

Finding Milky Way Progenitors and Their Orbiting Satellites To Explore Their Mutual Effect Using James Webb Space Telescope Data

James Rumbold

*Department of Physics,
University of Bath, UK, BA2 7AY*

(Dated: May 9, 2023)

The James Webb Space Telescope’s (JWST) increased sensitivity has enabled research into higher redshift and lower stellar mass galaxies. In this project the JWST data is used to find Milky Way (MW) progenitors and their candidate satellites up to redshift $z = 4$ and above stellar mass $M_* = 10^{7.16} M_\odot$ to explore their mutual effects for the first time. The satellites were tested for galactic conformity and increased likelihood of quenching over host galaxies, finding only those within the stellar mass range $10^{8.5} M_\odot < M_* < 10^{9.0} M_\odot$ to be quenched and only those with redshifts $0.3 < z < 0.6$ and $1.6 < z < 2.5$ to be weakly conforming. The stellar mass and rest-frame ($U - V$) colour evolution was calculated for the MW and Large Magellanic Cloud (LMC) for comparison with the progenitors and satellites, calculating a bluer rest-frame colour for the MW (0.43) than the median progenitor value (1.33) at $z = 0$, as well as finding the MW to have peaked in star formation rate earlier than the progenitors by 0.72 - 1.08 Gyr. This earlier peak agrees with the known mergers with the Gaia-Enceladus and Kraken satellite galaxies, demonstrating the importance of progenitor-satellite mergers on a galaxy’s evolution.

1. INTRODUCTION

The recent release of the Cosmic Evolution Early Release Science Survey (CEERS) (Bagley et al. 2023) from the James Webb Space Telescope (JWST) has for the first time enabled research into lower stellar mass galaxies at large lookback times. For this project, pointings 1,2,3, and 6 were used, probing galaxies up to redshift $z = 4$ (galaxies forming only 1.6 Gyr after the The Big Bang), with stellar mass as low as $10^7 M_\odot$. With the increased sensitivity of the telescope, Milky Way (MW) progenitors can now be found at greater redshifts, and the lower mass satellite population orbiting these galaxies can now be studied for the first time. In addition to this, the increased sensitivity enables the relationship between progenitors and their satellites to be explored in more detail, providing new insight into their mutual effect.

Initial efforts to identify higher-redshift MW progenitors (galaxies of higher redshift that will grow to be of MW mass at $z = 0$) of nearby MW-mass galaxies (hereafter MW progenitors) used an approach where at each epoch galaxies with the same cumulative comoving number density were selected. This approach corresponds to the assumption that the rank order of galaxy stellar mass is preserved over cosmic time, such that progenitors at high redshift have the same rank order as the MW does at $z = 0$. This method also therefore assumes the number of mergers is negligible. Both assumptions of the rank order and mergers are not strictly true, however numerous related projects have shown this method to be an effective way to find MW progenitors. Morishita et al. (2015) implemented this approach to data in CANDELS and 3D-*HST* for progenitors with redshift between $0.5 < z < 3$, determining a sudden

reddening of MW progenitors at $z \sim 1.6$ due to a shutdown of star-formation activities. Later refinements on the basis of cosmological simulations were developed to account for the net effects of mergers and variance in growth rates (Wellons & Torrey 2017). Due to these refinements, an evolving number density yields more accurate predictions for the evolution of galaxies than a constant number density, and hence was the chosen approach for this project.

To compare the MW’s evolution with the progenitors, the stellar mass and rest-frame ($U - V$) colour was calculated across the redshift range $0 \leq z \leq 4$. For the mass evolution, the star formation history (SFH) of the MW was obtained from Fantin et al. (2019). For the colour evolution, the star formation history was then combined with our own stellar population synthesis implementation, starting from a library of single stellar population (SSP) models from Bruzual & Charlot (2003). Papovich et al. (2015) used abundance matching techniques to identify MW progenitors in the ZFOURGE catalogue (Tomczak et al. 2014), observing a peak in progenitor sSFR at $z = 1.5$ with a steady decline afterwards and finding 85% of progenitors to be quiescent (possessing a low star formation rate) in the redshift range $0.2 < z < 0.7$. However, Mutch et al. (2011) observed the MW to have bluer optical colours than equivalent mass galaxies in the Sloan Digital Sky Survey, finding the MW to reside in the green valley on a colour-mass diagram. As a bluer colour indicates a greater star formation rate, the results suggest the MW has had an atypical evolution compared to its progenitors, which this project will explore.

Galaxies belong to one of two classes: either hosts or satellites. As the MW and its progenitors are host galaxies, their respective evolutions can be drastically altered by the presence of satellites (smaller galaxies which orbit a host) due to merging. Merging can induce immense bursts of star formation in the host galaxy, but can also expel cold gas used for star formation through stellar or black hole feedback

mechanisms and create pressure-supported, red elliptical galaxies. Conversely, studies in the nearby universe have shown satellites to demonstrate unique characteristics due to their hosts. Satellites are more likely to be quenched (shutdown of star formation) than host galaxies, and hence are observed to have redder colours (Knobel et al. 2015 and Wetzel et al. 2012). Knobel et al. (2015) also showed that the specific star formation rate (sSFR, star formation rate divided by stellar mass) for a satellite increased with increasing sSFR of the host. This is a phenomenon known as “galactic conformity” which is the tendency for satellite galaxies to mirror the properties of their host. However, the findings for the characteristics of satellites for both papers were limited to galaxies at low redshifts and high stellar masses ($z < 0.06$ and $M_* \geq 2.0 \times 10^9 M_\odot$). Kawinwanichakij et al. (2016) observed that the satellites were more quenched than similar-mass host galaxies across redshifts $0.3 < z < 2.5$, and that satellites demonstrated strong conformity to their hosts across the redshift range $0.6 < z < 1.6$. However, the paper found weak conformity across the redshift ranges $0.3 < z < 0.6$ and $1.6 < z < 2.5$, and was also limited to $M_* \geq 2.0 \times 10^9 M_\odot$. Despite the limitation to high galaxy stellar masses and lower redshifts, the characteristics of the host galaxies are shown to directly affect the characteristics of their orbiting satellites. With the CEERS data probing to $z = 4$ and stellar masses as low as $10^7 M_\odot$, many more satellites have been found, enabling the relationship between host and satellite galaxy evolution to be explored across much larger cosmic time and stellar mass range.

The satellites of the obtained progenitors can be found by determining the galaxies which reside within a progenitor’s gravitationally bound region. A suitable estimate for this region is given by the virial radius, which is dependent on the mass of a galaxy’s dark matter halo. The dark matter halo mass is usually unknown on an individual galaxy basis, but can be assessed statis-

tically from its relation to the stellar mass of the central. Multi-epoch abundance matching combines observational results on the galaxy stellar mass function with simulation results on how the halo mass function evolves with cosmic time to derive such a redshift-dependent stellar mass - halo mass relation (Moster et al. 2013), which will be used for this project. For a galaxy to be a candidate satellite of a progenitor, it must be found within the progenitor's virial radius on the 2D night sky coordinates and have a similar redshift probability density. As such, once the 2D progenitor-satellite pairs are obtained, thresholds were set to accept or reject pairs based on their respective peaks in redshift probability density.

The CEERS data was inspected for signatures of enhanced quenching by performing statistical tests on the rest-frame colours of the satellites with similar mass and redshift host galaxies, as well as comparing the probability densities of the rest frame ($U - V$) colour distributions for the satellites and host galaxies. To test galactic conformity, the strength of correlation was found between the sSFR of the satellites and the sSFR of hosts across the same redshift ranges as Kawinwanichakij et al. (2016). The MW's largest satellite, the Large Magellanic Cloud (LMC), also had a calculation of stellar mass and rest-frame colour performed for comparison with the properties of the candidate satellites. This was done so using a SFH obtained from Harris & Zaritsky (2009).

This project aims to find MW progenitors and determine any orbiting satellite galaxies. The satellites' colours and sSFR will be compared with their host's to investigate the quenching relation and galactic conformity. Then, the stellar mass and colour evolution of the progenitors and satellites will be compared with calculations of the MW's and LMC's, with discussion on the physical reasons for any differences. The project aims to test prior satellite relationships obtained from lower redshift, higher stellar mass data with the increased sensitivity and range of the JWST,

and to highlight any similarities or differences from the MW's and LMC's evolution to the obtained progenitors and satellites. This project is purely computational, and is coded using Python language.

2. METHOD

2.1. Finding Milky Way Progenitors

2.1.1. Evolving Stellar Mass

To find the MW progenitors, the median number density of MW progenitors at varying redshifts was calculated using the 'evolvingN' function obtained through the 'NDPredict' and 'torrey_cmf' modules in Python, created by Wellons & Torrey (2017). The function calculates the number density from an initial redshift value (z_0), a final redshift value, and the number density of galaxies with a MW-like stellar mass at z_0 (N_0). To obtain the number density of MW-like galaxies at the current time ($z_0 = 0$), N_0 was found through the 'getnum' function from 'NDPredict'. This function requires the stellar mass of the MW at $z_0 = 0$, which was taken to be $M_{*,MW} = 10^{10.78} M_\odot$ (Licquia & Newman 2015). From this, N_0 was calculated to be $10^{-3.14} \text{ Mpc}^{-3} \text{ dex}^{-1}$, where the number density is expressed in per unit comoving volume units.

The median number density of MW progenitors was calculated over the range $0 \leq z \leq 4$, corresponding to a lookback time of up to 12 billion years. The 'NDPredict' module includes the 'sigmaN' function which calculates the standard deviation of the number density. An estimate of the range of progenitor number densities was found using the median plus and minus the standard deviation. The density values were then converted to stellar mass values through the 'getmass' function from 'NDPredict', which requires a choice of a cumulative mass function. The ZFOURGE and Illustris data sets were both considered, with ZFOURGE chosen as it is an observa-

tional dataset of more than 60,000 galaxies from Tomczak et al (2014), whereas Illustris is a hydrodynamical simulation by Vogelsberger et al. (2014). Hence, the ZFOURGE mass function is more applicable for making observational predictions, and also relates to previous literature by Papovich et al. (2015). The ZFOURGE data are valid over the range $10^8 < M_*/M_\odot < 10^{11.5}$ and $0.2 < z < 3.0$, and hence to mitigate any nonphysical results a cut-off in the search for Milky Way progenitors was imposed at $z = 4$. The range of the calculated stellar masses was used to impose a cutoff on galaxies in the CEERS data, and so any CEERS galaxies within this region were accepted as progenitors. The stellar masses of the CEERS galaxies up to $z = 4$ are plotted in Figure 1, along with the calculated mass range where progenitors of MW-like galaxies can be found. From a total of 23,919 galaxies, 2,353 progenitors were found.

The (photometric) redshifts and stellar masses used in the CEERS data were derived by members of the CEERS team using the EAZY (Brammer et al 2008) and FAST (Kriek et al. 2009) template fitting codes using a multi-wavelength photometric catalogue described by Finkelstein (2023) as input.

2.1.2. Minimum Mass Function

Each galaxy’s flux measurement carries a certain amount of noise. The signal-to-noise ratio (SNR) represents the amount of noise a given measurement contains, and so a larger ratio indicates a more accurate measurement. Even with the greatly increased sensitivity of JWST’s CEERS, imaging the finite depth of the observations places a limit on how low in stellar mass galaxies can be detected. This stellar mass limit is a function of redshift as higher redshift galaxies tend to be fainter and thus have lower SNR ratios. By setting a minimum SNR threshold of 3σ , such that a galaxy with a smaller SNR ratio is not a statistically significant detection, the min-

imum flux measurement possible for each galaxy can be calculated. Assuming the luminosity of a galaxy, and therefore the received flux, is linearly proportional to the stellar mass of the galaxy, this threshold can be used to calculate the minimum stellar mass observable for each galaxy measurement. The minimum stellar was calculated through

$$M_{*,min} = M_* \times \frac{SNR_{threshold}}{SNR_{galaxy}} \quad (1)$$

Where $M_{*,min}$ is the minimum observable stellar mass based on a given galaxy’s SNR_{galaxy} , M_* is the stellar mass of a galaxy, $SNR_{threshold}$ is 3σ , and SNR_{galaxy} is the observed signal-to-noise of a galaxy. For the CEERS galaxies, the SNR was found by dividing the photometric flux of the galaxy in the F356W band divided by the flux error in the same waveband.

To obtain a curve fitting the minimum stellar masses (the minimum mass function), the calculated masses were separated into redshift bins of width $z_{width} = 0.01$ from $0 \leq z \leq 4$. With this, the stellar mass value for which 90% of the data in the bin is above was obtained. This percentile was chosen such that the minimum mass function (MMF) aligned with the apparent cutoff with lower stellar mass data in Figure 1. The curve connecting the redshift bins and 10th percentile values was approximated through an equation of the form

$$MMF = a \log_{10}(bz + c) + d \quad (2)$$

where z is the redshift, $a = 0.872 \pm 0.016$, $b = 10.633 \pm 5.942$, $c = 0.262 \pm 4.334$ and $d = 6.653 \pm 4.494$. The plot of the MMF is shown in Figure 1, and as expected, lower mass galaxies can generally only be observed at lower redshifts. As a result, there will be galaxies orbiting the obtained progenitors that have not been detected in the CEERS catalogue, affecting any observed trends in the data.

To ensure completeness in the CEERS data across redshifts $0 < z < 4$, a minimum mass was chosen such that galaxies could

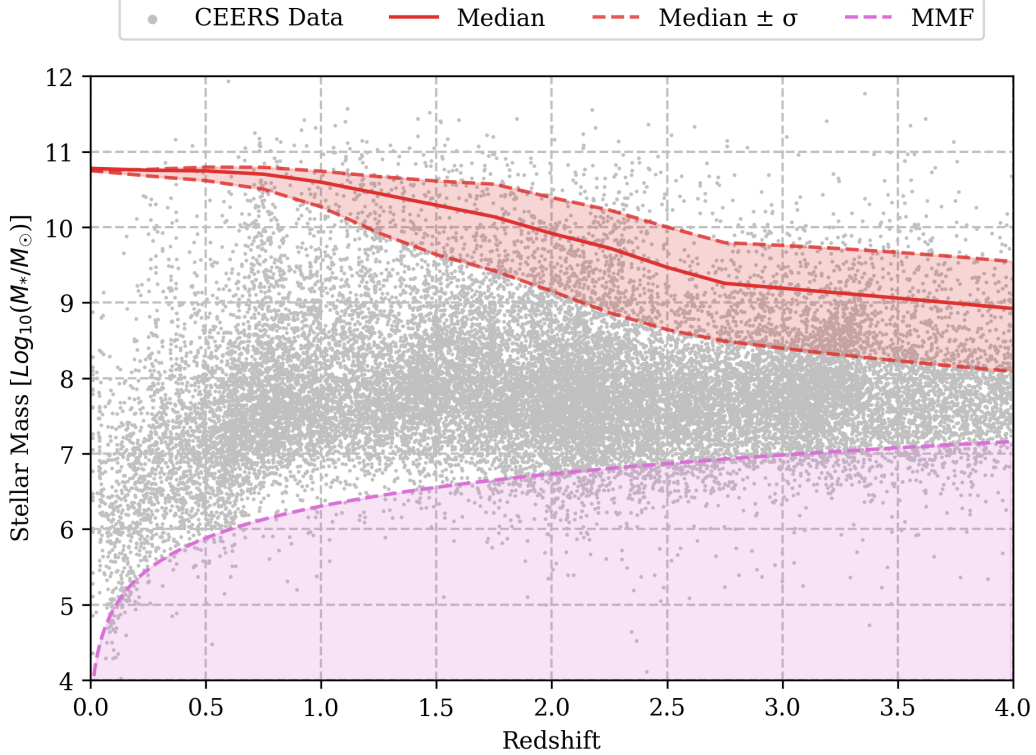


FIG. 1: The logarithm (base 10) of the stellar mass (solar mass units) of the 23,919 CEERS galaxies from $z = 0$ to $z = 4$. The red shaded area denotes the stellar mass range of MW progenitors obtained through the ZFOURGE cumulative mass function, with the bounds given by the median stellar mass \pm the standard deviation, σ . The CEERS galaxies that fall inside are labelled as MW progenitors. The purple shaded area gives the region below the minimum mass function, and illustrates the selection effect for observing lower stellar mass galaxies only at lower redshift.

be observed with this given stellar mass or higher across the entire redshift range. This corresponds to the stellar mass of the MMF at $z = 4$, which was calculated to be $M_{*,min} = 10^{7.16} M_{\odot}$. As such, for all satellite finding and subsequent analysis, only galaxies of stellar mass $M_* \geq 10^{7.16} M_{\odot}$ were considered.

2.2. Calculating the Progenitor's Gravitational Region

With the progenitors found, an estimate of their gravitationally bound region was calculated. To do so, the virial radius was found through the equation from Burkert

et al. (2016)

$$R_{virial} = \left(\frac{M_{DM}G}{100H(z)^2} \right)^{\frac{1}{3}} \quad (3)$$

Where

$$H(z) = H_0(\Omega_{\Lambda,0} + \Omega_{m,0}(1+z)^3)^{1/2} \quad (4)$$

Here R_{virial} is the virial radius, M_{DM} is the mass of the dark matter halo surrounding a galaxy, G is Newton's gravitational constant, $H(z)$ and H_0 are the Hubble constants at z and $z = 0$, and $\Omega_{\Lambda,0}$ and $\Omega_{m,0}$ are the energy densities of dark energy and total matter at $z = 0$. The values of the constants were set to $G = 6.674 \times 10^{-11} \text{ Nkg}^{-2}\text{m}^2$, $H_0 = 69.32 \text{ kms}^{-1}\text{Mpc}^{-1}$, $\Omega_{\Lambda,0} = 0.713$ and $\Omega_{m,0} = 0.287$.

The estimated mass of each galaxy's dark matter halo is not given with the

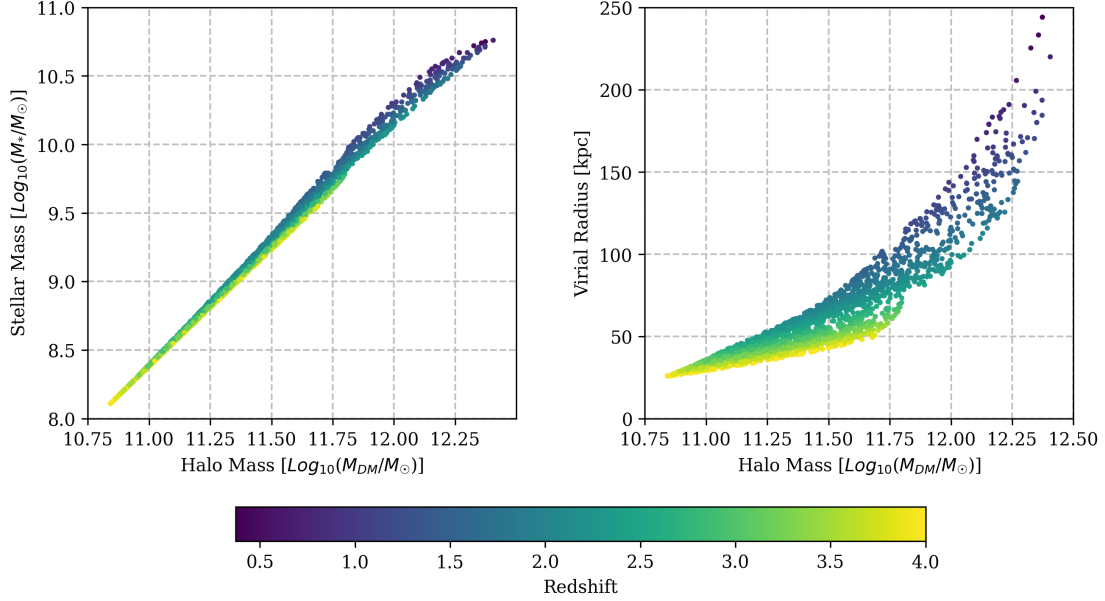


FIG. 2: Two plots displaying the relation between the dark matter halo mass and stellar mass of the progenitors, and the calculated virial radius (kpc) as a function of a progenitor’s dark matter halo mass. The color of the points denotes the redshift of the progenitor, with the scaling given in the shared colour-bar. The stellar mass-halo mass relation shows a strongly linear curve in the mass range of the MW progenitors, however at heavier stellar masses the relation begins to bend. The virial radius-halo mass displays a power-law fit as the radii are not in logarithmic form.

CEERS catalogue and so was calculated separately. The relationship between the stellar mass and dark matter halo mass of a galaxy is given through

$$M_* = 2NM_{DM} \left[\left(\frac{M_{DM}}{M_1} \right)^{-\beta} + \left(\frac{M_{DM}}{M_1} \right)^{\gamma} \right]^{-1} \quad (5)$$

Where

$$\log_{10} M_1(z) = M_0 + M_z \frac{z}{z+1}, \quad (6)$$

$$N(z) = N_0 + N_z \frac{z}{z+1}, \quad (7)$$

$$\beta(z) = \beta_0 + \beta_z \frac{z}{z+1}, \quad (8)$$

$$\gamma(z) = \gamma_0 + \gamma_z \frac{z}{z+1}. \quad (9)$$

The equations and constants are obtained from MEAM, a statistical model to determine the relationship between the stellar mass and dark matter halo mass of galaxies from Moster et al. (2013).

Equation 5 returns the stellar mass of a galaxy from the dark matter halo mass. To obtain the halo mass from the known stellar mass, a linear interpolation was performed due to the equation not being invertible. As the parameters $M_1(z)$, $N(z)$, $\beta(z)$ and $\gamma(z)$ are redshift dependent, this interpolation was performed for each progenitor in the dataset. As such, 2,353 different interpolations were performed for 10,000 equally spaced dark matter halo mass values in the range $10^{10} M_{\odot} \leq M_{DM} \leq 10^{16} M_{\odot}$. From this, the dark matter halo mass values for the progenitors were obtained.

With the dark matter halo mass of the progenitors found, the virial radius was calculated through Equation 3. The values were found in kiloparsec (kpc) and are plotted in Figure 2, alongside the dark matter halo mass and stellar mass relation. The radii range from $25.97 \text{ kpc} \leq R_{virial} \leq 244.04 \text{ kpc}$, and show an approximately exponential fit against halo mass (due to the radii plotted on linear-log scale). The

MW’s virial radius was calculated to be 242.08 kpc through the same method, using the known halo mass from Boylan-Kolchin et al. (2013) as $M_{DM} = 10^{12.20} M_{\odot}$ and $z = 0$. Dehnen et al. (2006) determined the MW’s virial radius as ≈ 200 kpc, although this value may vary for different papers. The calculated radius is of the same order of magnitude as the known, and hence the method for finding the virial radius of progenitors is sufficiently accurate for the paper’s goals.

2.3. Identifying Near Neighbours as Candidate Satellites

2.3.1. Near Neighbours in 2D

The CEERS catalogue contains the right ascension (RA) and declination (DEC) coordinates on the sky of the galaxies given in units of degrees. To find which galaxies are in the gravitational region of others, the calculated virial radii were converted from a distance measure to an angular size in degrees. With the angular size of the virial radii known in degrees, the angular distance between the progenitors and all the CEERS catalogue galaxies (including the progenitors and with the minimum mass limit of $10^{7.16} M_{\odot}$ applied) was calculated.

Once the angular separations were found, any pair with an angular separation larger than the angular size of the progenitor’s virial radius were removed. In addition to this, any pair with an angular separation of 0° were removed due to self-counting. Finally, those pairs with a satellite mass greater than the progenitor mass were removed. This is because any progenitor-satellite pair in which both are progenitors would be counted twice, and hence only the case of the larger progenitor mass should be included. The total number of pairs satisfying these conditions was 65,601, with the number of pairs satisfying each condition shown in Table I. Through this method, any galaxies that reside within the progenitor’s virial radius

have been found for the two-dimensional RA and DEC coordinates.

TABLE I: The number of progenitor-galaxy angular separation pairs after each subsequent condition. There is a total of 54,568,423 possible pairs from the product of 2,353, the number of progenitors, and 19,597, the number of CEERS galaxies with stellar mass above $M_{*,min} \geq 10^{7.16} M_{\odot}$.

Condition	Total Pairs
All Possible Pairs	46,111,741
Separation close to 0.0135°	1,827,016
Separation within R_{virial}	73,405
No-self counted	71,052
Progenitor mass is the heaviest	65,601

2.3.2. Near Neighbours in 3D

To determine the progenitor-satellite pairs, the three-dimensional case must be considered. The normalised redshift probability densities of the CEERS catalogue are available as by-products of the EAZY codes, and a true progenitor-satellite pair will have similar redshifts and hence overlapping densities. To set a threshold for the overlap of probability density, two approaches were combined. The product of two galaxy redshift probability densities gives the combined probability density, and so a highly peaked product density will give a very likely progenitor-satellite pair. For high peak thresholds, $P_{t,peak}$, accurate pairs are found. However, after a threshold of $P_{t,peak} = 0.30$ the number of pairs found is on the order 10^1 . To compare the evolutions and trends of satellites and their progenitor hosts, a sufficient amount of data is needed across a wide redshift and mass range. Upon reducing the peak threshold to obtain more pairs, the progenitor-satellite pairs found had probability densities consisting of one extremely peaked distribution and one broad, low probability distribution. These pairs are not likely to be true near-

neighbours, and so to find pairs with similar peak sizes and redshift, the cumulative distribution of the multiplied densities was found as well.

By combining two thresholds, one for the peak of the product probability density and another for the cumulative product distribution $P_{t,cum}$, progenitors and satellites pairs were found in greater number whilst retaining a high likelihood of being near-neighbours. To determine the threshold values, the number of pairs counted with varying $P_{t,peak}$ and $P_{t,cum}$ thresholds was found. The combination of $P_{t,peak} = 0.15$ and $P_{t,cum} = 0.40$ yielded 1,063 candidate progenitor-satellite pairs, and hence was chosen as the best compromise between accuracy and quantity of data. This represents 1.62 % of the pairs satisfying the conditions in Table I, and 0.0023 % of the total possible pairs. Figure 3 plots the probability densities, product density and cumulative product distribution for a pair found through this approach, and demonstrates a sufficient accuracy for determining the progenitor and satellites of a similar redshift.

To determine if the candidate satellites demonstrated quenching versus the field galaxies, the field galaxies with stellar mass and redshift closest to the satellites were found, and the two resultant $(U - V)_{rest}$ colour distributions were statistically tested with the two-sample Kolmogorov-Smirnov (KS) to determine if the satellites were of the same distribution or showed unique properties. By finding field galaxies closest, the stellar mass and redshift of both the satellite and field galaxy will be approximately the same, enabling any colour difference to be more accurately obtained. The colour probability density distributions were also compared to determine which distribution was more likely to exhibit redder colours. The KS test will be performed for satellites found with redshift $0 < z < 0.5$, to compare with the results of Knobel et al. (2015) and Wetzel et al. (2012, $0 < z < 2.5$ to compare with Kawinwanichakij et al. (2016), and for $2.5 < z < 4$.

To test galactic conformity between the

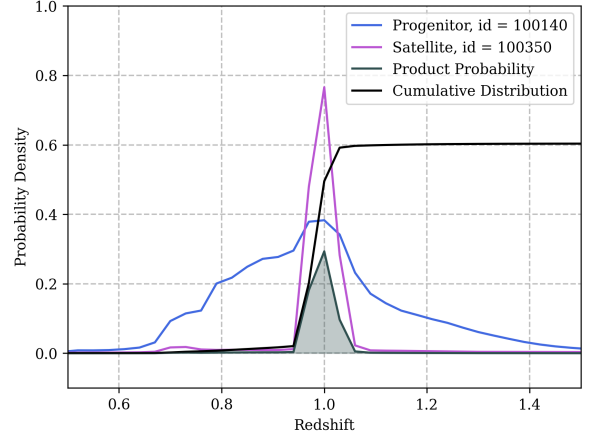


FIG. 3: A plot of the probability densities of a candidate progenitor-satellite pair using both cumulative product distribution and product density thresholds. The thresholds were set to $P_{t,cum} = 0.40$ and $P_{t,peak} = 0.15$, and the pair plotted are the fifth found with this approach. The progenitor and satellite's probability densities are denoted with a blue and pink curve respectively, with their multiplied probabilities given with dark grey. The galaxy identifiers denote their position in the first CEERS pointing.

satellites and progenitors, the Pearson correlation coefficient was found between a satellite's sSFR and its host progenitor's for the redshift ranges $0.6 < z < 1.6$, $0.3 < z < 0.6$, $1.6 < z < 2.5$, and $2.5 < z < 4.0$ for comparison with the results of Kawinwanichakij et al. (2016).

The extinction for the CEERS galaxies is known and was removed from the observed $(U - V)$ colour values. As a result, the colours referred to in this report will be rest colours in the AB magnitude system.

2.3.3. Line of Sight Considerations

Two galaxies with a separation larger than either of their virial radii may appear to be near neighbours in projection. For an observer with a line of sight close to the position vector connecting the two, it is reasonable to assume they would observe both as 2D near neighbours and determine the

redshift probability densities to successfully match. As a result, an incorrect conclusion of the two being a galaxy-satellite pair would be made. To understand the scale of this error for the candidate progenitor-satellite pairs, the separation between the Milky Way and Andromeda (also known as Messier 31, or M31) galaxies was modelled, and the likelihood for an observer to determine the two to be within a galaxy's virial radius was found. Vilardell et al. (2010) determined the distance to the Andromeda galaxy to be 744 ± 33 kpc, and Kafle et al. (2018) calculated the virial radius of Andromeda as 240 ± 10 kpc. Andromeda has a larger virial radius than the Milky Way (≈ 200 kpc), and so the worst case estimate was made with a separation of 711 kpc and virial radius of 250 kpc.

The angle θ defines the maximum angle at which the pair would be found as near neighbours, and was found through

$$\theta = \sin^{-1} \left(\frac{R_{\text{virial}}}{d} \right) \quad (10)$$

Where R_{virial} is Andromeda's virial radius and d is the separation between the two galaxies. Using the values above, θ was found to be 20.58° . For a 3D sphere of all viewing angles, this angle defines two cones from the galaxies for which an observer would determine the two to be near neighbours, as shown in Figure 4. To estimate the likelihood of being an observer in said cone, the ratio of the cone's surface area to the total surface area was found. The formula is given by

$$P_{\text{neighbour}} = 2 \times \frac{\tan^2(\theta)}{4} \quad (11)$$

And can be derived by considering the area of the cone equal to $\pi(R \tan(\theta))^2$ and the surface area of the sphere equal to $4\pi R^2$, where R is the distance to the galaxies. The factor of 2 arises from the two possible viewing cones either side of the galaxies. For the Milky Way-Andromeda system, $P_{\text{neighbour}}$ was found to be 7.05%.

As the result is independent of the distance to the galaxies, it is reasonable to

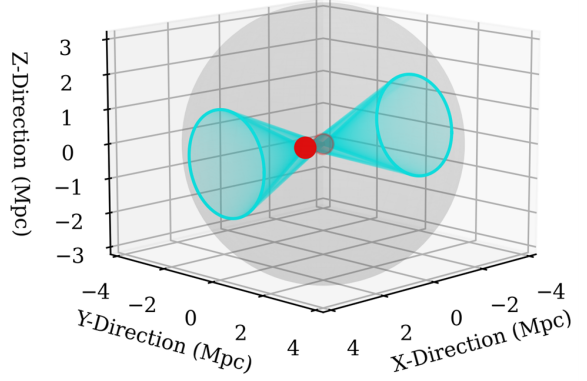


FIG. 4: An illustration of the viewing angles for which an observer would measure the Milky Way and Andromeda galaxies to be a progenitor-satellite pair. The two central red spheres denote the galaxies but are not representative of their size or shape. The grey sphere marks a sphere of radius 4 Mpc, which is an arbitrary distance from an observer to the two galaxies. The two light blue cones either side mark the viewing angles within which an incorrect determination is made.

estimate that approximately 7% of the obtained progenitor-satellite pairs (corresponding to approximately 92 pairs) are incorrectly assigned due to the line of sight angle to both. However, this projection effect is minor when compared to calculation of the photometric redshifts from FAST and EAZY, which carry larger uncertainties.

2.4. Stellar Mass and Colour Evolution of the Milky Way and Large Magellanic Cloud

To compare the MW's evolution to the progenitors found, and the Large Magellanic Cloud's (LMC) evolution to the satellites, the stellar mass growth history and rest-frame colour of both were computed over the range $0 \leq z \leq 4$. To do this, the following equations were used

$$M_*(t) = \int_0^t \Psi(t - t') dt' \quad (12)$$

$$L_U(t) = \int_0^t L_{U,SSP}(t') \Psi(t - t') dt' \quad (13)$$

$$M_{U,SSP} - M_\odot = -2.5 \log_{10} \left(\frac{L_{U,SSP}(t)}{L_\odot} \right) \quad (14)$$

$$(U - V)(t) = -2.5 \log_{10} \left(\frac{L_U(t)}{L_V(t)} \right) = M_U - M_V \quad (15)$$

Where $\Psi(t - t')$ is the star formation history (SFH) of the galaxy, $L_{U,SSP}(t')$ is the U-band luminosity of a single stellar population (SSP), L_U is the U-band luminosity of the galaxy, L_\odot and M_\odot are the luminosity and mass of the Sun, respectively, L_V is the V-band luminosity of the galaxy (calculated through Equation 13 with $L_{V,SSP}(t')$, the V-band magnitude of an SSP), $M_{U,SSP}$, M_U , and M_V are magnitudes of the SSP's U-band, and the galaxy's U and V band respectively, and $(U - V)(t)$ is a colour index of the galaxy at a given age.

The calculations for the MW and LMC were performed using the SFHs from Fantin et al. (2019) and Harris & Zaritsky (2009), respectively. As the MW's SFH was given in lookback time up to 12.33 Gyr, the MW was assumed to have formed at this point. The LMC's SFH was obtained in the range $5.6 \text{ Myr} \leq \text{age} \leq 14.75 \text{ Gyr}$, but was also assumed to have formed 12.33 Gyr ago.

With the SFHs of the MW and LMC obtained, an SSP was required for the colour calculations. This was obtained from Bruzual & Charlot (2003), which features a constant SFR for 1 Myr and then drops to zero. The magnitudes of the U and V-band are given in the AB magnitude system over the range of 1.259×10^5 to 2×10^{10} years. Figure 5 plots M_U , M_V , and $(U - V)$ for the SSP.

Using the obtained SFHs and SSP, L_U , L_V and M_* were calculated for the Milky Way and LMC over the range $100 \text{ Myr} \leq \text{age} \leq 12.33 \text{ Gyr}$ in steps of 100 Myr, resulting in 124 total age steps. To perform

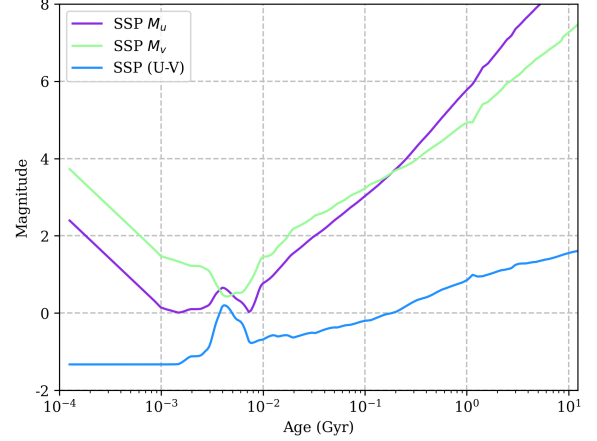


FIG. 5: The U and V band AB magnitude evolution for the Bruzual & Charlot (2003) SSP, and the $(U - V)$ colour index calculated using Equation 15. The plot displays the evolution over the range $125,900 \text{ yr} \leq \text{age} \leq 12.33 \text{ Gyr}$, the same used in the colour calculations of the Milky Way and LMC.

the integration, the trapezium rule was implemented. For each age, the stellar mass integration was calculated using a step size $h_{mass} = 100,000$ years. For the colour index calculations, a small step size on the scale 100,000 yr is required to ensure the more luminous, early ages of the SSPs is included in the integration. However, due to the greater number of calculations per step than the stellar mass integral, and the Bruzual & Charlot (2003) SSP ranging from age $\geq 125,900$, the step size $h_{colour} = 250,000 \text{ yr}$ was used.

For each step in the integration, the magnitudes of the U and V band of from the Bruzual & Charlot (2003) SSP were found using a linear interpolation. These values were converted into luminosities through Equation 14, with $M_\odot = 4.83$ and $L_\odot = 3.826 \times 10^{26} \text{ W}$ obtained from Willmer (2018). The star formation rate was also obtained from an interpolation of the SFHs for both galaxies.

With the the stellar mass integral performed, the resultant stellar mass calculation after 12.33 Gyr for both the MW and LMC was compared to the known values $M_* = 10^{10.78} M_\odot$ and $M_* = 10^{9.43} M_\odot$ (Marel

et al. 2002), respectively, and the factor between was used to scale the SFHs. The calculated values were $M_{*,MW} = 10^{10.59} M_{\odot}$ and $M_{*,LMC} = 10^{9.01} M_{\odot}$, respectively, and so the MW's SFH was scaled by 1.56, and the LMC's by 2.66 to match the known stellar masses. This was done to reduce the number of factors that make for differences that are hard to interpret (stellar mass loss, for example).

To compare with the found progenitors and satellites, the median and interquartile range of the progenitor and satellite colours and mass was found within redshift bins of size $z_{width} = 0.5$ from $0 \leq z \leq 4$. These values were plotted against the calculated results, are discussed further in the report.

3. RESULTS

3.1. Progenitor & Satellite Trends

A summary of progenitor and satellite galaxy properties is shown in Table II. The table shows the median stellar mass and rest-frame $(U - V)$ colour for each in redshift bins of width $z_{width} = 0.5$, along with the standard deviation of the result. As shown from the table, the progenitors evolve with increasing stellar mass and reddening colour. The $(U - V)_{rest}$ colour median demonstrates a large increase in value from 0.70 to 1.1 from redshift range $1.0 < z < 1.5$ to $0.5 < z < 1.0$, corresponding to the largest change in colour across any of the redshift bins (0.40). As such, the sudden reddening for progenitors observed by Morishita et al. (2015) is present in the CEERS data but, at a lower redshift value than $z \sim 1.6$. The fraction of host progenitors (progenitors with at least 1 candidate satellite) to total progenitor count increases with decreasing redshift, suggesting that older, redder progenitors are more likely to have orbiting satellites. The average number of satellites also increases with decreasing redshift, with the number increasing rapidly for $z < 1.0$. From the 2,353 total progenitors found, 586 were determined to be

host galaxies.

The satellite's median stellar mass remains relatively constant across the redshift range, with the standard deviation of the result generally increasing for decreasing redshift. The constant mass across the redshift range suggests the satellites observed at smaller redshift are not the evolved versions of galaxies at later redshift. The satellites demonstrate an increase in rest-frame colour and standard deviation for decreasing redshift, with the colour difference between each redshift range generally increasing with decreasing redshift.

The KS test performed on the colour distributions of the satellites and the closest stellar mass and redshift field galaxies for redshifts $0 < z < 4$ found the test statistic to be 0.022, with the p -value obtained as 0.968. As the p -value is much greater than the test statistic, the hypothesis that both are drawn from the same colour distribution is accepted. Figure 6 plots the binned probability densities and cumulative distribution for the satellite and field galaxy $(U - V)$ colour distributions across this redshift range. The plot confirms the results from the KS test and demonstrates the likeness between the two distributions. For satellites found within redshifts $0 < z < 0.5$, $0 < z < 2.5$, and $2.5 < z < 4$, all KS tests also showed the satellite colour distribution to be drawn from the same distribution as the closest stellar mass and redshift field galaxies. As the satellite galaxies were found to not be statistically redder than these host galaxies for each redshift range, the result that satellites are more likely to be quenched is not observed and disagrees with the observations from Knobel et al. (2015), Wetzel et al., (2012), and Kawinwanichakij et al (2016).

The Pearson correlation coefficient was found between the progenitor's sSFR and the satellites' sSFR within the redshift range of $0.3 < z < 0.6$ to be 0.21. For the redshift range $0.6 < z < 1.6$, the coefficient was found to be 0.06, and for redshifts $1.6 < z < 2.5$ the coefficient was found to be 0.11. Therefore, for redshift ranges $0.3 < z < 0.6$

TABLE II: A summary of MW progenitor and satellite. The values M_* and $U - V$ denote the median of the stellar mass and $(U - V)$ colour, respectively, and σ denotes the standard deviation of this result. All stellar mass values are the logarithm (base 10) of mass in solar mass units, M_\odot , and the colour is the rest colour in magnitudes (AB system). The count columns gives the number of galaxies found in a given redshift bin, whereas the hosts column denotes the number of progenitors with at least 1 orbiting satellite. The $\langle \text{Sats} \rangle$ column denotes the average number of satellites per host.

Redshift	Progenitors							Satellites				
	Count	M_*	σ_M	U-V	σ_{U-V}	Hosts	$\langle \text{Sats} \rangle$	Count	M_*	σ_M	U-V	σ_{U-V}
0.0 - 0.5	3	10.74	0.01	1.33	0.31	2	32	64	7.53	0.49	0.86	0.34
0.5 - 1.0	24	10.57	0.11	1.10	0.50	14	5.43	76	7.89	0.68	0.64	0.34
1.0 - 1.5	129	10.31	0.26	0.70	0.48	57	3.68	210	7.84	0.66	0.50	0.28
1.5 - 2.0	256	9.78	0.35	0.51	0.30	105	2.02	212	7.91	0.60	0.36	0.23
2.0 - 2.5	457	9.39	0.37	0.41	0.23	141	1.71	241	8.00	0.62	0.27	0.18
2.5 - 3.0	398	8.89	0.34	0.34	0.21	60	1.47	88	7.89	0.59	0.21	0.15
3.0 - 3.5	682	8.66	0.35	0.28	0.19	86	1.30	112	7.88	0.54	0.18	0.19
3.5 - 4.0	404	8.54	0.40	0.26	0.21	52	1.15	60	7.68	0.48	0.18	0.19
Total	2353					586		1063				

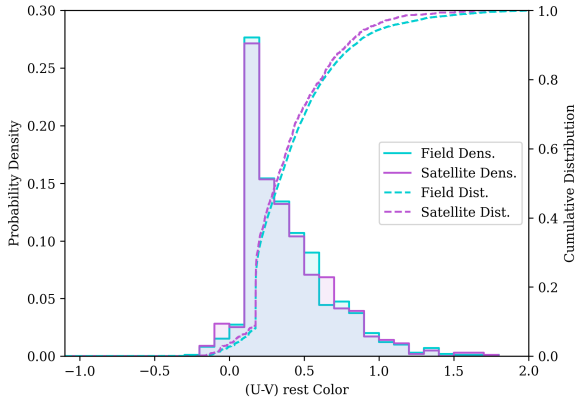


FIG. 6: The probability density and cumulative distribution of the $(U - V)$ colour for the satellites and closest stellar mass and redshift field galaxies across the redshift range $0 < z < 4$. The two distributions are extremely similar, suggesting satellites are not more likely to be quenched than field galaxies.

and $1.6 < z < 2.5$ the satellites are shown to have weak conformity to the host progenitors, agreeing with the results of Kawinwanichakij et al. (2016). However, for the range $0.6 < z < 1.6$ the correlation coefficient indicates extremely weak conformity, disagreeing with the strong conformity re-

sults from the same paper. For completeness, the correlation was found for satellites between $2.5 < z < 4.0$ as -0.003 , indicating no correlation between satellite sSFH and host sSFH in this redshift region.

The the log of the progenitor mass divided by the most massive satellite plot is shown in Figure 7. The plot shows that the majority of progenitors are on the scale of 10 times more massive than their most massive satellite, with the largest being 39.7 times more massive. The MW and LMC mass ratio is also plotted ($1.35 = \log_{10}(10^{10.78}/10^{9.43})$) for the galaxies), demonstrating the bulk of candidate satellites detected are closer in mass to their central than the LMC is to the MW. To determine the strength of the correlation between the progenitor mass and the mass ratio, the Pearson correlation coefficient was found to be 0.414 for the log-log scale shown in the figure. This indicates a weak, yet positive linear correlation between the two. As this result is determined on a log-log scale, for increasing stellar mass the relative mass of the progenitor increases with a power-law dependence over its most massive satellite.

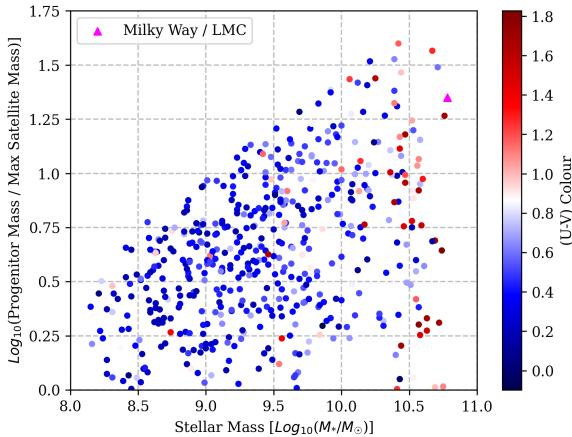


FIG. 7: The ratio of the progenitor’s stellar mass to its most massive satellite as a function of progenitor stellar mass. The colour-bar denotes the $(U - V)_{rest}$ colour of the progenitors, with the MW and LMC ratio shown through a magenta triangle (colour not illustrative of rest-frame colour). To better illustrate the trends, the logarithm (base 10) was taken of the mass ratios and the stellar masses.

3.2. Milky Way and LMC Mass and Colour Evolution

3.2.1. Mass Evolution

The MW and LMC stellar mass were found as a function of redshift and are plotted together with the median and upper and lower quartiles of the progenitor and satellite stellar mass values, in Figure 8. The SFH of both galaxies is also plotted for comparison with the calculated stellar mass evolution, and have both been scaled to return the known stellar mass values at $z = 0$. For the Milky Way, the figure shows that the galaxy is calculated to have a greater stellar mass than the median progenitor stellar masses for $z \leq 2.42$, or a lookback time of up to 10.81 Gyr. The shape of the evolution differs from the progenitor’s, with the MW growing immensely between $1.5 \leq z \leq 4.0$, or a period of 2.68 Gyr. Within this period the Milky Way is calculated to have grown in mass by a factor of 525 from $10^{7.49} M_{\odot}$ to $10^{10.21} M_{\odot}$. The Milky Way’s

SFH in this period of growth peaks at $14.00 M_{\odot} \text{ yr}^{-1}$ at redshift 1.68, or lookback time of 9.67 Gyr. The progenitor stellar mass bound approximately follows the shape of the ZFOURGE cumulative mass function, but with a smaller median and upper and lower quartile values due to fewer galaxies found at the higher stellar masses of the mass function. The stellar mass growth of the progenitors is shown to be much less than the MW’s at larger redshift, but follows a similar trend for $z < 1$.

For the LMC, the stellar mass calculation steadily increases over the redshift range, and is always greater than the median satellite mass. The median satellite stellar mass is approximately constant across the entire range, suggesting the satellites observed at earlier redshift are not the same galaxies as the satellites at later redshift, otherwise stellar mass growth would be present. The LMC’s stellar mass is also always greater than the MMF, and so the LMC would be observed in the CEERS data up to $z = 4$. From this calculation, it is also predicted that the LMC had a larger stellar mass than the Milky Way for $z \geq 2.79$, or a lookback time of more than 11.19 Gyr. If both were in the same region of space past this lookback time and the Milky Way resided in the LMC’s virial radius, then the Milky Way would have been determined as the LMC’s satellite. At $z = 4$, or a lookback time of 11.95 Gyr, the LMC’s stellar mass was calculated as $10^{8.06} M_{\odot}$, with the Milky Way’s found to be $10^{7.49} M_{\odot}$, thus the LMC was approximately 3.72 times more massive at this point.

3.2.2. Colour Evolution

The calculated evolving $(U - V)_{rest}$ colour of the Milky Way and LMC together with the median and interquartile range of progenitor and satellite colours is shown in Figure 9. For both galaxy calculations, the evolving rest frame colours are lower than the respective median values across the entire redshift range. As the rest-frame colour

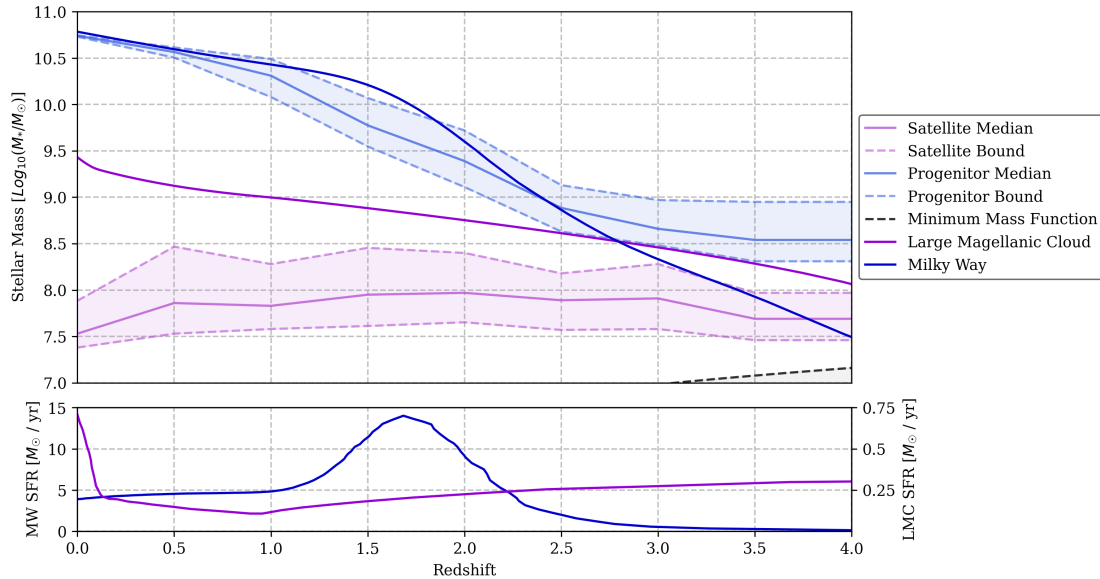


FIG. 8: The calculated stellar mass history of the MW and LMC using the SFH from Fantin et al. (2019) and Harris & Zaritsky (2009), respectively. The stellar mass is compared with the CEERS progenitors and candidate satellites through the median mass and upper and lower quartile range in redshift bins of with $z_{width} = 0.5$, with these bounds shaded in accordance with the legend. The median and interquartile range are plotted against the lower redshift of the bin range, and hence the median and range values at $z = 4.0$ were repeated from $z = 3.5$. The MMF is shown through the black dashed line. The plot below displays the SFH of the MW and LMC to compare with the shape of the evolving stellar mass history.

is directly affected by changes in the SFH, peaks and troughs in said rate over time show large changes in the evolution of the colours.

For the MW, the colour evolution is slowly increasing across redshifts $2 < z < 4$, and shows approximately the same change in rest-frame colour as the median progenitor values (0.10 for the MW, 0.15 for progenitors), but at a bluer colour. Between redshifts $1 < z < 2$ the MW greatly increases in colour from -0.27 to 0.30 due to the sharp rise and fall in SFR 9.67 Gyr ago. Past this point, the MW again slowly rises in rest-frame colour until a value of 0.43 at $z = 0$, much bluer than the progenitor median value of 1.33. This observation agrees with the bluer colours observed by the MW by Mutch et al. (2011) than similar mass galaxies. The progenitor median features a sharp rise in rest-frame colour over the redshifts $0 < z < 1.5$, increasing in colour by 0.83. In addition to this, across the entire

redshift range the progenitor median rest-frame colour increases by a larger amount (1.07 for the progenitors, 0.80 for the MW).

The evolution of the LMC’s rest-frame colour steadily increases until $z \sim 0.93$ where the colour drops due to the increase in SFR. At $z \sim 0.29$ the colour shows another drop due to another sharp increase in SFR. At $z = 0$, the rest-frame colour was calculated to be 0.06, much smaller than the MW’s and the median satellites (0.86). Unlike the MW, the LMC’s colour evolution does lie within the comparison data’s bounds (in this case, candidate satellites) for $0.80 < z < 1.56$, or 2.58 Gyrs, and without the dramatic increase in SFR would have remained within until $z = 0$. The candidate satellites’ rest-frame colour evolution is similar to the progenitors’, however features a lower overall increase of 0.69, and a slower overall reddening curve. Overall, whilst the LMC’s colour does reside within the range of colours for the candidate satel-

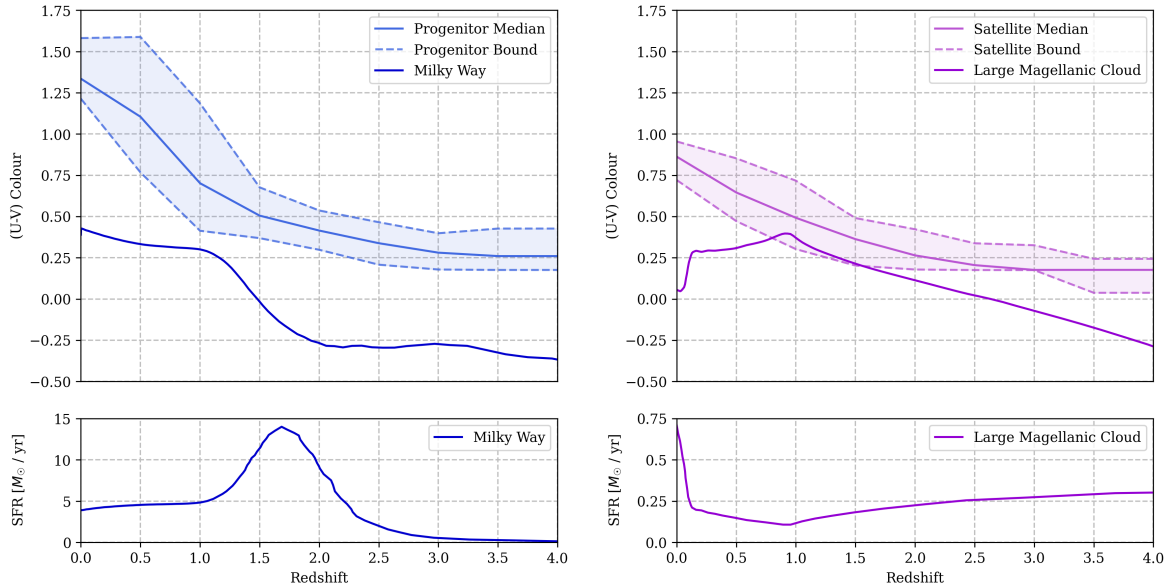


FIG. 9: The calculated $(U - V)_{rest}$ colour of the Milky Way and LMC using the SFH from Fantin et al. (2019) and Harris & Zaritsky (2009), respectively, and an SSP from Bruzual & Charlot (2003). The calculated colours are compared with the CEERS progenitors and candidate satellites through the median colour and upper and lower quartile range in redshift bins of with $z_{width} = 0.5$, with these bounds shaded in accordance with the legend. The median and interquartile range are plotted against the lower redshift of the bin range, and hence the median and range values at $z = 4.0$ were repeated from $z = 3.5$. The plot below displays the SFH of the MW and LMC to compare with the shape of the evolving colour curves.

lites at certain redshifts, its evolutionary path is less similar to the satellites than the MW’s is to the progenitors, suggesting a larger discrepancy in the SFH profile.

4. DISCUSSION

4.1. Candidate Satellite Quenching and Conformity

The results from Section 3.1 demonstrate the candidate satellites found in the CEERS data do not show a statistically significant reddening over similar mass and redshift host galaxies, indicating the satellites are not more likely to be quenched. This result was found for galaxies with stellar masses $M_* \geq 10^{7.16} M_\odot$, whereas previous literature observed a higher quenching likelihood for $M_* \geq 10^{9.30} M_\odot$ (Knobel et al. 2015; Wetzel et al. 2012; Kavinwanichakij et al. 2016). The result that satellites are shown to have higher quenching likelihoods

has physical explanation, and is the result of strangulation, tidal forces and harassment. Strangulation is the effect of a satellite’s cold gas supply halting after satellite falls into the host’s halo, strong gravitational tidal forces from the host can cause tidal stripping of stars and gas in the satellite, as well as heating, and harassment refers to repeated encounters with the host galaxy, which may strip gas from the satellite.

To probe the CEERS candidate satellites for signs of quenching, the candidate satellites were separated into stellar mass ranges of width $10^{+0.5} M_\odot$ from $10^{7.16} M_\odot$ to $10^{10} M_\odot$, across the entire redshift range $0 < z < 4$. Again, host galaxies with the most similar stellar mass and redshift were found, and the KS test was performed for their rest-frame colour distributions. Of the stellar mass ranges, all KS tests found both rest-frame colour distributions to be drawn from the same distribution except for satel-

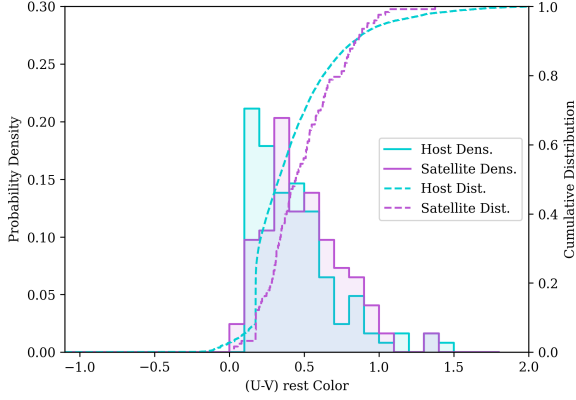


FIG. 10: The rest-frame colour densities for candidate satellites and the most similar stellar mass and redshift field galaxies between the stellar mass range $10^{8.5} < M_* < 10^{9.0}$. The cumulative distribution is also plotted, with the values given in the right axis.

lites with stellar masses between $10^{8.5} M_\odot < M_* < 10^{9.0} M_\odot$. For this range the p -value was found to be 0.018, much less than the test statistic of 0.195. Figure 10 plots the colour probability densities for the satellites and hosts within this stellar mass range, and shows the satellites have a peak colour probability density at redder colours than the closest mass and redshift host galaxies. This mass range for the satellites is approximately immediately above the upper quartile plotted in Figure 8, such that only the 25% most massive satellites demonstrate quenching.

For the galactic conformity, the results of the Pearson correlation coefficient demonstrated weak conformity for redshifts $0.3 < z < 0.6$ and $1.6 < z < 2.5$, and no conformity $0.6 < z < 1.6$ and $2.5 < z < 4$. The redshift ranges were applied for direct comparison with the results of Kawinwanichakij et al. (2016), and agreed with the results for $0.3 < z < 0.6$ and $1.6 < z < 2.5$. These results were determined using a minimum satellite stellar mass of $10^{7.16} M_\odot$, and hence for direct comparison the coefficients were redetermined for the minimum mass of $10^{9.30} M_\odot$. For redshifts $0.3 < z < 0.6$, no satellites were found. For redshifts 0.6

$< z < 1.6$, a coefficient of -0.05 was found, indicating no conformity. For $1.6 < z < 2.5$ a coefficient of -0.11 was found, suggesting weak anti-conformity, and for $2.5 < z < 4$ a coefficient of -0.74 was found, indicating strong anti-conformity. Whilst these results disagree entirely with Kawinwanichakij et al. (2016), the number of satellites found above $10^{9.30} M_\odot$ is on the scale of 10^1 , with only 3 found in the final redshift range. As such, these results are given less significance, but the coefficients determined for all candidate satellites suggests conformity is not strongly observed within the CEERS data, and weak conformity is only present in select redshift regimes.

4.2. Milky Way and Progenitor Evolutions

The results of Section 3.2 calculate the MW to exhibit a bluer rest-frame colour than similar mass galaxies, agreeing with the results of Mutch et al. (2011). The calculation also demonstrates that the MW possessed a bluer colour than the progenitors across the entire redshift range up to $z = 4$, or a total of 11.9 Gyrs. For the MW to possess bluer rest-frame colours, the SFR of the Milky Way must be greater than the progenitors observed. From Figure 9, the colour evolution of the progenitors across the entire redshift range $0 < z < 4$ has an approximately similar shape as the MW does for $z > 1$. For $z > 1$, the MW's SFH shows a peak and steady decline in star formation rate afterwards, similar to a declining tau model. The difference in the shape of the colour evolution between the MW and the progenitors arises due to the MW's continued, constant star formation after the peak (for $z < 1$), hence suggesting the progenitors' SFR continue to decline after their respective peaks. To further explore this trend, the SFR was found for each progenitor, with the median and interquartile range within redshift ranges of width $z_{width} = 0.5$ plotted against the MW's SFH from Fantin et al. (2019) in Figure 11.

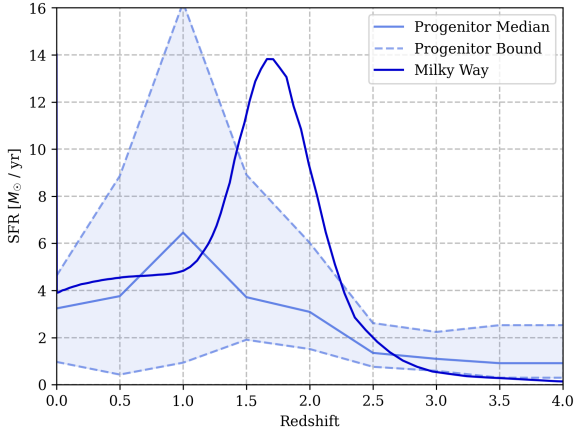


FIG. 11: The calculated $(U - V)_{rest}$ colour of the Milky Way and LMC using the SFH from Fantin et al. (2019) and Harris & Zaritsky (2009), respectively, and an SSP from Bruzual & Charlot (2003).

The figure displays a difference in redshift at which the SFR is peaked for the MW and the progenitors: the MW’s SFR peaked at $z \sim 1.68$, whereas the progenitors had a SFR peak between $1.0 < z < 1.5$. The MW’s SFR peak is also much more rapid than the progenitors, rising from $2 M_{\odot} \text{ yr}^{-1}$ at $z = 2.5$ to $\sim 14 M_{\odot} \text{ yr}^{-1}$ at $z \sim 1.68$, corresponding to a lookback time of 10.90 Gyr to 9.67 Gyr. Although the upper quartile of the progenitors peaks at a larger value than the MW ($16 M_{\odot} \text{ yr}^{-1}$), the MW displays a greater SFR than the progenitor median values for $z < 2.5$. Upon narrowing the redshift bin width to $z_{width} = 0.125$, the peak progenitor SFR was found to occur between redshifts $1.25 < z < 1.375$, hence the MW’s peak SFR occurred between 0.72 Gyr - 1.08 Gyr prior to its progenitors. These results are consistent with Papovich et al. (2015), who found progenitors to peak in SFR at $z \sim 1.5$. As such, the MW is shown to have a unique SFH due to a peak at a higher redshift whilst retaining a greater SFR today.

One explanation for the rapid onset of star formation experienced by the MW during redshifts $1.68 < z < 2.5$ may be due to the merging with satellite galaxies. The MW is known to have experienced mergers within this redshift range, the largest be-

ing with the Gaia-Enceladus (GE) galaxy and the Kraken galaxy, proposed by Kruijssen et al. (2018). Through observations of RR Lyrae stars, Helmi et al. (2018) derived results for the GE merger at $z \approx 1.8$ (10 Gyr ago) and stellar mass of $M_{*} = 6.0 \times 10^8 M_{\odot}$. For the Kraken galaxy, Kruijssen et al. (2020) used an artificial neural network trained on the E-MOSAICS cosmological simulations, determining the stellar mass and accretion redshifts (redshift when merger occurred) for the galaxy as $M_{*} = 1.9^{+1.0}_{-0.6} \times 10^8 M_{\odot}$ at $z = 2.26^{+0.39}_{-0.45}$ (10.61 Gyr ago). By using the EAGLE suite of cosmological simulations, Mackereth et al. (2019) found only 14% of simulated MW-haloes had accretion events the size of GE or Kraken for $z > 1.5$, further demonstrating the MW’s atypical accretion history and thus SFH. As a result, the discrepancy shown between the MW’s and the progenitors’ SFHs can be explained by considering the unique merger events that occurred in the MW’s history.

To compare these merging satellites with the obtained progenitor-satellites, the stellar mass ratio of the MW and the GE and Kraken galaxies at the point of merging was found to be 12.07 and 8.09, respectively, placing the ratios at ≈ 1 on the logarithmic scale in Figure 7. Of the 586 progenitors found with at least one candidate satellite, 436 ($\sim 75\%$) were observed to have a progenitor mass to maximum satellite mass ratio less than 1 on the logarithmic scale in Figure 7. This means that 436 progenitors have at least one relatively more massive satellite than the two merging galaxies GE and Kraken were for the MW. As the mass ratio decreases with decreasing progenitor stellar mass by a power law, less massive progenitors have satellites with a more similar stellar mass to themselves. Less massive progenitors also have much smaller virial radii (see Figure 2), and so these more-similar stellar mass satellites orbit much closer to the progenitors. Because of this, merging will thus be more likely at lower stellar mass progenitors, and using the knowledge of the rapid rise in

SFR experienced by the MW due to the GE and Kraken mergers, will be a much more dominant cause of stellar mass growth than higher stellar mass progenitors. The process of progenitors merging with their most massive satellites also explains the increase in mass ratio with increasing progenitor stellar mass. These results are consistent with Rodriguez-Gomez et al. (2015), who found the merging rate to strongly increase at higher redshift, and Martin et al. (2017) who found the star formation attributed to mergers was higher at $z \sim 3$ than at $z \sim 1$.

Interestingly, the LMC is calculated to have had a greater stellar mass than the GE and Kraken galaxies at the point they merged with the MW ($4.86 \times 10^8 M_\odot$ at $z = 2.26$ and $6.37 \times 10^8 M_\odot$ at $z = 1.8$). If the LMC was orbiting the MW at these times, then the LMC should have also merged with the MW. However, the LMC is observed to be gas rich at a distance from the MW where other satellites are poor, it has a greater velocity than other MW satellites, and has a stream of H_I gas trailing the galaxy. This is consistent with the LMC being on its first passage around the MW (Besla et al. 2007). Using a neural network algorithm trained on the EAGLE hydrodynamical simulation, Barmantloo & Cautun (2023) determined the infall time of the LMC to the MW to be 1.5 Gyr, corresponding a redshift of 0.12. If the LMC has only been inside of the MW’s virial radius for $z \leq 0.12$, then this explains why the LMC has been able to grow steadily in stellar mass without merging with other hosts and is now the largest stellar mass satellite of the MW.

5. CONCLUSION

This paper achieved the goal of finding MW progenitors in the CEERS data by implementing an evolving number density and determining the range of stellar mass values through the ZFOURGE cumulative mass function. In addition to this, candi-

date satellites were successfully determined by calculating the progenitors’ virial radii and implementing redshift probability density thresholds. To ensure completeness in the CEERS data across the entire redshift range $0 < z < 4$, a minimum stellar mass was set as $10^{7.16} M_\odot$ which was derived from the calculated MMF.

With this threshold set, the candidate satellites were tested for satellite conformity and increased likelihood of quenching. The Kolmogorov-Smirnov two-sample test was performed on the satellites and the closest stellar mass and redshift hosts, finding the two to have been drawn from the same distribution and hence displaying no satellite quenching. To compare more directly with the higher stellar masses of previous literature (Knobel et al. 2015; Wetzel et al. 2012; Kawinwanichakij et al. 2016), the stellar mass range of the satellites was varied, and the only mass range in which quenching was observed was found to be $10^{8.5} M_\odot < M_* < 10^{9.0} M_\odot$. Galactic conformity was successfully test by determining the Pearson correlation coefficient between the satellites’ sSFR and the progenitors’. The results found weak conformity for redshift ranges $0.3 < z < 0.6$ and $1.6 < z < 2.5$, agreeing with the results of Kawinwanichakij et al. (2016), but disagreed with the strong conformity between $0.6 < z < 1.6$ observed by the paper.

The MW’s and LMC’s stellar mass and rest-frame colour evolution was calculated and the results compared with the progenitors and candidate satellites. The MW displayed a much bluer rest-frame colour (0.43) than the median progenitor value (1.33) at $z = 0$, whilst possessing a similar stellar mass, agreeing with the results of Mutch et al. (2011). The shape of the MW’s SFH was compared with the obtained progenitor’s SFH median and interquartile range, finding the MW to have peaked in SFR 0.72 Gyr - 1.08 Gyr before the progenitors. However, the MW was shown to possess a larger SFR at $z = 0$ than the progenitors’ median value, explaining the bluer rest-frame colour. The

physical reasons behind the MW’s earlier SFH peak was explored, with the merger between the MW and Gaia-Enceladus and Kraken galaxies as potential causes for the rapid increase in SFR between redshifts $1.68 < z < 2.5$.

An exploration into the effect of satellites on the evolution progenitors was made by plotting the mass ratio of the progenitor’s stellar mass and its most massive satellite against the stellar mass of the progenitor. On the log-log scale, the Pearson correlation coefficient was found to be 0.41, demonstrating a power law dependence between the two. As lower mass progenitors were found with smaller mass fractions and smaller virial radii, their orbiting satellites will be found closer to them and with more similar-stellar masses. From this, progenitor-satellite merging is more likely at lower stellar masses (and hence higher redshift) and so may be the dominant force behind mass growth. Martin et al. (2017) observed the SFR of higher redshift galaxies to be more attributed to merging, agreeing with the hypothesis of progenitor-satellite merging.

The redshifts used for the CEERS galaxies were obtained by the photometric template fitting codes EAZY (Brammer et al 2008) and FAST (Kriek et al. 2009). Because of this, the redshifts used in find-

ing the candidate satellites were probabilistic, with many densities showing broad, low probabilities across large redshift ranges. A spectroscopic measurement would yield much more accurate redshifts, enabling the finding of satellites to be much more accurate and negate the need for arbitrary probability thresholds. With more accurate candidate satellites, the trends of satellite conformity and quenching can be recalculated, likely giving results more similar to previous literature.

6. ACKNOWLEDGEMENTS

I would like to thank Prof. Stijn Wuyts for supervising the project and providing expertise on how to approach the data as well as how to extract meaningful conclusions. I would also like to thank George Shingleton, my project partner, for his contribution to the code used to develop this project and background literature research. The Python packages ‘NumPy’, ‘SciPy’ and ‘astropy’ have also been integral for all of the calculations, and this project have not been possible without them.

7. REFERENCES

-
- Bagley, M., Finkelstein, S., Koekemoer, A., et al. 2023, [ApJL](#), **946**, L12
- Barmantloo, S. & Cautun, M. 2023, [MNRAS](#), **520**, 1704
- Besla, G., Kallivayalil, N., Hernquist, L., et al. 2007, [ApJ](#), **668**, 949
- Boylan-Kolchin, M., Bullock, J. S., Sohn, S. T., Besla, G., & van der Marel, R. P. 2013, [ApJ](#), **768**, 140
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, [ApJ](#), **686**, 1503
- Bruzual, G. & Charlot, S. 2003, [MNRAS](#), **344**, 1000
- Burkert, A., Schreiber, N. F., Genzel, R., et al. 2016, [ApJ](#), **826**, 214
- Dehnen, W., McLaughlin, D. E., & Sachania, J. 2006, [MNRAS](#), **369**, 1688
- Fantin, N., Girardi, L., Marigo, P., et al. 2019, [ApJ](#), **887**, 148
- Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, [ApJL](#), **946**, L13
- Harris, J. & Zaritsky, D. 2009, [AJ](#), **138**, 1243
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, [Nature](#), **563**, 85
- Kaffe, P. R., Sharma, S., Lewis, G. F., Robotham, A. S., & Driver, S. P. 2018, [MNRAS](#), **475**, 4043
- Kawinwanichakij, L., Quadri, R., Papovich, C., et al. 2016, [ApJ](#), **817**, 9

- Knobel, C., Lilly, S., Woo, J., & Kovač, K. 2015, [ApJ](#), **800**, 24
- Kriek, M., Van Dokkum, P. G., Labbé, I., et al. 2009, [ApJ](#), **700**, 221
- Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, [MNRAS](#), **498**, 2472
- Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2018, [MNRAS](#), **486**, 3180
- Licquia, T. C. & Newman, J. A. 2015, [ApJ](#), **806**, 96
- Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, [MNRAS](#), **482**, 3426
- Martin, G., Kaviraj, S., Devriendt, J. E. G., et al. 2017, [MNRAS](#), **472**, L50
- Morishita, T., Ichikawa, T., Noguchi, M., et al. 2015, [ApJ](#), **805**, 34
- Moster, B., Naab, T., & White, S. 2013, [MNRAS](#), **428**, 3121
- Mutch, S. J., Croton, D. J., & Poole, G. B. 2011, [ApJ](#), **736**, 84
- Papovich, C., Labbé, I., Quadri, R., et al. 2015, [ApJ](#), **803**, 26
- Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, [MNRAS](#), **449**, 49
- Tomczak, A., Quadri, R., Tran, K., et al. 2014, [ApJ](#), **783**, 85
- van der Marel, R., Alves, D., Hardy, E., & Suntzeff, N. 2002, [AJ](#), **124**, 2639
- Vilardell, F., Ribas, I., Jordi, C., Fitzpatrick, E., & Guinan, E. 2010, [AAS](#), **509**, 7
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, [MNRAS](#), **444**, 1518
- Wellons, S. & Torrey, P. 2017, [MNRAS](#), **467**, 3887
- Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, [MNRAS](#), **424**, 232
- Willmer, C. N. 2018, [ApJ](#), **236**, 47