 25 per cent of the growth in stellar mass between $z = 2$ and $z = 1$. These minor revisions aside, the central paradigm – of early formation of the stellar mass and late-time assembly of the BCG – remains intact.

Recent observations, however, suggest that the late-assembly picture may not be as concrete as once thought, and that early assembly may play a major role in the evolution of the BCGs. Several BCGs have been discovered at $z \sim 1 - 1.5$ that have stellar masses comparable to the most-massive galaxies in the Universe (Collins et al. 2009). If these BCGs were to grow at the rates theoretically expected via late-time hierarchical assembly, they would greatly exceed the predicted masses of those theoretical models. Additionally, the debate on the size evolution of BCGs is far from settled, as there is also evidence that the luminosities and sizes of the BCGs, as a population, do not evolve much past $z \sim 1$ (Whiley et al. 2008; Stott et al. 2011), suggesting that little growth through the hierarchical scenario is possible. The absence of observed evolution past $z \sim 1$ implies that these massive galaxies must grow via a combination of *in-situ* star formation and early assembly. With respect to the former, Webb et al. (2015) analyse a set of BCGs in the *Spitzer* Adaptation of the Red-Sequence Cluster Survey (SpARCS; see Muzzin et al. 2009; Wilson et al. 2009) and find that a large contribution to the overall growth of the BCGs must be due to *in-situ* star formation based on the estimated star formation rates (SFRs) of hundreds of BCGs (in the range $0.8 < z < 1.8$), and also find an increasing SFR with increasing redshift. As to the latter, there is growing evidence of highly overdense protocluster cores (e.g. Ishigaki, Ouchi & Harikane 2016; Jiang et al. 2018; Miller et al. 2018; Higuchi et al. 2019; also see Ito et al. 2019, and Overzier 2016 for a broad census review) at high redshifts ($z \gtrsim 4$). As we demonstrate in Section 2, protocluster cores with a high density of galaxies are the birthplaces of BCGs, and highly overdense systems should collapse rapidly in the ΛCDM theory – casting into doubt whether the theoretical consensus of late assembly is valid for the entire population of BCGs.

In this paper, we investigate the above tension between the current theoretical picture and accumulation of observational results in order to gain insight into the evolution of the BCG population. Specifically, we use a bespoke non-cosmological simulation based on the observed parameters of the SPT2349–56 protocluster (Miller et al. 2018) to track its forward evolution. Our interest lies in determining the future evolution of the protocluster core – including the fate of the observed galaxies – and the time-scale of its evolution. We then use the MultiDark Planck 2 Bolshoi simulation (Riebe et al. 2013; Klypin et al. 2016) a large-volume non-baryonic simulation, to estimate the frequency of similar events in the universe. While dark-matter simulations exist that provide ample resolution and population statistics (through their large volumes) for discovering overdense protoclusters at high redshift, simulating the equivalent volumes in tandem with the hydrodynamical equations of motion and galactic-baryonic physical processes is, at present, not feasible due to computational constraints. These constraints force us to study the forward evolution of SPT2349–56 in the bespoke simulation. In Section 2, we describe our set-up and initial conditions for the SPT2349–56 simulation. In Section 3, we discuss the assembly and growth of the system. In Section 4, we analyse a large volume dark-matter-only simulation in order to determine how frequent such highly overdense events may be. Lastly, we synthesise our findings and present a revised paradigm for the formation and the evolution of the BCGs in Section 5.

2 METHODOLOGY

We start by constructing a bespoke simulation of the SPT2349–56 system in order to specifically study its forward evolution. Observations indicate that the 14 galaxies that comprise the core of the protocluster are within a 130 kpc (physical) projected region on the sky at a mean redshift of $z \approx 4.3$, and we show the observed physical properties of each of the 14 galaxies in leftmost two columns of Table 1. The observed line-of-sight velocity distribution, $\Delta V_{\rm LOS}$, was found to approximate a Gaussian distribution with $\sigma_{\rm LOS} = 408\,{\rm km\,s^{-1}}$ (Miller et al. 2018). The cold gas masses, $M_{\rm gas,\,gal}$, were estimated by converting the estimated[1] CO(1–0) line

[1]Miller et al. (2018) measured the CO(4–3) line luminosity and then converted to a CO(1–0) luminosity by using the ratio of the line brightness temperatures, calculated from a sample of sub-millimetre galaxies with measurements of both lines.

Table 1. Observed and estimated physical properties of SPT2349–56 (Miller et al. 2018). We use only the observed line-of-sight offset velocities ($\Delta V_{\rm LOS}$) and cold gas masses ($M_{\rm gas,\,gal}$) to estimate the remaining physical parameters in this table. $M_{*,\,\rm gal}$ is the estimated galactic stellar mass, $C_{\rm vir}$ is the estimated NFW halo concentration, $M_{\rm vir}$ is the estimated virial mass of each halo, and $V_{\rm vir}$ is the estimated virial velocity of the galaxy's host halo.

Label	$\Delta V_{\rm LOS}$ (km s^{-1})	$M_{\rm gas,\,gal}$ (10^{10} M$_\odot$)	$M_{*,\,\rm gal}$ (10^{10} M$_\odot$)	$M_{\rm vir}$ (10^{10} M$_\odot$)	$C_{\rm vir}$	$V_{\rm vir}$ (km s^{-1})
A	−90	12.0	5.14	514	1.67	537
B	−124	11.2	4.79	479	1.69	524
C	603	6.7	2.87	287	1.81	442
D	−33	8.4	3.6	360	1.75	477
E	84	4.8	2.05	205	1.89	395
F	395	3.4	1.46	146	1.97	353
G	308	1.6	0.685	68.5	2.18	274
H	−719	4.4	1.88	188	1.91	384
I	310	2.2	0.942	94.2	2.09	305
J	481	2.2	0.942	94.2	2.09	305
K	631	3.1	1.33	133	2.00	342
L	−379	3.3	1.41	141	1.98	348
M	34	1.2	0.514	51.4	2.26	249
N	90	1.0	0.428	42.8	2.31	234

luminosity to a gas mass using a conservative estimate of the $\alpha_{\rm CO}$ conversion factor, $\alpha_{\rm CO} = 0.8\ {\rm M}_\odot\ /\ ({\rm K\,km\,s^{-1}\,pc^2})$.

In order to simulate the forward evolution of the system, we use a modified version of `GIZMO`[2] (Hopkins 2015), a publicly available gravity plus hydrodynamics simulation program, that is equipped with an implementation of the mesh-free finite mass method (Lanson & Vila 2008a, b; Gaburov & Nitadori 2011).

We simulate the protocluster in isolation with vacuum boundary conditions. Specifically, we simulate the initial coalescence phase of the separate systems (i.e. galaxies with their own dark matter halo, gas, and stellar components) that make up the SPT2349–56 protocluster core. At this early phase, we do not expect the cluster-scale dark matter envelope to be in place. The simulation is non-cosmological, and evolves the equations of motion for 1 Gyr. Gas and star particles in the simulation have an initial mass of $M_{\rm gas} = M_* = 10^6\ {\rm M}_\odot$. Gas properties are calculated using the cubic spline kernel with 32 neighbouring particles in our simulations. The dark matter particle mass is $M_{\rm dark} = 5 \times 10^6\ {\rm M}_\odot$. Additionally, we seed each galaxy with a black hole of mass $M_{\rm BH} = 10^5\ {\rm M}_\odot$. We use adaptive gravitational softening for all gravitationally interacting particles (Hopkins et al. 2018), which requires minimum softening parameters. Baryonic particles have a minimum softening of $\epsilon_{\rm b,min} = 50$ pc and, for dark matter, we use a minimum softening of $\epsilon_{\rm dark,min} = 200$ pc.

2.1 Initial conditions

We assume a Planck Collaboration XVI (2014) cosmology throughout the following procedure, and generate initial conditions using the `makegalaxy` software (Hernquist 1993; Springel & White 1999; Springel 2000; Springel, Di Matteo & Hernquist 2005). Each galaxy in our synthetic SPT2349–56 system consists of a dark matter halo, gas disc and stellar disc, with no stellar bulge or surrounding gaseous circumgalactic medium. Given the high overdensity of the SPT2349–56 system, we expect that the large-scale background dark matter field of the surrounding region contributes much less to the mass budget within 130 kpc compared to the individual systems. Therefore, we do not include an extended dark-matter field as the dynamics within the 130 kpc region should not be affected. `makegalaxy` employs the methods of Springel et al. (2005) to create stable spiral galaxies, which we summarize here. We model the dark matter distribution in each galaxy with a Hernquist (1990) profile where the scale length is related to the corresponding Navarro–Frenk–White (NFW) concentration $C_{\rm vir}$ of the halo (Navarro, Frenk & White 1997). We model the gas and stellar discs with exponentially declining surface densities with the scale length H related to the angular momentum (through the spin parameter λ) of the system. We follow Robertson et al. (2006) and use $\lambda = 0.033$, which is the mode of the spin distribution from cosmological simulations (Vitvitska et al. 2002). The combined gas + stellar disc is also under the condition that the total disc mass is a fixed fraction of the total mass of the system, i.e. $M_{\rm disc} = m_{\rm d} M_{\rm vir}$. To ensure disc stability, we choose $m_{\rm d} = 0.03$ for each disc as values between $0.03 \lesssim m_{\rm d} \lesssim 0.05$ lead to stable discs in the ΛCDM cosmology (Mo, Mao & White 1998). The vertical structure of the stellar disc is that of an isothermal sheet with a radially constant scale height z_0 given as a free parameter proportional to the scale length of the disc, which we assume to be $z_0 = 0.1H$. For the vertical structure in the gas disc, we set the scale height such that hydrostatic equilibrium is enforced. We tested each galaxy in isolation to ensure that the system is physically and numerically stable, and find an average SFR of $\approx$75 M$_\odot$ yr^{-1} over 1 Gyr, with a peak of $\approx$140 M$_\odot$ yr^{-1}, for the most-massive galaxy.

In order to set-up each galaxy, we require estimates of their dark matter halo, stellar, and gas masses as well as the concentrations of the dark matter haloes. We base our following calculations on the observed total galactic cold gas mass, $M_{\rm gas,\,gal}$, shown in Table 1. We prepare and simulate three separate realisations of the SPT2349–56 system.

2.1.1 *Masses*

We estimate the dark matter halo mass of each individual galaxy in the protocluster core by first assuming a reasonable gas fraction $f_{\rm gas,\,gal}$, computing the corresponding galactic stellar mass $M_{*,\,\rm gal}$, and then estimating the halo virial mass $M_{\rm vir}$ from $M_{*,\,\rm gal}$.

We define the gas fraction as

$$f_{\rm gas,gal} \equiv \frac{M_{\rm gas,gal}}{M_{\rm gas,gal} + M_{*,\rm gal}} \quad (1)$$

[2] http://www.tapir.caltech.edu/ phopkins/Site/GIZMO.html

and, therefore, the stellar mass is

$$M_{*,\mathrm{gal}} = \left(\frac{1}{f_{\mathrm{gas,gal}}} - 1\right) M_{\mathrm{gas,gal}}. \quad (2)$$

We assume all of our galaxies have the same gas fraction, $f_{\mathrm{gas,\,gal}} = 0.7$, estimated from the results in Narayanan, Bothwell & Davé (2012) and Tadaki et al. (2019). This assumption is reasonable for high-redshift galaxies, where gas fractions $f_{\mathrm{gas}} > 0.4$ are routinely inferred, with a large spread above this value (Carilli et al. 2010; Daddi et al. 2010; Tacconi et al. 2010, 2013). In fact, the median gas fraction of protocluster galaxies measured in Tadaki et al. (2019) is $f_{\mathrm{gas}} \approx 0.77$, comparable to our value, with a spread from $f_{\mathrm{gas}} \approx 0.4$ to $f_{\mathrm{gas}} \approx 0.9$. We discuss the impact of varying f_{gas} in Section 3.

As for relating M_{vir} and $M_{*,\,\mathrm{gal}}$, Behroozi, Wechsler & Conroy (2013b) show that the stellar-to-halo mass fraction is $M_{*,\,\mathrm{gal}}/M_{\mathrm{vir}} \sim 0.01$, within a factor of 2 for a wide range of haloes at $z > 4$ that eventually become clusters of mass $\sim 10^{15}\,\mathrm{M}_{\odot}$, the predicted $z = 0$ mass of the SPT2349–56 system (see Fig. 2 of Miller et al. 2018). We expect that altering the stellar mass to halo mass ratio by a factor of $\sim$2 will change our results by the same factor. We do note that the results in Behroozi et al. (2013b) apply to central galaxies; however, we assume that the 14 galaxies in the SPT2349–56 are the centrals of their own haloes before commencing to merge, coalesce, and form the protocluster core region. Using these results, we estimate $M_{*,\,\mathrm{gal}} \approx 0.428 M_{\mathrm{gas,\,gal}}$ and $M_{\mathrm{vir}} \approx 42.8 M_{\mathrm{gas,\,gal}}$ to within a factor of 2. We show the results of our calculations in Table 1.

2.1.2 Halo properties

Having an estimate of the virial mass of each system allows us to calculate the halo concentrations and virial velocities. Using the universal halo concentration model from Bullock et al. (2001), the concentration parameter is

$$C_{\mathrm{vir}}(M_{\mathrm{vir}}, z) \approx 9\left(\frac{M_{\mathrm{vir}}}{M_{\mathrm{coll},0}}\right)^{-0.13} (1+z)^{-1}, \quad (3)$$

where $M_{\mathrm{coll},\,0}$ is the typical collapsing halo mass at $z = 0$, with $M_{\mathrm{coll},0} = 1.18 \times 10^{13}\,\mathrm{M}_{\odot}$. Diemer & Kravtsov (2015) show that, for halo masses in our regime of interest ($M_{\mathrm{vir}} \lesssim 4.42 \times 10^{12}\,\mathrm{M}_{\odot}$), equation (3) is an excellent approximation to their more general models. The virial velocity follows from the virial mass as

$$V_{\mathrm{vir}} \approx (10 G M_{\mathrm{vir}} H(z))^{1/3}, \quad (4)$$

where $H(z) = H_0\sqrt{\Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{\Lambda,0}}$. We show the results of equations (3) and (4), applied to each halo, in Table 1.

2.1.3 System dynamics

For the dynamical evolution of the entire system, we require the initial positions and velocities of each galaxy. We select initial positions randomly within a sphere of physical radius 65 kpc (the observed maximal separation) for each galaxy with no dependence on the true separations between the observed objects. Fig. 1 shows a qualitative view of the positioning of each galaxy in one of the realisations of our synthetic SPT2349–56 system. The observed velocity distribution in Table 1 provides the initial velocities of our simulated galaxies. We fit a Gaussian to the distribution and then sample 14 new velocity components for each spatial direction, assuming the velocity distribution is isotropic. We randomly select the orientation of the spin axes for each galaxy.

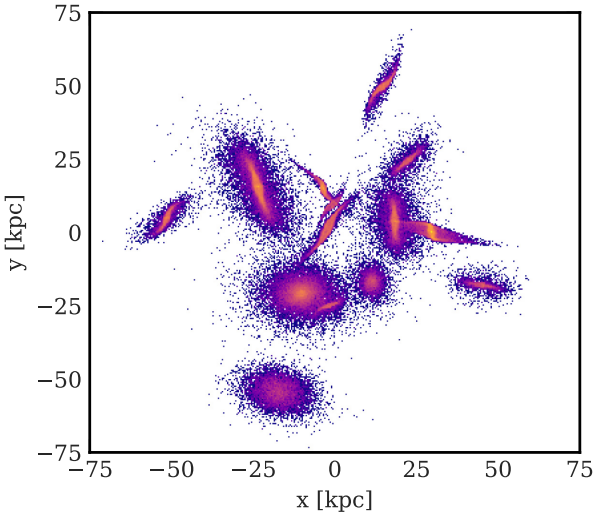

Figure 1. A schematic view of one of the realisations of our synthetic SPT2349–56 system. We bin the *xy*-plane positions of each stellar particle with an arbitrary logarithmic scaling at 1 Myr after the initial condition. We constrain the mass-weighted centres of each galaxy to lay within a sphere of radius 65 kpc (physical) based on the observed projected separations.

2.2 Galactic physics

Our sub-grid physics models are the same as those described in Rennehan et al. (2019) (based on the model in Davé, Thompson & Hopkins 2016), except that we now include a model for supermassive black hole (SMBH) growth and feedback. We briefly describe the models below and point the reader to the aforementioned reference for more information.

2.2.1 Cooling and star formation

For radiative cooling, we calculate the cooling rates in the presence a ultraviolet background (Faucher-Giguère et al. 2009) using the `GRACKLE-3.1` cooling library[3] (Smith et al. 2017).

Our star formation implementation follows that in the MUFASA simulations (Davé et al. 2016, 2017). We determine the conversion rate of gas into stars based on the estimated fraction of molecular hydrogen ($f_{\mathrm{H_2}}$) in the gas based on the approximations in Krumholz, McKee & Tumlinson (2009). We convert gas at densities above the threshold $n_{\mathrm{crit}} = 0.2\,\mathrm{cm}^{-3}$ into stars at a rate $\mathrm{d}\rho_*/\mathrm{d}t = \epsilon_* f_{\mathrm{H_2}} \rho_{\mathrm{gas}}/t_{\mathrm{dyn}}$ where ρ_* is the stellar density, $\epsilon_* = 0.02$ is star formation efficiency (Kennicutt 1998), and $t_{\mathrm{dyn}} = (G\rho_{\mathrm{gas}})^{-1/2}$ is the local dynamical time. We also force gas on to an artificial equation of state, $T_{\mathrm{EoS}} = 10^4 (n_{\mathrm{gas}}/n_{\mathrm{crit}})^{1/3}$ K, where n_{gas} is the gas hydrogen number density, above the star formation critical density ($n_{\mathrm{gas}} > n_{\mathrm{crit}}$) to suppress numerical fragmentation (Teyssier et al. 2011; Davé et al. 2016).

2.2.2 Stellar feedback

We include energetic feedback from supernova Types Ia and II (SNIa and SNII, respectively), stellar radiation, and stellar winds from asymptotic giant branch (AGB) stars based on the MUFASA cosmological simulation model. We also include mass injection from SNIa, SNII, and AGB stars, which is important for enriching

[3] https://grackle.readthedocs.io

the gas in the simulation (Davé et al. 2016; Liang et al. 2016). We account for the effects of both prompt and delayed SNIa (Scannapieco & Bildsten 2005).

Metals are vital in determining the balance of gas cooling and heating in astrophysical gas, and therefore we include their production and account for their role in cooling. We consider metal production by SNIa, SNII as well as AGB stars (Iwamoto et al. 1999; Nomoto et al. 2006; Oppenheimer & Davé 2008). For details, we refer the reader to Liang et al. (2016), Davé et al. (2016), and Rennehan et al. (2019).

2.2.3 *AGNs*

High luminosity galaxies often host AGNs concurrently with intense starburst episodes in the local Universe (Nardini et al. 2008), and at early epochs (Alexander et al. 2005). Therefore, we also include the effects of AGN feedback into our investigation. AGN are important in determining the correct estimate of stellar mass growth. Our model is that of Springel et al. (2005), which we briefly describe below.

We initially place black holes of mass 10^5 $M_\odot$ in the centres of each galaxy, and allow them to grow via Eddington-limited Bondi accretion. We use the unboosted Bondi model because the mesh-free finite mass method can resolve higher densities at the same mass resolution, compared to the common smoothed particle hydrodynamics implementations (Hopkins 2015). We also include energetic feedback, and assume that each AGN generates energy in the gas at a rate $\dot{E} = \epsilon_r \epsilon_f \dot{M}_{BH} c^2$, where $\epsilon_r = 0.1$ is the radiative efficiency, $\epsilon_f = 0.05$ is the coupling fraction to the gas, and c is the speed of light. The accretion rate $\dot{M}_{BH} = 4\pi G^2 M_{BH}^2 \rho_{gas}/(c_s^2 + v_{rel}^2)^{3/2}$ is the Bondi accretion rate on to the black hole, where c_s is the surrounding gas sound speed, v_{rel} is the relative velocity of the SMBH with respect to the gas, and ρ_{gas} is the surrounding gas density. The gas properties – density, sound speed, and relative velocity – are calculated over the nearest 128 neighbouring gas particles. The energy is deposited to the surrounding gas in a kernel-weighted manner, over the same nearest 128 neighbouring particles. To follow the dynamical evolution of the SMBHs, we use the model from Tremmel et al. (2015) in which the dynamical friction force is calculated by using the approximation from Chandrasekhar (1943).

3 STELLAR ASSEMBLY AND GROWTH

To gain a qualitative understanding of the protocluster assembly, we examine one realisation of the system visually in Fig. 2. We bin the positions of each star particle in the simulated *xy*-plane at three times: $t = 0.12$, 0.5, and 1 Gyr, from top to bottom, respectively.

In the top panel of Fig. 2, there are several stellar streams protruding through the system as the galaxies undergo the initial collapse after ∼120 Myr. These streams are due to tidal stripping from the companion galaxies as the initial velocity dispersion of the system, combined with the close proximity of galaxies, is unable to prevent imminent merging. At maximum distance, the tidal tails extend approximately 90 kpc.

In the middle panel, 500 Myr after the start of the simulation, several streams are visible in addition to shell-like structures surrounding the core of the galaxy. It is difficult to distinguish any of the original structure as the stellar populations begin to mix. At this point during the simulation, many of the stellar particles are launched out to ≈75 kpc from the centre-of-mass of the system, building up the diffuse stellar envelope and intracluster light.

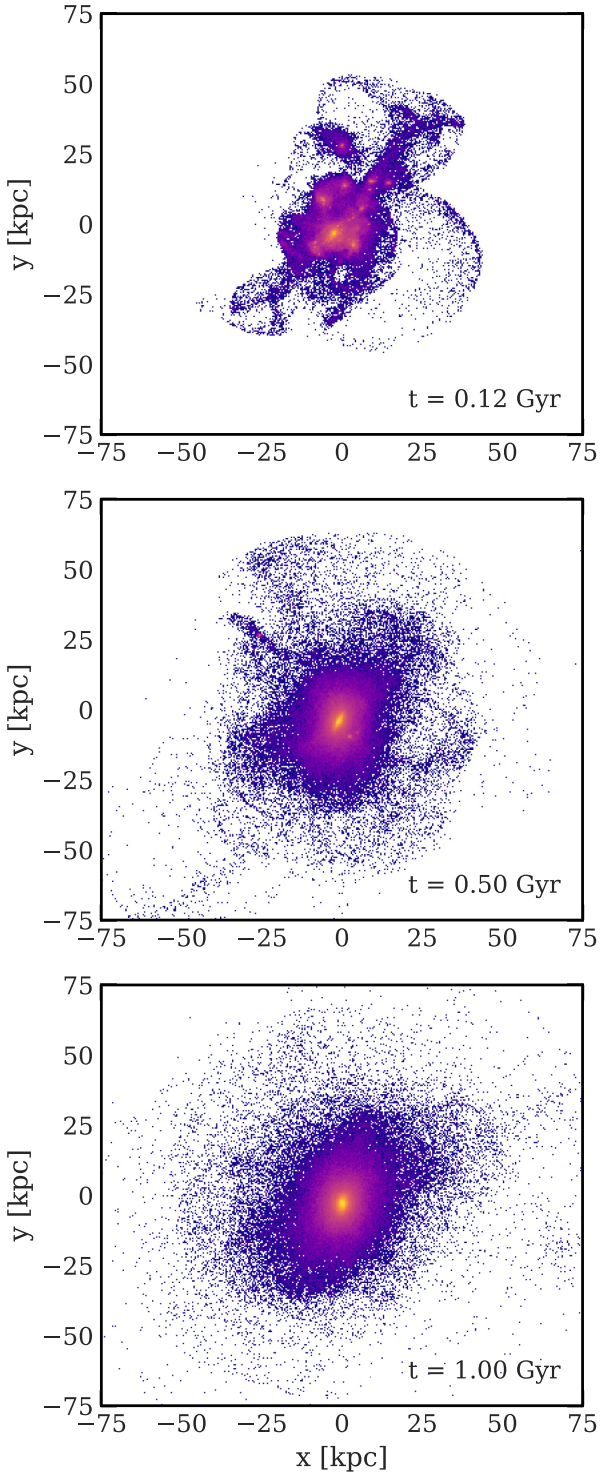

Figure 2. A particle view of the assembly of one of the realisations of our synthetic SPT2349–56. We bin the positions of the stellar particles in the simulation in the *xy*-plane (150 kpc × 150 kpc) using an arbitrary logarithmic scaling. These panels represent $t = 0.12$, 0.5, and 1 Gyr from the initial condition, from top to bottom, respectively. There are obvious shell-like structures and stellar streams throughout the short assembly period, until the system resembles a massive elliptical galaxy at ∼500 Myr.

At 1 Gyr in the bottom panel of Fig. 2, the system resembles a massive elliptical galaxy. Many of the stars from the initial galaxies, as well as those formed *in situ*, were kicked out and formed an extended diffuse stellar halo. There is no longer any visible structure in this halo. Using radiative transfer we are able to further investigate the evolution of the system.

Fig. 3 shows colour-composite mock *James Webb Space Telescope* (*JWST*) observations of one of our synthetic SPT2349–56 realisations. We generated these observations using the SKIRT radiative transfer code (Baes et al. 2011; Baes & Camps 2015; Camps & Baes 2015) and an adapted version of the observational realism suite described in Bottrell et al. (2017a), Bottrell et al. (2017b). To produce these images, we first assigned star particles spectral templates based on the STARBURST99 (Leitherer et al. 1999) spectral energy distribution set for old stellar populations and Groves et al. (2008) templates, which include emission from HII and photodissociation regions, for young (<10 Myr-old) star particles. We use a multicomponent dust model (Zubko, Dwek & Arendt 2004) with a constant dust-to-metal ratio of 0.3 and do not limit dust to star-forming gas particles. SKIRT produced rest-frame optical data cubes that we processed into noiseless, idealised photometric images in the *JWST* NIRCam *F*150*W*, *F*200*W*, *F*277*W*, *F*356*W* band passes[4] at redshifts corresponding to the simulation snapshots. In the right-hand column of Fig. 3, we show idealised images. The idealised images are noiseless and are neither rebinned down to the NIRCam angular resolution nor convolved with the NIRCam point-spread function. At each of the redshifts, we consider $z \in \{4, 3.3, \text{and } 2.7\}$, corresponding to $t \in \{0.12, 0.5, \text{and } 1\}$ Gyr, at least three of these filters reside in the protocluster's redshifted rest-frame optical domain.

We assigned the filtered light at the smallest wavelengths to the blue values, the mid-range wavelengths to the green, and the longest to the red to construct a qualitative, visual representation of the system from $t \sim 120$ Myr to $t \sim 1$ Gyr (from top to bottom in Fig. 3, respectively).[5]

In the top right-hand panel of Fig. 3, we show $t \sim 120$ Myr after the initial condition. Already by this point in the system's evolution, we see the long stellar streams from the top panel of Fig. 2 as a spatially extended low surface brightness web surrounding the remnant cores of each initial galaxy. Several bright star-forming cores remain visible in blue and white. Initial gravitational torques cause the fragmentation we observe in the protocluster early on due to the close proximity of the 14 galaxies.

At 500 Myr in the centre row on the right-hand side, there are a few remnant cores of the original galaxies and the system assumes a more spherical shape in the diffuse stellar halo. By this point in time, the gas settles into an extended disc in the centre of the gravitational potential, and we see the system shift toward a redder appearance. We analysed the simulation data and discovered that the disc only exists from $t \sim 300$ Myr to $t \sim 600$ Myr, before the remaining cores of the original galaxies dynamically pummel the disc, leading to its demise. It is not surprising that there is a disc given the high gas fraction of the system (Hopkins et al. 2009).

At 1 Gyr, the system is elliptical and resembles a typical low-redshift cD galaxy. Although the disc vanishes, star formation remains at an absolute rate of $\sim 40\ \mathrm{M_\odot\ yr^{-1}}$, or a specific rate of $5 \times 10^{-11}\ \mathrm{yr^{-1}}$, much lower than values predicted in Webb et al. (2015) for BCGs at slightly lower redshift ($z \sim 2$).

In addition to the idealised mock observations in the right-hand column of Fig. 3, we show mock observations of the protocluster in the left-hand column. We rebin the idealised images to the NIRCam CCD angular scale corresponding to the appropriate channel (0.031 arcsec pixel^{-1} for *F*200*W* and 0.063 arcsec pixel^{-1} for the others), point-spread function convolution, and added noise corresponding to the predicted NIRCam surface-brightness sensitivity for our observing strategy. Using the simulated ramp-optimization feature of the *JWST* pynrc package (Leisenring et al., 'pyNRC: A NIRCam ETC and Simulation Toolset', in preparation.), we determined the optimal observing strategy for the protocluster assuming 10 ks of observing time. For more details, see Appendix A. We analyse these results at three selected redshifts: $z \in \{4, 3.3, \text{and } 2.7\}$, from top to bottom in Fig. 3, respectively. These leftmost panels correspond to the same redshift as the idealised images in the same row in Fig. 3.

In the top left-hand panel, the substructure at $z \approx 4$ is clearly visible in the *JWST* composite image. We find that the absolute AB-magnitude in the *F*277*W* NIRCam band is $M_{\mathrm{AB},\,F277W} \approx -28.7$, making our synthetic SPT2349–56 an extremely bright object. In an analysis of 430 BCGs in Donzelli, Muriel & Madrid (2011), the brightest galaxy (object A0401 in the table 2 of that study) has an absolute *R*-band magnitude of −27.25. Although our filters are not equivalent, we find the same approximate magnitude order in all of the *JWST* filters we apply and, therefore, we can conclude that SPT2349–56 would end up being one of the brightest galaxies in the observable Universe. Not only is our realisation bright, it is also very blue. We calculated the rest-frame SDSS $g - r$ colours for the system and find $g - r \approx -0.05$ at $z \approx 4$ for our dust model. A typical BCG has a redder colour at $g - r \approx 0.6$ (Cerulo, Orellana & Covone 2019).

The qualitative view of the BCG at a redshift of $z \approx 3.3$ corresponds to 500 Myr. At this time, the stars remain blue in the system ($g - r \approx 0.13$), which is interesting since we do not expect fully assembled, highly star-forming BCGs at $z \approx 3.3$ with a large fraction of young stellar populations. If the predictions of substantial (i.e. factors of ∼2 to ∼4) growth past $z \sim 1$ in the literature (De Lucia & Blaizot 2007; Lavoie et al. 2016; Ragone-Figueroa et al. 2018) are correct, then systems such as SPT2349–56 would go on to become the most-massive galaxies in the Universe at $z = 0$, reaching stellar masses up to $\sim 3 \times 10^{12}\ \mathrm{M_\odot}$. We do not include any of the substructure that may exist outside of the core of the SPT2349–56 protocluster, which could increase our estimate even further if those galaxies were to accrete between $z \sim 3$ and $z \sim 1$ via dynamical friction (i.e. galactic cannibalism; Ostriker & Tremaine 1975; White 1976). However, since they grow rapidly to a large mass, the dynamical friction time-scale for lower mass satellites to merge into the system becomes large, which could hamper late-time growth.

In the top panel in Fig. 4, we show the forward-evolution of the SFR of our synthetic SPT2349–56 system (i.e. our fiducial simulation with $f_{\mathrm{gas,\,gal}} = 0.7$) as a function of time, averaged over three realisations, up to 1 Gyr ($z \approx 2.7$) after the observation ($z \approx 4.3$).[6] The SFR peaks at $\sim 3000\ \mathrm{M_\odot\ yr^{-1}}$ approximately 5 Myr after the initial condition and decays exponentially to $\sim 40\ \mathrm{M_\odot\ yr^{-1}}$ at 1 Gyr. Our simulated SFR is comparable to the observed results of Miller et al. (2018) who find a total SFR of $\approx 6500\ \mathrm{M_\odot\ yr^{-1}}$ for the

[4] https://jwst-docs.stsci.edu/near-infrared-camera/nircam-instrumentation/nircam-filters

[5] Specifically, we use the NIRCam filters {*F*200*W*, *F*277*W*, *F*356*W*} at 120 Myr, and {*F*150*W*, *F*200*W*, *F*277*W*} at both 500 Myr and 1 Gyr.

[6] In terms of redshift, our simulation began at $z \approx 4.3$ and ended 1 Gyr later at $z \approx 2.7$, assuming a Planck Collaboration XVI (2014) cosmology.

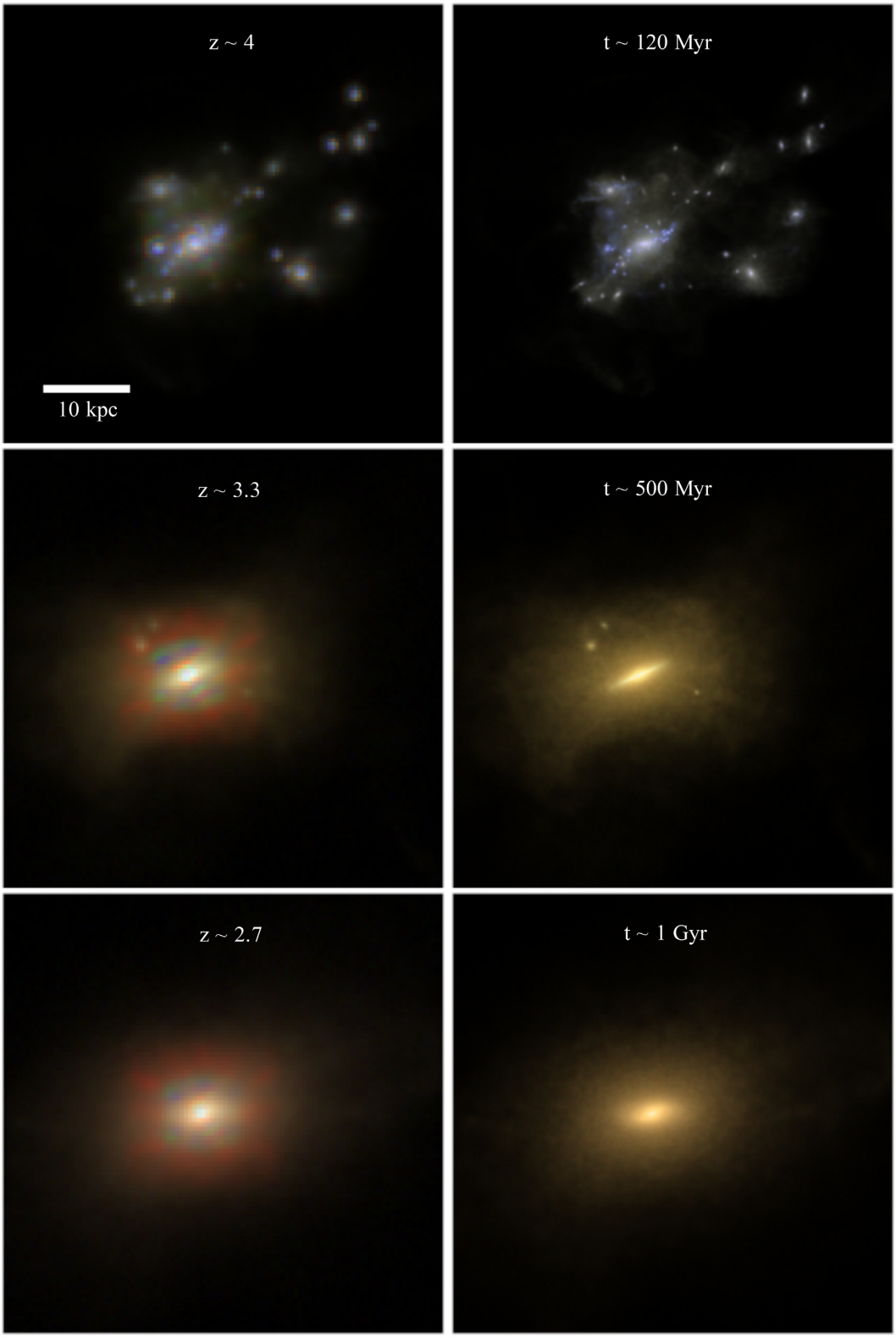

Figure 3. (left-hand column) *JWST* mock observations of the forward evolution of SPT2349–56. Redshifts are $z \sim 4$, 3.3, and 2.7, from top to bottom, respectively. (right-hand column) Synthetic false colour images made from three of the *JWST* NIRCam filters without the CCD angular scale or noise. The snapshots correspond to the redshift labels in the left-hand column, for each row. The scale marker shows 10 kpc (physical) and each panel is 50 kpc per side.

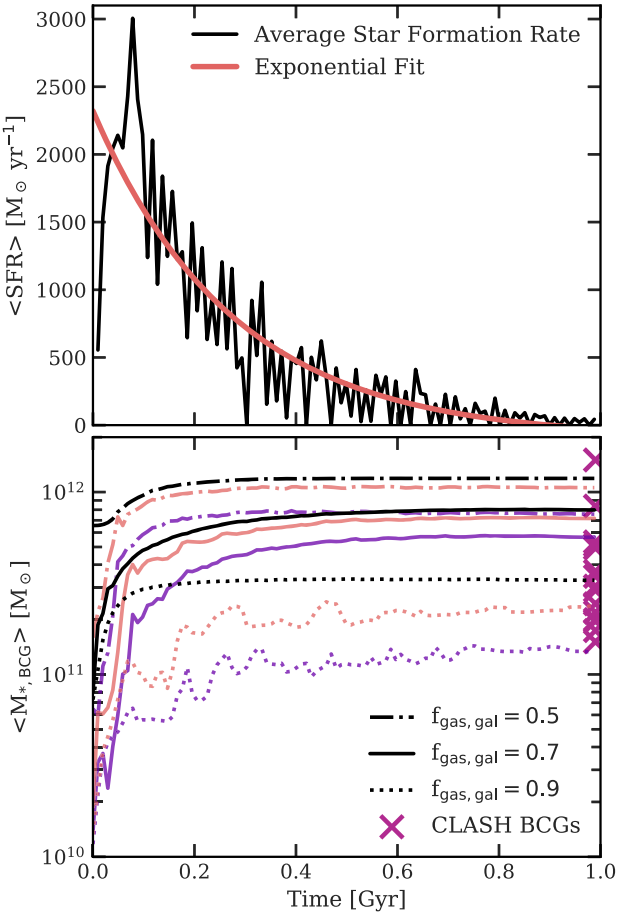

Figure 4. Growth and assembly history of the protocluster simulations. (Top panel) SFR over the first giga-year of evolution averaged over three independent realisations of our fiducial simulation with initial gas fraction $f_{\rm gas,\,gal} = 0.7$. We used a non-linear least squares method to fit a decaying exponential and found a decay time-scale of $\tau \approx 200$ Myr. (Bottom panel) The black curves show the total mass of all stellar particles in the simulation volume for each system with different initial gas fraction. The purple and coral curves show the stellar mass within 5 and 15 kpc apertures, centred on the peak projected stellar mass, respectively. We show the observed masses of BCGs from the CLASH survey as × symbols at 1 Gyr which were derived using MAG_AUTO magnitudes. After 1 Gyr ($z \approx 2.7$), the total stellar masses across the three simulations with varying gas fractions is larger than most of the lower redshift CLASH BCGs.

14 galaxies in the SPT2349–56 system. We use a non-linear least squares fitting method to fit an exponential curve to the SFR, and find that the decay time is $\tau \sim 200$ Myr. Initially, the SFR peaks due to high compression from the proximity of each individual galaxy, and the fact that they are rapidly collapsing under their mutual gravity. Additionally, strong fluctuations in SFR occur as the discs undergo tidal interactions. Stellar feedback is the main cause of the declining SFR, as the black holes provide little feedback past the initial ∼50 Myr due to their low accretion rates. They are unable to accrete because of their high velocities relative to the medium – which suppresses accretion. On the other hand, stellar feedback is most powerful when the SFR is the highest, as the system coalesces, and the supernovae and stellar winds begin to excavate the gas out of the original galaxies.

The balance of stellar feedback and star formation causes a smooth decline in SFR, and consequently a smooth increase in the stellar mass, as we show in the bottom panel in Fig. 4. The total stellar mass of our fiducial simulation is given as the solid black line with initial and final (at 1 Gyr) stellar masses, $M_{\rm initial} = 2.81 \times 10^{11}$ M$_\odot$ and $M_{\rm final} = 8.14 \times 10^{11}$ M$_\odot$, respectively. The total stellar mass is the sum of all of the individual stellar particle masses in the entire simulation volume. The assembly mass, which we define as 90 per cent of the total final mass, is $M_{\rm assembly} = 7.33 \times 10^{11}$ M$_\odot$. The time at which assembly occurs is $t_{\rm assembly} \sim 370$ Myr. Assembly occurs rapidly in the system, and our results show that the system nearly quadruples size within a giga-year.

Apart from the total stellar mass, we also examine the growth of the stellar mass of largest galaxy in the simulation, which we label as the evolving protocluster's BCG. Observationally, the stellar masses of BCGs are usually determined within an aperture with a specified radius since only the projected stellar light distribution is known. Therefore, to compare to observational stellar masses, it is more reasonable that we choose an aperture centred on the peak stellar brightness and sum the stellar mass along the corresponding cylinder. Kravtsov et al. (2018) advocate apertures of fixed radii (they used 30, 50, and 70 kpc) for comparing simulated BCG masses to observed masses. Burke, Hilton & Collins (2015), whose masses we compare to below, adopted aperture sizes for each BCG in the Cluster Lensing and Supernova Survey with Hubble (CLASH: Postman et al. 2012) sample such that it encompassed ∼ 90 per cent of the flux. Stott et al. (2010) used the same method to determine the BCG stellar masses for the systems in the *XMM* Cluster Survey (XCS; Romer et al. 2001; Sahlén et al. 2009), and showed that a fixed 50 kpc aperture gave similar results. A 50 kpc aperture is, however, too large for our high-redshift protocluster since the diameter of the system is only 135 kpc. A 50 kpc aperture centered on the peak in the projected stellar mass distribution[7] – approximating the location of the peak surface brightness – encompasses 95 per cent of the total stellar mass (black curves in Fig. 4) over course of the simulation. To isolate the most massive galaxy in our simulated protocluster and track the growth of its stellar mass over time, we use smaller apertures of 5 and 15 kpc since the effective radius (i.e. the half-light radius) of the stellar distribution is ≈2.2 kpc at 1 Gyr in our fiducial simulation and, therefore, the apertures enclose a few effective radii.

The purple and coral solid curves in the bottom panel of Fig. 4 show the stellar mass growth within the aforementioned cylindrical apertures of radii 5 and 15 kpc apertures, respectively. The smaller apertures more closely track the mass growth in the most massive galaxy across the entire simulation although the merging process is so rapid that distinguishing between the BCG and the debris from infalling, merging, disrupting galaxies past ∼100 Myr is impractical, and the differences between the stellar masses within each aperture become less significant. The masses with the apertures rise rapidly as the galaxies coalesce over the course of the first ∼100 Myr (see the top panel of Fig. 2) and converge to their final values by 200 − 250 Myr. At 1 Gyr, the 5 and 15 kpc apertures contain ≈ 70 per cent and ≈ 90 per cent of the total stellar mass, respectively.

We also show the observed masses of BCGs in Burke et al. (2015) from the CLASH survey with × symbols at 1 Gyr. These BCGs belong to clusters with masses $\gtrsim 10^{15}$ M$_\odot$, and range in redshift from $0.187 < z < 0.890$. Although our system is at $z \sim 2.7$ after 1 Gyr, the total stellar mass and the mass within the smaller apertures are already more massive than most of the BCGs from the CLASH survey. We emphasise that we only mean to compare the observed

[7]We tested multiple viewing angles and found that our results do not change significantly.

BCG masses with our simulated BCG at 1 Gyr, after the system resembles a fully formed BCG (see the bottom panel of Fig. 2).

It is pertinent that we emphasise that we designed our numerical experiment to be simple in nature to determine the SFR and assembly time-scale of the protocluster. We did not tune the SFR to match the observations and, given that we are within a factor of ∼2, the broad assumptions appear reasonable. The main sources of error are (i) our choice for the gas fraction and (ii) the abundance matching results. To test the dependence on gas fraction, we simulated two additional realisations with a lower ($f_{\rm gas,\,gal} = 0.5$) and higher ($f_{\rm gas,\,gal} = 0.9$) gas fraction. We show the total stellar mass evolution in the protocluster system for each initial gas fraction in the bottom panel of Fig. 4.

In the case of $f_{\rm gas,\,gal} = 0.5$, the total initial stellar mass and, therefore, dark matter mass are a factor of ∼2.3 larger than in our fiducial simulation. This is because we fix our gas masses based on the observations in Miller et al. (2018). We find that the total stellar mass of this low gas fraction system at 1 Gyr is $\sim 1.2 \times 10^{12}$ M$_\odot$, which is a factor of ∼1.5 higher than in our fiducial simulation but still much more massive than the majority of CLASH BCGs. The dashed purple and coral curves in Fig. 4 show that the growth within 5 and 15 kpc is rapid in this system and the most massive galaxy quickly grows in size. At 1 Gyr, ≈ 99 per cent of the total stellar mass is contained within the 5 kpc aperture and ≈ 90 per cent is contained within the 15 kpc aperture. Based on our definition of the stellar assembly time, we find $t_{\rm assembly} \sim 150$ Myr which is a factor of ∼2.5 smaller than in the fiducial simulation. Our low gas fraction system therefore reaches a slightly higher stellar mass and does so more quickly.

Next, we consider the simulation with a high gas fraction of $f_{\rm gas,\,gal} = 0.9$. This results in an initial stellar mass a factor of ∼4 lower than our fiducial simulation and, therefore, the same factor lower in halo mass. The total stellar mass at 1 Gyr of the simulated system is $\sim 3.3 \times 10^{11}$ M$_\odot$ – a factor of ≈2.5 lower than our fiducial simulation. This BCG is much less concentrated than the other BCGs with lower gas fractions as we find ≈ 50 per cent and ≈ 85 per cent of the total stellar mass in the apertures of radii 5 and 15 kpc, respectively. We find that the assembly time is a factor of ∼3 shorter than in our fiducial simulation at $t_{\rm assembly} \sim 120$ Myr.

In summary, the above variations in the initial gas fractions result in only factors of 2–3 in the final stellar mass and assembly time. Therefore, we argue that our results are relatively robust; we expect that all scenarios involving factors of ∼2 changes in $f_{\rm gas,\,gal}$ or $M_{*,\,\rm gal}/M_{\rm vir}$ will result in a massive BCG at high-redshift that assembles its stellar mass very quickly (within ∼1 Gyr). We attribute the differences in assembly time-scales between the simulations with different initial gas fractions as due to the varying potential well depth. In the lowest gas fraction case, the potential is deeper and the interactions are stronger, therefore the assembly process occurs on shorter time-scales. In the high gas fraction case, the potential well is shallower and stellar feedback, following the initial intense star burst, leads to the expulsion of gas – quenching the system more rapidly than in the fiducial case.

As for the choice of $M_{\rm vir} \sim 0.01 M_{*,\,\rm gal}$ (see fig. 8 of Behroozi et al. 2013b), for haloes at $z > 4$ which become haloes of mass $\sim 10^{15}$ M$_\odot$ at $z = 0$, the stellar mass to halo mass ratio is ∼0.01 to within a factor of 2. Decreasing (increasing) the ratio would lower (raise) the virial mass of the halo, which, in turn, will alter the dynamics of the system at the level of a factor ∼2−3, similar to what we discussed above.

As an additional source of uncertainty, Granato et al. (2015) point out that it is difficult to reproduce high SFRs in high-redshift protoclusters in numerical simulations using standard sub-grid models of star formation, stellar feedback, and AGNs feedback. However, the sub-grid models we used in GIZMO have been shown to reproduce broad galaxy population properties (Davé et al. 2017, 2019) at a wide-range of redshifts and in high-redshift ($z \sim 6$) galaxies at similar particle mass resolutions (Olsen et al. 2017).

We posit that recently discovered systems similar to SPT2349–56 (e.g. Ishigaki et al. 2016; Jiang et al. 2018; Higuchi et al. 2019) are the proto-cores of the massive galaxy clusters. The speed of assembly and growth of stellar mass in our simulated realisations of SPT2349–56 is obvious from Figs 2–4. The SFR declines exponentially with an *e*-folding time of ∼200 Myr while simultaneously, an object that qualitatively looks like many observed BCGs forms by 500 Myr. From these results, we predict that the observed high-redshift over-dense protoclusters (Ishigaki et al. 2016; Jiang et al. 2018; Miller et al. 2018) will undergo a similar evolution and therefore form massive BCGs as early as $z \approx 4$. As we demonstrated, the *JWST* will be able to clearly see the massive BCGs out to redshift of $z \approx 3$, and their progenitors out to $z \approx 4.3$, opening up a new frontier for exploration, especially in collaboration with survey telescopes such as the *Wide-Field Infrared Survey Telescope* (*WFIRST*).

4 IMPLICATIONS FOR GALAXY CLUSTERS

In the hierarchical structure formation scenario, galaxy clusters are the youngest and most-massive objects in the Universe. In this theory, the more massive the cluster, the younger and rarer the system. Hence, the most-massive clusters should continue to assemble the bulk of their mass until late times. What we have demonstrated in the previous section is the possibility of a downsizing effect (Bower et al. 2006; Cimatti, Daddi & Renzini 2006; Neistein, van den Bosch & Dekel 2006; Fontanot et al. 2009; Oser et al. 2010) on the cluster scale, where the cores of these massive clusters could be much older than the cores of less-massive clusters, beginning to assemble at redshifts $z \gtrsim 2$ at the minimum. Our results on the rapid assembly of the BCG are then not unexpected as these high-redshift protocluster cores would be the strongest relative overdensities in the early Universe. This begs the question: why has the theoretical community not predicted high-redshift fully-assembled BCGs?

One of the issues is that the main tools of contemporary theoretical astrophysicists are numerical simulations. These are necessary because of the strong non-linearity of the structure formation process after the linear perturbation theory breaks down. However, a lack of computational power limits the spatio-temporal dynamic range of the numerical simulations. We are therefore forced into a compromise. We can model a small comoving volume of the Universe and resolve the galaxies in this volume (e.g. Schaye et al. 2015; Pillepich et al. 2018; Davé et al. 2019) but such volumes generally do not contain the rare massive clusters we would expect to host an object such as SPT2349–56. Or, we can sacrifice resolution in favour of large volumes but then galaxy formation must be introduced in an ad hoc fashion (Ruszkowski & Springel 2009), which can introduce biases. Despite these difficulties, we do expect that these protoclusters are present in the largest dark-matter-only simulations, such as the MultiDark Bolshoi (Klypin, Trujillo-Gomez & Primack 2011) and Millennium XXL (Angulo et al. 2012) simulations.

To test our theory of early, rapid assembly of the cores of the most-massive clusters with regards to the general population of galaxy clusters, we shift focus from SPT2349–56 and now investigate the

assembly history of the most massive clusters (at $z = 0$) in the MultiDark Planck 2 (MDPL2) simulation (Riebe et al. 2013; Klypin et al. 2016) – a child of the MultiDark Bolshoi suite of simulations. Our goal is to determine if there is a population of BCGs that assemble early and their relationship to their host clusters. The MDPL2 simulation consists of 3840^3 particles within a simulation volume of side length 1 cGpc h^{-1}, and has a mass resolution of 1.51×10^9 $M_\odot$ h^{-1}. All of the halo data is publicly available online[8] and we specifically use the `MDPL2.rockstar` database in the following analysis. This data base contains halo properties that were determined using the `rockstar` halo finder (Behroozi, Wechsler & Wu 2013a), and includes the substructure trees for each host halo.

First, we require a set of criteria for the occurrence of highly overdense massive collapse events at high redshift. We analysed the formation histories of all galaxy clusters that had final masses[9] of $M_{\rm vir} \geq 5 \times 10^{14}$ $M_\odot$ at $z = 0$, and narrowed our search to those with a large number of relatively massive haloes entering their progenitor's virial radius across cosmic time. Specifically, we define an overdense collapse event to be when $N \geq 5$ haloes of individual mass $M_{\rm vir} \geq 2 \times 10^{11}$ $M_\odot$ all enter the virial radius of a more massive halo within a time Δt, where Δt is the time between two consecutive simulation snapshots. For example, at $z \approx 4.5$ this time-scale corresponds to ∼40 Myr in the `MDPL2.Rockstar` database. We also ensure that there are at least five lower mass haloes that are no more than 20 times less massive than the more massive halo. The latter condition considers only those events in which the substructures have a chance of merging within a *Hubble* time at their respective redshift due to dynamical friction, estimated at the virial radius of the more massive halo (Mo, van den Bosch & White 2010),

$$t_{\rm df}|_{r_{\rm i}=R_{\rm vir}} \approx \frac{1.17}{\ln(M_{\rm large}/M_{\rm sub})} \left(\frac{M_{\rm large}}{M_{\rm sub}}\right) \frac{1}{10H(z)}, \quad (5)$$

where $M_{\rm large}$ is the virial mass of the more massive system, $M_{\rm sub}$ is the mass of the substructure, and $H(z) = H_0\sqrt{\Omega_{\rm m,0}(1+z)^3 + \Omega_{\Lambda,0}}$. Each structure is within a factor of ∼20 of the more massive halo and, therefore, $t_{\rm df}(z) \sim 3t_{\rm H}(z)/4$ where $t_{\rm H}(z) = 1/H(z)$ is the *Hubble* time. We consider the dynamical friction time-scale to be a conservative estimate of how quickly the systems will merge since our criteria captures all mergers with $M_{\rm large}/M_{\rm sub} \lesssim 20$ and the merging haloes are not on circular orbits at the virial radius; therefore, they should merge faster (Poole et al. 2006; Boylan-Kolchin, Ma & Quataert 2008).

We also assume that the substructures and the host each contain a galaxy with stellar mass given by the abundance matching relations in Behroozi et al. (2013b). One important caveat to note is that the results of Behroozi et al. (2013b) apply to central galaxies and not satellites. However, since the substructures are independent haloes before merging, this implies that our criteria is equivalent to a minimum of six central galaxies merging. For comparison, the most massive galaxy in our scenario plays a similar role to galaxy A in Table 1 and in the analysis of Section 3.

We determine the first redshift at which an overdense event occurs for each of the galaxy clusters above our mass limit. The least massive events we find are a factor of ∼20 less massive than our estimated mass for SPT2349–56 and ought to rapidly (if found above $z \gtrsim 3$) collapse into elliptical galaxies of mass $M_* \gtrsim 10^{11}$ $M_\odot$, if we assume the same abundance matching estimates from Section 2 combined with the aforementioned dynamical friction time-scale constraint.

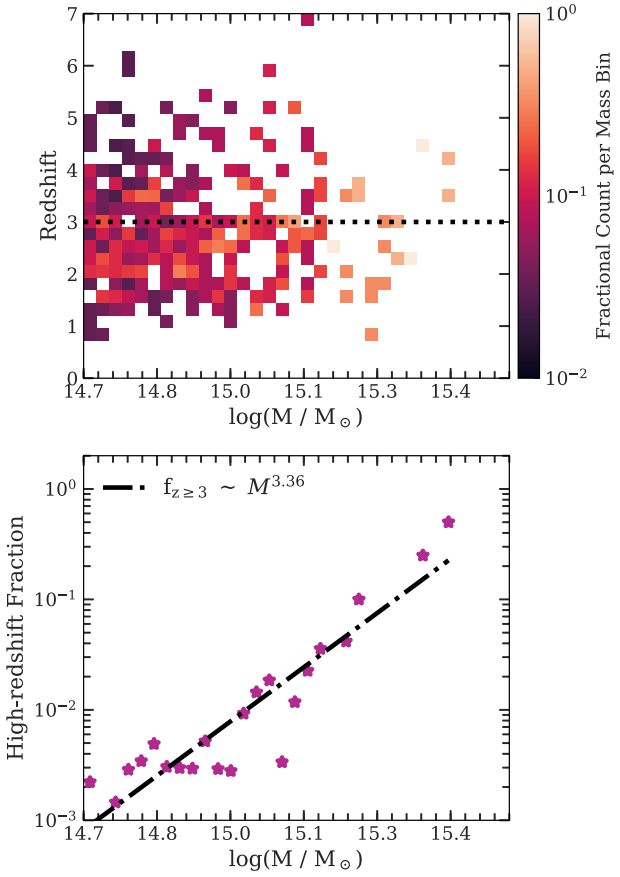

Figure 5. (Top panel) A two-dimensional histogram showing the probability of the first occurrence of a highly overdense region collapsing at a specific redshift, in a given bin of final cluster virial mass (i.e. the virial mass at $z = 0$) spanning from 5×10^{14} $M_\odot$ to 3.25×10^{15} $M_\odot$ (the most-massive cluster). We remove non-events from this histogram so that the fraction is truly the probability of finding an event at a specific redshift, given that such an event occurs. There is a slight trend for more instances of the events at higher final cluster masses but the scatter in the distributions shows that many events occur for all of the clusters here above $z = 3$. (Bottom panel) The number of overdense collapse events above $z = 3$, normalized to the total number of clusters in each mass bin. The bins are identical to those in the top panel. The dash-dot line shows a power-law fit with $f \sim M^{3.36}$ – high-redshift overdense collapse events are more frequent as a function of increasing final cluster mass.

In the top panel of Fig. 5, we present a two-dimensional histogram that encodes the occurrence of high-density collapse regions for a cluster of a given final mass at $z = 0$. Specifically, we count the number of first occurrences of an over-dense event in a grid of redshift and final cluster mass coordinates. We normalise to the total number of events in each mass bin, so that the colouring shows the probability of finding the first event for that final cluster mass at a given redshift, compared to all other events that occur. We do not include non-events in the top panel as our interest lies in the probability of an event occurring at a specific redshift, given all of the events that occur.

From the top-panel in Fig. 5, lower-mass clusters at $z = 0$ are less likely to have a high-density collapse at redshifts $z \gtrsim 3$ (above the dashed line in Fig. 5) compared to clusters with masses $M_{\rm vir} \gtrsim$

[8] https://www.cosmosim.org

[9] Henceforth we assume a Planck Collaboration XVI (2014) cosmology for h in our quoted masses and distances.

10^{15} $M_\odot$. There are, however, clusters that did not experience an overdense collapse at any redshift.

The trend of increasing event frequency with increasing final cluster mass is more evident in the bottom panel of Fig. 5, where we show the number of events above $z \geq 3$ in each mass bin, normalized to the total number of clusters in each mass bin. We show a fit to the fraction as a function of $z = 0$ cluster mass, and find that the fraction scales as $f \sim M^{3.36}$. There is a clear power-law trend where the fraction of events increases with increasing $z = 0$ mass of the clusters. In the highest mass bin, there is a ≈ 50 per cent chance of all clusters having a high-redshift over-dense collapse event occur at $z \geq 3$.

For the present purposes, our interest lies in the objects that collapse in the range $z \gtrsim 3$ since these would end up as the most-massive, blue, elliptical galaxies before $z \sim 1.5$ by our predictions. We find that 16.4 per cent of clusters with final masses $M_{\rm vir} > 10^{15}$ $M_\odot$ have an overdense collapse event occur above $z = 3$, and 9.8 per cent of clusters with final masses 5×10^{14} $M_\odot \leq M_{\rm vir} < 10^{15}$ $M_\odot$. In total, we find 155 high-density collapse events at $z \geq 3$. Given the volume of the simulation, $V \approx 3.2$ cGpc^3, we expect a comoving number density of corresponding massive, blue, elliptical cluster BCGs at redshifts $z \gtrsim 1.5$ to be $n \approx 48$ cGpc^{-3} (or $n \approx 4.8 \times 10^{-8}$ cMpc^{-3}).

In Fig. 6, we examine the spatial, mass, and velocity distribution (from top to bottom, respectively) of an example cluster protocore in the MDPL2 simulation.[10] This protocluster is the most-massive progenitor (MMP) at $z = 4.266$ (the approximate redshift of SPT2349–56) of the second most-massive cluster at $z = 0$.

In the top panel of Fig. 6, we show the spatial distribution of the substructure entering the virial region of the MMP. The dotted line shows the observed extent of the SPT2349–56 object for comparison. Each halo is marked with a triangle that we colour based on the logarithmic mass of the halo, and we arbitrarily scale the size of each triangle to the virial radius of each halo. Furthermore, we mark the centre of mass of the system with a $\times$ symbol. The virial radius of the MMP is $R_{\rm vir,\,phys} \approx 120$ kpc and we see that most of the massive substructure is within the SPT2349–56 sphere, or approximately half of the virial radius of the MMP. In this particular system, there are eight haloes with masses $M_{\rm sub} > 10^{11}$ $M_\odot$, and we assume that they will merge on time-scales shorter than the dynamical friction time-scale taken at the virial radius of the host following equation 5. At $z = 4.266$ the *Hubble* time is $t_{\rm H} \approx 1.43$ Gyr assuming a Planck Collaboration XVI (2014) cosmology and, therefore, $t_{\rm df} \approx 1$ Gyr. Of course, this is a reasonable estimate since the substructure is within half of the virial radius of the more massive system and, therefore, will merge much more quickly. Evidently, overdense systems such as the example in the top panel assemble rapidly in the early universe. However, will they go on to form objects as massive and bright as our synthetic SPT2349–56?

To address this question, we show the mass distribution of our cluster protocore in the middle panel of Fig. 6. As we mention above, there are eight objects above $M_{\rm sub} > 10^{11}$ $M_\odot$, and we first assume that each of these haloes host a galaxy. Using the same methodology as in Section 2.1.1, we estimate that each halo should host a galaxy with stellar mass that is 1 per cent of the halo mass. Therefore, each galaxy in the system should contain, on average, a stellar mass of $M_* \sim 5 \times 10^{10}$ $M_\odot$. If we use a conservative estimate

[10]Specifically, we examine the halo with `rockstarId` = 923730455 in the `MDPL2.Rockstar` table.

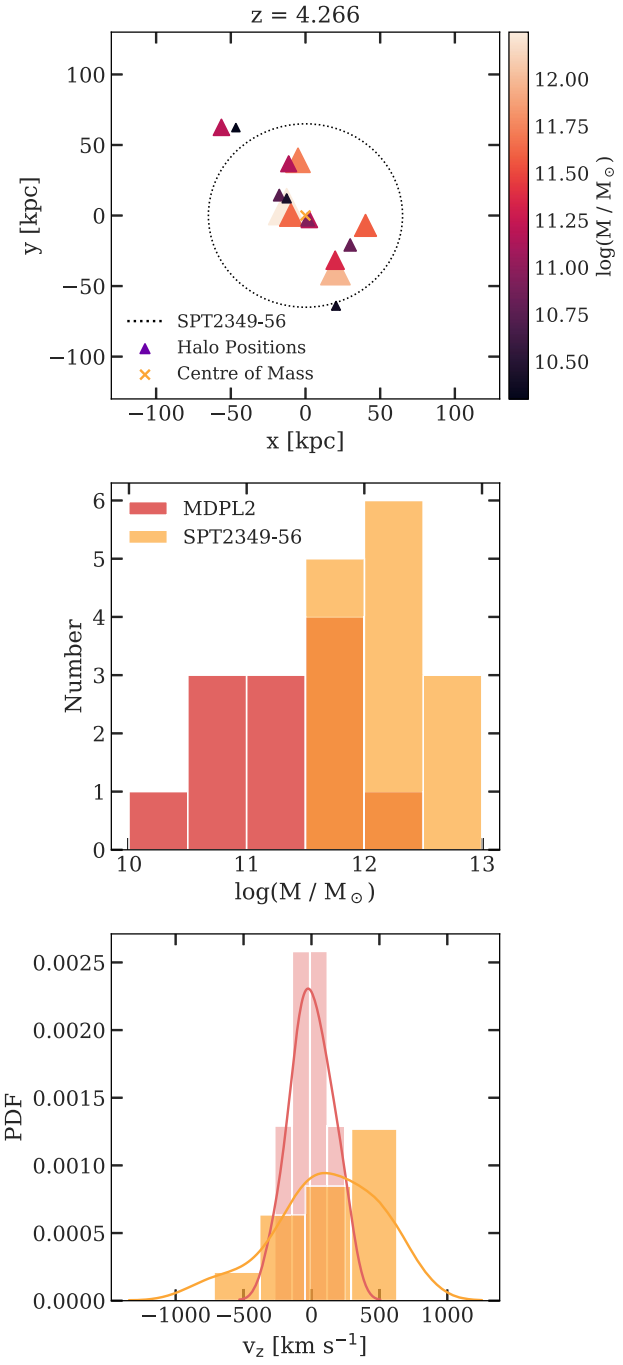

Figure 6. An example of one of the many high-redshift overdense regions we found in the Multi-Dark Planck 2 simulation. This system is the progenitor of the second-most-massive galaxy cluster at $z = 0$, and is approximately an order of magnitude less in mass than our estimation of SPT2349–56. (Top panel) Spatial (physical) positions of substructure within a more-massive halo ($M_{\rm vir} \sim 10^{13}$ $M_\odot$). The physical projection of the observed SPT2349–56 region is overlaid with a dotted line. (Middle panel) Virial mass histogram of substructure compared with the estimated virial masses of the dark matter haloes in SPT2349–56. We consider only those masses that lay within the dotted line. (Bottom panel) Kernel density estimate of the substructure velocities along the line of sight. We show the SPT2349–56 observed values as a comparison. As above, we only consider substructure within the dotted circle in the top panel for the velocity distribution.

of mass doubling within 1 Gyr of evolution based on Section 3, we would expect this object to become an elliptical galaxy with stellar mass $M_* \sim 10^{11}$ M$_\odot$ at $z \approx 2.7$ and, like the BCG in our simulation, it would be extremely bright and blue at that redshift.

In the bottom panel of Fig. 6, we show a comparison of the velocity differences between SPT2349–56 and our example system. We calculate the velocity offset in the perpendicular direction to the top panel, relative to the centre of mass of the system, in order to mimic how Miller et al. (2018) determined the velocities of their object.[11] The solid lines show the kernel density estimate of the underlying distribution for both our system and SPT2349–56. The velocity dispersion of our example system is $\sigma \approx 590$ km s^{-1} and we find a tightly peaked distribution within the extent of the velocity dispersion, indicating that the system is indeed bound, with no signs of filamentary structure within the virial radius. In SPT2349–56, we see that there is a bias toward approaching velocities, indicating that the system is most likely being fed by a filament.

Our investigation into the MDPL2 simulation suggests that there is a non-negligible fraction ($\sim$ 10 per cent) of BCGs that assemble their stars extremely rapidly in the early universe. This counters the notion of late-time assembly for the entire population of BCGs, and especially counters the idea of treating the entire set of BCGs as a homogeneous population to begin with. What we have not addressed is the overall trend of later assembly with decreasing cluster mass and what this means for the BCG population. Using the same aforementioned time-scale estimates, systems that we identify as highly overdense regions that collapse at $z \sim 3$ should go on to form BCGs with stellar masses $M_* \gtrsim 10^{11}$ M$_\odot$ by $z \sim 1.5$. Therefore, our results suggest that there is a continuous population of bright, star-forming BCGs down to $z \sim 1.5$ – where the *Hubble* time is still short, allowing for the rapid assembly of substructure. Recent observations have indeed begun to undercover this population of star-forming galaxies (McDonald et al. 2016), especially with Webb et al. (2015) finding that the SFR in their BCG sample increases to $\sim$1000−3000 M$_\odot$ yr^{-1} as a function of increasing redshift. Similarly, the core of the recently investigated galaxy cluster SpARCS104922.6 + 564032.5 at $z = 1.71$ (Webb et al. 2017; Trudeau et al. 2019) appears qualitatively similar to our results from Section 3, and the scenario we propose above could explain the extended star formation and morphology within that system.

There is an additional connection of these high-redshift BCGs to galaxy clusters – one that could help ameliorate discrepancies in the determination of cosmological parameters. Specifically, the value of the density fluctuation power spectrum amplitude, σ_8, derived from galaxy cluster counts is known to disagree at $\sim 2\sigma$ with the value derived from the cosmic microwave background in the Planck mission (Planck Collaboration XX 2014; Douspis, Salvati & Aghanim 2019; Salvati, Douspis & Aghanim 2018). The number of massive clusters is highly sensitive to σ_8 and we expect these BCGs to preferentially exist in those clusters. In this study, through our use of the MDPL2 simulation, we use the Planck Collaboration XVI (2014) value of σ_8 which is higher than those derived from cluster counts. As it is, events such as SPT2349–56 are rare in this cosmology and ought to be rarer still if σ_8 were even slightly lower. A combination of surveying with the *WFIRST* and follow-up confirmation with the *JWST* would provide a lower bound on the number density, hence providing an additional constraint on σ_8 that could settle the issue. While we provide only an approximate estimate of the number density of massive BCGs, a more detailed study could provide an exact constraint on the value.

One major caveat is that for the BCGs to be highly star forming, they must not exhaust their gas supply, and they must not have their star formation halted by AGNs feedback. However, given that the regions we discovered at $z \gtrsim 3$ are highly overdense, they undergo what could be considered an effectively *monolithic* formation scenario where the substructure gives rise to galaxies that are gas rich, which then rapidly merge to form the BCG. And only afterwards, after the newly formed BCG's gas content depletes due to feedback and consumption, would star formation quench and the galaxy age passively.

5 CONCLUSIONS

The cores of galaxy clusters are home to the most luminous galaxies in the Universe – the BCGs. These galaxies have unique properties, such as their velocity dispersion and luminosity profiles, that set them apart from other galaxies at the high-end of the galaxy luminosity function. The contemporary picture of their formation and growth scenario is that their stars are old and formed at high redshift ($z \gtrsim 4$) in separate individual galaxies that, at late times ($z \lesssim 1$), hierarchically assemble to form the massive galaxies we observe today. There are, however, recent observations of highly overdense protoclusters at $z \gtrsim 4$ (Ishigaki et al. 2016; Jiang et al. 2018; Miller et al. 2018) that muddle this simple picture, since we expect these to go on to form the BCGs (Ito et al. 2019).

The hierarchical assembly picture of structure formation predicts the gradual build-up of objects through successive mergers of small objects. Under this theory, we expect the most-massive galaxies to be the youngest in terms of their assembly time. However, observations of massive ellipticals show a *downsizing* effect whereby the mass of the systems and the stellar ages are anticorrelated (Bower et al. 2006; Cimatti et al. 2006; Fontanot et al. 2009). In other words, the most-massive systems appear to have the oldest stellar populations. We propose a new paradigm wherein a similar downsizing effect occurs on the scale of galaxy clusters themselves. In our proposal, the cores of the most-massive clusters – the BCGs – assemble earlier than the cores of lower-mass clusters, on average. A subset of cluster-cores are assembled at very high redshift ($z \gtrsim 3$), with the probability of high-redshift assembly decreasing as a function of decreasing mass (at $z = 0$) of the clusters.

In order to determine the rapidity of assembly and growth of the BCGs, we studied the forward-evolution of a recently discovered, highly overdense protocluster at $z \approx 4.3$, SPT2349–56 (Miller et al. 2018) using a non-cosmological hydrodynamic simulation informed by the observations. We found that the star formation peaks at $\sim$3000 M$_\odot$ yr^{-1} very early in our simulation, and decays exponentially with a time-scale of $\tau \sim 200$ Myr. By 1 Gyr, we found that the system remains at a stable SFR of $\sim$40 M$_\odot$ yr^{-1}. We found that the system assembles 90 per cent of its mass at $t_{\rm assembly} \sim 370$ Myr after the initial condition and has a SDSS rest-frame $g - r$ colour of $g - r \approx 0.13$. In terms of redshift, 370 Myr corresponds to $z \sim 3.3$ and implies that fully formed, highly star forming, blue BCGs should exist at redshifts $z \gtrsim 3.3$, given that there are observations of systems similar to SPT2349–56 above $z \gtrsim 4.3$ (Ishigaki et al. 2016; Jiang et al. 2018). We demonstrated that new observational tools such as the *JWST* will be able to easily image such systems, given that we estimated their absolute magnitudes to be $M_{\rm AB,\ F277W} \sim -28.7$ in the *F277W* NIRCam band.

[11] The velocity distribution of the simulated haloes does not change appreciably with respect to the projection plane.

Of course, these estimations depend on the assumptions we made when constructing our numerical experiment. The main sources of uncertainty are the gas fractions and the applicability of the abundance matching results (Behroozi et al. 2013a) to the $z \approx 4.3$ protocluster. However, all high-redshift galaxies are found with relatively high gas fractions and while we are near the boundary of applicability in the abundance matching results, we are still within the halo mass range where the results apply.

We expect that systems such as SPT2349–56 go on to form the cores of the most-massive clusters in the Universe. We used the Multi-Dark Planck 2 simulation[12] – a child of the Multi-Dark Bolshoi simulations – to investigate the occurrence of highly over-dense assembly events in the early universe for all clusters with $z = 0$ masses $M_{\rm vir} \geq 5 \times 10^{14}$ M$_\odot$. Our analysis revealed that there is a clear trend of less-massive clusters having over-dense events occur at lower redshifts (compared to more-massive clusters), indicating that there is a *downsizing* effect. Specifically, we found that the fraction of events above $z \geq 3$ for a given bin of final cluster mass scales strongly with the final cluster mass, $f \sim M^{3.36}$. We also determined that ~ 10 per cent of all the clusters we investigated begin to assemble their cores rapidly at high redshift, with a higher percentage of 16.4 per cent for clusters with final masses above $M_{\rm vir} \geq 10^{15}$ M$_\odot$. Based on these estimates, we predict that there is a population of bright, blue BCGs above $z \gtrsim 1.5$ with a comoving number density of $n \approx 48$ cGpc^{-3}. Additionally, we predict that there is a similar star-forming BCG population that extends down to $z \sim 1.5$ – although those at lower redshift will be up to an order-of-magnitude less massive than their high-redshift counterparts, given the downsizing trend. At redshifts lower than $z \sim 1.5$, there is insufficient gas to support high SFRs and those BCGs that begin assembling late will assemble through dry mergers, in keeping with the conventional picture. We emphasise the distinction between core assembly and assembly of the rest of the cluster, and that our results do not suggest that the entire cluster assembles at high-redshift. We expect protocluster cores to be embedded in an extended lower density (compared to the protocore) galaxy distribution, with these galaxies eventually forming the satellite population of the assembling cluster.

Given that extraordinary infrared observational tools such as *WFIRST* and the *JWST* will launch in the upcoming decade, we expect the census of interesting astronomical objects to broaden significantly. Based on our arguments in this paper, we anticipate that some of those high-redshift objects will be the cores of the most-massive clusters in the Universe – the BCGs. Not only are these interesting objects in their own right but their discovery and census would also help to constrain the discrepancy in the measurements of the cosmological parameter σ_8, given their rarity and association with the most-massive clusters.

ACKNOWLEDGEMENTS

This research was enabled in part by support provided by WestGrid and Compute/Calcul Canada. Some of the computations in this work were performed on facilities supported by the Scientific Computing Core at the Flatiron Institute. The Flatiron Institute is supported by the Simons Foundation. DR acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number 534263]. DR and AB acknowledge support from NSERC (Canada) through the Discovery Grant program. CB acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number 504189]. MHH acknowledges the receipt of a Vanier Canada Graduate Scholarship. We thank Chervin Laporte, Carlos Frenk, Julie Hlavacek-Larrondo, Laura Salvati, Axel Weiss, Mark Brodwin, Gabriella De Lucia, Michael McDonald, Belaid Moa, and Ondrea Clarkson for insightful discussions during the course of the study. We also thank the anonymous reviewer for their helpful comments which helped us improve our study.

The CosmoSim database used in this paper is a service by the Leibniz-Institute for Astrophysics Potsdam (AIP). The MultiDark database was developed in cooperation with the Spanish MultiDark Consolider Project CSD2009-00064. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE, www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de). The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance AstroComputing Center (UC-HiPACC) and were run at the NASA Ames Research Center.

Our analysis was performed using the Python programming language (Python Software Foundation, https://www.python.org). The following packages were used throughout the analysis: `h5py` (Collette 2013), `numpy` (Oliphant 2006; van der Walt, Colbert & Varoquaux 2011), and `matplotlib` (Hunter 2007). Part of the analysis was performed in the IPython environment (Perez & Granger 2007). This research additionally made use of `astropy` (http://www.astropy.org) – a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018).

[12]A (1 cGpc h^{-1})^3 volume with 3840^3 dark matter elements.

REFERENCES

Alexander D. M., Smail I., Bauer F. E., Chapman S. C., Blain A. W., Brandt W. N., Ivison R. J., 2005, Nature, 434, 738
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Aragon-Salamanca A., Baugh C. M., Kauffmann G., 1998, MNRAS, 297, 427
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
Baes M., Camps P., 2015, Astron. Comput., 12, 33
Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, ApJS, 196, 22
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57
Bottrell C., Torrey P., Simard L., Ellison S. L., 2017a, MNRAS, 49, 1033
Bottrell C., Torrey P., Simard L., Ellison S. L., 2017b, MNRAS, 467, 2879
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93
Brough S., Couch W. J., Collins C. A., Jarrett T., Burke D. J., Mann R. G., 2008, MNRAS, 385, L103
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Burke C., Hilton M., Collins C., 2015, MNRAS, 449, 2353
Camps P., Baes M., 2015, Astron. Comput., 9, 20
Carilli C. L. et al., 2010, ApJ, 714, 1407
Cerulo P., Orellana G. A., Covone G., 2019, MNRAS, 487, 3759
Chandrasekhar S., 1943, ApJ, 97, 255
Cimatti A., Daddi E., Renzini A., 2006, A&A, 453, L29
Ciotti L., Ostriker J. P., 1997, ApJ, 487, L105

Collette A., O'Reilly Media, 2013, Python and HDF5
Collins C. A. et al., 2009, Nature, 458, 603
Conroy C., Wechsler R. H., Kravtsov A. V., 2007, ApJ, 668, 826
Contini E., De Lucia G., Villalobos Á., Borgani S., 2014, MNRAS, 437, 3787
Cowie L. L., Binney J., 1977, ApJ, 215, 723
Daddi E. et al., 2005, ApJ, 626, 680
Daddi E. et al., 2010, ApJ, 713, 686
Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265
Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, MNRAS, 467, 115
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Donzelli C. J., Muriel H., Madrid J. P., 2011, ApJS, 195, 15
Douspis M., Salvati L., Aghanim N., 2019, preprint (arXiv:1901.05289)
Dubinski J., 1998, ApJ, 502, 141
Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748
Fabian A. C., Nulsen P. E. J., 1977, MNRAS, 180, 479
Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416
Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, MNRAS, 397, 1776
Gaburov E., Nitadori K., 2011, MNRAS, 414, 129
Granato G. L., Ragone-Figueroa C., Domínguez-Tenreiro R., Obreja A., Borgani S., De Lucia G., Murante G., 2015, MNRAS, 450, 1320
Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, ApJS, 176, 438
Hernquist L., 1990, ApJ, 356, 359
Hernquist L., 1993, ApJS, 86, 389
Higuchi R. et al., 2019, ApJ, 879, 28
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, ApJ, 691, 1168
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ishigaki M., Ouchi M., Harikane Y., 2016, ApJ, 822, 5
Ito K. et al., 2019, ApJ, 878, 68
Iwamoto K., Brachwitz F., Nomoto K., Kishimoto N., Umeda H., Hix W. R., Thielemann F., 1999, ApJS, 125, 439
Jiang L. et al., 2018, Nat. Astron., 2, 962
Kennicutt R. C., 1998, ApJ, 498, 541
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 699, 850
Lanson N., Vila J.-P., 2008a, SIAM J. Numer. Anal., 46, 1912
Lanson N., Vila J.-P., 2008b, SIAM J. Numer. Anal., 46, 1935
Laporte C. F. P., White S. D. M., 2015, MNRAS, 451, 1177
Laporte C. F. P., White S. D. M., Naab T., Gao L., 2013, MNRAS, 435, 901
Lauer T. R., Postman M., Strauss M. A., Graves G. J., Chisari N. E., 2014, ApJ, 797, 82
Lavoie S. et al., 2016, MNRAS, 462, 4141
Leitherer C. et al., 1999, ApJS, 123, 3
Liang L., Durier F., Babul A., Davé R., Oppenheimer B. D., Katz N., Fardal M., Quinn T., 2016, MNRAS, 456, 4266
Lidman C. et al., 2012, MNRAS, 427, 550
Lidman C. et al., 2013, MNRAS, 433, 825
Lin Y., Mohr J. J., 2004, ApJ, 617, 879
Loubser S. I., Hoekstra H., Babul A., O'Sullivan E., 2018, MNRAS, 477, 335
Martizzi D., Hahn O., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2016, MNRAS, 459, 4408
McDonald M. et al., 2016, ApJ, 817, 86
Miller T. B. et al., 2018, Nature, 556, 469
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Mo H., van den Bosch F., White S., 2010, Galaxy Formation and Evolution. Cambridge Univ. Press, Cambridge
Muzzin A. et al., 2009, ApJ, 698, 1934
Narayanan D., Bothwell M., Davé R., 2012, MNRAS, 426, 1178
Nardini E., Risaliti G., Salvati M., Sani E., Imanishi M., Marconi A., Maiolino R., 2008, MNRAS, 385, L130
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neistein E., van den Bosch F. C., Dekel A., 2006, MNRAS, 372, 933
Nomoto K., Tominaga N., Umeda H., Kobayashi C., Maeda K., 2006, Nucl. Phys. A, 777, 424
Oliphant T. E., 2006, A guide to NumPy. Trelgol Publishing, USA
Olsen K., Greve T. R., Narayanan D., Thompson R., Davé R., Rios L. N., Stawinski S., 2017, ApJ, 846, 105
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Ostriker J. P., Tremaine S. D., 1975, ApJ, 202, L113
Overzier R. A., 2016, A&A Rev., 24, 14
Peebles P. J. E., 1968, ApJ, 153, 13
Perez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Perrin M. D., Soummer R., Elliott E. M., Lallo M. D., Sivaramakrishnan A., 2012, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, Vol. 8442, SPIE. p. 1193–1203 http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.925230
Perrin M. D., Sivaramakrishnan A., Lajoie C.-P., Elliott E., Pueyo L., Ravindranath S., Albert L., 2014, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, Vol. 9143, SPIE. p. 1174. http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2056689
Pillepich A. et al., 2018, MNRAS, 475, 648
Pipino A., Szabo T., Pierpaoli E., MacKenzie S. M., Dong F., 2011, MNRAS, 417, 2817
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration XX, 2014, A&A, 571, A20
Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, MNRAS, 373, 881
Postman M. et al., 2012, ApJS, 199, 25
Ragone-Figueroa C., Granato G. L., Ferraro M. E., Murante G., Biffi V., Borgani S., Planelles S., Rasia E., 2018, MNRAS, 479, 1125
Rennehan D., Babul A., Hopkins P. F., Davé R., Moa B., 2019, MNRAS, 483, 3810
Riebe K. et al., 2013, Astron. Nachr., 334, 691
Robertson B., Cox T. J., Hernquist L., Franx M., Hopkins P. F., Martini P., Springel V., 2006, ApJ, 641, 21
Romer A. K., Viana P. T. P., Liddle A. R., Mann R. G., 2001, ApJ, 547, 594
Ruszkowski M., Springel V., 2009, ApJ, 696, 1094
Sahlén M. et al., 2009, MNRAS, 397, 577
Salvati L., Douspis M., Aghanim N., 2018, A&A, 614, A13
Sandage A., 1976, ApJ, 205, 6
Scannapieco E., Bildsten L., 2005, ApJ, 629, L85
Schaye J. et al., 2015, MNRAS, 446, 521
Shankar F. et al., 2015, ApJ, 802, 73
Silk J., Rees M. J., 1998, MNRAS, 324, 128
Smith B. D. et al., 2017, MNRAS, 466, 2217
Springel V., 2000, MNRAS, 312, 859
Springel V., White S. D. M., 1999, MNRAS, 307, 162
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Stott J. P. et al., 2010, ApJ, 718, 23
Stott J. P., Collins C. A., Burke C., Hamilton-Morris V., Smith G. P., 2011, MNRAS, 414, 445
Tabor G., Binney J., 1993, MNRAS, 263, 323
Tacconi L. J. et al., 2010, Nature, 463, 781
Tacconi L. J. et al., 2013, ApJ, 768, 74
Tadaki K.-. ichi . et al., 2019, PASJ, 71, 40
Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195

Tremaine S. D., Richstone D. O., 1977, ApJ, 212, 311
Tremmel M., Governato F., Volonteri M., Quinn T. R., 2015, MNRAS, 451, 1868
Trudeau A. et al., 2019, MNRAS, 487, 1210
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
van der Wel A., Holden B. P., Zirm A. W., Franx M., Rettura A., Illingworth G. D., Ford H. C., 2008, ApJ, 688, 48
Vitvitska M., Klypin A. A., Kravtsov A. V., Wechsler R. H., Primack J. R., Bullock J. S., 2002, ApJ, 581, 799
Webb T. M. A. et al., 2015, ApJ, 814, 96
Webb T. M. A. et al., 2017, ApJ, 844, L17
Whiley I. M. et al., 2008, MNRAS, 387, 1253
White S. D. M., 1976, MNRAS, 174, 19
Wilson G. et al., 2009, ApJ, 698, 1943
Zubko V., Dwek E., Arendt R. G., 2004, ApJS, 152, 211

APPENDIX A: *JWST* OBSERVING STRATEGY

To generate mock NIRCam images, we select an optimal observing strategy for our synthetic SPT2349–56 system assuming 10 ks of observing time. We used the MUTLIACCUM MEDIUM 8 observing read-mode using Module B, in which each 10 × 1 ks integration is divided non-destructively into 10 × 100 s groups with 8 × 10 s frames/group and 2 × 10 s drop frames/group. This MULTIACCUM 'up-the-ramp' observing strategy enables cosmic-ray rejection, reduces the readout noise (roughly by the square root of the total number of frames) and increases the dynamic range by preventing saturation by bright sources. We generate point spread function (PSF) images for this observing strategy in each band pass using `pynrc` and the `WebbPSF` package (Perrin et al. 2012, 2014) and convolve with the idealised image. The 1 σ AB surface brightness sensitivities determined by `pynrc` for the NIRCam detector with this observing strategy were $m_{\rm AB,\,band} = 27.38$, 27.62, 28.52, and 28.65 AB mag arcsec^{-2} for *F*150*W*, *F*200*W*, *F*277*W*, *F*356*W*, respectively. These sensitivity estimates are based on characterization data for the detectors including readout and 1/f noise, dark current, and background levels equal to 1.2 times the minimum zodiacal light background. The current development version of `pynrc` does not allow us to convert between calibrated and non-calibrated images (i.e. AB zero points, effective gain and exposure time), so we do not incorporate Poisson shot noise from the source light. For now, we naively model the total noise contributions as a single Gaussian process with standard deviation equal to the reported sensitivity in each band pass using our observing strategy.

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