Globular cluster numbers in dark matter haloes in a dual formation scenario: an empirical model within EMERGE

Lucas M. Valenzuela,1★ Benjamin P. Moster,1,2 Rhea-Silvia Remus,1 Joseph A. O’Leary,1 and Andreas Burkert1,3,4
1Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstraße 1, D-81679 München, Germany
2Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching, Germany
3Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße 1, D-85748 Garching, Germany
4Excellence Cluster ORIGINS, Boltzmannstraße 2, D-85748 Garching, Germany

Accepted XXX. Received YYY; in original form ZZZ.

ABSTRACT

We present an empirical model for the number of globular clusters (GCs) in galaxies based on recent data showing a tight relationship between dark matter halo virial masses and GC numbers. While a simple base model forming GCs in low-mass haloes reproduces this relation, we show that a second formation pathway for GCs is needed to account for observed younger GC populations. We confirm previous works that reported the observed linear correlation as being a consequence of hierarchical merging and its insensitivity to the exact GC formation processes at higher virial masses, even for a dual formation scenario. We find that the scatter of the linear relation is strongly correlated with the relative amount of smooth accretion: the more dark matter is smoothly accreted, the fewer GCs a halo has compared to other haloes of the same mass. This scatter is smaller than that introduced by halo mass measurements, indicating that the number of GCs in a galaxy is a good tracer for its dark matter mass. Smooth accretion is also the reason for a lower average dark matter mass per GC in low-mass haloes. Finally, we successfully reproduce the observed general trend of GCs being old and the tendency of more massive haloes hosting older GC systems. Including the second GC formation mechanism through gas-rich mergers leads to a more realistic variety of GC age distributions and also introduces an age inversion in the halo virial mass range \( \log M_{\text{vir}} / M_\odot = 11–13 \).

Key words: galaxies: star clusters – galaxies: evolution – galaxies: statistics – galaxies: haloes – dark matter

1 INTRODUCTION

Globular clusters (GCs) are among the oldest objects observed in our Universe and can be found in all galaxies but the smallest dwarfs. It has generally been found that the more massive a system is, the larger the GC population in it is (e.g. Blakeslee et al. 1997). In fact, as reported by Harris et al. (2017b), there exists a tight correlation between the total dark matter halo mass of a galaxy and its number of GCs. For the most massive galaxies, GC systems of more than 30,000 GCs have been reported (Harris et al. 2016). On the other hand, our own galaxy, the Milky Way (MW), has a rather poor system of GCs with only about 150 of them (Harris 1996). Recently, Forbes et al. (2018) followed this trend between dark matter halo mass and GC system richness down to dwarf galaxies with total masses around \( 10^9 M_\odot \), including systems in which only one GC was found as expected from this relation. Even for ultra-diffuse galaxies (UDGs) in the COMA galaxy cluster, GC systems of similar richness have been reported (van Dokkum et al. 2017; Amorisco et al. 2018; Forbes et al. 2020), albeit with a tendency for the UDG GC systems to be richer than those of normal dwarf galaxies of similar stellar masses (van Dokkum et al. 2017; Forbes et al. 2020). Interestingly, while estimates of the dark matter masses in UDGs are extremely difficult, the results from dynamical modelling point towards UDGs having a similar halo mass to GC richness as normal galaxies of the same dark matter mass (van Dokkum et al. 2016, 2017). However, there is still ongoing debate about this topic, as recently Saifollahi et al. (2021) concluded that the number of GCs in some UDGs might actually be strongly overestimated.

While GCs have been studied in detail, especially in our own MW, their origin and formation is still a matter of debate. Measurements of the ages of GCs are extremely difficult, in particular for GCs outside the MW, but they all indicate GCs generally being old (e.g. Salaris et al. 1997; Cohen et al. 1998; Brodie et al. 2005; Strader et al. 2005; Brodie & Strader 2006; Sharina et al. 2006; ...
Chies-Santos et al. 2012; VandenBerg et al. 2013; Leaman et al. 2013), at 10 Gyr and older. However, a few galaxies with younger populations of GCs are also reported: for example, our neighbouring galaxy, the small Magellanic Cloud, seems to have a rather young GC system with the oldest GC being only 8 Gyr and the youngest less than 5 Gyr old (Parisi et al. 2014; see also Sharina et al. 2006 for young GCs in other local group dwarf galaxies). Especially young GCs have also been reported for NGC 1316 (Goudfrooij et al. 2001; Sesto et al. 2018). More recently, Usher et al. (2019) measured the age distributions for three galaxies outside the local group in comparison with the MW, and found a tendency for more massive galaxies to host older GC systems, while less massive galaxies feature populations of younger globular clusters (see also Beasley et al. 2008 for the GC system of NGC 5128). Similarly, Chies-Santos et al. (2011) found that, on average, S0 galaxies have younger GC systems than elliptical galaxies.

These age differences are speculated to be related to the observation that GCs usually come in two different categories of colours and metallicities: metal poor, blue GCs, and metal-rich, red GCs (e.g. Usher et al. 2012). Most massive galaxies have GCs of both kinds, with the metal-rich red GCs also occurring in the inner parts of galaxies and tracing the inner light distribution (e.g. Peng et al. 2004; Schuberth et al. 2010; Pota et al. 2013), while the metal-poor blue GCs are usually distributed at larger radii and can have very different density slopes compared to the observed stellar light (e.g. Forbes et al. 1997; Schuberth et al. 2010). This gave rise to the idea that blue GCs are usually formed in small, metal-poor galaxies at very early times and successively assembled onto the larger galaxies (e.g. Côté et al. 1998; Schuberth et al. 2010; Chies-Santos et al. 2011; Forbes & Remus 2018), while red, metal-rich GCs are formed later in larger galaxies when the gas has already been enriched with metals, for example during a starburst of gas-rich major mergers (e.g. Goudfrooij et al. 2001; Schuberth et al. 2010; Harris et al. 2017a). This is also supported by several observations that found that the red GCs are younger while the blue GCs are older (Peng et al. 2004; Beasley et al. 2008), even though other studies see no such strong age differences between the blue and red GC systems (e.g. Strader et al. 2005).

While this idea of a dual formation scenario for GCs in galaxies has been tested extensively in observations and is still a matter of debate, GCs have proven to be rather difficult to capture from the modelling and simulation sides: simulations of GC systems in dark matter haloes, implemented within semi-analytic model assuming that GCs are formed either in proto-structures and thus find two populations of metal-rich and metal-poor GCs with ages similar to the observations. Choksi et al. (2018) assumed that GCs are usually formed in massive merger events and thus used an empirical model to populate their galaxies (based on the Millennium simulation runs) with GCs in such events. They found a remarkable agreement between their results and observations in terms of ages, metallicities, and stellar masses of the GC systems with the host galaxies’ halo masses. Kruisjes (2015) modelled the GC distribution in galaxies assuming that GCs are formed at redshifts $z > 2$ inside high-pressure disks and redistribute later on inside the galaxies through hierarchical merging. The idea of hierarchical merging to establish the present-day observed relation between the numbers of GCs and the total mass of a galaxy has also been studied by Boylan-Kolchin (2017) in a phenomenological model, finding that the observed relations can be well reproduced by forming GCs by $z = 6$ in dark matter haloes with masses above $1.07 \times 10^{9} M_{\odot}$. In a subsequent study, El-Badry et al. (2019) showed that this GC-to-halo mass relation is a natural consequence of the central limit theorem in the case of hierarchical merging, making the richness of a galaxy’s GC system a good tracer for its total halo mass. Both Boylan-Kolchin (2017) and El-Badry et al. (2019) make use of halo merger trees generated by the extended Press & Schechter (1974) model with the algorithm by Parkinson et al. (2008), which reproduces merger trees from N-body simulations well (Jiang & van den Bosch 2014). The connection to the central limit theorem was further supported by a simple hierarchical merging model by Burkert & Forbes (2020), who showed that with an average seeding halo virial mass per GC of $M_{\text{seed}} = 5 \times 10^{8} M_{\odot}$ and hierarchical merging, the observed linear relation between GC richness and the dark matter halo virial mass of a galaxy can be reproduced.

In this work we present an empirical model for the number of GCs in dark matter haloes, implemented within \textsc{emerge} and building on the investigations and models by Boylan-Kolchin (2017), Choksi et al. (2018), and Burkert & Forbes (2020), to further analyse the hierarchical nature of GC system assembly. In Section 2, we introduce the simulations, the components of the base and extended model, and the observational data used in our analysis. Section 3 describes how the model was fit to the observations, matching the general relation between halo virial masses and globular cluster numbers and its scatter. In Section 4, we present a variety of features that can be explored in the fitted model and their dependence on redshift. Addressing the drawbacks found in the first sections for the base model, in Section 5 we present how an extended model with a second formation pathway for GCs through gas-rich merger events changes and improves on the previously shown results. Finally, we discuss our findings in Section 6 and summarise the key results in Section 7.

2 METHOD

2.1 \textsc{emerge} & Backbone Simulations

For the globular cluster formation model, we extracted halo merger trees from a cosmological dark matter-only N-body simulation with a side length of 30 Mpc. This simulation adopted cosmological parameters consistent with Planck Collaboration XIII (2016): $\Omega_m = 0.308$, $\Omega_{\Lambda} = 0.692$, $\Omega_b = 0.0484$, $H_0 = 67.81 \text{km s}^{-1} \text{Mpc}^{-1}$, $n_s = 0.9677$, and $\sigma_8 = 0.8149$. To generate the initial conditions, MUSIC (Hahn & Abel 2011) was used with a power spectrum computed by \textsc{cmak} (Lewis et al. 2000). The simulation contains $512^3$ dark matter particles with particle mass $m_{\text{DM}} = 7.90\times 10^{4} M_{\odot}$. Using the TreePM code \textsc{gadget3} (Springel 2005) with periodic boundary conditions, the simulation was run from $z = 63$ to 0, creating 132 snapshots ranging from $a = 0.0457$ to 1.0000, equally spaced in scale factor. The halo finder \textsc{rockstar} (Behroozi et al. 2013a) was used to identify the dark matter haloes in each snapshot.
which were linked to construct halo merger trees with consistent trees (Behroozi et al. 2013b). In this work, the term ‘main halo’ is used for haloes that are not contained within any larger halo, whereas ‘subhalo’ is used for all other haloes.

We implemented the GC formation model within the code of the empirical model Emerge (Moster et al. 2018), which populates the simulated haloes with galaxies to reproduce observed statistical galaxy data. For more details on the merger histories of the galaxies found in Emerge, see O’Leary et al. (2021). In all the GC model runs, we used the Emerge model parameters found in their table 1.

The halo mass function at $z = 0$ (Fig. 1) that resulted from the simulation shows that the chosen resolution allows us to make reasonable statistics for halo masses above $\sim 10^{9.5} M_{\odot}$ (see also Knebe et al. 2011 for a comparison and discussion of halo finders, including their resolution limits). We can also see that the sample includes halo masses of up to $10^{13.5} M_{\odot}$. This upper limit will suffice for comparison with the samples used by Burkert & Forbes (2020) and Harris et al. (2017b), which contain masses up to $10^{14}$ and $10^{14.5} M_{\odot}$, respectively (excluding the galaxy clusters).

2.2 Globular Cluster Formation Model

Motivated by the fact that GCs come in two different flavours, i.e. red and blue GCs, we considered two GC formation models: one that we refer to as the base model, which forms GCs in small haloes, and a second one that forms GCs in situ during gas-rich mergers.

2.2.1 Base Model

For the base model, we used a simple phenomenological model for creating GCs in low-mass haloes, for example through gravitational collapse of a single gas cloud (Peebles & Dicke 1968). Assuming that the observed correlation between the number of GCs in a halo, $N_{\text{GC}}$, and the virial mass of a halo, $M_{\text{vir}}$, follows from hierarchical merging, the goal is to find the halo virial mass at which GCs have to be seeded, $M_{\text{seed,GC}}$, to reproduce observations at $z = 0$. We made some simplifying assumptions for the seeding process to keep the number of free parameters at a minimum: (1) Each halo has a one-time opportunity to form GCs as soon as it reaches $M_{\text{seed,GC}}$ for the first time. (2) GCs are seeded according to a predetermined distribution $p(x)$ with expectation value 1. Additionally, two further assumptions were made with respect to already existing GCs: (3) A halo that becomes a subhalo immediately turns over all its GCs to the respective main halo. (4) A GC cannot disappear. The latter assumptions are introduced to avoid additional model parameters, but these aspects provide good starting points for future refinements and expansions of the model.

From assumption 1 it follows that merging events do not have a direct influence on the formation of GCs in this model and that a distribution of different GC ages is obtained. For assumption 2, we tested three different seeding functions: the first seeding distribution implemented, $p_1(x)$, simply seeds exactly one GC in every seeding process. The second distribution, $p_2(x)$, seeds zero, one, or two GCs with equal probability. Finally, the third distribution we tested is a so-called ‘geometric distribution’, $p_3(x) = 1/2^x+1$, which seeds GCs according to an exponential law. Assumption 3 leads to the possibility of backslash haloes having very few or no GCs at all. This does not affect our conclusions because all the fits were performed on the simulated main haloes with at least one GC. Even when considering all main haloes, the difference at high masses is negligible (e.g. Fig. 8). Assumption 4 means that GC disruption is neglected in our model. Note that for example for the relation between GC mass and halo virial mass, Bastian et al. (2020) come to the conclusion that GC disruption plays a dominant role, though for example Burkert & Forbes (2020) point out that GC disruption does not affect the linear correlation between $M_{\text{vir}}$ and $N_{\text{GC}}$.

2.2.2 Model Extension

With the aim to address issues found in the results of the base model and motivated by the suggestion that the observed blue and red GC populations may originate from different formation scenarios (e.g. Côté et al. 1998; Schuberth et al. 2010), we added an extension to the base model (discussed in Section 5): a second formation mechanism triggered by gas-rich merger events (e.g. Ashman & Zepf 1992). To keep the number of additional parameters to a minimum and taking into account the framework that Emerge provides, we decided to follow the approach used by Choksi et al. (2018) for the GC formation model. GC formation is triggered on the condition that the logarithmic halo mass accretion rate, $A_m$, surpasses a certain threshold, $A_m^\text{min}$, which is generally the case for major mergers:

$$A_m = \frac{M_{\text{vir}} - M_{\text{vir,prog}}}{M_{\text{vir}}\Delta t} > A_m^\text{min}. \quad (1)$$

where $M_{\text{vir}}$ and $M_{\text{vir,prog}}$ are the virial masses of the halo considered for GC formation and its progenitor, respectively, $\Delta t$ is the timestep between their two snapshots, and the threshold $A_m^\text{min}$ is one of the two free parameters of the model. Note that this corresponds to the parameter $p_3$ of Choksi et al. (2018).

When the formation criterion is met, the mass of the present cold gas is computed according to $M_{\text{gas}} = \text{SFR} \times t_{\text{dep}}$, SFR is the star formation rate computed by Emerge (according to the best fit to observations, Moster et al. 2018; O’Leary et al. 2021) and $t_{\text{dep}}$ is the gas depletion time, which we calculate from the redshift according to Tacconi et al. (2020):

$$t_{\text{dep}} = \varepsilon_H^{-1} \times t_H \times (0.7 + 0.3 \times (1 + z)^3)^{-1/2}, \quad (2)$$

with $\varepsilon_H = 0.1$ and $t_H = 13.98$ Gyr. The total mass of GCs formed in the formation event, $M_{\text{GC}}$, is then determined from the cold gas mass, $M_{\text{gas}}$:

$$M_{\text{GC}} = 1.8 \times 10^{-4} \eta_{\text{GC}} M_{\text{gas}}. \quad (3)$$

where $\eta_{\text{GC}}$ is the second and last free parameter of the model, corresponding to $p_2$ of Choksi et al. (2018). The same conversion was
previously also applied by Li & Gnedin (2014), based on predictions by Kravtsov & Gnedin (2005). We adapted the same cluster initial mass function and normalisation as Choksi et al. (2018),
\[
\frac{dN}{dM} = M_{max}^{-2},
\]
with \(M_{max}\) being the most massive GC formed. \(M_{max}\) depends on \(M_{GC}\) according to
\[
M_{GC} = \int_{M_{min}}^{M_{max}} dM M \frac{dN}{dM} = M_{max} \ln \frac{M_{max}}{M_{min}},
\]
where \(M_{min} = 10^5 M_\odot\) is the minimum GC mass that is expected to survive for a few Gyr (Li & Gnedin 2014; Choksi et al. 2018). This can be solved for \(M_{max}/M_{min}\) with the Lambert W function:
\[
\frac{M_{max}}{M_{min}} = \exp \left( W \left( \frac{M_{GC}}{M_{min}} \right) \right),
\]
from which the average number of formed GCs can finally be obtained:
\[
\langle N_{GC} \rangle = \int_{M_{min}}^{M_{max}} dM M \frac{dN}{dM} = \frac{M_{max}}{M_{min}} - 1.
\]
Like for the approach of El-Badry et al. (2019), if an average of for example 5.2 GCs would be formed, the halo forms 5 GCs with a probability of 80 per cent and 6 GCs otherwise. No GCs are formed if \(M_{GC} < M_{min}\).

### 2.3 GC Studies & Obtaining Mock Observations
Having specified our model, the free parameter \(M_{seedGC}\), and the available seeding distributions, we populated the haloes from the merger trees with GCs using the base model. To obtain the best value for \(M_{seedGC}\) and its uncertainty, the computed numbers of GCs per halo were compared to the observed samples used by Burkert & Forbes (2020) and Harris et al. (2017b) (hereafter B&F20 and H+17). Both are compilations of literature work, where the former sample focuses on high-quality studies, i.e., the corrections used to determine \(N_{GC}\) are required to be small (see Harris et al. 2015 for a brief overview of the typical approach to getting \(N_{GC}\)) and includes a variety of low-mass galaxies. The latter study includes a larger sample size, of which the large majority have masses above \(10^{11} M_\odot\). Because there are overlaps between the two studies and to prevent systematic differences in their approaches of obtaining \(N_{GC}\) and \(M_{vir}\), from distorting our results, we decided to fit our model to the two observational studies separately. For the comparison with the former sample, we used the mean observed values of log \(N_{GC}\) from their equation 1,
\[
\langle \log N_{GC} \rangle = (-9.58 \pm 1.58) + (0.99 \pm 0.13) \times \log M_{vir}/M_\odot,
\]
and the scatters, \(\sigma_{obs}\), in the mass bins used in fig. 2 of B&F20. For the sample from H+17, we determined the mean linear correlation,
\[
\langle \log N_{GC} \rangle = (-8.86 \pm 0.89) + (0.91 \pm 0.07) \times \log M_{vir}/M_\odot.
\]
and the scatters from it in the same bins.

To get mock observations of the simulated virial masses, we added normally distributed error values to their logarithms (see Section 3.1 about how the variance of the error distribution was determined). We then binned the haloes in mock virial mass and calculated the mean of log \(N_{GC}\) and its scatter, \(\sigma_{sim}\), in each bin. The scatter was calculated as the 1σ-deviation of log \(N_{GC}\) – \(\log N_{GC}\) within a bin. Using a spline interpolation, we could then compute the values for each observed data point. To quantify the likelihood of the chosen parameters, we used the \(\chi^2\) goodness-of-fit test.

### 2.4 Fitting Model Parameters
To determine the parameter \(M_{seedGC}\), we required the model to reproduce the observed relation between \(M_{vir}\) and \(N_{GC}\). To this end, we used a Markov chain Monte Carlo (MCMC) method, taking into account the best mock observation error distribution (see Sections 3.1 and 3.2). The sampler is implemented according to one of the methods described by Goodman & Weare (2010), an affine invariant ensemble sampler for MCMC (see Moster et al. 2018 for more details on the implementation). \(M_{seedGC}\) was the only free parameter, with all other parameters held fixed.

### 3 GLOBULAR CLUSTER SEEDING MASS

#### 3.1 Direct Simulation Results
As a first step, we manually fitted the obtained GC numbers with the seeding distribution \(p_1\) to the observed ones by adjusting \(M_{seedGC}\) without using the MCMC method. This way, we were able to obtain linear correlations that are very similar to the observed ones (Fig. 2). Note that the virial masses here are not the masses that result from mock observations, but the original ones from the merger trees. The best results were obtained with \(M_{seedGC} = 4.7 \times 10^5 M_\odot\) and \(M_{seedGC} = 7.2 \times 10^6 M_\odot\) for the samples of B&F20 and H+17, respectively. After putting the haloes with at least one GC at \(z = 0\) into 20 mass bins, we calculated the mean of log \(N_{GC}\) for each of them. The mean value becomes near linear when plotted against log \(M_{vir}\) around \(10^{10.5} M_\odot\), so we computed the least-square fitting line for the haloes above that mass, which resulted in the following respective relations:
\[
\langle \log N_{GC} \rangle = -9.27 + 0.96 \times \log M_{vir}/M_\odot,
\]
\[
\langle \log N_{GC} \rangle = -9.36 + 0.96 \times \log M_{vir}/M_\odot.
\]
Both y-intercepts and slopes are well within the error ranges of the fits to the observed samples.

However, one can clearly see that the scatter in log \(N_{GC}\) for a given mass is much larger for the observed values than for the simulated ones, in particular at high virial masses. While it decreases strongly with increasing mass for the simulated haloes, it remains roughly constant for the observed values (Fig. 3). This was first noticed by B&F20 when they compared their observational sample to the results of their simple Monte Carlo halo merging simulation.

There are two straightforward ways of changing the scatter in our model: manipulating the scatter of \(N_{GC}\) and manipulating the scatter of \(M_{vir}\). The former can be addressed by using one of the other seeding distributions, \(p_2\) or \(p_3\), since they have a larger variance than \(p_1\). However, when comparing the scatter of the three resulting distributions of GCs in the haloes, we found only a slight increase of the scatter with \(p_2\) and \(p_3\), but no significant flattening (Fig. 3). B&F20 argued that through hierarchical merging the central limit theorem would predict the scatter of log \(N_{GC}\) to follow the relation \(\sigma \propto 1/\sqrt{M_{vir}}\) (dashed line in Fig. 3), which their simulation closely resembles. Interestingly, in our simulated cases the scatter decreases at a smaller rate than the prediction of the central limit theorem. We suggest that this may be due to the variance in smoothly accreted dark matter for haloes (discussed in Section 4.3), which could also be the reason for the flattening of the scatter curve at large virial masses by creating a lower limit for the inherent scatter of \(N_{GC}\). In Fig. 4 one can see that the seeding distribution only visibly affects the scatter of \(N_{GC}\) at virial masses below \(10^{11} M_\odot\). The distribution above this mass remains the same, regardless of how exactly the GCs were seeded. This agrees very well with the
results from El-Badry et al. (2019), who stated that for \( M_{\text{vir}}(z = 0) \gtrsim 10^{11.5} M_\odot \), the ratio between the halo virial mass and the total GC mass in the respective halo is constant, independent of the GC-to-halo mass relation at the time of GC formation.

This clearly does not resolve the discrepancy between the simulated and observed scatter at high virial masses. We therefore took a look at the second possibility, following B&F20: manipulating the scatter of the virial masses. We addressed this by introducing an observation error to the virial mass to obtain mock observations, as described in Section 2.3. The random error is drawn from a normal distribution and added to \( M_{\text{vir}} \). We tested a variety of variances for the random errors, which gave us significantly different results for the scatter (Fig. 5). We clearly get the best match for an observation error of \( 0.5 \log M_{\text{vir}}/M_\odot = 0.3 \) dex, in particular for the scatter above a mass of \( 10^{11} M_\odot \). This is also the range we are interested in since the linear relationship in Fig. 2 was fitted to the haloes with virial masses above \( 10^{10.5} M_\odot \). We also found the same result when comparing the model to the observational sample of H+17. The value is slightly higher than the findings of B&F20 who inferred an error of 0.25 dex. The logarithmic error of 0.3 dex corresponds to an error of a factor 2.0 for determining virial masses, which is a realistic value for true observations. This value is used to obtain mock observations for fitting \( M_{\text{seedGC}} \) with the MCMC method.

3.2 Mock Observations of the Simulation Results

Having determined the mock observation error that allows us to reproduce the observed scatter of \( N_{\text{GC}} \), we ran the MCMC model (Section 2.4) with the seeding distribution \( p_1 \) to get the most likely value for the parameter \( M_{\text{seedGC}} \) and its error for both observational samples. We set up 100 walkers around the expected best fit and evolved each of them for 130 steps. The error of the walker distribution had converged by that time and could be extracted from the final walker samples.

For B&F20, we obtained a best fit for \( \log M_{\text{seedGC}}/M_\odot = 8.50_{-0.39}^{+0.19} \), with the uncertainties being the upper and lower 1σ parameter errors. For H+17, the best fit was found for \( \log M_{\text{seedGC}}/M_\odot = 8.67_{-0.21}^{+0.16} \). We then ran the simulation with \( M_{\text{seedGC}} = 10^{8.50} M_\odot \), again adding mock observation errors to the halo virial masses, and placed the haloes with at least one GC at \( z = 0 \) in the mass bins. As in Section 3.1, we calculated the mean of \( \log N_{\text{GC}} \) in each of them and computed the least-square fitting line for the haloes with \( M_{\text{vir}} > 10^{10.5} M_\odot \), where the mean value becomes near linear (Fig. 6). In comparison to Figs. 2 and 4, the scatter of the simulated number of GCs matches the scatter of the observed galaxies much better when using mock observations, as expected.

Figure 2. Correlation between dark matter halo virial masses, \( M_{\text{vir}} \), and numbers of globular clusters, \( N_{\text{GC}} \), of the simulated haloes compared to the observed galaxies from B&F20 (top) and H+17 (bottom). GCs were seeded at a virial mass of \( 10^{8.67} \) and \( 10^{8.86} M_\odot \), respectively. The distribution used for the number of GCs to seed was \( p_1 \), where exactly one GC is seeded in a seeding event. The fits to the simulations were computed by the least-square method for the bins masses above \( 10^{10.5} M_\odot \). The fits to the observations correspond to Equations 8 and 9, respectively.

Figure 3. Scatter of \( N_{\text{GC}} \) around the mean correlation compared between the simulated haloes and observations from B&F20 and H+17. GCs were seeded at a virial mass of \( 10^{8.67} \) and \( 10^{8.86} M_\odot \) (top and bottom, respectively). The scatter of the simulated values decreases at a smaller rate than \( \sigma = 1/\sqrt{M_{\text{vir}}} \) (dashed line), which corresponds to the prediction of the central limit theorem (B&F20). Neither the different GC seeding distributions nor the two values of \( M_{\text{seedGC}} \) have a significant effect on the slope. In contrast, the observed values from the sample of B&F20 have a near constant scatter around the mean correlation. The sample of H+17 follows a very similar correlation, although the scatter varies more around it.
from Fig. 5. Using one of the other two seeding distributions, $p_2$ or $p_3$, we can see the same effect as in Fig. 4: The exact seeding process is irrelevant for the distribution of the plotted simulated haloes with virial masses above $10^{11} \, M_\odot$. The same conclusions are drawn for the observational sample from H+17.

Even having the mock observation errors, it is not yet possible to tell which of the seeding distributions is best for describing the observations at lower masses. We decided to use $p_2$ for the rest of this work, with the reason that it is not realistic that exactly one GC will form in every single dark matter halo with the virial mass $M_{\text{seedGC}}$. At the same time, it is also not realistic that a theoretically unlimited amount of GCs can form in a halo of that mass, which is why we also decided against $p_3$. Since the results obtained for the fits to both observational samples are qualitatively the same, we only show the plots featuring the simulated globular cluster numbers fitted to the sample used by B&F20 in the following.

4 EVOLUTION WITH REDSHIFT

4.1 Correlation at Higher Redshifts

Having established a value for $M_{\text{seedGC}}$, a mock observation error, and a seeding distribution, we determined our model’s prediction of the distribution of GCs in galaxies at higher redshifts. For this, we picked three redshift values, $z = 0, 3$, and $6$ (Fig. 7). The shape at low virial masses remains the same at any of the selected redshifts. Only the upper limit of $N_{\text{GC}}$ and $M_{\text{vir}}$ is lower for higher redshifts, which is to be expected since high virial masses require many merging events, so these did not yet exist at high $z$ in the simulated box. Even at the higher redshifts the scatter of $N_{\text{GC}}$ around the observed linear relation matches that of the observations very well, just as for $z = 0$. Our model clearly predicts that on average, the number of GCs and its scatter will be the same in haloes of equal virial mass, independent of redshift.

4.2 Dark Matter Halo Mass per GC

B&F20 introduced the characteristic dark matter mass in virialised haloes per GC at $z = 0$ as the arithmetic mean, $M_{\text{DM,GC}} = \langle M_{\text{vir}}/N_{\text{GC}} \rangle$, for haloes with at least one GC. They also remark that the linear correlation breaks down for haloes with virial masses below $M_{\text{DM,GC}}$ and predict that it extends statistically down into the dwarf galaxy regime. Since our model forms GCs well below...
this value, at $10^{8.50} \, M_\odot$, we needed to find a quantification that addresses this statistical extension. For that, we redefined $M_{DM,GC}$ as the harmonic mean, such that haloes with no GCs in them can be included:

$$M_{DM,GC} = \left( \frac{N_{GC}}{M_{vir}} \right)^{-1} \cdot (12)$$

Using this definition, we binned the haloes in virial mass and computed the value for $M_{DM,GC}$ in every bin at three different redshifts, $z = 0, 3,$ and $6$ (Fig. 8). When only considering haloes with at least one GC, there is no significant difference at different redshifts. Unsurprisingly, taking all haloes into account results in higher values of $M_{DM,GC}$, particularly visible for virial masses below $\sim 10^{10.5} \, M_\odot$, which appears to also decrease slightly with redshift. This may be caused by a higher number of low-mass backsplash haloes with no GCs at lower redshift. A feature seen in all cases is the decrease of $M_{DM,GC}$ towards lower virial masses below $\sim 10^{10} \, M_\odot$, which means that the correlation between $M_{vir}$ and $N_{GC}$ is not entirely linear across the entire mass range. This effect is likely caused by ‘smooth accretion’, which we define as any dark matter mass not accreted through mergers with haloes above $M_{seedGC}$, therefore generally adding no GCs to the main halo (this is discussed in more detail in Section 4.3). Note that there is no significant difference when using a different seeding distribution, except that when only considering haloes with at least one GC, the value of $M_{DM,GC}$ is higher at low virial masses for $p_1$ and lower for $p_3$.

The values of $M_{DM,GC}$ for haloes with at least one GC match the values of the observational sample used by B&F20 well, even at low virial masses, which is remarkable because the focus of the fit was the linear correlation seen at higher masses. Note that the small number of low-mass haloes in the observational sample lead to large uncertainties in that region, however. The sample from H+17 features an overall higher value of $M_{DM,GC}$, which is to be expected because of the larger GC seeding mass necessary to fit the model to it. Unsurprisingly, using the simulations with $M_{seedGC} = 10^{6.67} \, M_\odot$ shifts the simulated lines upwards to better match that sample.

Figure 6. Correlation between dark matter halo virial masses $M_{vir}$ and numbers of globular clusters $N_{GC}$ of the simulated haloes compared to the observational sample from B&F20. GCs were seeded at a virial mass of $10^{8.50} \, M_\odot$ and the seeding distribution $p_1$ was used. The values of the simulated virial masses are mock observations of the originally simulated values. The fit to the simulation was computed by the least-square method for the bin masses between $10^{10.5}$ and $10^{13} \, M_\odot$, averaged over 1000 mock observations of the simulated virial masses. The fit to the observations corresponds to Equation 8.

Figure 7. Correlation between dark matter halo virial masses, $M_{vir}$, and numbers of globular clusters, $N_{GC}$, of the simulated haloes at redshifts $z = 0, 3,$ and $6$, from top to bottom respectively, compared to the observed linear relation determined by B&F20 (Equation 8). GCs were seeded at a virial mass of $10^{8.50} \, M_\odot$ and the seeding distribution $p_2$ was used. The values of the virial masses are mock observations of the originally simulated values.

1 This definition of ‘smooth accretion’ is not physically motivated, but arises from the GC model itself to quantify the amount of dark matter accreted containing no potential GCs. It is for this reason that it breaks down at low virial masses near $M_{seedGC}$.
4.3 Smooth Accretion

To take a closer look at the smoothly accreted mass in haloes, we determined the total mass accreted through mergers with haloes more massive than $M_{\text{seedGC}}$ by iterating the merger trees along the most massive progenitors. For this, we removed all substructures from the trees by immediately adding subhaloes to the respective main haloes. The remaining mass of a halo, $M_{\text{smooth}}$, is considered to have been smoothly accreted. It is therefore the mass that remains when the sum of all mergers above $M_{\text{seedGC}}$ is subtracted from a halo’s mass at $z = 0$. The average fraction of the total virial mass that is smoothly accreted is shown in Fig. 9 as a function of virial mass. At virial masses below $10^{10} \, M_{\odot}$, our definition of smooth accretion breaks down due to resolution effects near $M_{\text{seedGC}}$. The overall trend shows that a larger fraction of $M_{\text{vir}}$ of low-mass haloes was smoothly accreted compared to that of high-mass haloes. There appears to be a small increase of the smoothly accreted mass fraction at virial masses above $10^{11} \, M_{\odot}$ at higher redshifts, which could be related to them having to grow faster in mass via larger mergers to reach the respective virial mass at earlier times. Note that there are inherent difficulties of extracting accurate information from halo merger trees, such as consistency problems caused by fly-throughs or losing track of haloes between snapshots. Since we only iterate the trees without following the individual particles, we are not able to identify the loss of matter if the total virial mass increases from one snapshot to the next. For this reason, it is possible that the total mass accreted by mergers with haloes above $M_{\text{seedGC}}$ surpasses the total virial mass (in which case we assume $M_{\text{smooth}}$ to be zero), reflecting a bias towards low values of $M_{\text{smooth}}$. Still, the method suffices to find an important trend with globular cluster numbers.

The intrinsic scatter of $N_{\text{GC}}$ at fixed $M_{\text{vir}}$ is correlated with the deviation of smoothly accreted mass from the mean at $M_{\text{vir}}$, without including mock observation errors (Fig. 10). This relation can be described as follows: The more dark matter is smoothly accreted, the fewer GCs a halo has compared to other haloes of the same mass. However, for this scatter to be observable, it will first have to be possible to determine dark matter virial masses with greater accuracy, since this is still the dominant cause of scatter in observations.

Our definition of smooth accretion is comparable to the diffuse accretion introduced by Fakhouri & Ma (2010), for which they either count none of the resolved haloes that are accreted or set a limit via the merger mass fraction, $\epsilon_{\text{limit}} = 0.04$. It is for this reason that they find larger diffusely accreted mass fractions compared to this work, never dropping below 40 per cent. This is compatible with our findings because our absolute limit of $M_{\text{seedGC}}$ allows the merger mass fraction to become much smaller for high-mass main haloes, leading to fewer mergers being included in the smooth accretion. However, a direct comparison with their findings is complicated by the different quantity considered: the overdensity within the surroundings instead of $M_{\text{vir}}$. Still, the general trend of a decreasing smooth accretion fraction with increasing overdensity matches the one we find with $M_{\text{vir}}$ very well. Genel et al. (2010) perform a more detailed analysis by looking at the individual dark matter particle fractions coming from mergers versus being smoothly accreted or
stropped from another halo. Particles of the latter two types would roughly correspond to our smoothly accreted mass. Taking those two together, they also obtain higher values of the smoothly accreted mass fraction, above 50 per cent for high-mass haloes. In this case, their lower resolution leads to more unresolved mergers being counted towards smooth accretion. As they state, their results of close to 100 per cent at lower masses towards $10^{10} M_{\odot}$ are due to resolution effects, much like we found for our analysis near the smooth accretion limit $M_{\text{seedGC}}$. Their results do not differ across the redshift range they looked at, $z = 0-2$, which is in agreement with the small differences seen in our simulations in the range $z = 0-6$ (Fig. 9). They also find a decrease of the smooth accretion fraction with $M_{\text{seed}}$ as we do.

4.4 Age Distributions

By extracting the redshift values at which GCs are seeded, we are able to make predictions on their age distribution across all simulated haloes (Fig. 11). More than half of all GCs are formed more than 12 Gyr ago, with a peak formation rate above 4 Gyr$^{-1}$ Mpc$^{-3}$ at 12.9 Gyr ($z = 6.4$). At later times, the formation rate drops quickly. The first 10 per cent of GCs are formed by $z = 8.8$, the first third by $z = 5.9$, and the last third after $z = 3.0$. Only 10 per cent of all GCs are younger than 8.6 Gyr ($z = 1.2$). For the higher seeding mass $M_{\text{seedGC}} = 10^{7.67} M_{\odot}$, the total number of GCs in the simulation and the formation rate decrease by a third. However, the redshift values until which the specified fractions of GCs formed increase only slightly by $\Delta z = 0.1-0.2$.

Binning the dark matter haloes in virial mass, we determined the average GC age distributions for galaxies in those bins (see upper panel of Fig. 12). In general, we find that most of the GCs are old, in good agreement with observations (e.g. Forbes et al. 2001; Strader et al. 2005; Brodie & Strader 2006). More specifically, we find that more massive haloes generally contain older GCs, while less massive haloes contain relatively more younger GCs. Between 30 and 40 per cent of GCs in haloes more massive than $10^{11} M_{\odot}$ are older than 13 Gyr, with formation times between redshifts $z = 4$ and 7. But also for the smaller halo masses between $10^9$ and $10^{10} M_{\odot}$, about half of the GCs are older than 12 Gyr.

The strongest difference can be found between the mass bins log $M/M_{\odot} = 8.5-9$ and log $M/M_{\odot} = 9-10$: While there already is a lack of GCs younger than 6 Gyr in haloes of the more massive bin, about 20 per cent of the GCs in haloes of the lowest mass bin are younger than 6 Gyr, and about 50 per cent are actually younger than 9 Gyr. On the one hand, this is in agreement with what is seen for the small Magellanic Cloud, where the oldest GCs are only 8 Gyr old, and the youngest have ages below 5 Gyr (Parisi et al. 2014). Sharina et al. (2006) even report two GCs in local dwarf galaxies with ages below 2 Gyr. On the other hand, these haloes have masses barely above our seeding mass of $M_{\text{seedGC}} = 3.2 \times 10^8 M_{\odot}$, and are thus strongly affected by the formation criterion implemented in our model. Furthermore, it is also not clear from observations whether GCs are still born in low-mass haloes at low redshifts in the same way as at high redshifts. The reason for the question is that there are also observations of dwarf galaxies with purely old GC systems (Strader et al. 2005), and several with mixed ages (Sharina et al. 2006). However, given that all those dwarf galaxies only have very few GCs, the picture is not conclusive yet.

We compare our results with the GC age distributions presented by Usher et al. (2019) for four different galaxies, NGC 3377, NGC 3115, NGC 1407, and the MW, with the GC ages of the latter taken from VandenBerg et al. (2013) and Leaman et al. (2013). Their dark halo virial masses are on the order of $10^{12}, 10^{12}, 10^{13}$, and $10^{12} M_{\odot}$, respectively (Forbes et al. 2016 for the former three, using their estimate of the total dark matter virial mass being a factor 10 higher than within 8 effective radii; and Karukes et al. 2020 for the MW). As can be seen in the lower panel of Fig. 12, the age distributions of the MW and NGC 1407 are similar to what is found in our model for galaxies with similar halo masses, in particular with respect to the older GCs. However, both observed galaxies have a younger GC population with ages of 8–12 Gyr, which is much less present in our model. The same age deviation for middle-aged GCs can also be seen for NGC 3115, but to a larger degree. NGC 3377 shows the strongest deviation from our results, with an age distribution similar to what we find for the smallest mass bin, though it is much more massive than the haloes in the smallest mass bin. There is not a single simulated halo in the same mass bin as NGC 3377 with a similar GC age distribution. We specifically looked at the haloes with the youngest average GC systems with halo masses above $10^{10} M_{\odot}$ and could not find a single GC age distribution similar to NGC 3377. While we could find haloes with GC systems that have a young age tail, the GCs with ages less than 6 Gyr were always below 15 per cent. We also looked at the formation histories of the haloes with the youngest and oldest GC systems and could not find a trend with either the amount of mergers nor with the amount of smooth accretion.

We conclude that the simple assumption of GC assembly through hierarchical merging with GCs being born in low mass haloes of about $M_{\text{seedGC}} = 3.2 \times 10^8 M_{\odot}$ can well reproduce the general observed trend for galaxies of larger halo masses to also have older GC systems as well as the peaks at old ages above 12 Gyr, as seen for NGC 1407 (Usher et al. 2019) and M87 (Cohen et al. 1998). This clearly indicates that hierarchical merging and mixing is sufficient to explain the overall population of old GCs in galaxies. We can also find rather young GC populations in the dwarf halo range, albeit this is certainly closely linked to the seeding criterion. However, we also clearly see that our simple model is not able to account for the observed amount of medium-aged GCs with ages of 8–12 Gyr (Peng et al. 2004; Usher et al. 2019), and especially fails to explain the younger GC populations observed in some non-dwarf galaxies like NGC 3377 (Usher et al. 2019) or NGC 1316 (Goudfrooij et al. 2001; Sesto et al. 2018). To account for these, an additional GC formation mechanism is necessary, for example.
forming GCs in gas-rich (major) mergers, which is discussed in the following section.

5 DUAL FORMATION PATHWAYS

5.1 Merger Model Properties

Because of the clear lack of younger GCs compared to observations (Section 4.4), and the suggestions in literature of GC formation in gas-rich mergers (e.g. Ashman & Zepf 1992), we extended the base GC model by an additional formation mechanism generally triggered by gas-rich mergers (model description in Section 2.2.2). Before analysing the full combined model, we first inspected the properties of the merger model only, i.e., of the GCs formed through the merger formation pathway alone. Using $A_{\text{min}} = 0.5 \, \text{Gyr}^{-1}$ and $\eta_{\text{GC}} = 7.0$, which lie close to the best fit of Choksi et al. (2018) ($A_{\text{min}} = 0.5 \, \text{Gyr}^{-1}$ and $\eta_{\text{GC}} = 6.75$), leads to a good resemblance with the observational abundances of B&F20 above $log M_{\text{vir}} \sim 12.25$ (Fig. 13), in agreement with Choksi et al. (2018) and Boylan-Kolchin (2017). However, it is evident that the simulated number of GCs is too low in less massive haloes and GCs are almost nonexistent below $\sim 10^{10.5} M_\odot$ compared to what is observed (Forbes et al. 2018). This is true for all parameter combinations of $A_{\text{min}}$ and $\eta_{\text{GC}}$

An important feature of the merger model is that the GCs formed in mergers are generally younger than those formed in the base model, which precisely addresses the weak point of the base model in the comparisons with observations. This is what makes the combination of the base model and merger model appealing: GCs are formed both in low-mass haloes and in mergers, and the age distribution is made younger by the GCs formed in the merger model.

5.2 Parameter Space & GC Numbers

The merger model introduces two further free parameters, $A_{\text{min}}$ and $\eta_{\text{GC}}$. Together with $M_{\text{seedGC}}$, the parameters are too degenerate with respect to the GC numbers to be able to fit all three of them simultaneously to the observed abundances. Unfortunately, the scatter of $N_{\text{GC}}$ does not help in constraining these since the scatter varies too little between parameter sets and is always smaller than the observed scatter, likely due to observational errors of $N_{\text{GC}}$ and $M_{\text{vir}}$. Because of the small number of full observed GC age distributions, these can also only be used as an indicator, but not for any precise fitting procedure. For these reasons, we explored the parameter space by fixing $A_{\text{min}}$ and $\eta_{\text{GC}}$ to then adapt $M_{\text{seedGC}}$ to best match the observations of B&F20 and H+17 (Tables 1 and 2).
Lower values of the logarithmic halo mass accretion rate threshold, $A_{\min}$, and higher values of the gas mass conversion parameter, $\eta_{GC}$, lead to more GCs being formed through mergers. The models in which these parameters alone lead to too many GCs being formed compared to the respective observational sample are excluded and denoted by dashes in the tables. Models at the border of having too many of these GCs do not allow for enough GCs to be formed through the base model, thus underpopulating the GCs in low-mass haloes and having an S-shaped $N_{GC}$-$M_{vir}$-relation instead of a linear one. Those models are denoted by values put in parentheses. All other models are good options when only considering the abundances. For some of these models there are gas-rich mergers with up to ~100 GCs being formed at once.

Simultaneously, we determined the mock observation errors of each model that lead to the best match of the scatter of $N_{GC}$ with the observations. The GC abundances of a selection of six models fit to B&F20 are plotted in Fig. 14 with the values of $A_{\min}$ and $\eta_{GC}$ denoted in the respective plots, where the respective mock observation errors are applied to the halo virial masses. $M_{seedGC}$ decreases from models M1 to M6 towards the value determined for the base model in Section 3.2. In the same way, the fraction of GCs formed in mergers decreases approximately equally-spaced from 55 down to 10 per cent in haloes with $M_{vir} > 10^{12.5} M_\odot$. We found the trend that having more GCs formed in mergers leads to a steeper linear fit that deviates more from the observational fit. Given that we fit the data starting at $10^{10.5} M_\odot$, this shows how the slope of the linear fit approaches that of the dash-dotted line in Fig. 13 for the merger model alone when the number of GCs formed through merger events is increased.

It is also interesting to note that having more GCs formed in mergers leads to a larger intrinsic scatter of $N_{GC}$ in the models. This can be seen in the required mock observation errors that have to be applied to the simulated virial masses to reproduce the observed scatter (Fig. 15). While it needs to be 0.3 dex for the model with the least GCs formed in mergers (M6), it drops down to 0.2 dex for the models with the largest fraction of them (M1, M2). The same also applies to the models fitted to H+t+17. In no case does the scatter come close to the observed one, however.

This shows that further properties that introduce a range of differences in GC formation naturally lead to a larger intrinsic scatter of $N_{GC}$. This means that our findings simply indicate a likely upper limit of the actual observational uncertainties of 0.2 to 0.3 dex (ignoring any systematic errors).

5.3 Red GC Fractions

We compared the fractions of GCs formed in mergers with the observed red GC fractions from Harris et al. (2015). Under the assumption that GCs formed in the merger model tend to correspond to observed red GCs and that those formed in the base model tend to correspond to blue GCs, we decided to use these fractions to further constrain the parameter space to a rough sensible region. For this, we used the observed feature that the mean fraction of red GCs as a function of $M_{vir}$ appears to converge at around 40–45 per cent, which is also consistent with recent observations on a small sample by Kang & Lee (2021). All simulated models also have such a feature, converging towards higher values for larger $A_{\min}$ and smaller $\eta_{GC}$. The models that lead to mean fractions of GCs formed in mergers between 39 and 46 per cent in haloes with $M_{vir} > 10^{12.5} M_\odot$ are highlighted in grey in Tables 1 and 2. A comparison of one of those models with Harris et al. (2015) is shown in Fig. 16, generally demonstrating a good agreement between observational and simulated values. The simulated fractions are lower than those observed for virial masses below ~$10^{12} M_\odot$. Excluding the haloes with fractions of zero results in higher mean values for both the observations and simulations at virial masses below ~$10^{11.5} M_\odot$, but it still does not resolve the difference in the lower mass range. This is also the case for the other models highlighted in the tables. Clearly, the current merger model is not entirely capable of representing the red GC population, especially at low halo masses, although the similar behaviour of the fractions indicate that this approach has its merit. Overall, the best-fitting models feature base model seed massings of around $M_{seedGC} = 10^{8.80} M_\odot$ for B&F20 and $M_{seedGC} = 10^{8.95} M_\odot$ for H+t+17, which are about twice the values found for the base model alone.

5.4 Age Distributions

As for the base model, we extracted the redshift values at which GCs are formed in the different models to obtain the GC age distributions. Gas-rich mergers take place in haloes orders of magnitude more massive than $M_{seedGC}$, therefore leading to an on average later formation time of GCs. Because of this, GCs in massive haloes tend to be younger for larger mean fractions of GCs formed in mergers at the high mass end. Among the selected models with the appropriate fractions of GCs formed in mergers compared to observed red GC fractions (Section 5.3), those with lower values of $A_{\min}$ and $\eta_{GC}$ have younger GC systems on average. This means that forming many GCs at a time (high $\eta_{GC}$) with a higher accretion threshold (low $A_{\min}$) leads to GCs being formed at earlier times on average. Comparing the GC age distributions in the respective virial mass bins with the observations from Usher et al. (2019), we find that the six highlighted models from Tables 1 and 2 with $\eta_{GC} \leq 0.5$
Figure 14. Correlation between dark matter halo virial masses $M_{\text{vir}}$ and numbers of globular clusters $N_{\text{GC}}$ created through a selection of models (composed of the base and merger models) compared to the observational sample from B&F20. GCs were seeded at the virial masses indicated in Table 1. In all cases, the base model seeding distribution $p_2$ was used. The virial masses of the simulations are mock observations of the originally simulated ones, as indicated by the values given in dex. The fits to the simulations were computed by the least-square method for the bin masses between $10^{10.5}$ and $10^{13} M_\odot$, averaged over 1000 mock observations of the simulated virial masses. The fits to the observations correspond to Equation 8.
appearance of GCs formed in mergers at around $M_{\text{vir}} = 10^{11} \, M_\odot$, which generally have younger ages and therefore drop the mean GC age. Interestingly, this inversion is overridden at virial masses above $\sim 10^{13} \, M_\odot$, likely due to the decreased rate of gas-rich major mergers in those haloes at late times (O'Leary et al. 2021).

In contrast to the base model, the GC age distributions now vary a lot more among haloes with virial masses above $10^{11} \, M_\odot$. In fact, we were able to find haloes in some of the models that feature GC populations with similar age distributions to that of NGC 3115 (bottom panel of Fig. 17), which we could not solely with the base model. Particularly in the models fit to H+17, where the GC ages are on average younger because of the larger base model seeding mass, there are some haloes that have age distributions very close to the observed one of NGC 3115. A closer look at the assembly history of the simulated halo best matching the GC age distribution of NGC 3115 revealed that it has many minor mergers around $z \sim 1–2$ and a rather slow growth before and after. Observers have determined that a likely formation scenario for NGC 3115 is composed of two phases with an early gas-rich history and a secular evolution thereafter (Arnold et al. 2011; Guevara et al. 2016; Poci et al. 2019; Buzzo et al. 2021). It is interesting that the selected simulated halo has a history with mergers that appear to take place slightly later than inferred from observations, but still follows the overall picture. This suggests that GC models may be used to extract additional information on the assembly histories of galaxies once their origin is better understood.

However, even in these extended models, we could still not find one halo with GCs similar in age to NGC 3377, proving that even though this model is now capable of producing GC populations with realistic ages on average, even leading to distributions with younger GCs like NGC 3115, it still does not cover more extreme cases.

5.5 Smooth Accretion & Mean GC System Ages

Introducing the merger model does not destroy the correlation between smooth accretion and the scatter of $N_{\text{GC}}$ discussed in Section 4.3. Additionally, it introduces a further trend visible for the more massive haloes: the older the GC population, the more dark matter was smoothly accreted onto the respective halo compared to similarly massive haloes (Fig. 18). An explanation for this is that the added GC merger model traces haloes with gas-rich major mergers, which also tend to have a smaller amount of smooth accretion, as the model forms more young GCs through the merger events. In contrast, any haloes with more smooth accretion and fewer mergers will have a larger relative amount of GCs formed in low-mass haloes, generally corresponding to the older GC population. Combining the two trends of smooth accretion with $N_{\text{GC}}$ and the mean GC age, this means that there is also a trend for haloes with on average younger GCs having slightly more GCs in total.

Fig. 18 also shows how the large downward scatter of mean GC system ages at around $M_{\text{vir}} \gtrsim 10^{11.5} \, M_\odot$ leads to the age inversion discussed in the previous section. The indicated observed galaxies from Usher et al. (2019) lie in this interesting mass range and all but NGC 3377 have mean GC system ages consistent with the simulated values. The downward scatter is more pronounced the more GCs are formed in mergers and therefore depends on the parameters used for the GC merger model. Further estimations of mean GC system ages in observed galaxies could prove helpful in further analysing the properties of such a downward scatter feature at $M_{\text{vir}} = 10^{11.5} \, M_\odot$ to better understand GC formation processes in mergers.
5.6 Dark Matter Halo Mass per GC

Because of the two different pathways that trigger GC formation at different times and in haloes of different virial masses, the dark matter halo mass per GC for the extended model changes compared to the base model (Fig. 19). Due to the typically later formation of GCs formed through mergers, the number of GCs in the lower virial mass range is smaller than previously, leading to a peak of \( M_{\text{DM,GC}} \) at around \( 10^{11} \, M_\odot \). The peak is less pronounced when only taking haloes with at least one GC into account. Because of this, \( M_{\text{DM,GC}} \) is consistently too high compared to the dwarf galaxy regime from B&F20. These values decrease towards what is found in the base model the smaller the fraction of GCs formed in mergers is. Another difference is that now the evolution of \( N_{\text{GC}} \) with redshift changes for high-mass haloes, decreasing with \( z \), such that the peak of \( M_{\text{DM,GC}} \) at \( 10^{11} \, M_\odot \) is intensified at higher redshifts. This means that haloes in the upper virial mass range tend to have slightly more GCs on average at higher redshifts. The general redshift evolution at low halo virial masses remains the same as for the base model (Section 4.2). The deviation from observations in the dwarf galaxy regime shows that larger observational samples in this halo mass range will be helpful to further constrain the dual formation models. As before, the models fit to H+17 simply lead to higher values of \( M_{\text{DM,GC}} \), but continue to show the same qualitative features as for the models fit to B&F20.
6 DISCUSSION

Our presented base model links the number of globular clusters in a dark matter halo to the halo’s virial mass. The resulting distribution of GCs and the linear relationship with halo virial masses is in good agreement with recent observations by H+17 and Forbes et al. (2018). We used the data from B&F20 and H+17 to obtain the seeding mass, i.e., the halo virial mass at which GCs are seeded, and obtained \( M_{\text{seedGC}} = 3.2 \times 10^8 \) and \( 4.7 \times 10^8 \) M\(_{\odot}\), respectively. We show that the linearity between GC numbers and halo virial masses results from hierarchical merging, and that the distribution of GCs at virial masses above \( 10^{11} \) M\(_{\odot}\) does not depend on the details of GC formation, in agreement with models by El-Badry et al. (2019) and B&F20. Varying the seeding distribution, which determines how many GCs are seeded in a single dark matter halo when it reaches \( M_{\text{seedGC}} \), therefore proved to have no effect on the resulting relation. Differences only arise for lower virial masses, in particular in the dwarf galaxy regime, for which observational data are still too limited to draw definite conclusions. Our choice to use the distribution that seeds 0, 1, or 2 GCs with equal probabilities therefore only has little impact on our findings. We expect that further observations in the dwarf galaxy regime will give us a better understanding of when GC populations are formed and in what quantities.

Because of generally too old GC populations compared to observed GC age distributions, we extended this base model by a second GC formation pathway following the approach taken by Choksi et al. (2018). In this merger model, GCs are formed when a dark matter halo’s virial mass increases at a higher rate than an accretion rate threshold, \( A_{\text{min}} \), according to an assumed cluster initial mass power law function and a gas-to-GC-mass conversion factor, \( \eta_{\text{GC}} \). This results in GCs generally being formed in gas-rich major mergers for the merger model. These GCs are generally found in haloes with virial masses above \( 10^{10.5} \) M\(_{\odot}\). Because of a lack of better observational constraints for the properties analysed in this work, we limited ourselves to a parameter space analysis of the combined models. Constraining the models with observations from Harris et al. (2015) by assuming a correlation between observed red GCs and the GCs formed through mergers in the simulation, we suggest that values of \( M_{\text{seedGC}} \sim 6.3 \times 10^8 \) and \( 8.9 \times 10^8 \) M\(_{\odot}\) are the most likely when fitting \( N_{\text{GC}} \) to B&F20 and H+17, respectively, approximately twice as large as for the base model alone. Using the observed GC age distributions from Usher et al. (2019), we constrained \( A_{\text{min}} \) to values below 0.5 Gyr\(^{-1}\) and \( \eta_{\text{GC}} \leq 4.0 \) and 2.0, with the upper limit of \( \eta_{\text{GC}} \) depending on the respective observational sample fitted to. While the value of \( A_{\text{min}} \) matches the best fit found by Choksi et al. (2018) of 0.5 Gyr\(^{-1}\), they find a higher conversion factor of \( \eta_{\text{GC}} = 6.75 \). This deviation is a result of our model additionally consisting of the base model, therefore requiring fewer GCs formed through the merger pathway, while their model tries to capture all observed GCs solely through the merger formation pathway.

There has recently been discussion about the relationship between GCs and the virial masses of their host haloes in the dwarf galaxy regime. Forbes et al. (2018) found that the relation between the total GC mass in a halo and the halo’s virial mass extends linearly down to \( M_{\text{vir}} \sim 10^8 \) M\(_{\odot}\), almost two orders of magnitude lower than observations had previously allowed us to investigate. They noted that they found this proportionality for galaxies with at least one GC. Boylan-Kolchin (2017) presented a model to reproduce the same linear relationship, but predicted that the minimum halo virial mass hosting GCs at \( z = 0 \) is \( \sim 10^{10} \) M\(_{\odot}\), which is not in agreement with recent data. The semi-analytical model introduced by El-Badry et al. (2019) leads to very little GC formation in haloes less massive than \( M_{\text{vir}} \sim 10^{10.5} \) M\(_{\odot}\). They did note, however, that increasing the formation efficiency of GCs in low mass haloes would extend the linear relationship to lower masses. All of these models do not investigate ‘mean’ properties of all haloes, irrespective of having GCs. Instead, they focus on the properties of GC ‘populated’ haloes. In contrast, our model is capable of both obtaining GCs in the dwarf galaxy regime and also making predictions for the ‘mean’ properties of GC populations in dark matter haloes. However, it fails to agree with the prediction made by Forbes et al. (2018), that the mean total GC mass per halo virial mass would decrease below \( M_{\text{vir}} \sim 10^{10} \) M\(_{\odot}\) if galaxies without GCs were to be included in the observational data. Instead, our model indicates that the opposite is true: dark matter haloes with low virial masses contain more GCs per halo mass. It should be noted that one difference is that we only count the number of GCs, whereas they used the total GC mass in a halo. Still, we do not see any way in which our model would agree with their prediction, even if we were to model the total mass of GCs instead of using GC numbers.

We find that the scatter of the simulated number of GCs in haloes lies between the prediction of the central limit theorem and the actual observed scatter for all seeding distributions. Varying the seeding distribution does not noticeably increase the scatter of the GC numbers for high virial masses, however. This means that not only the linear relation, but even the scatter of \( N_{\text{GC}} \) at higher virial masses follows from hierarchical merging, independent of the GC formation details. The merger model adds its own intrinsic scatter in \( N_{\text{GC}} \), however only slightly increasing the total scatter when combining the base and the merger model. Adding a mock observation error to the simulated halo virial masses results in a scatter of the GC numbers that matches the observed scatter more closely. This agrees with the findings of B&F20 that the larger scatter in the observations can be explained by an error of dark matter halo virial masses of a factor of 2. Adding more properties and features with their own intrinsic scatter to \( N_{\text{GC}} \) would naturally lower this value, meaning that our findings simply indicate a likely upper limit of the actual observational errors (ignoring any systematic errors).

For the first time, we show that the inherent scatter of the number of globular clusters is closely related to the amount of smooth accretion, i.e., a halo’s mass that was not accreted through mergers with haloes more massive than \( M_{\text{seedGC}} \). This relation shows that for \( M_{\text{vir}} \geq 10^{10.5} \) M\(_{\odot}\), haloes with above-average smooth accretion generally contain a below-average number of GCs. Since this inherent scatter is much smaller than that introduced by the mock observations, it would mean that the number of GCs is actually a very precise indicator for the virial mass of the corresponding dark matter halo, particularly for those in the high-mass regime. This is still the case when adding the second GC formation pathway, which even introduces a further correlation with the mean GC age in high-mass haloes: the younger the GC population is, the less dark matter mass was smoothly accreted compared to haloes at the same virial mass, and therefore the more GCs the halo contains. However, these relations could change when allowing for GC disruption or when applying a different mechanism for the second GC formation pathway.

\( M_{\text{seedGC}} \) is an order of magnitude lower than the characteristic dark matter mass in virialised haloes per GC at \( z = 0 \) found when fitting the linear relation to the observations. This means that the amount of dark matter mass per GC increases with higher masses, reaching a factor of almost ten. This is caused by the smoothly accreted dark matter, which clearly shows that it is important and...
necessary to account for smooth accretion when modelling GC formation and the relationship between GCs and dark matter halo virial masses. Even for the extended models, the base model seeding masses are still more than half an order of magnitude below the virial masses. Even for the extended models, the base model seeding masses are still more than half an order of magnitude below the virial masses. Additionally, they introduce a peak of the characteristic dark matter mass per GC at $M_{\text{vir}} \approx 10^{11} \, M_\odot$. This feature will allow further constraining the models with the help of more GC observations in galaxies with haloes in this mass range, as the peak becomes more pronounced the more GCs are formed through mergers.

For the base model, we used the best-fitting values for the GC seeding halo virial mass and the mock observation error to predict what the GC number distribution looks like at higher redshifts. We find that there is no significant difference between redshifts, except that we find less massive galaxies at higher redshifts, attributing to the fact that more massive galaxies need more time to grow through merging and thus only appear at lower redshifts. The similarities of the relation between $N_{\text{GC}}$ and $M_{\text{vir}}$ and its scatter at different redshifts suggest that the merging histories of haloes with equal virial masses, but at different times, is self-similar, which agrees well with the findings of Wechsler et al. (2002). Adding the GC merger model, GCs are formed at different times and in different environments, depending on their formation pathway. This affects the relation between $M_{\text{vir}}$ and $N_{\text{GC}}$ at higher redshifts, leading to a slightly larger number of GCs in haloes with virial masses above $10^{11} \, M_\odot$. Note that such observations of GCs at high redshifts are very difficult and may not be possible.

Finally, we also looked at the age distributions of the simulated GCs in the base model, both across all simulated haloes and also for different host halo virial mass ranges. Generally, most of our GCs are old, in agreement with observations (e.g. Forbes et al. 2001; Strader et al. 2005; Brodie & Strader 2006). We find a clear formation peak at $12$–$13$ Gyr, with two thirds of GCs already formed by redshift $z = 3$. We could also successfully reproduce the tendency for more massive haloes to host older GC systems, indicating that this is caused by the hierarchical assembly and not by internal galaxy formation processes. In addition, there is also a tail of middle aged and young GCs in the simulated sample, albeit the GCs with ages below 6 Gyr are extremely rare. Even when looking at individual haloes, we could not find any with a significant middle-aged or even young GC population. This is in contrast to observations, which report middle-aged GCs for almost all GC populations for which age distributions are measured (e.g. Usher et al. 2019).

Introducing the second GC formation pathway through the merger model instead led to generally younger GC populations, simultaneously increasing the variance of GC age distributions. Some of the models in the constrained parameter space feature mean age distributions in the appropriate halo virial mass bins very similar to those found in two of the galaxies analysed by Usher et al. (2019), the MW and NGC 1407. Even galaxies with somewhat outlying age distributions, such as NGC 3115, are matched by GC populations in selected individual haloes from the simulations. The suggested two-phase formation scenario of NGC 3115 composed of an early gas-rich merger history and a secular evolution thereafter (Arnold et al. 2011; Guérout et al. 2016; Poci et al. 2019; Buzzo et al. 2021) is overall consistent with that of the matching simulated halo. We suggest that understanding the origin of GCs better will allow using GC models to extract additional information on the assembly histories of galaxies in the future. However, even the extended model is still incapable of explaining young GC systems, as for example reported for the small Magellanic Cloud (Parisi et al. 2014), NGC 3377 (Usher et al. 2019), and NGC 1316 (Goudfrooij et al. 2001; Sesto et al. 2018).

An interesting feature that emerges in the extended model is an age inversion in the virial mass range $\log M_{\text{vir}}/M_\odot = 11$–$13$, where more massive haloes contain overall younger GC populations. Obtaining more observational data on mean GC ages of galaxies’ GC populations will help place further constraints on the details of GC formation and their pathways by searching for such an age inversion feature and inspecting how pronounced it might be.

Since this is an empirical model, we have chosen to base it on few but simple assumptions. It currently restricts us to solely making statements about the number of GCs in dark matter haloes, their formation times, and, when including the extension to the base model, the ratios between GCs of different formation pathways. For this work, we decided against trying to model other GC properties, such as masses, metallicities, and any sorts of dynamics, including positions and velocities within the host halo. We also believe it is of importance to more accurately model the relation between GC numbers and virial masses in lower mass galaxies before extending the model with additional properties. It may be worthwhile to investigate possible effects of GC disruption as well in the future. We conclude that our base model, which assumes GCs to be born in low-mass haloes, is sufficient to explain the origin of the old GCs and a low number of younger GCs, but that an additional channel of GC formation is needed to account for the observed amount of middle-aged and young GCs, as shown by the GC merger model through the formation of GCs during gas-rich merger events (see also Ashman & Zepf 1992 for more details).

7 CONCLUSION

In this work, we presented an empirical model for the number of globular clusters in galaxies by using observational data on the relationship between dark matter halo virial masses and GC numbers to fit the model. We iterate halo merger trees that are extracted from cosmological dark matter simulations and seed GCs at a certain halo virial mass, $M_{\text{seedGC}}$, according to a prespecified GC number seeding distribution. GCs are only formed through this process and cannot disappear once seeded. Haloes that become subhaloes turn over their GC population to the respective main halo.

We also presented an extended model that for the first time combines two different GC formation pathways. The additional formation mechanism depends on a dark matter mass accretion rate threshold, $A_{\text{min}}$, and typically occurs in gas-rich mergers. GCs are formed according to a conversion prescription from gas to GC mass and a cluster initial mass power law function, meaning that multiple GCs can be formed simultaneously.

We fitted $M_{\text{seedGC}}$ to achieve the best agreement between our model and the observational data compiled by Burkert & Forbes (2020) and Harris et al. (2017b). The best fit is obtained for $M_{\text{seedGC}} = 3.2 \times 10^8$ and $4.7 \times 10^8 \, M_\odot$, respectively. We found that we can reproduce the observed linear correlation between halo virial masses and GC numbers very well, in particular for high-mass haloes. The observed scatter of GC numbers around the linear relationship can be explained by an observation error of a factor of 2 for dark matter halo virial masses. For the extended model, constraining the parameter space leads to base model seeding masses that are approximately twice as large.

The model has allowed us to come to the following key conclusions:

- The GC number seeding distribution has no impact on the
linear relation between $M_{\text{vir}}$ and $N_{GC}$ or its scatter for $M_{\text{vir}} \geq 10^{11} \, M_\odot$ because of hierarchical merging.

- The correlation and scatter between $M_{\text{vir}}$ and $N_{GC}$ does not change with redshift, except for massive haloes only appearing at lower redshifts. Introducing the second formation pathway introduces a peak of the mean dark matter mass per GC in haloes at around $10^{11} \, M_\odot$ and decreases at higher redshifts in more massive haloes.

- Smooth accretion leads to a characteristic dark matter per GC that is one order of magnitude larger than $M_{\text{seed}}$ (half an order of magnitude for the extended model), meaning that GCs form in less massive haloes than expected from the linear relation between $M_{\text{vir}}$ and $N_{GC}$ observed for higher-mass haloes.

- The average amount of smooth accretion does not change with redshift at fixed halo virial mass, clearly showing that the different appearance of galaxies at different redshifts is purely driven by gas and star formation physics and not by dark matter accretion.

- The inherent scatter of $N_{GC}$ at fixed halo virial mass is strongly correlated with the deviation of smoothly accreted mass from the mean at the given virial mass. However, this scatter is smaller than that introduced by halo mass measurements, making $N_{GC}$ a good tracer for the halo virial mass.

- Including the second GC formation pathway introduces a further correlation between the mean GC age and the deviation of smoothly accreted mass, and therefore also between the mean GC age and the scatter of $N_{GC}$ at fixed halo virial mass.

- The base model achieves to form old GC populations, but under-populates the young ones. The extended model achieves much better comparability with observed GC ages, however still fails to reproduce very young GC populations found in some particular galaxies.

- The extended model with two GC formation pathways is capable of producing numbers of GCs formed through mergers that follow the observed relation between $M_{\text{vir}}$ and red GC fractions in galaxies very well.

We conclude that our presented base model is already sufficient to explain several observed properties of GC systems in galaxies, clearly highlighting the importance of hierarchical merging for the number assembly of GCs. We can also successfully explain the age peak of GCs at high redshifts due to the formation of GCs in low-mass haloes. However, to more accurately reproduce observations with respect to GC ages, the additional second GC formation pathway is necessary. None of the formation mechanisms on their own covers the broad range of properties observationally found for the numbers and ages of GCs in galaxies of different masses. We believe that our presented model lays the groundwork for even more sophisticated empirical models that are also capable of fine-tuning the distribution of GC numbers in low-mass galaxies and further improving the comparison to red GC populations to best match observations.

ACKNOWLEDGEMENTS

We thank Duncan Forbes for helpful discussions. We also thank the anonymous referee for their helpful suggestions. This research was supported by the Excellence Cluster ORIGINS, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – MO 2979/1-1. Finally, we thank the developers of astropy (Astropy Collaboration et al. 2013, 2018), jupyter (Kluyver et al. 2016), matplotlib (Hunter 2007), numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), and pandas (Wes McKinney 2010; Reback et al. 2020) for their very useful free software. This research has made use of NASA’s Astrophysics Data System Bibliographic Services and of the arXiv preprint repository.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Amorisco N. C., Monachesi A., Agnello A., White S. D. M., 2018, MNRAS, 475, 4235
Arnold J. A., Romanowsky A. J., Brodie J. P., Chomiuk L., Spitler L. R., Strader J., Benson A. J., Forbes D. A., 2011, ApJ, 736, L26
Ashman K. M., Zepf S. E., 1992, ApJ, 384, 50
Astropy Collaboration et al. 2013, A&A, 558, A33
Astropy Collaboration et al. 2018, AJ, 156, 123
Bastian N., Pfeffer J., Kruijssen J. M. D., Cruijil-Gomez S., Reina-Campos M., 2020, MNRAS, 498, 1050
Beasley M. A., Baugh C. M., Forbes D. A., Sharples R. M., Frenk C. S., 2002, MNRAS, 333, 383
Beasley M. A., Bridges T., Peng E., Harris W. E., Harris G. L. H., Forbes D. A., Mackie G., 2008, MNRAS, 386, 1443
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busa M. T., Klypin A. A., Primack J. R., 2013b, ApJ, 763, 18
Blakeslee J. P., Troup J. L., Meitgen M. R., 1997, AJ, 114, 482
Boylan-Kolchin M., 2017, MNRAS, 472, 3120
Brodie J. P., Strader J., 2006, ARA&A, 44, 193
Brodie J. P., Strader J., Demichelis G., Beasley M. A., Cenarro A. J., Larsen S. S., Kuntschner H., Forbes D. A., 2005, AJ, 129, 2643
Burkert A., Forbes D. A., 2020, AJ, 159, 56
Buzzo M. L., Cortesi A., Werle A., Mednes de Oliveira C., 2021, IAU Symposium, 359, 421
Cies-Santos A. L., Larsen S. S., Kuntschner H., Anders Wehrm E., Strader J., Brodie J. P., Santos J. F. C., 2011, A&A, 525, A20
Chies-Santos A. L., Larsen S. S., Klessi-Patir M., 2012, MNRAS, 427, 2349
Choksi N., Gnedin O. Y., Li H., 2018, MNRAS, 480, 2343
Cohen J. G., Blakeslee J. P., Ryzhov A., 1998, ApJ, 496, 808
Côté P., Marzke R. O., West M. J., 1998, ApJ, 501, 554
El-Badary K., Quataert E., Weisz D. R., Choksi N., Boylan-Kolchin M., 2019, MNRAS, 482, 4528
Fakhouri O., Ma C.-P., 2010, MNRAS, 401, 2245
Forbes D. A., 2018, MNRAS, 479, 4760
Forbes D. A., Brodie J. P., Grillmair C. J., 1997, AJ, 113, 1652
Forbes D. A., Beasley M. A., Brodie J. P., Klessi-Patir M., 2001, ApJ, 563, L143
Forbes D. A., Alabi A., Romanowsky A. J., Brodie J. P., Strader J., Usher C., Pota V., 2016, MNRAS, 458, L44
Forbes D. A., Read J. I., Gieses M., Collins M. L. M., 2018, MNRAS, 481, 5592
Forbes D. A., Alabi A., Romanowsky A. J., Brodie J. P., Arimoto N., 2020, MNRAS, 492, 4874
Genel S., Bouché N., Naab T., Stornberg A., Genzel R., 2010, ApJ, 719, 229
Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65
Goudloof P., Mack J., Klessi-Patir M., Meylan G., Minniti D., 2001, MNRAS, 322, 643
Valenzuela et al.

Guérou A., Emsellem E., Krajnović D., McDermid R. M., Contini T., Weilbacher P. M., 2016, A&A, 591, A143

Hahn O., Abel T., 2011, MNRAS, 415, 2101

Harris W. E., 1996, AJ, 112, 1487

Harris W. E., Harris G. L. H., Hudson M. J., 2015, ApJ, 806, 36

Harris W. E., Blakeslee J. P., Whitmore B. C., Gnedin O. Y., Geisler D., Rothberg B., 2016, ApJ, 817, 58

Harris W. E., Ciccone S. M., Eadie G. M., Gnedin O. Y., Geisler D., Rothberg B., Bailin J., 2017a, ApJ, 835, 101

Harris W. E., Blakeslee J. P., Harris G. L. H., 2017b, ApJ, 836, 67

Harris C. R., et al., 2020, Nature, 585, 357

Hunter J. D., 2007, Computing in Science & Engineering, 9, 90

Jiang F., van den Bosch F. C., 2014, MNRAS, 440, 193

Kang J., Lee M. G., 2021, arXiv e-prints, p. arXiv:2104.00672

Karukes E. V., Benito M., Iocco F., Trotta R., Geringer-Sameth A., 2020, J. Cosmology Astropart. Phys., 2020, 033

Kim J.-h., et al., 2018, MNRAS, 474, 4232

Kluyver T., et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press, Netherlands, pp 87–90, https://eprints.soton.ac.uk/403913/

Knebe A., et al., 2011, MNRAS, 415, 2293

Kravtsov A. V., Gnedin O. Y., 2005, ApJ, 623, 650

Kruijssen J. M. D., 2015, MNRAS, 454, 1658

Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019, MNRAS, 486, 3134

Lahén N., Naab T., Johansson P. H., Elmegreen B., Hu C.-Y., Walch S., 2019, ApJ, 879, L18

Leaman R., VandenBerg D. A., Mendel J. T., 2013, MNRAS, 436, 122

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Li H., Gnedin O. Y., 2014, ApJ, 796, 10

Moster B. P., Naab T., White S. D. M., 2018, MNRAS, 477, 1822

O’Leary J. A., Moster B. P., Naab T., Somerville R. S., 2021, MNRAS, 501, 3215

Parisi M. C., et al., 2014, AJ, 147, 71

Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557

Peebles P. J. E., Dicke R. H., 1968, ApJ, 154, 891

Peng E. W., Ford H. C., Freeman K. C., 2004, ApJ, 602, 705

Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, MNRAS, 475, 4309

Planck Collaboration XIII 2016, A&A, 594, A13

Poci A., McDermid R. M., Zhu L., van de Ven G., 2019, MNRAS, 487, 3776

Pota V., et al., 2013, MNRAS, 428, 389

Press W. H., Schechter P., 1974, ApJ, 187, 425

Reback J., et al., 2020, pandas-dev/pandas: Pandas 1.1.3, doi:10.5281/zenodo.4067057, https://doi.org/10.5281/zenodo.4067057

Saifollahi T., Trujillo I., Beasley M. A., Peletier R. F., Knapen J. H., 2021, MNRAS, 502, 5921

Salaris M., Degl’Innocenti S., Weiss A., 1997, ApJ, 479, 665

Schuberth Y., Richtler T., Hilker M., Dirsch B., Bassino L. P., Romanowsky A. J., Infante L., 2010, A&A, 513, A52

Sesto L. A., Faifer F. R., Smith Castelli A. V., Forte J. C., Escudero C. G., 2018, MNRAS, 479, 478

Sharina M. E., Afanasiev V. L., Puzia T. H., 2006, MNRAS, 372, 1259

Springel V., 2005, MNRAS, 364, 1105

Strader J., Brodie J. P., Cenarro A. J., Beasley M. A., Forbes D. A., 2005, AJ, 130, 1315

Tacconi L. J., Genzel R., Sternberg A., 2020, ARA&A, 58, 157

Usher C., et al., 2012, MNRAS, 426, 1475

Usher C., Brodie J. P., Forbes D. A., Romanowsky A. J., Strader J., Pfeffer J., Bastian N., 2019, MNRAS, 490, 491

VandenBerg D. A., Brogaard K., Leaman R., Casagrande e L., 2013, ApJ, 775, 134

Virtanen P., et al., 2020, Nature Methods, 17, 261

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52

Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, Proceedings of the 9th Python in Science Conference. pp 56 – 61, doi:10.25080/Majora-92bf1922-00a

van Dokkum P., et al., 2016, ApJ, 828, L6

van Dokkum P., et al., 2017, ApJ, 844, L11

This paper has been typeset from a TeX/LaTeX file prepared by the author.