Tracing galaxy populations through cosmic time: A critical test of methods for connecting the same galaxies between different redshifts at $z < 3$

Carl J. Mundy$^1$, Christopher J. Conselice$^1$ and Jamie R. Ownsworth$^1$

$^1$School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK

23 April 2015

ABSTRACT
Connecting galaxies with their descendants (or progenitors) at different redshifts can yield strong constraints on galaxy evolution. Observational studies have historically selected samples of galaxies using a physical quantity, such as stellar mass, either above a constant limit or at a constant cumulative number density. Investigation into the efficacy of these selection methods has not been fully explored. Using a set of four semi-analytical models based on the output of the Millennium Simulation, we find that selecting galaxies at a constant number density (in the range $-4.3 < \log n [\text{Mpc}^{-3} h^3] < -3.0$) is superior to a constant stellar mass selected sample, although it still has significant limitations. Recovery of the average stellar mass, stellar mass density and average star-formation rate is highly dependent on the choice of number density but can all be recovered to within $< 50\%$ at the commonly employed choice of $\log n [\text{Mpc}^{-3} h^3] = -4.0$, corresponding to $\log M_\odot/h \sim 11.2$ at $z = 0$, but this increases at lower mass limits. We show that there is a large scatter between the location of a given galaxy in a rank ordering based on stellar mass between different redshifts. We find that the inferred velocity dispersion may be a better tracer of galaxy properties, although further investigation is warranted into simulating this property. Finally, we find that over large redshift ranges selection at a constant number density is more effective in tracing the progenitors of modern galaxies than vice-versa.

Key words: galaxies: evolution - galaxies: formation

1 INTRODUCTION
In the now commonly accepted paradigm, galaxies form in the gravitational wells of collapsed cold dark matter halos, which themselves are seeded by primordial quantum fluctuations in the Universe’s first moments. In this hierarchical picture of galaxy formation and evolution, galaxies build up stellar mass through both in-situ star-formation and galaxy mergers, where more massive galaxies merge with smaller systems which result in more massive descendants. Over time, these processes produce the array of galaxies and the environments within which they are observed. In order to understand how galaxies form and evolve, the evolution in their properties (e.g. stellar mass, size) must be observed. As the most massive, and thus the brightest, galaxies are the easiest systems to observe out to high redshift, it is these galaxies and their properties that observations attempt to study. This has typically been achieved by selecting galaxies in two ways.

Historically, selecting galaxies across a redshift range of interest using a constant stellar mass cut has been used to study the evolution of the most massive galaxies (e.g. Conselice et al. 2003; Mortlock et al. 2013). Use of this selection method intrinsically assumes galaxies have more or less been a passively evolving population from high redshift. However, processes such as major galaxy mergers and bursts of star-formation interfere with these assumptions and contaminate the selection by changing the rank order of galaxies. Thus, the wider the redshift range this method is applied to, the less accurately it may trace the galaxies of the original selection.

Selecting galaxies at a constant cumulative comoving number density, when ranked by some physical property such as stellar mass or luminosity, has proven a popular alternative in the recent literature when observing both field and cluster galaxy evolution (e.g. Lin et al. 2013). Using this technique, the averaged star-formation history of a constant number density selected sample of galaxies over the redshift range $3 < z < 8$ has been shown to be able to account for the stellar mass growth of these galaxies (Papovich et al. 2010).
The average stellar mass of the most massive galaxies (those with \( M_*/ > 11.0 \) at \( z = 0 \), and \( \log n \left[ \text{Mpc}^{-3} \right] = -4.0 \)) has been found to increase by a factor of \( \sim 4 \) over the redshift range \( 0.3 < z < 3 \). However, the integrated star-formation history appears unable to account for the growth in stellar mass at \( z < 1.5 \). Therefore, the influence of both major and minor galaxy mergers is required to account for this discrepancy at low redshift (Ownsworth et al. 2014).

Furthermore, studies have observed the evolution in H\( \alpha \) equivalent width, structural properties and colours of galaxy populations selected at various number densities (Fumagalli et al. 2012; Marchesini et al. 2014). Stellar mass measured inside a radius of \( r = 5 \text{kpc} \) on stacked images of massive galaxies (selected at \( \log n \left[ \text{Mpc}^{-3} \right] = -3.7 \) corresponding to galaxies with \( \log M_*^1 \left[ M_\odot \right] > 11.4 \) and \( \log M_*^2 \left[ M_\odot \right] > 11.1 \) at \( z = 0.1 \) and \( z = 2.0 \) respectively) is found to be approximately constant over the redshift range \( 0.6 < z < 2.0 \). On the other hand, the stellar mass content beyond this radius is found to increase by a factor of \( \sim 4 \) (van Dokkum et al. 2010). Conselice et al. (2013) compare the derived gas fractions of massive galaxies (\( M_*/ > 10^{11} M_\odot \)) with their star-formation histories in the redshift range \( 1.5 < z < 3 \), selecting galaxies at a merger-adjusted constant number density. They conclude that gas accretion is the dominant source of observed stellar mass production for these galaxies over this redshift range.

Investigation into the efficacy of either selection method has not been fully explored. Numerical calculations presented in van Dokkum et al. (2010) suggest the influence of galaxy mergers has little influence on stellar mass growth when measured using a constant number density selected sample. Papovich et al. (2010) used dark matter halo merger trees from the Millennium Simulation (MS; Springel et al. 2005; Lenson & Consortium 2006) to show that the recovery fraction of descendant halos at redshifts \( 3 < z < 8 \) is \( \sim 50\% \). Behroozi et al. (2013) found a small change in the cumulative number density of the most massive (\( M_*/ > 10^{11.5} M_\odot \)) \( z = 0 \) progenitor galaxies of \( +0.22 \text{dex} \) per unit \( \Delta z \). Furthermore, they find that this change and thus the mass histories of descendants and progenitors are different. More directly, Leja et al. (2013) used the Guo et al. (2011) G11 semi-analytical model (SAM) applied to the MS in order to test the validity of the underlying assumptions of constant number density selection. They find that, within this particular SAM, the median stellar mass of descendant galaxies can be recovered over the redshift range \( 0 < z < 3 \) to within \( \sim 40\% \) of the true value. Corrections for stellar mass growth rate scatter, galaxy mergers and quenching are found to reduce this discrepancy to within \( \sim 12\% \) - well within typical observational error attributed to the calculation of stellar masses. These results, however, are model and cosmology dependent and are sensitive to the dark matter merger trees and the recipes used to determine galaxy properties. How the stellar mass is calculated is a prime example. Different methods of calculating this may introduce different levels of scatter into the rank order of galaxies across redshifts. Sensibly investigating the efficacy of these techniques requires a mixture of SAMs, merger trees and cosmology to gauge the amount of variance in the results.

While stellar mass is used in the studies mentioned previously, it may not be the most appropriate property with which to rank and select galaxies in order to trace their properties. Increasing evidence has suggested that the central velocity dispersion of a galaxy is a good predictor of galaxy properties, including star-formation rate (SFR) and colour across large redshift ranges. Furthermore, it is thought to be a more stable quantity with redshift compared to, for example, stellar mass (B succès et al. 2012; Wake et al. 2012), partly due to the weak dependence of velocity dispersion on both stellar mass and galaxy size, with \( \sigma \propto (M_*/R_e)^{2} \). Leja et al. (2013) briefly investigated the change in velocity dispersion for descendants of \( z = 3 \) galaxies selected at a constant cumulative number density. They found a small change (\( < 0.15 \text{dex in log} \sigma \)) in the average inferred velocity dispersion from \( 0 < z < 3 \) in the G11 SAM. Similarly, simulations of massive galaxies (\( \log M_* > 10^{8.6} \)) presented in Oser et al. (2012) find an increase in velocity dispersion of \( \Delta \sigma = 0.2 \text{dex} \) over the redshift range from \( z = 2 \) to \( z = 0 \), consistent with observational estimates (e.g. Javier Cenarro & Trujillo 2008; Martinez-Manso et al. 2011). These observations warrant a detailed study into the use of inferred velocity dispersion in place of stellar mass.

To generate an accurate framework for how galaxies form and evolve, one must observe the evolution of their properties over time. With these observations, models can be crafted to explain them. If, however, the evolution is not traced correctly, these frameworks can deviate from the truth. Although the aforementioned literature works provide some arguments to support the use of their selection methods, no study has attempted to quantify the recovery efficiency of these methods. To this end, we study in detail the ability of these selection methods to trace individual galaxies, as well as their stellar mass and star-formation properties over cosmic time.

In this paper we use models of galaxy evolution to investigate the ability of galaxy selections which are a) above a constant stellar mass limit; b) at a constant cumulative comoving number density in stellar mass; c) above a constant stellar velocity dispersion limit; and d) at a constant cumulative comoving number density in stellar velocity dispersion. We test how well these methods trace the true evolution of progenitor and descendant populations initially selected at redshifts \( z = 0 \) and \( z = 3 \) respectively.

This paper is constructed in the following way. In §2 we discuss the data and the metrics we use to test the efficacy of the observational selection methods. In §3 we present the results and analysis of our work when selecting galaxies using their stellar mass. In §4 we present results when selecting galaxies using their inferred stellar velocity dispersions. Finally, we discuss and conclude our main results in §5 and §6. We describe the simulations and associated cosmologies we use in section §2.1.

2 DATA & SELECTION METHODS

In this paper, we determine the efficacy of two different methods in recovering the direct (i.e. most massive) progenitors and descendants of the most massive galaxies from initial selections at redshifts of \( z = 0 \) and \( z = 3 \), respectively. The first selection method is at a constant limit (in either

---

1 http://www.mpa-garching.mpg.de/galform/virgo/millennium/
 Connecting the most massive galaxies at \( z < 3 \)  

stellar mass or velocity dispersion), above which galaxies are selected. The second selection is at a constant cumulative comoving number density. This is achieved by integrating the galaxy stellar mass function (GSMF), or the galaxy velocity dispersion function (GVDF) if velocity dispersion is used, to obtain the integrated number density as a function of stellar mass (or velocity dispersion). From this we obtain the stellar mass limit above which all the galaxies are below a certain number density. The sample we examine at that redshift thus contains all galaxies with a stellar mass greater than this value. Additionally, we want to quantify how well each selection method recovers both the average and sum total stellar mass in the descendant or progenitor populations, as well as the average SFR. The combination of cosmological dark matter simulations and semi-analytical recipes, as well as cosmological hydrodynamical simulations, continue to provide the only environments in which to conduct such an investigation.

2.1 Simulated Data

To test this, we utilise the output of the Millennium Simulation (MS) and the catalogues of four SAMs applied to it and its variants. The original simulation consists of \( 2160^3 \) dark matter particles of mass \( 8.6 \times 10^8 \, M_\odot \) within a comoving box of size \( 500 \, h^{-1} \, \text{Mpc} \) on a side. The MS uses a \( \Lambda \)CDM cosmological model with \( \Omega_m = 0.25, \Omega_b = 0.045, h = 0.73, \Omega_{\Lambda} = 0.75, n_s = 1 \) and \( \sigma_8 = 0.9 \), where the Hubble constant is parameterized as \( H_0 = 100 \, h \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1} \). The Guo et al. (2013, G13) SAM utilises a subsequent simulation which follows \( 2160^3 \) particles of mass \( 9.3 \times 10^8 \, m_\odot \) using an updated WMAP7 cosmology with \( \Omega_m = 0.272, \Omega_b = 0.0455, h = 0.704, \Omega_{\Lambda} = 0.728, n_s = 0.967 \) and \( \sigma_8 = 0.81 \). The main difference between these two simulations is the value of the linear power spectrum amplitude on scales of \( 8 \, h^{-1} \, \text{Mpc} \), which affects the dark matter halo merger rate, and thus the galaxy merger rate, inside the simulation.

Bower et al. (2006, B06) presents an updated variant of the Durham semi-analytical model of galaxy formation (Cole et al. 2000) in which the treatment of active galactic nuclei (AGN) and stellar feedback on halo quenching is improved. They find that these updated treatments reduce the number densities of higher mass galaxies and remove cooling flows from rich clusters. De Lucia & Blaizot (2007, D06) applied their model to the output of the MS with updated treatments for stellar populations, dust attenuation and cooling flow suppression via AGN feedback. They find that supernovae and AGN feedback processes play a vital role in the early quenching of star-formation in the progenitors of local brightest cluster galaxies (BCGs). G11 describe an updated model of galaxy formation and evolution with new recipes for supernovae feedback and galaxy bulge sizes among others. They find that the simulated abundance of massive galaxies, with log \( M_* \, [M_\odot] > 11.0 \), are consistent with observations out to \( z \sim 1.0 \). However, they over predict galaxies of lower stellar mass beyond \( z \sim 0.6 \) and under predict massive galaxies at \( z > 1.0 \) by at least an order of magnitude (see Fig. 23 in G11). Finally, G13 describe the results of implementing their SAM in a WMAP7 cosmology. They find a requirement for weaker feedback and star-formation efficiency than a WMAP1 cosmology in order to reproduce the observed local GSMF. Merger trees used by B06 are described in Harker et al. (2006) while those employed by the remaining SAMs are presented in Springel et al. (2005). It is these models based upon these merger trees extracted from the Millennium Simulation from which we study the observational selection methods.

Physically motivated semi-analytical models of galaxy formation (see, for example, Bower et al. 2006, Vogelsberger et al. 2014) applied to cosmological dark matter simulations, such as the MS, provide an unparalleled tool to probe the evolution of dark matter halos and the galaxies that reside within them. Simulations and observations at low redshift (\( z < 2 \)) are found to be consistent in many respects (e.g. luminosity functions), however SAMs have varying degrees of success in matching observational quantities beyond this. Comparison of different SAMs and other models show simulated galaxy stellar mass functions are generally consistent with most observations out to \( z \sim 2 \) if feedback mechanisms from AGN and supernovae are included (Croton et al. 2006, Lu et al. 2013). This agreement extends to comparison between simulated and observed major mergers for the most massive galaxies inside the MS (Bertone & Conselice 2003), but not for lower mass systems.

2.2 Selection Method Metrics

It is prudent to measure how accurately each selection method samples the progenitors or descendants of the galaxy population being studied. In this work, descendants of an initial \( z = 3 \) selection are identified by following the ‘descendantId’ property in the SAM output catalogues. At each step, duplicate descendants (due to mergers between two or more galaxies) are removed such that the number of true descendants decreases with time. On the other hand, progenitors of an initial \( z = 0 \) selection are defined as the most massive galaxy in the previous redshift snapshot that came to be the galaxy in the current snapshot. These definitions allow traversing of different branches along merger trees, depending on the direction we take. In this work we measure the ability of each selection method to recover galaxy properties using various metrics.

Firstly, the recovery fraction quantifies how many of the available progenitors or descendants are recovered at different redshifts in the sample obtained using a given selection method such that

\[
\frac{f_{rec} = N_s}{N_{tot}},
\]

where \( N_s \) is the number of descendants/progenitors included in the observational selection, and \( N_{tot} \) is the total number of descendants/progenitors available to be selected.

Although helpful, it may not strictly be necessary to sample the descendants or progenitors of interest to reproduce the true evolution of galaxy properties - sampling different galaxies from the true progenitors or descendants might be sufficient if the galaxies replacing those lost have similar properties. Therefore a low recovery fraction may not necessarily correspond to an inability to recover the true evolution in, for example, average stellar mass or SFR. Because of this, we also consider the fraction of the observed sample that is not a galaxy of interest, and call this the contamination fraction, defined as
where \(N_{\text{sel}}\) is the number of galaxies within the observed selection and \(N_s\) is the number of true (i.e. most massive) descendants/progenitors included in the selection. In this paper we also compare the true mean stellar mass, \(\bar{m}_{\text{true}}\), of the progenitors or descendants to that observed using each selection method, \(\bar{m}_{\text{obs}}\), defined as

\[
\kappa_{m^*} = (\bar{m}_{\text{obs}} - \bar{m}_{\text{true}}^*)/\bar{m}_{\text{true}}^* = \Delta \bar{m}^*/\bar{m}_{\text{true}}^*.
\]

In a similar fashion, the ability to trace the evolution of the stellar mass density, or sum of the stellar mass, is important to understand the build up of stellar mass in galaxies over time. We compare the observational selection techniques’ abilities to return the true mean stellar mass density, quantified as

\[
\kappa_\Psi = \Delta \Psi^*/\bar{\Psi}_{\text{true}}^*.
\]

where \(\bar{\Psi}^*\) is the mean SFR. We choose to study these galaxy properties in particular because they are the most fundamental, and the most used in the literature thus far. Additionally, we investigate whether each selection method is best applied to tracing progenitor or descendant galaxy populations, i.e. whether the selection methods are best applied forwards or backwards in time.

### 2.3 Velocity Dispersion Selection

As central velocity dispersion has been shown to exhibit a shallow evolution over time, it is prudent to investigate this physical property as a tracer. From scalar virial theory the stellar velocity dispersion of a system can be estimated by

\[
\sigma^2 = \frac{GM_\star}{5R_{\frac{1}{2}}}
\]

where \(G\) is Newton’s gravitational constant, \(M_\star\) is the total stellar mass and \(R\) is the half-mass radius (see, e.g., Cappellari et al. 2006). Using the reported total and bulge stellar masses and sizes, we follow Leja et al. (2013) in estimating the half-mass radius of each simulated galaxy as

\[
R_{\frac{1}{2}} = \frac{M_d R_d + M_b R_b}{M_b + M_d}.
\]

As before, \(M\) is stellar mass, \(R\) is the half-mass radius and subscripts \(b\) and \(d\) correspond to the bulge and disk components respectively. We correct the disk scale radius provided in the SAM catalogues to convert it to a half-mass radius such that \(R_d = 1.678 R_{\text{scale}}\). This relation is ideally obtained numerically however approximations are available for a range of Sédov (1962) indices. We refer the interested reader to Graham & Driver (2005) and references therein for more information.

We calculate these metrics using both stellar mass and velocity dispersion at four constant number density selections and four constant limits. Number density values are chosen to cover a wide range, allowing comparison with previous work, and to be representative of what number densities are currently applicable to observational studies. Constant limits in both stellar mass and velocity dispersion are chosen to enable comparison with the number density selections. Constant stellar mass limits are defined as the mass limit of a number density selection at either \(z = 0\) or \(z = 3\), depending on whether progenitors or descendants are being investigated. In short, the limits are chosen such that the initial selections, whether at \(z = 0\) or \(z = 3\), are the same. In Fig 3 we report our findings using stellar mass as the ranking property and in Fig 4 we report the results of using inferred velocity dispersion.

### 3 STELLAR MASS SELECTIONS

We investigate the ability of two different galaxy selection methods, using two different galaxy properties, to recover the mean stellar mass, stellar mass density, mean velocity dispersion and average SFR of progenitor and descendant populations. All our results are available to download online.

#### 3.1 Descendants

In Figure 1 we show the results of selecting descendants at four constant stellar mass limits, described in Table 1. This is defined as the stellar mass limit for a number density selection beginning at \(z = 3\) and examining evolution at lower redshift. As one might expect, all constant stellar mass selections recovered all available descendant galaxies at every redshift (top panel), as galaxies typically experience a net gain in stellar mass over these redshift ranges. However, the fraction of the selected sample that are not descendants of interest, \(f_{\text{contam}}\), increases to \(\sim 50\%\) by \(z = 2\). By \(z = 0\), samples selected above each stellar mass limit are almost completely contamination. This shows that using a constant mass cut at \(z = 3\) selects essentially none of the same galaxies (descendants) at lower redshift.

Contrary to what this metric might suggest, the difference between the observed and true mean stellar mass is underestimated by only \(50\%\) by \(z = 0\), decreasing linearly at lower redshift. This seems to be largely independent of the stellar mass limit in all but the B06 SAM which fares relatively better at higher limits. Recovery of the median stellar mass is indistinguishable from the mean stellar mass for the two smallest number densities. At the two largest choices however, the median stellar mass is further underestimated towards lower redshift such that at \(z = 0\), this property is underestimated by \(\sim 60\%\). The stellar mass density is overestimated by a factor of \(\sim 4\) \((\sim 20)\) at the lowest (highest) mass limit in all SAMs by \(z = 0\). As the B06 SAM does not report SFRs in its catalogues, we instead consider only the remaining three SAMs in recovery of the mean SFR. At the lowest stellar mass limit, this is recovered to within \(\sim 10\%\) down to \(z = 1\). At lower redshifts, however, the SFR begins to be increasingly underestimated and by \(z = 0\) it is underestimated by \(\sim 50\%\).

2 http://goo.gl/X6Q2Sc

© 2015 RAS, MNRAS 000, 1–13
Connecting the most massive galaxies at $z < 3$

Figure 1. Recovery fraction of individual descendants, their average stellar mass, stellar mass density and average SFR (rows) for four constant stellar mass limit selections (columns) covering the stellar mass range at $z = 3$ of $\log M_\ast > 10.7, 10.6, 10.4, 10.2$. Initial stellar mass selection limits for each SAM are given in the third column of Table 1. SAMs used are Bower et al. (2006, B06), De Lucia & Blaizot (2006, D06), Guo et al. (2011, G11) and Guo et al. (2013, G13), represented by solid green, dashed orange, dashed-dotted magenta and dotted blue lines respectively. (A colour version of this figure is available in the online journal.)

Figure 2. Recovery fraction of individual descendants, their average stellar mass, stellar mass density and average SFR (rows) for four constant number density selections (columns) covering the number density range $-4.3 < \log n < -3.0$. These approximately correspond to stellar masses at $z = 3$ of $\log M_\ast > 10.7, 10.6, 10.4, 10.2$. Initial stellar mass selection limits for each SAM are given in the third column of Table 1. SAMs used are Bower et al. (2006, B06), De Lucia & Blaizot (2006, D06), Guo et al. (2011, G11) and Guo et al. (2013, G13), represented by solid green, dashed orange, dashed-dotted magenta and dotted blue lines respectively. (A colour version of this figure is available in the online journal.)

3.2 Progenitors

Now we take an initial selection at $z = 0$ and trace the most massive progenitors back to higher redshifts. As Figure 3 shows, the recovery of individual progenitors using a constant stellar mass limit (detailed in Table 1) deteriorates exponentially such that at $z = 1$, only 30% are recovered in the selection at the smallest stellar mass limit and less than 5% at the largest stellar mass limit (smallest number density). The sample’s contamination fraction increases immediately and, at all stellar mass limits, the sample is > 95% contamination by $z = 3$.

The discrepancy between the true and observed mean stellar mass of the progenitors increases approximately linearly with redshift, overestimating the mean mass by a factor of three at $z = 3$, independent of stellar mass limit and weakly dependent on the choice of SAM. Recovery of the median stellar mass is again indistinguishable from the mean recovery at the two largest stellar mass limits. At the two highest, the median mass is overestimated by factors of $3 \rightarrow 7$ times. Furthermore, observed stellar mass density is increasingly underestimated with redshift in all SAMs. Finally, the mean SFR is recovered to within a factor of $\sim 4$ by $z = 3$. 

We show in Figure 2 that by using a constant number density selection the recovery fraction, $f_{\text{rec}}$, decreases exponentially with decreasing redshift such that by $z = 0$ between 30% and 60% of the available descendants are selected. The contamination fraction is found to vary between half and three quarters of the sample at the highest and lowest number densities respectively. Mean stellar mass is overestimated by a factor that increases with both number density and redshift, overestimating the true value by a factor of 1.3 and 1.6 by $z = 0$. While not plotted, the median stellar mass is qualitatively similar above $z \sim 0.5$. Below this, the median stellar mass is further overestimated by up to 30% more at $z = 0$. Similarly, the stellar mass density is eventually overestimated by a factor of 1.5 (1.8) at the largest (smallest) number densities. Finally, over the entire redshift range the SFR is recovered to within $\sim 50\%$ in all SAMs and at all number density choices.

Compared to a constant stellar mass selection, a constant number density selection recovers far fewer of the true descendants at lower redshift. However, at all number density choices, the lower redshift selections have considerably less contamination.
Selecting progenitors at a constant number density fares relatively better, as shown in Figure 4. Out to \(z = 3\), no less than \(\sim 50\%\) \((\sim 30\%)\) of progenitors are recovered at the largest \((\)smallest\) selections. The mean stellar mass is recovered to within a factor of \(\sim 1.5\) at \(z = 3\) in all SAMs and choices of \(n\) and the observed stellar mass density follows a very similar trend. Median stellar mass recovery is indistinguishable from the mean recovery except at the largest number density. Here up until \(z \sim 2\) where the overestimation becomes larger by approximately \(\sim 50\%\). Lastly, the mean SFR is recovered to within \(\pm 20\%\) at all number density selections except for D06, which overestimates the SFR at a peak of \(\sim 50\%\) at \(z = 1\).

Comparing these results with the descendant population \((\Sect{2.3})\), we find that a constant cumulative comoving number density selection recovers all descendant population properties within a factor of two of the true value. Similarly, all the progenitor properties are recovered to within a factor of 1.5 of the true value. Furthermore, a constant number density selection appears to trace the ensemble progenitor properties of \(z = 0\) massive galaxies better than the descendants of those at \(z = 3\).

### 3.3 Fitting Forms

For convenience, we fit parametric functions to the metrics described above for a constant number density selected sample of galaxies. These fits are strictly valid over the redshift range \(0 < z < 3\). For each metric, we take the average value of all SAMs at each redshift. Next, at each redshift we randomly sample a Gaussian of width equal to the spread between SAMs, centred on the mean value. The function is then fit to these points. These steps are repeated \(10^4\) times to obtain the average parameters and their associated errors for each metric at each number density. Our detailed results of fitting for all number densities are reported in Table A1.

| SAM | \(n [\text{Mpc}^{-3} \text{h}^3]\) | Descendants \((z = 3)\) | Progenitors \((z = 0)\) |
|-----|-------------------------------|------------------|------------------|
|     | \(\log M_{\text{lim}}/h\)    | \(\log M_{\text{lim}}/h\) |                  |
| B06 | \(1 \times 10^{-3}\)         | 10.32            | 10.87            |
|     | \(5 \times 10^{-4}\)         | 10.51            | 11.01            |
|     | \(1 \times 10^{-4}\)         | 10.76            | 11.22            |
|     | \(5 \times 10^{-5}\)         | 10.82            | 11.30            |
| D06 | \(1 \times 10^{-3}\)         | 10.31            | 10.92            |
|     | \(5 \times 10^{-4}\)         | 10.46            | 11.04            |
|     | \(1 \times 10^{-4}\)         | 10.68            | 11.29            |
|     | \(5 \times 10^{-5}\)         | 10.74            | 11.39            |
| G11 | \(1 \times 10^{-3}\)         | 10.21            | 10.84            |
|     | \(5 \times 10^{-4}\)         | 10.33            | 10.97            |
|     | \(1 \times 10^{-4}\)         | 10.52            | 11.24            |
|     | \(5 \times 10^{-5}\)         | 10.59            | 11.35            |
| G13 | \(1 \times 10^{-3}\)         | 10.00            | 10.77            |
|     | \(5 \times 10^{-4}\)         | 10.15            | 10.89            |
|     | \(1 \times 10^{-4}\)         | 10.40            | 11.12            |
|     | \(5 \times 10^{-5}\)         | 10.48            | 11.21            |

The descendant galaxy recovery and contamination fraction is fit with a function of form

\[
f = a + b \times \exp(c \times (1 + z)),\]

where \(a\), \(b\) and \(c\) are free parameters. We find that forcing \(a\) to zero for the recovery fraction, and \(c\) to unity for the contamination fraction, gives better fits. The recovery of average stellar mass \((\kappa_\text{m}^\star)\) and stellar mass density \((\kappa_\rho^\star)\) are then parametrised as

\[
\kappa = a + b \times (1 + z).\]

To fit the progenitor galaxy contamination fraction, we modify the parametrisation to the form

\[
f' = a + b \times (1 + z) + (1 + z)^c.\]

Finally, we do not fit the recovery of average SFR recovery. Fitted parameters and associated errors are available in Table A1 for constant cumulative number density selected samples.
Connecting the most massive galaxies at \( z < 3 \)

Figure 4. Recovery of individual progenitors, their average stellar mass, stellar mass density and average SFR (rows) for four constant number density selections (columns) covering the number density range \(-4.3 < \log n < -3.0\). The B06, D06, G11 and G13 models represented by solid green, dashed orange, dashed-dotted magenta and dotted blue lines respectively. (A colour version of this figure is available in the online journal.)

4 INFERRED VELOCITY DISPERSION SELECTIONS

In an era of ever larger and deeper spectroscopic surveys, the extra information these observations afford of internal properties could possibly be employed as a better tracer of progenitor or descendant galaxy properties. One product of such a survey is the measurement of the central stellar velocity dispersion of a galaxy.

As the SAM catalogues do not report the velocity dispersions of galaxies, we infer this quantity using Equations 6 and 7. The inferred velocity dispersion, as defined here, is a relatively direct observable at the redshifts probed because the stellar mass and half-mass radius are observable. In this section we report the results of using this property in place of stellar mass. As a reminder, this would be achieved observationally by integrating the GVDF to obtain the cumulative number density of galaxies as a function of their velocity dispersion. The velocity dispersion limit, above which all galaxies are at a number density \( n \), can simply be read off.

4.1 Descendants

As displayed in Figure 5, selecting galaxies above a constant inferred velocity dispersion limit, given in Table 2, results in slowly losing descendants with decreasing redshift in the B06 SAM. However, in the G11 and G13 SAMs, the recovery fraction increases below \( z < 2 \). At \( z = 0 \), 90% and 60% of descendant galaxies are selected above the lowest and highest inferred velocity dispersion limits respectively. As with selection above a constant stellar mass limit, the contamination fraction increases exponentially towards lower redshift. At the lowest (highest) velocity dispersion limits there is significant contamination in the observed sample at the level of 70% (90%). Recovery of the descendant mean stellar mass is increasingly underestimated. The true value is maximally underestimated at \( z = 0 \) at all inferred velocity dispersion limits by 30%. Conversely, the stellar mass density is increasingly overestimated with time by up to a factor of \( \sim 5 \) times the true value. Finally, a constant inferred velocity dispersion selection recovers the descendants’ average SFR to within 10% at all limit choices and redshifts.

In Figure 6 we show the result of selecting at a constant cumulative number density in inferred velocity dispersion. Inferred velocity dispersion is, as defined in this paper, a function of and proportional to stellar mass for each galaxy type (early and late) and so it is not surprising that the results are similar to those obtained in §3. Comparing with the stellar mass selections described in §3.1, these results suggest inferred velocity dispersion is just as competent a tracer as stellar mass, and even more accurate in some cases. However, any improvements are small over the use of stellar mass.
4.2 Progenitors

Figures 7 and 8 display the results of attempting to trace the progenitors of $z = 0$ galaxies via selection above a constant inferred velocity dispersion and at a constant cumulative number density, ordered by inferred velocity dispersion, respectively. Selection above a constant limit loses progenitor galaxies from the sample with increasing redshift. By $z = 3$, only 10% of true progenitors are sampled at all velocity dispersion limits and SAMs. The average stellar mass is increasingly overestimated with redshift, by 50 – 100% at $z = 3$. The stellar mass density is increasingly underestimated. In the B06 SAM, it is even more underestimated at higher inferred velocity dispersion limits than at smaller limits. However, in the G11 and G13 SAMs, it is underestimated by 70% at all limits by $z = 3$. Similarly, the SFR is recovered to within 50% in the two SAMs considered.

Selection at a constant cumulative number density, ordered by inferred velocity dispersion, results in recovery fractions similar to the number density selection using stellar mass. Decreasing exponentially from $z = 0$, 50% (30%) at the largest (smallest) number densities. Both the average stellar mass and the stellar mass density are recovered to within 40% of the true value across all redshifts, SAMs and number densities investigated in this paper. Finally, the SFR is recovered to within 20% at all times.
Furthermore, lower mass systems from below the selection at high redshift increase their stellar mass at a higher rate than those more massive systems and become most of the most massive galaxies in the local Universe.

It is worth noting that all of the SAMs used in this work fail to match observed galaxy stellar mass functions beyond some redshift (typically \(z \sim 1.5\)) meaning that they also fail to reproduce the observed evolution of certain galaxy populations. A known problem with the original MS is the cosmology used. Use of a larger \(\sigma_8\) than currently observed \cite{Komatsu2011, Collaboration2015} will increase the merger rate and therefore the scatter in the rank order of galaxy stellar mass. Furthermore, this cosmology produces a larger population of quenched galaxies earlier than observed. This requires the SAMs to build up the low mass end of the GSMF at early times in order to match the observed local stellar mass function. See \cite{Leja2013} and \cite{Genel2011} for an in-depth discussion into this and other issues. We would thus expect less scatter in the real Universe, and therefore better recovery of galaxy properties, compared to the results obtained here. Cosmological hydrodynamical simulations, e.g. \cite{Furlong2014} and \cite{Genel2014}, show closer agreement with observed galaxy stellar mass functions out to high redshift and may offer a better environment in which to conduct these tests.

5.1 Where are the progenitors of \(z = 0\) massive galaxies at high redshift?

A simple question we can ask is where exactly in the ranking (in either stellar velocity dispersion or stellar mass) are the progenitors of the most massive galaxies at \(z = 0\) at earlier times. Taking the two extremes of the number density choices in this work, we show in Figure \ref{f9} the stellar mass distributions of the most massive progenitors of \(z = 0\) galaxies in the G13 SAM selected at two extreme number densities (top row). At each redshift we fit the stellar mass distributions with a Gaussian function and show in the middle panels that fit residuals are minimal (< 5%). This is done to quantify the changes in these distributions as a function of redshift. In the bottom row, we show how the properties of these distributions change with redshift. The mean and widths (distribution standard deviation) are shown in the bottom left and bottom right panels respectively.

We find that at higher redshift, the mass distributions move systematically towards smaller mass galaxies and the stellar mass distribution widens. At the highest redshift, the distributions are found to have standard deviations of \(\sigma = 0.32 \, [\log M_{\odot} h^{-1}]\) and \(\sigma = 0.21 \, [\log M_{\odot} h^{-1}]\) for the largest and smallest number density selections. In both cases, the distributions increase in width by a factor of \(\sim 3\) since \(z = 0\).

Furthermore, the selections made by a constant number density (the mass limits of which are indicated by the vertical dashed lines in the top panels of Figure \ref{f9}) show that beyond \(z > 1\), the majority of the progenitors are below this limit (i.e. the the peak of the actual progenitor stellar mass distribution is found at a lower stellar mass than the selection’s stellar mass limit). Therefore, within the Millennium Simulation at the very least, it can be said the progenitors of the most massive local galaxies are not only the most massive galaxies at higher redshifts - they span a wide range of masses at higher redshifts. For example, the most massive

\begin{figure}
\centering
\includegraphics[width=\textwidth]{example}
\caption{Recovery of individual progenitors, their average stellar mass, stellar mass density and average SFR (rows) for four constant number density selections (columns) covering the range \(-4.3 < \log n < -3.0\). The B06, G11 and G13 models are represented by solid green, dashed-dotted magenta and dotted blue lines respectively. (A colour version of this figure is available in the online journal.)}
\end{figure}
progenitors span more than an order of magnitude ($>1$ dex) in stellar mass for the largest initial number density selection (top right panel) and $\sim 0.5$ dex for the smallest number density at $z = 3$. As shown in this work however, this does not appear to significantly impact the ability of a number density selected sample to recover the average stellar mass, star-formation rate or stellar mass density.

5.2 How do mergers affect our selections?

The number of progenitors or descendant galaxies can change over time due to mergers between objects within the initial sample. Selecting galaxies at a constant number density ignores these changes, and potentially contributes to the over- or underestimation of ensemble properties.

To determine the extent of this, we calculate the number of mergers between the descendants of an initial selection at $z = 3$ in the B06 and G13 SAMs, as these are based on different dark matter merger trees from the MS. In B06, 1.3% and 15.5% of galaxies in our initial selection are lost due to mergers from $z = 3$ to $z = 0$ at the smallest and largest number density selections. In the G13 SAM however, these measurements are higher at 10.9% and 29.6%, respectively. For the most massive galaxies, this translates to approximately 3–5 mergers per massive galaxy (see Table 1 for mass limits) over the redshift range $0 < z < 3$. It must be noted that these numbers represent all (total) mergers, and are not major mergers as they may include some mergers with mass ratios greater than $1:4$, the most widely used definition. These measurements are slightly higher compared with pair fraction and morphological observations of major mergers in comparably massive galaxies (see, e.g., Bluck et al. 2004, Conselice et al. 2006, Man et al. 2012). At increasingly larger number densities, mergers within the descendant population may become increasingly important. As such, selection at a constant number density may not be applicable over such a redshift range. It may be appropriate to correct the number density between redshift bins to account for mergers that have occurred within the sample. However, the reduction of the number density in response to descendant galaxy mergers does not result in the desired effect. Qualitatively, reducing the number density at each redshift results in higher stellar mass limits. The average stellar mass of the observed sample would therefore increase. As the average stellar mass and stellar mass density are already overestimated, this discrepancy would only increase. On the other hand, the ‘un-merging’ of galaxies going backwards in time would increase the number density used to trace progenitor galaxies. This would lower the stellar mass limit used to select the samples and thus decrease the measured average stellar mass of the observed samples. As this quantity is also overestimated, this discrepancy would be reduced. However, as we have increased the number density contamination would also increase.

5.3 Can we infer velocity dispersion in a semi-analytical model?

We investigate inferred velocity dispersion in place of stellar mass as a ranking property due to evidence of a shallower and more stable evolution with redshift. It is prudent to ask whether this quantity can be accurately obtained from the SAMs used in this work. Use of Equation 5 implicitly assumes a spherically symmetric system and would correspond to a system with a Sérsic index of $n \approx 5.5$ (Cappellari et al. 2008). However, the factor in the denominator doesn’t account for multiple components (i.e. a bulge and disk) and is influenced not only by Sérsic index but also galaxy black hole mass. Thus the value calculated using this equation is a simplistic estimate at best and not strictly applicable to every type of galaxy. Furthermore, disk-dominated systems are not spherical and isotropic and thus this equation is not strictly applicable to these types of systems.

Using the bulge-to-total stellar mass ratio ($B/T$) as a proxy for disk and bulge dominated morphologies, we find that the most massive galaxies at $z = 0$ ($z = 3$) in the B06 SAM are typically bulge dominated with only 30% (40%) having $B/T < 0.5$. While this suggests Equation 5 is applicable at these redshifts, this SAM does not reproduce observations of larger disk-dominated fractions at high redshift (Bluck et al. 2014, Bruce et al. 2014). The G13 SAM reproduces observations more closely with 50% (95%) of systems having $B/T < 0.5$ at $z = 0$ ($z = 3$). Because of this, the velocity dispersions inferred within this SAM at the highest redshifts probed can be considered discrepant with observations only at the highest redshifts. While these caveats must be taken into consideration, the values of velocity dispersion inferred are physical and generally in agreement with observations of spheroidal/passive systems (Bernardi et al. 2010, Oser et al. 2012), with $1.8 < \log \sigma_\star [\text{km s}^{-1}] < 2.7$ depending on the SAM.

We must also consider whether the physical sizes of the simulated galaxy components can be used to infer the velocity dispersion. In G11 and G13, the resulting mass-size relations are shallower than the observations (see, e.g., Lani et al. 2013, van der Wel et al. 2014) with both masses and sizes larger (smaller) at low (high) redshift (Guo et al. 2011). For the purposes of this work however, we only care if the evolution in $M/R$ is represented correctly, and it is possible that, at least within G11 and G13, this may not be the case. We conclude that inferred velocity dispersion could provide a useful property with which to trace the evolving properties of the most massive galaxies. However, more detailed future simulations that accurately reproduce the evolution in both stellar mass and galaxy component sizes, or that report a value for velocity dispersion directly, are needed to confirm these findings.

5.4 Comparison with previous works

Our results are consistent with the work of Leja et al. (2013) who investigate cumulative number density selection of descendant galaxies over $0 < z < 3$ in the range $0.5 < n [10^{-4} \text{ Mpc}^{-3}] < 8.0$ using the G11 SAM. Uncorrected for mergers and growth scatter, they show that for the two smallest number densities, the median stellar mass evolution is overestimated by between $0.05 – 0.15$ dex ($12 – 41\%$) by $z = 0$. Using the mean stellar mass, we find an overestimate at $n = 5 \times 10^{-5} \text{Mpc}^{-3}$ of 40% in the G11 and G13 SAMs. Most recently, Henriques et al. (2014, H06) contrasted the mean stellar mass of progenitors derived from a constant number density selection with the values obtained from their SAM using a Planck (Ade et al. 2014) cosmology. They
Figure 9. Progenitor mass distributions in the Guo et al. (2013) SAM for \( z = 0 \) galaxy selections at constant number densities \( \log n = -4.3 \text{ Mpc}^{-3} \) (top left) and \( \log n = -3.0 \text{ Mpc}^{-3} \) (top right). These number densities correspond to galaxies with a \( z = 0 \) stellar mass of \( \log M^* > 11.21 [M_\odot h^{-1}] \) and \( \log M^* > 10.77 [M_\odot h^{-1}] \), respectively. Stellar mass distributions at \( z = 0, 1, 2, 3 \) are given in black, red, orange and green solid lines. Dashed vertical lines represent the stellar mass cuts inferred from a number density cut at each redshift. Residuals from Gaussian fits to these distributions are displayed in the middle panels. Parameters from Gaussian fits to the progenitor masses are shown in the bottom row, with the mean stellar mass and standard deviation at bottom left and bottom right respectively with the largest (blue squares) and smallest (red circles) number densities plotted as a function of redshift. (A colour version of this figure is available in the online journal.)

found that mean stellar mass evolution is overestimated by a factor of \( 3 - 5 \) for the progenitors of galaxies with a \( z = 0 \) stellar mass between \( 10.25 < \log M_* [M_\odot h^{-2}] < 11.25 \). This is a larger increase than we find for similar mass galaxies. We suggest this discrepancy is due to the ability of H06 to correctly reproduce the abundance of massive galaxies out to higher redshift. With less massive and passive galaxies, H06 must produce a larger evolution in stellar mass from high redshift to match the local stellar mass function.

6 SUMMARY

We have compared the use of two popular galaxy selection methods and contrasted the use of galaxy stellar mass and inferred velocity dispersion in semi-analytical models based on the output of the Millennium Simulation over the redshift range \( 0 < z < 3 \). We select galaxies above a constant limit of stellar mass in this redshift range, and conclude that:

- Descendants can be fully recovered over the entire redshift range regardless of mass limit choice. However, progenitors of \( z = 0 \) galaxies are lost from the selection with at least 80% below the stellar mass limit at \( z = 1 \).
- Average descendant ensemble stellar mass is increasingly underestimated with increasing redshift by an amount that varies between SAMs but is, on average, around 50% at \( z = 0 \). Similarly, the average stellar mass of progenitors is increasingly overestimated. At \( z = 0 \), average progenitor stellar mass is overestimated by a factor of \( \sim 5 \).

Additionally, selecting galaxies at a constant cumulative number density in stellar mass, we conclude that:

- Recovery of individual descendant galaxies falls exponentially with a time scale dependent on choice of number density. Just 30% of the most massive galaxies (selected at \( \log n = -4.3 \text{ Mpc}^{-3} \) at \( z = 3 \)) are at the same cumulative number density at the lowest redshift. For the largest number density selection, this increases to 60%. Recovery of progenitors is similar, but with 50 – 30% recovered at the highest redshifts, depending on choice of number density.
- The average stellar mass of descendants is overestimated by 15% (70%) at the highest (smallest number densities) by \( z = 0 \), increasing linearly from \( z = 3 \). Furthermore, independent of number density, progenitors’ average stellar mass is overestimated by \( \sim 50% \) at the highest redshift.

Finally, we have investigated inferred velocity dispersion as a property with which to trace galaxies over the same redshift range. We have found that a constant number density in velocity dispersion recovers average stellar mass,
stellar mass density and average SFR to within ±50% for both descendants and progenitors. Furthermore, selecting galaxies at a constant velocity dispersion limit recovers the aforementioned properties to within ±80% of the true values. However, these results are based on inferring velocity dispersion of galaxies which may not be strictly applicable to some SAMs and redshift ranges. Further study is required into velocity dispersion through dedicated simulations that predict these properties.

In conclusion, selecting galaxies at a constant cumulative number density is found to trace the true evolution of average stellar mass and the average SFR of the progenitors and descendants of galaxies in initial selections at \( z = 0 \) and \( z = 3 \). However, it does not trace the exact same galaxies but rather galaxies with very similar properties. Furthermore, it is found that selecting galaxies above a constant stellar mass with redshift returns the actual evolution within a larger factor of between two and thirty.

ACKNOWLEDGEMENTS

We wish to thank Kenneth Duncan and Frazer Pearce for useful discussions related to this work, and Joel Leja for his rapid response in request for information. We thank the anonymous referee for their constructive comments that improved the contents of this paper. We gratefully acknowledge support from the Science and Technology Facilities Council (STFC) and the Leverhulme Trust. The Millennium Simulation databases used in this paper and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory (GAVO).

REFERENCES

Ade P. A. R., et al., 2014, Astronomy & Astrophysics 571, A16
Behroozi P. S., Marchesini D., Wechsler R. H., Muzzin A., Papovich C., Stefanon M., 2013, The Astrophysical Journal 777, L10
Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, Monthly Notices of the Royal Astronomical Society 404, 2087
Bertone S., Conselice C. J., 2009, Monthly Notices of the Royal Astronomical Society 396, 2345
Bezanson R., van Dokkum P., Franx M., 2012, The Astrophysical Journal 760, 62
Bluck A. F. L., Conselice C. J., Bouwens R. J., Daddi E., Dickinson M., Papovich C., Yan H., 2009, Monthly Notices of the Royal Astronomical Society: Letters 394, L51
Bluck A. F. L., Mendel J. T., Ellison S. L., Moreno J., Simard L., Patton D. R., Starkenburg E., 2014, Monthly Notices of the Royal Astronomical Society 441, 599
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, Monthly Notices of the Royal Astronomical Society 370, 645
Bruce V. A., et al., 2014, Monthly Notices of the Royal Astronomical Society 444, 1001
Cappellari M., et al., 2006, Monthly Notices of the Royal Astronomical Society 366, 1126
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, Monthly Notices of the Royal Astronomical Society 319, 168
Collaboration P., 2015, eprint arXiv:1502.01589
Conseilce C. J., Bershady M. A., Dickinson M., Papovich C., 2003, The Astronomical Journal 126, 42
Conseilce C. J., Yang C., Bluck A. F. L., 2009, Monthly Notices of the Royal Astronomical Society 394, 1956
Conseilce C. J., Mortlock A., Bluck A. F. L., Grutzauch R., Duncan K., 2013, Monthly Notices of the Royal Astronomical Society 430, 1051
Conseilce C. J., Bluck A. F. L., Mortlock A., Palamara D., Benson A. J., 2014, Monthly Notices of the Royal Astronomical Society 444, 1125
Croton D. J., et al., 2006, Monthly Notices of the Royal Astronomical Society 365, 11
De Lucia G., Blaizot J., 2006, Monthly Notices of the Royal Astronomical Society 375, 2
Fumagalli M., et al., 2012, The Astrophysical Journal 757, L22
Furlong M., et al., 2014, p. 19
Genel S., et al., 2014, eprint arXiv:1405.3749, 445, 175
Graham A. W., Driver S. P., 2005, Publications of the Astronomical Society of Australia 22, 118
Guo Q., et al., 2011, Monthly Notices of the Royal Astronomical Society 413, 101
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, Monthly Notices of the Royal Astronomical Society 428, 1351
Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, Monthly Notices of the Royal Astronomical Society 367, 1039
Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2014, eprint arXiv:1410.0365
Javier Cenarro A., Trujillo I., 2009, The Astrophysical Journal 696, L43
Komatsu E., et al., 2011, The Astrophysical Journal Supplement Series 192, 18
Lani C., et al., 2013, Monthly Notices of the Royal Astronomical Society 435, 207
Leja J., van Dokkum P., Franx M., 2013, The Astrophysical Journal 766, 33
Lemson G., Consortium t. V., 2006, eprint arXiv:astro-ph/0608019
Lin Y.-T., Brodwin M., Gonzalez A. H., Bode P., Eisen...
APPENDIX A: FITTING PARAMETERS

Here we present tabulated fitting parameters for descendant and progenitor galaxy populations, as described in §3.3. These are obtained by taking the mean of individual SAM data points and taking the variance between SAMs as the $1\sigma$ of a Gaussian distribution on each data point. Fitting is then processed on randomly sampled values $10^4$ times to calculate the most likely fitting parameters and their associated uncertainties.
Table A1. Fitting parameters, described in [8] for descendants and progenitors selected at a constant cumulative number density across the redshift range $0 < z < 3$.

| Metric | Descendants | Progenitors |
|--------|-------------|-------------|
|        | $a$ | $b$ | $c$ | $a$ | $b$ | $c$ |
| $f_{\text{rec}}$ | $0.000 \pm 0.000$ | $0.480 \pm 0.030$ | $0.175 \pm 0.018$ | $0.000 \pm 0.000$ | $1.212 \pm 0.055$ | $-0.260 \pm 0.027$ |
| $f_{\text{contam}}$ | $0.511 \pm 0.019$ | $-0.010 \pm 0.001$ | $1.000 \pm 0.000$ | $-0.708 \pm 0.070$ | $-0.235 \pm 0.056$ | $0.563 \pm 0.064$ |
| $\kappa_{m^*}$ | $0.191 \pm 0.066$ | $-0.044 \pm 0.019$ | - | $-0.131 \pm 0.030$ | $0.125 \pm 0.017$ | - |
| $\kappa_{p^*}$ | $0.592 \pm 0.050$ | $-0.156 \pm 0.015$ | - | $-0.132 \pm 0.031$ | $0.126 \pm 0.018$ | - |

$n = 1 \times 10^{-3}$ [Mpc$^{-3}$ $h^3$]

| Metric | Descendants | Progenitors |
|--------|-------------|-------------|
|        | $a$ | $b$ | $c$ | $a$ | $b$ | $c$ |
| $f_{\text{rec}}$ | $0.000 \pm 0.000$ | $0.372 \pm 0.044$ | $0.240 \pm 0.032$ | $0.000 \pm 0.000$ | $1.234 \pm 0.057$ | $-0.302 \pm 0.029$ |
| $f_{\text{contam}}$ | $0.568 \pm 0.032$ | $-0.011 \pm 0.001$ | $1.000 \pm 0.000$ | $-0.670 \pm 0.035$ | $-0.242 \pm 0.017$ | $0.580 \pm 0.009$ |
| $\kappa_{m^*}$ | $0.294 \pm 0.070$ | $-0.071 \pm 0.020$ | - | $-0.134 \pm 0.058$ | $0.134 \pm 0.036$ | - |
| $\kappa_{p^*}$ | $0.596 \pm 0.061$ | $-0.156 \pm 0.018$ | - | $-0.134 \pm 0.058$ | $0.134 \pm 0.036$ | - |

$n = 1 \times 10^{-4}$ [Mpc$^{-3}$ $h^3$]

| Metric | Descendants | Progenitors |
|--------|-------------|-------------|
|        | $a$ | $b$ | $c$ | $a$ | $b$ | $c$ |
| $f_{\text{rec}}$ | $0.000 \pm 0.000$ | $0.180 \pm 0.047$ | $0.430 \pm 0.071$ | $0.000 \pm 0.000$ | $1.516 \pm 0.155$ | $-0.519 \pm 0.079$ |
| $f_{\text{contam}}$ | $0.725 \pm 0.050$ | $-0.014 \pm 0.001$ | $1.000 \pm 0.000$ | $-0.623 \pm 0.036$ | $-0.216 \pm 0.019$ | $0.582 \pm 0.005$ |
| $\kappa_{m^*}$ | $0.691 \pm 0.099$ | $-0.151 \pm 0.029$ | - | $-0.107 \pm 0.072$ | $0.144 \pm 0.045$ | - |
| $\kappa_{p^*}$ | $0.742 \pm 0.062$ | $-0.194 \pm 0.018$ | - | $-0.107 \pm 0.073$ | $0.145 \pm 0.046$ | - |

$n = 1 \times 10^{-5}$ [Mpc$^{-3}$ $h^3$]

| Metric | Descendants | Progenitors |
|--------|-------------|-------------|
|        | $a$ | $b$ | $c$ | $a$ | $b$ | $c$ |
| $f_{\text{rec}}$ | $0.000 \pm 0.000$ | $0.128 \pm 0.044$ | $0.522 \pm 0.096$ | $0.000 \pm 0.000$ | $1.723 \pm 0.321$ | $-0.626 \pm 0.144$ |
| $f_{\text{contam}}$ | $0.787 \pm 0.054$ | $-0.014 \pm 0.001$ | $1.000 \pm 0.000$ | $-0.621 \pm 0.058$ | $-0.204 \pm 0.026$ | $0.582 \pm 0.009$ |
| $\kappa_{m^*}$ | $0.747 \pm 0.136$ | $-0.198 \pm 0.041$ | - | $-0.083 \pm 0.062$ | $0.139 \pm 0.036$ | - |
| $\kappa_{p^*}$ | $0.868 \pm 0.093$ | $-0.228 \pm 0.028$ | - | $-0.085 \pm 0.064$ | $0.141 \pm 0.037$ | - |

$a$ Descendants use Equation [8] and progenitors use Equation [10] for fitting.