Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **527,** 3692–3708 (2024) https://doi.org/10.1093/mnras/stad3345
Advance Access publication 2023 November 2

Catalogue of model star clusters in the Milky Way and M31 galaxies

Yingtian Chen★ and Oleg Y. Gnedin
Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

Accepted 2023 October 30. Received 2023 October 26; in original form 2023 September 21

ABSTRACT

Detailed understanding of the formation and evolution of globular clusters (GCs) has been recently advanced through a combination of numerical simulations and analytical models. We employ a state-of-the-art model to create a comprehensive catalogue of simulated clusters in three Milky Way (MW) and three Andromeda (M31) analogue galaxies. Our catalogue aims to connect the chemical and kinematic properties of GCs to the assembly histories of their host galaxies. We apply the model to a selected sample of simulated galaxies that closely match the virial mass, circular velocity profile, and defining assembly events of the MW and M31. The resulting catalogue has been calibrated to successfully reproduce key characteristics of the observed GC systems, including total cluster mass, mass function, metallicity distribution, radial profile, and velocity dispersion. We find that clusters in M31 span a wider range of age and metallicity, relative to the MW, possibly due to M31's recent major merger. Such a merger also heated up the *in-situ* GC population to higher orbital energy and introduced a large number of *ex-situ* clusters at large radii. Understanding the impacts of galaxy mergers and accretion on the GC populations is crucial for uncovering the galaxy assembly histories.

Key words: Galaxy: formation – globular clusters: general – galaxies: individual: M31 – galaxies: star clusters: general.

1 INTRODUCTION

Over the past decade, significant observational efforts have been made to uncover the origins of globular clusters (GCs) in the local Universe. Spectroscopic surveys like the Apache Point Observatory Galactic Evolution Experiment (APOGEE, Majewski et al. 2017) have offered valuable insights into the chemical compositions of these ancient stellar clusters. Particularly, the advent of the *Gaia* mission (Gaia Collaboration 2016, 2018, 2023) has revolutionized our understanding of the GC spatial and kinematic properties, and their stellar populations. Studies such as Massari, Koppelman & Helmi (2019) and Malhan et al. (2022) took advantage of the *Gaia* data to paint a comprehensive picture of the origins of Galactic GCs, shedding light on key questions related to the formation and evolutionary history of GCs, including where GCs formed and how they were brought to their current locations in the Galaxy.

At the same time, numerous observational and theoretical studies have improved our understanding of the formation history of the Milky Way (MW) and Andromeda (M31) galaxies. For example, Deason, Belokurov & Weisz (2015) investigated the ratio of blue straggler stars to blue horizontal branch stars in the MW halo and suggested the accretion of massive satellite galaxies as progenitors of the stellar halo. Inspired by this, subsequent chemical and kinematic studies focused on disc and halo stars (Belokurov et al. 2018; Deason et al. 2018; Helmi et al. 2018) and confirmed the likely major merger at lookback time $\gtrsim$10 Gyr. The progenitor galaxy of this merger is referred to as the Gaia–Sausage/Enceladus (GS/E). Analysis of the kinematics of MW stars (Belokurov & Kravtsov 2022) showed that after this early bursty star forming epoch, the Galaxy transitioned to a steady stage of disc formation within 1–2 Gyr. Cosmological simulations of galaxy formation also supported the early bursty stage and the subsequent transition to a steady state (Semenov et al. 2023; Yu et al. 2023). On the other hand, recent studies found that M31 galaxy had a distinct assembly history from the MW, characterized by a massive merger with an M32-like progenitor around 2 Gyr ago (D'Souza & Bell 2018).

Our understanding of the formation and evolution of star clusters throughout cosmic time has also been greatly advanced over the past decade. Some of the young massive clusters formed in high-redshift galaxies may survive tidal disruption, becoming GCs at the present. In this framework, theoretical modelling of GCs has become possible by embedding necessary recipes of cluster formation and evolution as sub-grid prescriptions into cosmological simulations. For example, Li et al. (2017), Li, Gnedin & Gnedin (2018), and Li & Gnedin (2019) treated forming star clusters effectively as sink particles centred on giant molecular clouds in a suite of zoom-in hydrodynamical simulations. The MOdelling Star cluster population Assembly In Cosmological Simulations within EAGLE project (E-MOSAICS, Pfeffer et al. 2018; Kruijssen et al. 2019) applied the MOSAICS model for cluster formation and evolution (Kruijssen & Lamers 2008; Kruijssen 2009; Kruijssen et al. 2011) to a rerun of the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulations (Crain et al. 2015; Schaye et al. 2015). The EMP-*Pathfinder* project (Reina-Campos et al. 2022) further updated the E-MOSAICS physics recipes with new prescriptions for cluster

---

★ E-mail: ybchen@umich.edu

formation and the multiphase interstellar medium (ISM). These works were able to match various observational properties of both young and old clusters. They also offered insights to the links between GC properties and the assembly history of their host galaxy (Kruijssen et al. 2019; Pfeffer et al. 2020; Reina-Campos et al. 2020; Trujillo-Gomez et al. 2021).

In addition to these direct simulations, post-processing methods have also gained traction in the field. Such methods employ analytical prescriptions for GC formation and evolution, applying them to the merger trees and particle outputs of existing cosmological simulations. Unlike full galaxy formation simulations, these models are not sensitive to the specific baryonic prescriptions used in simulations, and do not require costly reruns. Such robustness and flexibility make them particularly advantageous when applied to large samples of galaxies. Examples of these methods are Renaud, Agertz & Gieles (2017), Creasey et al. (2019), Halbesma et al. (2020), Phipps et al. (2020), Valenzuela et al. (2021, 2023), including previous versions of our model (Muratov & Gnedin 2010; Li & Gnedin 2014; Choksi, Gnedin & Li 2018; Choksi & Gnedin 2019a, b).

Our model takes the halo merger tree as the input and triggers GC formation when the specific mass accretion rate exceeds a predefined threshold. Utilizing a sequence of scaling relations, the model analytically calculates the mass and metallicity of the newly formed GCs. It also accounts for mass-loss due to stellar evolution and tidal disruption. Starting in Chen & Gnedin (2022, hereafter CG22), we have improved the model by including the spatial and kinematic information for GCs using tracer star particles from the simulations. This enabled a comprehensive comparison between model predictions and the nine-dimensional (mass + age + metallicity + 3D positions + 3D velocities) characteristics of MW GCs. A subsequent extension by Chen & Gnedin (2023, hereafter CG23) incorporated dark matter (DM) particles to represent GCs in collisionless simulations. Our refined model successfully replicated key statistics of observed GC systems, including the distributions of GC mass, metallicity, distance from the galaxy centre, velocity dispersion, and anisotropy. By applying this model across a wide range of galaxy mass, we reproduced global scaling relations such as the effective radius–galaxy mass and the nearly linear GC mass–galaxy mass correlations.

In this work, we use the model results to produce a catalogue of model GCs with the properties that match the observations of the GC systems in MW and M31. The primary use of this catalogue is to help the analysis of observations to reveal the relationship between GC features and the assembly histories of the MW and M31. To achieve this, we start by selecting a set of simulated galaxies that meet all major observational constraints, including the virial mass, circular velocity profile, and the defining assembly events, such as the GS/E-like merger for the MW and the M32-progenitor-like merger for M31. Next, we calibrate our model parameters by comparing observed GC attributes with those in our model systems. These attributes include total GC number, mass function, metallicity distribution, radial profile, and velocity dispersion. The resulting catalogue can enhance our understanding of how galaxy assembly events influence the current distribution of GCs in the property space, such as the age–metallicity plane and the integral of motion (IOM) space. The catalogue enables us to assess the accuracy of various classification algorithms in these spaces to identify the original progenitors for *ex-situ* GCs and, ultimately, reconstruct the host galaxy's assembly history.

The rest of the paper is organized as follows. We outline the background simulations and detail all updates made to the CG22 and CG23 models in Section 2. Next, we define the criteria used to select the best-matching MW and M31 analogues within our simulations in Section 3. In Section 4, we present the GC systems in the selected analogues and compare them with the observational data. We also explore the influence of the last major merger in M31 on the spatial distribution of GCs. The discussion in Section 5 focuses on the potential application of the model to uncover galaxy assembly histories and addresses known caveats. We conclude and summarize the study in Section 6. In addition, we present a new functional form for the time-dependent galactic stellar mass–metallicity relation (MZR), used to determine the metallicity of model GCs, in Appendix A.

2 MODEL SET-UP

To create model catalogues of GC systems in the MW and M31, we apply our GC formation model to six carefully selected galaxies from the cosmological Illustris TNG50-1 simulation (hereafter TNG50; Nelson et al. 2019; 2021 Pillepich et al. 2019) and from a suite of collisionless simulations of a Local Group (LG) environment (CG23). These simulated galaxies have assembly histories similar to the MW or M31 and reproduce most of the observable properties of their GC systems. In this section, we provide an overview of the simulations and the model set-up. We describe the criteria used to select the best MW/M31 analogues in the next section.

2.1 Background cosmological simulation

We apply our model on two suites of simulations. The first suite is TNG50, performed with the moving mesh hydrodynamic code AREPO (Springel 2010). TNG50 employs the IllustrisTNG model of galaxy formation (Pillepich et al. 2018) in a (51.7 Mpc)^3 comoving box, adopting a flat Lambda cold dark matter cosmology with $\Omega_{\rm b} = 0.0486$, $\Omega_{\rm m} = 0.3089$, $\Omega_\Lambda = 0.6911$, $h = 0.6774$, $\sigma_8 = 0.8159$, and $n_{\rm s} = 0.9667$ (Planck Collaboration 2016). TNG50 is initiated with 2160^3 DM particles and the same number of gas particles, yielding an average DM particle mass $\sim 4.5 \times 10^5\,{\rm M}_\odot$ and an average gas cell mass $\sim 8.5 \times 10^4\,{\rm M}_\odot$. The gas cells in star-forming regions have typical sizes ~100 pc (Pillepich et al. 2019).

TNG50 applies the Friends-of-Friends algorithm and the SUBFIND code (Springel et al. 2001) to identify haloes and subhaloes. We use the term 'galaxy' to refer to subhaloes throughout the paper. Additionally, TNG50 provides galaxy merger trees using the SUBLINK code (Rodriguez-Gomez et al. 2015).

The second suite contains two collisionless zoom-in simulations, which we ran in CG23. We performed the simulations with the adaptive refinement tree (ART; Kravtsov, Klypin & Khokhlov 1997) code on the modified initial conditions (ICs) from the Exploring the Local Volume in Simulations (ELVIS, Garrison-Kimmel et al. 2014). The ICs are `Thelma & Louise` and `Romeo & Juliet`, each producing a galaxy environment similar to the LG. We refer to this simulation suite as the 'LG simulations' hereafter.

We run the ROCKSTAR (Behroozi, Wechsler & Wu 2013a) halo finder and the CONSISTENT TREE code to construct halo catalogues and merger trees for the LG simulations, respectively. The two LG simulations each produces two main galaxies. We find the `Louise`, `Romeo`, and `Juliet` galaxies have a quiescent 'MW-like' mass assembly history after $z \sim 5$ with no major merger with a mass ratio less than 4:1. This feature is similar to the formation history of the MW (Hammer et al. 2007). On the other hand, `Thelma` has more major mergers at later times.

2.2 Modelling cluster formation and evolution

Our GC formation and evolution model is based on CG22, where we modified the previous versions of the model (Muratov & Gnedin 2010; Li & Gnedin 2014; Choksi, Gnedin & Li 2018) to include positional and kinematic information by tagging simulation particles as ‘GC tracer particles’. In this section, we briefly describe the set-up of the CG22 model and further modifications we make in this work.

Our model consists of four steps: (1) cluster formation, (2) cluster sampling, (3) particle assignment, and (4) cluster evolution. The formation of GCs is triggered by rapid mass growth of the host galaxy (e.g. major mergers or intense mass accretion), quantified by the specific mass accretion rate $\dot{M}_{\rm h}/M_{\rm h}$ exceeding a threshold value p_3, which is an adjustable model parameter. We then apply the empirical stellar mass–halo mass (SMHM) relation (Behroozi, Wechsler & Conroy 2013b) to compute the stellar mass $M_\star$ from the halo merger history; the gas mass–stellar mass relation (Lilly et al. 2013; Genzel et al. 2015; Tacconi et al. 2018; Wang et al. 2022) to calculate gas mass $M_{\rm gas}$ from stellar mass; and finally the linear gas mass–cluster mass relation (Kravtsov & Gnedin 2005), $M_{\rm tot} = 1.8 \times 10^{-4} p_2 M_{\rm gas}$, to compute the total GC mass $M_{\rm tot}$, where p_2 is another model parameter quantifying the cluster formation rate. We also use the time-dependent MZR to assign the host galaxy metallicity to its population of GCs forming at a given epoch. We make some modifications to the scaling relations employed in CG22 to match the updates in the theoretical and observational results and to improve the modelling of scatter evolution. We provide details of these modifications in Section 2.2.1.

Next, we sample the masses of individual clusters from the Schechter (1976) initial cluster mass function (ICMF) derived from observations of young star clusters. We extend the range of cluster masses down to $10^4\,{\rm M}_\odot$ instead of $10^5\,{\rm M}_\odot$ in CG22. This allows us to capture some surviving low-mass GCs in the outskirts of the galaxies. We do not model clusters with initial mass below $10^4\,{\rm M}_\odot$ for two primary reasons. First, the lowest mass galaxies we consider have halo masses around $10^8\,{\rm M}_\odot$ (resolved by ∼200 particles in TNG50), and typically produce clusters with masses close to $10^4\,{\rm M}_\odot$, see fig. 5 in CG23. Secondly, our tidal disruption model is effective in disrupting clusters of this mass within a few Gyr. According to calculations in CG23, a $10^4\,{\rm M}_\odot$ cluster located at a galactocentric radius =3 kpc would likely survive less than 1 Gyr. Therefore, extending the mass function below this limit is meaningless.

After obtaining the list of newly formed clusters, we assign them to certain types of simulation particles depending on the simulation. For the hydrodynamic simulations like TNG50, we first select young (age < 10 Myr) stellar particles within an initial radius of 3 kpc from the galactic centre. We have tested different initial radii and found that a larger radius leads to a final spatial distribution more extended than observations. However, reducing the initial radius does not significantly impact the final spatial distribution of the clusters. When there are not enough newly formed stellar particles, we also use older stellar particles formed between the adjacent snapshots. In the rare cases (∼10 per cent), when there are still not enough stellar particles, we use DM particles near the galactic centre as GC tracer particles. We only use the positions and velocities of these particles to passively track individual GCs but calculate all other properties analytically. This minimizes the model dependence on the baryon physics employed in the hydrodynamic simulation.

On the other hand, for the collisionless LG simulations we follow CG23 to select collisionless particles in local density peaks near the galaxy centre, corresponding to surviving dense cores of satellite galaxies or other galactic structure with deep potential wells where massive clusters are more likely to form than elsewhere. We identify peaks within the scale radius of the best-fitting Navarro–Frenk–White halo profile. We also require the peak to be denser than any of the 16 closest grid cells and 30 times the mean density enclosed within the scale radius.

Finally, we compute the GC mass-loss due to the stellar evolution and dynamical disruption based on the local tidal field along their orbits. Different from CG22, we update the prescription for tidal disruption following Gieles & Gnedin (2023), where the disruption rate is a multivariate power-law function of the cluster initial mass and current mass. In Section 2.2.2, we describe the modifications made to the cluster evolution step in greater detail.

To choose the best model parameters, we calibrate the model to match the metallicity distribution, mass function, total GC mass–halo mass relation, and the spatial distribution. We describe the calibration in Section 3.2.

2.2.1 *Modifications to the scaling relations*

Some of the scaling relations in the CG22 model are modified. First, we modify the gas mass–stellar mass relation by updating the upper bound of the gas mass constrained by the extragalactic ultraviolet background after reionization. Following the more recent CG23 model, we require the sum of the gas fraction $f_{\rm gas} = M_{\rm gas}/M_{\rm h}$ and the stellar fraction $f_* = M_\star/M_{\rm h}$ to be less than the total accreted baryon fraction $f_{\rm in}$ parametrized by Kravtsov & Manwadkar (2022):

$$f_{\rm in} = f_{\rm b}\, s(M_{\rm ch}(z)/M_{\rm h}, 2), \tag{1}$$

where $f_{\rm b} = \Omega_{\rm b}/\Omega_{\rm m}$ is the universal baryon fraction, $s(x, y) = [1 + (2^{y/3} - 1)x^y]^{-3/y}$ is a soft step-function, and $M_{\rm ch}$ is the characteristic mass scale at which $f_{\rm in} = 0.5 f_{\rm b}$:

$$M_{\rm ch}(z) = 1.69 \times 10^{10}\,{\rm M}_\odot \frac{\exp(-0.63z)}{1 + \exp[(z/\beta)^\gamma]}, \tag{2}$$

where

$$\beta = z_{\rm rei}\left[\ln\left(1.82 \times 10^3 \exp(-0.63 z_{\rm rei}) - 1\right)\right]^{-1/\gamma}. \tag{3}$$

We adopt the reionization epoch at $z_{\rm rei} = 6$ and $\gamma = 15$ as in Kravtsov & Manwadkar (2022). If $f_{\rm gas} + f_* > f_{\rm in}$, we enforce $f_{\rm gas} = f_{\rm in} - f_*$ by setting $M_{\rm gas} = (f_{\rm in} - f_*)M_{\rm h}$.

Next, although we still use the Behroozi, Wechsler & Conroy (2013b) SMHM relation, we change the way we model and evolve the scatter. In our previous models since Choksi, Gnedin & Li (2018), the scatter was modelled as the cumulative scatter of short-term star formation periods. Such a technique guarantees the stellar mass $M_\star$ to be a monotonic function of time but may underestimate the resultant scatter at present day. Also, this method computes the final $M_\star$ as the summation of a series of lognormal distributions at each epoch, which may not result in a final lognormal distribution to match the present-day observations. Instead, in this work we model the evolution of scatter by a Gaussian process,

$$\log M_\star(M_{\rm h}, z) \sim \mathcal{GP}\left\{\log {\rm SMHM}(M_{\rm h}, z),\ K(t_1, t_2)\right\}. \tag{4}$$

For simplicity, we drop the ‘10’ subscript in the base-10 logarithm for this expression and hereafter. A Gaussian process is a probabilistic model that represents a function as a probability distribution over all possible functions that are consistent with a given relation. It is fully specified by its mean function (in our case, $\log {\rm SMHM}(M_{\rm h}, z)$) and covariance function, or kernel, $K(t_1, t_2)$. The mean function provides the expected value of the function at any epoch, and the covariance function determines how the values at different epochs are related to each other. We choose a squared exponential kernel

$$K(t_1, t_2) = \xi^2(z)\ \exp\left[-\frac{(t_1 - t_2)^2}{2\tau^2}\right] \tag{5}$$

with the scatter of the Behroozi, Wechsler & Conroy (2013b) relation: $\xi(z) = 0.218 + 0.023\,z/(1+z)$. The parameter τ characterizes the autocorrelation time-scale. The limit of $\tau \to 0$ represents a pure Gaussian noise. We set $\tau = 2$ Gyr to reflect a typical gas depletion time-scale in galaxies. However, we have verified that any value in the range $\tau = 1$–4 Gyr does not noticeably affect any property we analyse. The Gaussian process technique solves all the caveats of the previous methods. In addition, the non-zero autocorrelation preserves memory of $M_\star$ at previous epochs, leading to a smoother evolution of $M_\star$. This provides a monotonic function $M_\star(z)$ during the peak of GC formation at $z = 1$–6. At $z \lesssim 1$, when the galactic star formation rate drops and the growth of M_h slows, the calculated value of $M_\star(z)$ may also decrease. In this case, we set $M_\star$ to equal the previous value as we do not expect the stellar mass to decrease in reality (aside from the mass-loss due to stellar evolution, which is already incorporated in the SMHM relation).

Lastly, we update the galactic MZR to the following expression:

$$[\mathrm{Fe/H}](M_\star, z) = 0.3 \log \frac{M_\star}{10^9\,\mathrm{M}_\odot} - 1.0 \log(1+z) - 0.5. \tag{6}$$

This relation is updated from the previous version (Choksi, Gnedin & Li 2018) by recalibrating with the new observational data spanning a broad range of mass ($M_\star = 10^7$–$10^{11}\,\mathrm{M}_\odot$) and redshift ($z = 0.7$–12). The low-$M_\star$ and high-z data are mainly from the recent programmes obtained with the *JWST*, see Appendix A for a detailed description.

Similarly to SMHM, we model the scatter of [Fe/H] and its evolution via the Gaussian process,

$$[\mathrm{Fe/H}]_\mathrm{g}(M_\star, z) \sim \mathcal{GP}\left\{[\mathrm{Fe/H}](M_\star, z),\ K_\mathrm{g}(t_1, t_2)\right\}, \tag{7}$$

where K_g is also a squared exponential function,

$$K_\mathrm{g}(t_1, t_2) = \sigma_\mathrm{g}^2\ \exp\left[-\frac{(t_1 - t_2)^2}{2\tau^2}\right]. \tag{8}$$

The scatter of MZR is characterized by the parameter σ_g, which we set $\sigma_\mathrm{g} = 0.3$ dex to be consistent with the observed scatter (Appendix A). Note that $[\mathrm{Fe/H}]_\mathrm{g}$ refers to the mean metallicity of the galaxy. Clusters formed within this galaxy do not necessarily inherit exactly the same metallicity because of spatial variation and gradients within the ISM. We add an additional Gaussian noise to equation (7) to account for the internal metallicity dispersion:

$$[\mathrm{Fe/H}]_\mathrm{c}(M_\star, z) \sim \mathcal{N}\left\{[\mathrm{Fe/H}]_\mathrm{g}(M_\star, z),\ \sigma_\mathrm{c}^2\right\}, \tag{9}$$

where $[\mathrm{Fe/H}]_\mathrm{c}$ stands for the metallicity of individual clusters formed within a galaxy of stellar mass $M_\star$ at redshift z. The internal scatter is quantified by the parameter σ_c. While σ_g can be measured directly from galaxy surveys, we must calibrate σ_c to match the observations of nearby GC systems. We achieve this by running the model with different σ_c to search for the model realizations that reproduce the total metallicity dispersion observed in the Virgo Cluster Survey (Peng et al. 2006). We find that $\sigma_\mathrm{c} = 0.2$–0.35 dex can match observations within one standard deviation. We adopt the smaller $\sigma_\mathrm{c} = 0.2$ dex in this work.

2.2.2 *Modifications to cluster evolution*

In the cluster evolution step, we follow the trajectories of GC particles taking into account two main processes of mass evolution: stellar mass-loss and tidal disruption. Since most mass-loss due to stellar evolution happens in the first tens of Myr, which is much shorter than the lifetime of a typical GC (∼10 Gyr), we treat stellar evolution as an instantaneous mass-loss at formation:

$$M_\mathrm{i} = \mu_\mathrm{sev}\ M_0, \tag{10}$$

where M_0 is the cluster mass at formation and M_i is the cluster mass after stellar evolution. The remaining mass fraction μ_sev is a function of the stellar initial mass function (IMF) and metallicity. Here, we assume that the IMF is constant for all clusters. Metallicity affects the exact duration of the stellar evolution but not the final remaining fraction (only by ∼1 per cent for $-4 < [\mathrm{Fe/H}] < 1$). Therefore, μ_sev is close to a constant. We follow Gieles & Gnedin (2023) to set $\mu_\mathrm{sev} = 0.55$.

After accounting for stellar evolution, we can consider M_i as the 'initial' mass before tidal disruption. Following CG23, we express the tidal disruption rate of a cluster with mass M as

$$\frac{\mathrm{d}M(t)}{\mathrm{d}t} = -20\ \frac{\mathrm{M}_\odot}{\mathrm{Myr}} \left[\frac{M_\mathrm{i}}{2\times10^5\,\mathrm{M}_\odot}\right]^{1-x} \left[\frac{M(t)}{M_\mathrm{i}}\right]^{1-y} \left[\frac{\Omega_\mathrm{tid}(t)}{150\ \mathrm{Gyr}^{-1}}\right] \tag{11}$$

with the parameters $x = 2/3$ and $y = 4/3$ that match the low-density *N*-body models of Gieles & Gnedin (2023). This is different from CG22, where we used $x = y = 2/3$ (which is appropriate only for very concentrated clusters).

The disruption rate also depends directly on the local tidal field strength (Gieles & Baumgardt 2008). We parametrize the tidal field strength by the angular frequency Ω_tid via the effective eigenvalue $\lambda_{1,\mathrm{e}}$ that takes into account the centrifugal, Euler, and Coriolis forces (Renaud, Gieles & Boily 2011):

$$\Omega_\mathrm{tid}^2 \simeq \lambda_{1,\mathrm{e}} \simeq \lambda_1 - \lambda_3, \tag{12}$$

where λ_1, λ_2, and λ_3 are the eigenvalues of the tidal tensor in descending order. This expression describes the mass-loss rate more accurately than the expression in CG22.

We derive the tidal tensor numerically following the same method as in CG22: we compute the second-order derivatives of the gravitational potential on a 3 × 3 × 3 cubic grid centred on the GC tracer particle, with side length = 300 pc. Although the side length is still too large compared to a typical tidal radius of GCs, we cannot adopt a lower value because of limited spatial resolution of simulations. To distinguish between the true eigenvalues λ and those we derive numerically, we use the notation $\hat{\lambda}$ for the latter. To correct for the systematic deviation of the derived value from the true Ω_tid, we use the third adjustable model parameter κ as a correction:

$$\Omega_\mathrm{tid}^2 = \kappa\,(\hat{\lambda}_1 - \hat{\lambda}_3). \tag{13}$$

3 GALAXY SELECTION

To find GC systems in TNG50 and the LG simulations that can match the observational properties of the GC systems in the MW and M31, we first select two samples of galaxies that have properties similar to the MW and M31, respectively. Next, we apply our model to these galaxies to obtain the model GC systems at present day. Finally, by comparing the model GC systems to the MW and M31 systems specifically, we rank the model GC systems and output the three best analogues to the MW and M31. In the following subsections, we describe the criteria to select the galaxy samples and rank model GC systems.

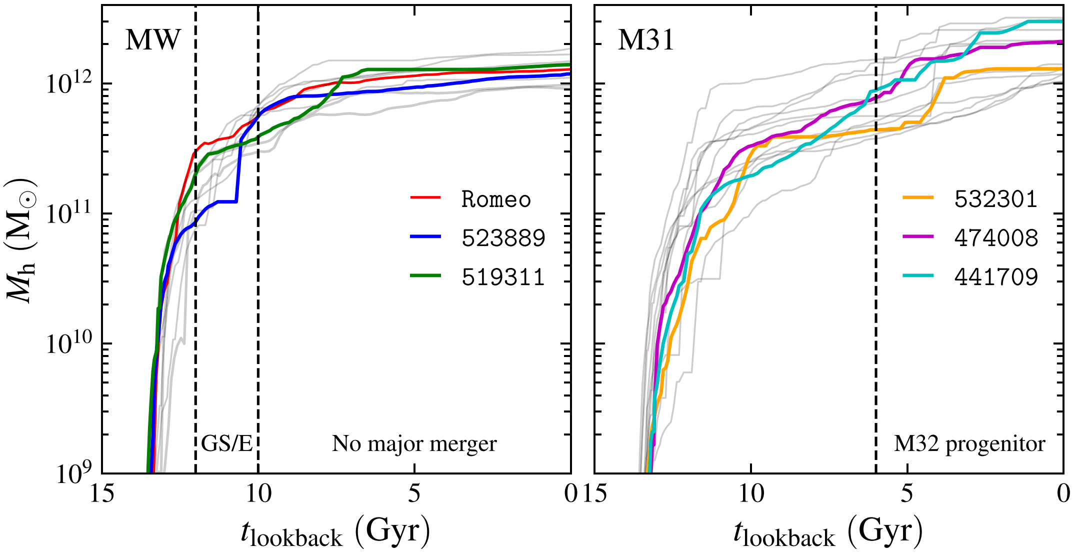

Figure 1. Mass growth histories of the main progenitor branch for the samples of 10 MW analogues (*left*) and 14 M31 analogues (*right*). The dashed vertical lines label the epochs for key assembly events required by our selection criteria, including the GS/E-like merger (10–12 Gyr) and the subsequent quiescent stage ($<$10 Gyr) for MW, and the M32-progenitor-like merger ($<$6 Gyr) for M31. We show each galaxy as a thin grey curve and highlight the three best-matching galaxies in each set with thick coloured curves. We keep the same colour scheme in all plots hereafter when referring to these galaxies.

3.1 Galaxy samples

We apply the following criteria to select MW analogues in TNG50:

(i) Galaxies with total mass[1] between $10^{11.9}$ and $10^{12.3}\ \mathrm{M_\odot}$.

(ii) Galaxies with maximum circular velocity $V_\mathrm{c,max}$ between 210 and 270 km s^{-1}.

(iii) Galaxies with at least one major merger between 10 and 12 Gyr ago to match the accretion of the GS/E satellite.

(iv) Galaxies with no major merger in the last 10 Gyr.

(v) Galaxies formed 25–35 per cent of their present-day stellar mass (calculated from halo mass via the SMHM relation of Behroozi, Wechsler & Conroy 2013b) at $t_\mathrm{lookback} = 10$ Gyr.

Our virial mass range includes the result ($9 \times 10^{11}\ \mathrm{M_\odot}$) of the most recent modelling of the Sagittarius stream (Vasiliev, Belokurov & Erkal 2021), but some deviation at the virial radius is expected because the simulated haloes are not likely to have exactly the same density profile as the MW. We keep the mass variation within a factor of ∼2. We also add the circular velocity criterion to match the inner mass distribution, which may be more relevant for modelling star and cluster formation. The range of $V_\mathrm{c,max}$ is chosen to match the observed rotation curve (Eilers et al. 2019) with a variation ± 30 km s^{-1}. The circular velocity value at Sun's location at 8.5 kpc from the centre is constrained to be within 200–240 km s^{-1}. The stellar mass constraint at 10 Gyr is from Leitner (2012).

We define a major merger as the mass ratio less than 4:1. That is, the total mass of the incoming galaxy is greater than 1/4 of the main galaxy when the incoming galaxy reaches its maximum mass. Our two merger criteria select galaxies that assembled early. To illustrate this, we plot the mass growth histories of the main progenitor branch of these galaxies in the left panel of Fig. 1. In TNG50, there are six galaxies that match the above criteria. Additionally, we include one more galaxy (ID `519311`[2]) from the sample of 'early-spin up' MW-like disc galaxies by Semenov et al. (2023). Our best MW analogue (`523889`) is also included in their sample.

We include the `Louise`, `Romeo`, and `Juliet` galaxies from the LG simulations since they have MW-like mass assembly histories. The small group environment of the LG simulations provides a more realistic background for the satellite accretion. In total, the MW sample has 10 (7 from TNG50 + 3 from LG simulations) galaxy candidates with similar properties to the MW.

For the M31 sample, we follow the criteria below:

(i) Galaxies with total mass between 10^{12} and $10^{12.5}\ \mathrm{M_\odot}$.

(ii) Galaxies with at least one major merger in the recent 6 Gyr to resemble the recent merger with the M32 progenitor as suggested by D'Souza & Bell (2018).

Since M31 contains a much richer system of GCs and its total mass is less certain than in the MW, we allow a wider range of mass. We find 14 analogues of M31 satisfying the criteria. In the right panel of Fig. 1, we show the mass growth histories of the main progenitor branch of these galaxies. The M31 analogues are typically assembled later and have more variable mass growth at late times compared to the MW analogues. This is directly linked to the last selection criterion.

When selecting samples of the MW and M31 analogues, we only focus on the mass assembly history and do not take into account any baryonic properties such as the luminosity, surface brightness, or metallicity (even the stellar mass criterion for selecting MW analogues is derived from M_h using SMHM relation). This minimizes the dependence on the specific prescriptions used in the TNG50 model. However, as we show later, even the differences only in the mass assembly history can lead to a wide range of properties of the GC systems. Some of these realizations correctly match the observed properties of GC systems in the MW and M31, suggesting that GC formation is strongly related to the hierarchical assembly of galaxies.

---

[1] The galaxy total mass refers to the `SubhaloMass` in TNG50 SUBFIND catalogue. This corresponds to the total mass of all member particles bound to this galaxy.

[2] The galaxy ID refers to the `SubhaloID` in TNG50 SUBFIND catalogue.

3.2 Model calibration

Before going into details about how we calibrate the model parameters, we introduce the observational data with which we compare the model systems.

The mass and spatial distributions for the MW GCs are from the third version of the Hilker et al. (2019) catalogue.[3] The metallicities are from the 2010 version of the Harris (1996) catalogue.

The magnitude and position data for the M31 GCs are from the Revised Bologna Catalogue[4] (RBC) of Galleti et al. (2004; version V.5, August 2012). We compute the cluster mass from the V magnitudes using a mass-to-light ratio $M/L_{\rm V} = 1.83\,{\rm M}_\odot/L_{{\rm V},\odot}$ (Baumgardt, Sollima & Hilker 2020). The metallicities are from the LAMOST spectroscopy survey of star clusters in M31 (Chen et al. 2016) using spectral fitting with the models of Vazdekis et al. (2010).

We limit the MW sample to clusters with $M > 10^4\,{\rm M}_\odot$ to avoid the potentially incomplete low-mass clusters. For the M31 sample, we require the clusters to have $M > 10^{4.5}\,{\rm M}_\odot$ and locate outside the central 1 kpc (projected radius) annulus to avoid contamination near the centre of M31. We also apply the same criteria to the model GC systems to consistently compare them with the observations. Since the inclination angle of the M31 galaxy is 77° (Simien et al. 1978), we project the model coordinates for M31 analogues using the same inclination angle with respect to the disc plane. We define the disc plane using the direction of the angular momentum of all stellar particles in the galaxy. We always refer to this inclination angle for projected radius throughout the rest of the work, unless specified otherwise.

Our model has three adjustable parameters: p_2, p_3, and κ. To find the best values for these parameters, we need to calibrate the model to match important observable features. For this purpose, we use a merit function that quantifies the discrepancy between the model and observations:

$$\mathcal{M} \equiv \frac{1}{N}\sum_{i=1}^{N} \mathcal{G}_i, \tag{14}$$

where $\mathcal{G}_i$ is a gauge function of the i-th simulated galaxy. We calculate the merit function separately for the MW and M31 samples. We take into account five features of the GC system: total GC mass, the metallicity distribution, mass function, radial distribution, and velocity dispersion. They correspond to the five independent terms:

$$\mathcal{G} \equiv -\chi^2_M - \chi^2_\sigma + \sum_x \sum_{\rm th} \theta(p_x - p_{\rm th}). \tag{15}$$

The first term is the χ^2 function to evaluate the total mass of surviving clusters,

$$\chi^2_M = \frac{\left(\log M_{\rm GC} - \log M_{\rm GC,obs}\right)^2}{\sigma^2_M}, \tag{16}$$

where $M_{\rm GC}$ represents the total GC mass of the i-th galaxy, and $M_{\rm GC,obs} = 3.7\times10^7\,{\rm M}_\odot$ when compared with the MW. This number comes from the sum of individual MW cluster masses greater than $10^4\,{\rm M}_\odot$. However, since the M31 data are incomplete below $10^{4.5}\,{\rm M}_\odot$, we cannot simply sum up the mass of GCs greater than $10^4\,{\rm M}_\odot$. Instead, we fit the M31 mass function greater than $10^{4.5}\,{\rm M}_\odot$ with a lognormal function, and integrate this function down to $10^4\,{\rm M}_\odot$ to get the total GC mass. This gives a correction $\lesssim$2 per cent and yields $M_{\rm GC,obs} = 9.5\times10^7\,{\rm M}_\odot$ for M31. The denominator is the uncertainty of the number of clusters. We set $\sigma_M = 0.2$ dex to approximate the uncertainty of total GC mass.

The second term is similar to the first term but evaluates the velocity dispersion of GCs,

$$\chi^2_\sigma = \frac{\left(\log \sigma_{\rm GC} - \log \sigma_{\rm GC,obs}\right)^2}{\sigma^2_\sigma}, \tag{17}$$

where $\sigma_{\rm GC}$ represents the 3D velocity dispersion of all surviving GCs in i-th galaxy, and $\sigma_{\rm GC,obs} = 200\ {\rm km\,s^{-1}}$ (calculated from the Hilker et al. 2019 catalogue of Galactic GCs) when compared with to MW. In M31, due to background contamination near the galaxy centre, most velocity measurements are limited to the outer region. For example, Mackey et al. (2019) studied the kinematics of clusters $\gtrsim$25 kpc and found $\sigma_{\rm GC,obs} = 134\ {\rm km\,s^{-1}}$ at $R = 25$ kpc. To properly compare the model M31 analogues with the observations, we only calculate $\sigma_{\rm GC}$ of the 32 GCs closest to 25 kpc. The intrinsic scatter of the dispersion can be approximated as $\sigma_\sigma = 0.3$ dex.

Inside the summation, $p_x \in \{p_M,\ p_Z,\ p_R\}$ stands for the Kolmogorov–Smirnov (KS) p-value for the mass function, metallicity distribution, and radial distribution. The radial distribution refers to the 3D galactocentric radius when calibrating for the MW. It refers to the projected radius for the M31. The Heaviside θ function returns 1 if p_x is greater than a threshold value $p_{\rm th}$, otherwise 0. We employ $p_{\rm th} \in \{0.1, 0.03, 0.01\}$. In other words, $p_x > 0.1$ contributes 3 points to the final $\mathcal{G}$, $0.03 < p_x < 0.1$ contributes 2 points, and $0.01 < p_x < 0.03$ contributes 1 point. For example, the greatest possible $\mathcal{G}$ is 9 if all $p_M, p_Z, p_R > 0.1$, $M_{\rm GC} = M_{\rm GC,obs}$ and $\sigma_{\rm GC} = \sigma_{\rm GC,obs}$. We employ three threshold values instead of one (as in our previous model versions) to provide a finer grading hierarchy with more possible grades contributed by the three KS tests. This finer grading system avoids the discreteness that too many galaxies receive the same grade, and thus allows more nuanced calibration and selection.

For any set of model parameters, we can compute $\mathcal{G}$ of all MW analogues by comparing their properties to the MW. The average of $\mathcal{G}$ yields the merit function for the MW sample. Similarly, we can obtain the merit function for the M31 sample by comparing to M31. We then span the parameter space and select the parameters that maximize the merit functions.

It is worth noting that due to the different numerical resolution and output frequencies in TNG50 and the LG simulations, the best model parameters for the two sets of simulations are not necessarily the same. Therefore, we calibrate the model parameters separately for the two simulation sets. For TNG50, we first compute the merit functions on a grid of model parameters. The best parameters for the MW and M31 differ slightly but systematically: the merit function of the M31 maximizes at greater p_2 and smaller κ compared to that of the MW. However, the merit function flattens near the maximum, allowing us to vary the parameters near the peak values without significantly affecting the model performance. We then select a single set of best parameters between the peak values of both merit functions for the MW and M31 samples: $(p_2, p_3, \kappa) = (18, 0.5\,{\rm Gyr}^{-1}, 1.5)$.

Since the LG simulations only contain MW analogues, we simply compute the merit function for the three MW-like galaxies and select the parameters that maximize the merit function. This gives $(p_2,\ p_3,\ \kappa) = (14,\ 0.5\,{\rm Gyr}^{-1},\ 1.5)$. These values differ from our previous values (CG23, which has $p_3 = 0.7\,{\rm Gyr}^{-1}$) because we have updated the model set-ups, especially the new prescription for cluster evolution, which effectively leads to stronger tidal disruption. To balance the decrease of clusters due to disruption, we need a smaller p_3 to enable more GC formation at later times.

---

[3] https://people.smp.uq.edu.au/HolgerBaumgardt/globular/

[4] http://www.bo.astro.it/M31/

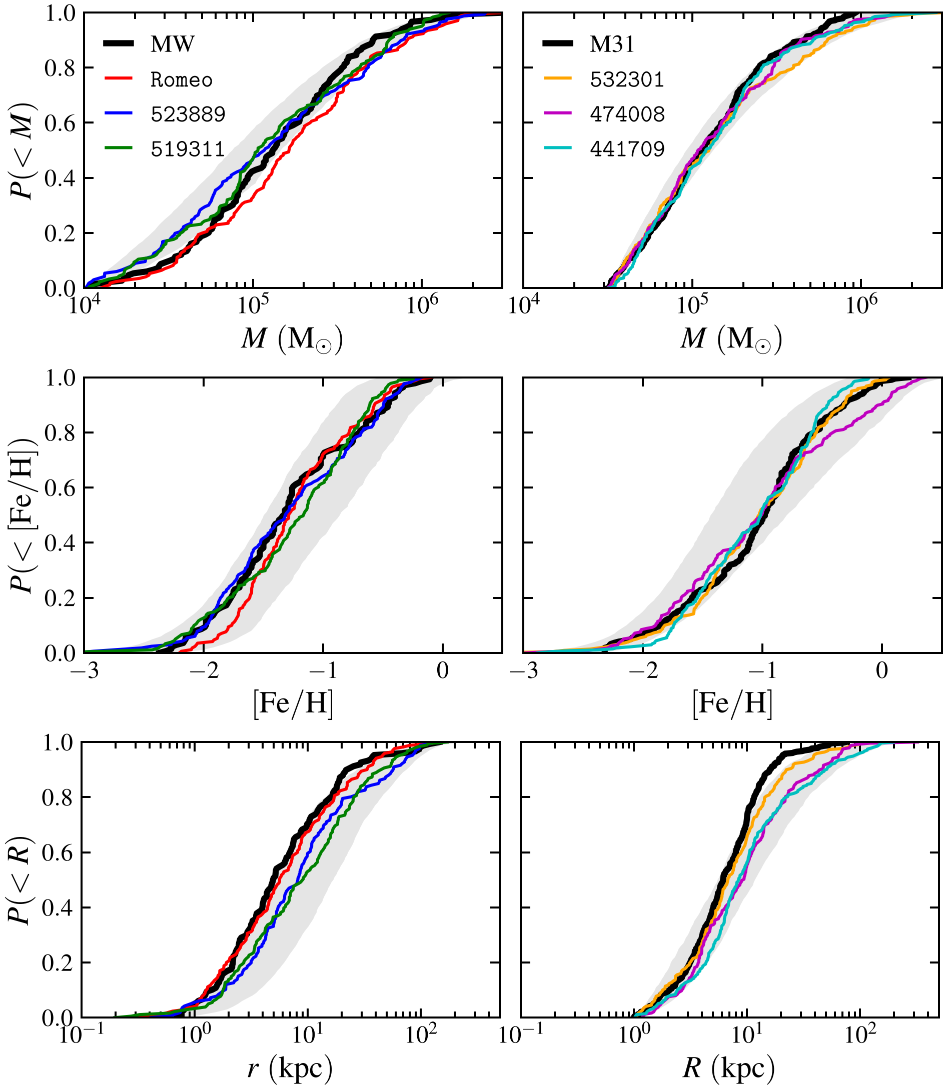

Figure 2. Cumulative distributions of the cluster mass (*upper row*), metallicity (*middle row*), and distance from the centre (projected for M31; *lower row*) for the 50 realizations of 10 MW analogues (*left column*) and 14 M31 analogues (*right column*), represented by the grey shaded regions as the 16–84th percentiles. We highlight the three best-matching galaxies in each panel with the coloured curves. For comparison, we overplot the observational data for the MW and M31 systems as thick black curves.

4 SELECTED MW AND M31 ANALOGUES

We run our model 50 times on each model galaxy with different random seeds to generate an ensemble of 50 realizations. Having multiple random realizations offers two major advantages. First, the model randomness, which includes the scatter in scaling relations and the stochasticity involved in sampling clusters from the ICMF and assigning clusters to simulation particles, can result in GC systems with diverse properties when the same model is rerun for the same galaxy. Analysing the average properties of these 50 realizations reveals their systematic dependence on the model's input and the model itself.

Furthermore, having multiple random realizations enables us to search for the best MW/M31 analogues from significantly larger samples. For each MW/M31 analogue, we calculate $\mathcal{G}$ for the 50 realizations and identify the one with the highest $\mathcal{G}$ as the best representative. In this section, we present the three best-matching representatives from all sample galaxies, considering these galaxies as the best MW/M31 analogues.

4.1 Properties of globular clusters in model galaxies

The three best MW analogues are `Romeo` from the LG simulations and `523889` and `519311` from TNG50. The IDs of the three best M31 analogues are `532301`, `474008`, and `441709`. In Fig. 1, we plot the mass growth history of these analogues as coloured curves. The $\mathcal{G}$ function of the three M31 analogues is generally lower than that of their MW counterparts, indicating that our model is more effective in matching the MW system compared to M31. We note that the two TNG50 MW analogues and three M31 analogues are also included in the TNG50 MW/M31 sample by Pillepich et al. (2023), who selected 198 galaxies based on stellar mass, stellar morphology, and environment.

Next, we study the distributions of important GC properties. In the upper panel of Fig. 2, we plot the cumulative mass function for clusters in each galaxy. The GC systems of MW analogues in TNG50 have slightly more abundant low-mass clusters ($M \lesssim 10^5\,\mathrm{M}_\odot$) compared to the observations. This is likely because the tidal disruption is underestimated for the low-mass clusters due to the limited spatial resolution of TNG50 ∼100 pc. We do not find such a deviation for `Romeo` because the LG simulations have higher spatial resolution ∼60 pc, and because the surviving GCs in `Romeo` are formed mainly *in-situ*, which subjects them to stronger tidal disruption.

On the other hand, the mass functions of all M31 analogues are consistent with the mass function of M31 for $M > 10^{4.5}\,\mathrm{M}_\odot$.

In the middle panel of Fig. 2, we plot the cumulative cluster metallicity distribution. Our model can match the observed metallicity distributions of both MW and M31 very well. The observed MW metallicity distribution completely falls into the 1σ region of all MW sample galaxies. Our model also predicts that the M31 clusters have systematically broader metallicity distribution with higher abundance of metal-rich clusters compared to the MW. This is likely due to the younger GC population in M31 caused by the last major merger.

The abundance of $[\mathrm{Fe/H}] \lesssim -1.5$ clusters in the full M31 sample is slightly higher but still within the 1σ region of the observations. Since $[\mathrm{Fe/H}] \lesssim -1.5$ corresponds to GCs formed in smaller galaxies and at early epochs ($z \gtrsim 3$, see equation 6) such a mismatch may indicate the need for additional constraints on the early assembly history of M31 (our M31 selection criteria have no constraint at $t_{\mathrm{lookback}} > 6$ Gyr).

Finally, we plot the cumulative distribution of distance from the galaxy centre in the lower panel of Fig. 2. Although the radial profile of `Romeo` can closely match the observed profile of the MW GCs with KS p-value > 0.1, our model applied to TNG50 tends to produce more spatially extended GC distribution than the MW system. The half-number radii for the two best-matching MW analogues in TNG are 8–9 kpc, which is close to twice the observational value ∼5 kpc. Similarly, we find that only `532301` can match the radial profile of M31 GCs. The other two best-matching M31 analogues are also spatially more extended than the M31 GC system: the half-number radius of M31 GCs is ∼6 kpc, whereas the two TNG50 galaxies yield 8–9 kpc. The only model setting directly related to the spatial distribution is the initial boundary radius to distribute newly formed clusters. However, as we have verified in Section 2.2, changing this parameter does not improve the final result significantly. As we discuss later in Section 5.2, this mismatch is likely because TNG50 galaxies have more abundant mergers than the galaxies in the LG environment.

4.2 Cluster catalogues

Based on the above comparison, the three MW analogues and three M31 analogues are consistent with all the observable properties of their GC systems. We release the catalogues of model GCs on our model site www.github.com/ognedin/gc_model_mw. The catalogues include the following properties: galaxy ID, cluster formation/disruption/accretion time, cluster mass at formation/current, position, velocity, orbital actions, pericentre/apocentre radii, gravitational potential, metallicity, progenitor galaxy IDs, galaxy total mass at formation/current, and galaxy stellar mass at formation/current. We summarize these properties in Table 1 with a brief description for each entry.

Except for the orbital parameters (orbital actions, pericentre/apocentre radii, and gravitational potential), all other properties are direct outputs of the model. We describe the calculation of orbital parameters in the following section.

4.2.1 *Orbital properties*

We calculate the orbital properties of GCs using the same methods as in CG22. We first employ the AGAMA code (Vasiliev 2019) to model the galactic potential with the multipole expansion and spline approximation methods. Since the potential of MW can be described by spheroids and discs (e.g. McMillan 2017), we model the present-day potentials of TNG50 galaxies with these two components. A largely spheroidal DM potential is modelled by the axisymmetric spherical harmonic expansion,

$$\Phi(r,\theta) = \sum_{l=0}^{l_{\max}} \Phi_l(r)\, Y_l^0(\theta), \tag{18}$$

where Y_l^m are the real-valued spherical harmonics. We only take into account the axisymmetric terms with $m = 0$. The r-dependent coefficient $\Phi_l(r)$ is calculated on 20 grid points evenly spaced in $\log r$. The AGAMA code employs a quintic spline to connect $\Phi_l(r)$ values on grid points.

The discy baryonic (stars + gas) potential is modelled as an axisymmetric form in a cylindrical coordinate system $\Phi(R, z)$. We use a 2D quintic spline to approximate the potential on a 20 × 20 grid evenly spaced in the $\log R$–$\log z$ space.

We find $l_{\max} = 2$ is sufficient to describe the simulation-provided potential to < 2 per cent accuracy. Such an error is so small that we can ignore its influence on the subsequent calculation of the orbital parameters.

Next, we input the present-day positions and velocities of model GCs to AGAMA to calculate the orbital actions and pericentre/apocentre radii. The orbital actions of a closed orbit are defined as

$$J_q \equiv \frac{1}{2\pi} \oint v_q\, \mathrm{d}q, \tag{19}$$

where $q \in \{r, \phi, z\}$ is the radial, azimuthal, and vertical coordinates. J_r and J_z characterize the oscillations in radial and azimuthal directions, whereas J_ϕ equals the z component of angular momentum L_z. In separable potentials,[5] the actions are functions of the IOMs, $J_i = J_i(I_1, I_2, I_3)$, where $I_1 = E$ is the total energy. The AGAMA code computes actions using the Stäckel fudge method (see section 3.2 in Vasiliev 2019). This approach assumes the potential is a Stäckel form (which is not necessarily true) so that we can analytically find the second and third integrals: the second integral in this case is L_z, while the third integral is non-classical. AGAMA speeds up the calculation by using a pre-computed interpolation table of $J_i = J_i(I_1, I_2, I_3)$. Vasiliev (2019) showed that this approximation is accurate to 90–99 per cent even for extremely eccentric orbits.

We calculate the pericentre/apocentre radii by computing the closest/farthest possible distances the cluster can reach with the current energy and angular momentum. They are the two roots of the equation $E = \Phi(r, \theta = 0) + L^2/2r^2$, where $\Phi(r, \theta = 0)$ is the in-plane axisymmetric potential.

We do not calculate orbital properties for `Romeo` because it is from a collisionless simulation with no baryon physics. Since the baryonic matter should become dominant near the galaxy centre

---

[5] A separable potential allows solving the Hamilton–Jacobi equation with separation of variables.

Table 1. Cluster properties in the catalogue.

Key	Unit	Description
`galaxy_id`		SUBFIND (TNG) or CONSISTENT TREE (LG) ID of the current central galaxy.
`t_form`	Gyr	Formation time (lookback) of the cluster.
`t_disrupt`	Gyr	Disruption time (lookback) of the cluster; −1 for surviving clusters.
`t_accrete`	Gyr	Accretion time (lookback) of the cluster's host satellite galaxy; −1 for *in-situ* clusters.
`log_m_form`	$\mathrm{M}_\odot$	$\log_{10}$ of the cluster mass at formation.
`log_m_gc`	$\mathrm{M}_\odot$	$\log_{10}$ of the current cluster mass; −1 for disrupted clusters.
`x`, `y`, `z`	kpc	Galactocentric coordinates, with `z` as the stellar disc (TNG) or total (LG) angular momentum axis.[a]
`vx`, `vy`, `vz`	$\mathrm{km\,s^{-1}}$	Galactocentric velocity components.[b]
`Jr`, `Jz`, `Jp`	$\mathrm{kpc\,km\,s^{-1}}$	Orbital actions J_r, J_Z, and J_ϕ (equation 19); 0 for LG.
`rapo`, `rperi`	kpc	Apocentre and pericentre radii; 0 for LG.
`Ep`	$\mathrm{km^2\,s^{-2}}$	Gravitational potential by multiple expansion (equation 18) and spline approximation; 0 for LG.
`feh`	dex	Iron abundance [Fe/H].
`host_id_form`		SUBLINK or CONSISTENT TREE ID (main leaf) for the host galaxy at cluster formation.
`host_id_accrete`		SUBLINK or CONSISTENT TREE ID (main leaf) for the host satellite galaxy at accretion; −1 for *in-situ* clusters.
`log_mh_form`	$\mathrm{M}_\odot$	$\log_{10}$ of host galaxy total mass at cluster formation.
`log_ms_form`	$\mathrm{M}_\odot$	$\log_{10}$ of host galaxy stellar mass at cluster formation, from Behroozi, Wechsler & Conroy (2013b) SMHM relation.[c]
`log_mh`	$\mathrm{M}_\odot$	$\log_{10}$ of current central total mass.
`log_ms`	$\mathrm{M}_\odot$	$\log_{10}$ of current central stellar mass, from Behroozi, Wechsler & Conroy (2013b) SMHM relation.[c]
`log_ms_k18`	$\mathrm{M}_\odot$	$\log_{10}$ of current central stellar mass, from Kravtsov, Vikhlinin & Meshcheryakov (2018) SMHM relation.

Notes. [a] The galactocentric coordinates are the simulation particle positions relative to the current central galaxy, oriented *face-on*.
[b] The galactocentric velocity components are the simulation particle velocities relative to the current central galaxy, oriented *face-on*.
[c] Scatter from equation (4) is also included in stellar mass.

where most *in-situ* clusters reside, the orbital parameters would be severely miscalculated. Therefore, all entries for the orbital properties in the `Romeo` catalogue are set to zero.

4.2.2 *Progenitor branch ID*

In the simulation merger trees, it is common for the same galaxy at different snapshots to be assigned different IDs. It is thus more appropriate to refer to this galaxy as a 'branch' rather than a 'galaxy' within the terminology of merger trees. To establish a clear and unique identifier for each branch, an effective approach is to label this branch with the ID of the 'main leaf' galaxy. The main leaf galaxy corresponds to the first instance of the galaxy in this branch. This labelling technique guarantees a one-to-one mapping between each branch and its main leaf ID without any ambiguity.

We provide the `host_id_form` and `host_id_accrete` entries to identify the progenitor branches for the *ex-situ* clusters. The `host_id_form` specifies the main leaf ID of the host galaxy at the time of cluster formation. This host galaxy may either merge directly into the central galaxy or first merge with another satellite, which is subsequently accreted on to the central galaxy. In the first case, the host galaxy at cluster formation is the same as the galaxy that ultimately delivered the cluster to the central galaxy. However, in the second scenario, the two are different. To account for this, we introduce `host_id_accrete` as the main leaf ID of the last satellite responsible for bringing the cluster into the central galaxy. We also provide the `t_accrete` entry specifying its accretion time.

In a major merger event, the accreted satellite is likely to have its own merger history prior to joining the central galaxy. Since we expect all GCs from such a satellite to exhibit similar kinematics at present, regardless of their specific progenitors, the `host_id_accrete` entry provides a more straightforward way to query all GCs brought to the main galaxy by that satellite. Conversely, during the early stages of galaxy formation when the main galaxy was not dominant in its local environment, distinguishing all satellites becomes important. In such a scenario, the `host_id_form` entry is more relevant. This entry is also useful for studying the multiple populations of GCs in the same progenitor galaxy, under the assumption that GCs originating from different satellites lead to distinct populations.

4.3 Impact of the last major merger of M31 on the cluster spatial distribution

In our selection, the M31 sample is characterized by a major merger in the last 6 Gyr, whereas the MW analogues did not experience any major merger in the last 10 Gyr. The former correspond to the 'late-assembled' galaxies, while the latter correspond to the 'early-assembled' galaxies. In this section, we investigate the impact of the last major merger on the observable properties of the resulting GC population. Without loss of generality, we study the recent merger event of the best-matching M31 analogue `532301` as an example. All other M31 analogues lead to similar conclusions.

In Fig. 3, we plot the *face-on* (not 77° inclination angle) projection of surviving GCs (defined as $M > 10^{4.5}\,\mathrm{M}_\odot$) before the merger, at the satellite apocentre after the first passage, and at the end of the merger. The merger happens at $t_{\mathrm{lookback}} = 3$–$4$ Gyr, with a merger ratio close to unity. It can be easily noticed that the clusters brought by the recent merger have broader spatial distribution than the *in-situ* clusters at the end of the merger.

To quantify such a difference in the spatial distribution, we split the clusters into three categories: clusters formed in the satellite prior to the merger (or 'satellite' clusters), clusters formed or accreted into the central galaxy prior to this merger ($t_{\mathrm{form}} > 6$ Gyr, or 'central' clusters), and new clusters formed during this merger ($t_{\mathrm{form}} < 6$ Gyr, or 'new' clusters). We plot the evolution of the *face-on* radial profiles for the three categories in Fig. 4. The radial profiles are obtained by kernel density estimation using an Epanechnikov (1969) kernel: $K(x) = 0.75[1 - (x/h)^2]$ with $h = 0.25$ dex. The peak at $R \sim 1$ kpc in the middle panel is contributed by just one cluster captured by the central galaxy during the first passage. Since most satellite clusters

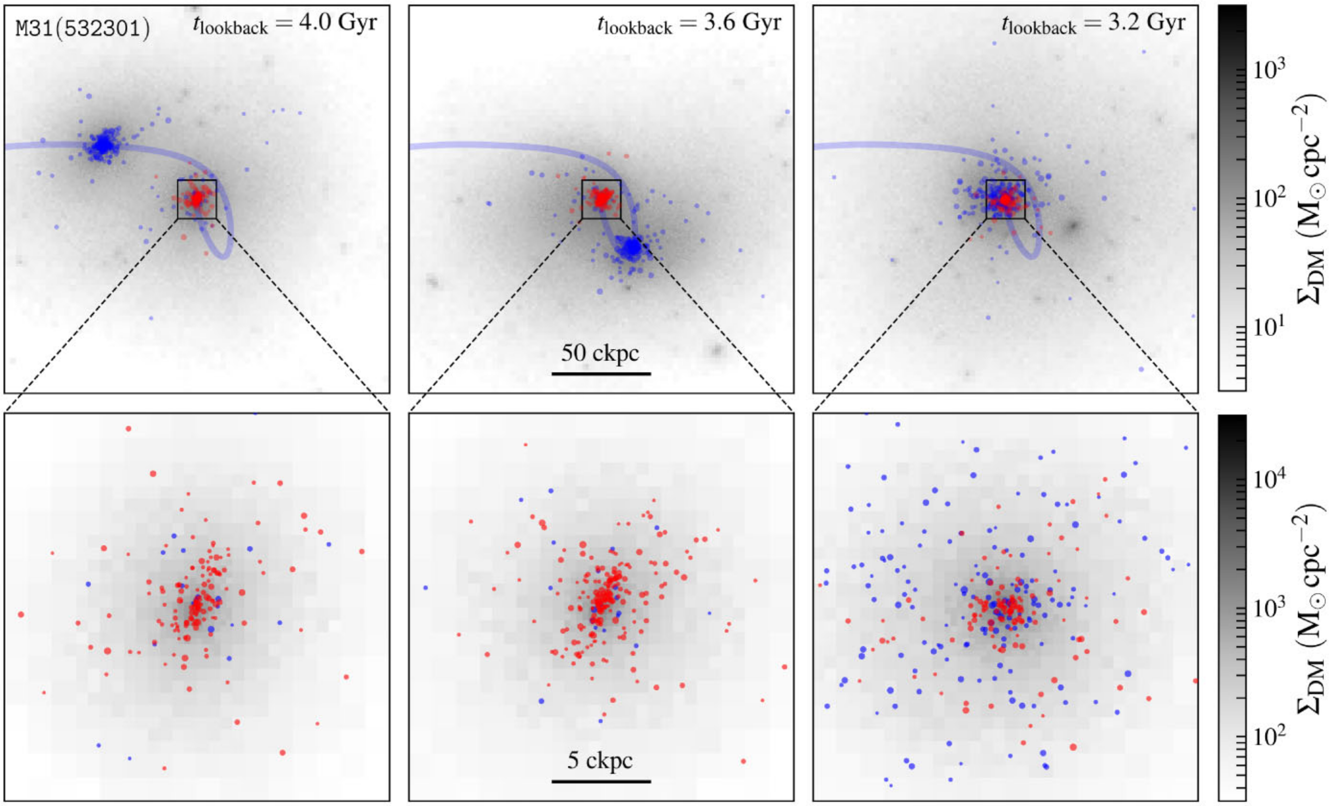

Figure 3. *Face-on* projection plots of GCs during the recent major merger of the best-matching M31 analogue `532301`. From left to right, the three panels correspond to the epochs before the merger, at the satellite apocentre after the first passage, and at the end of the merger, respectively. Each panel in the *top row* corresponds to a (200 comoving kpc)^3 cube centred on the main galaxy. We plot the DM column density as background using the quadtree-based projection code PRJ_PLOTTER (Chen 2023). The *in-situ* and *ex-situ* clusters are shown as red and blue circles, respectively. The size of the circles increases with cluster mass. We also plot the trajectory (smoothed with cubic splines) of the incoming satellite as the blue curve. In the *bottom row*, we zoom-in to the central (20 ckpc)^3 region of each galaxy to show the GC distribution near the galactic centre.

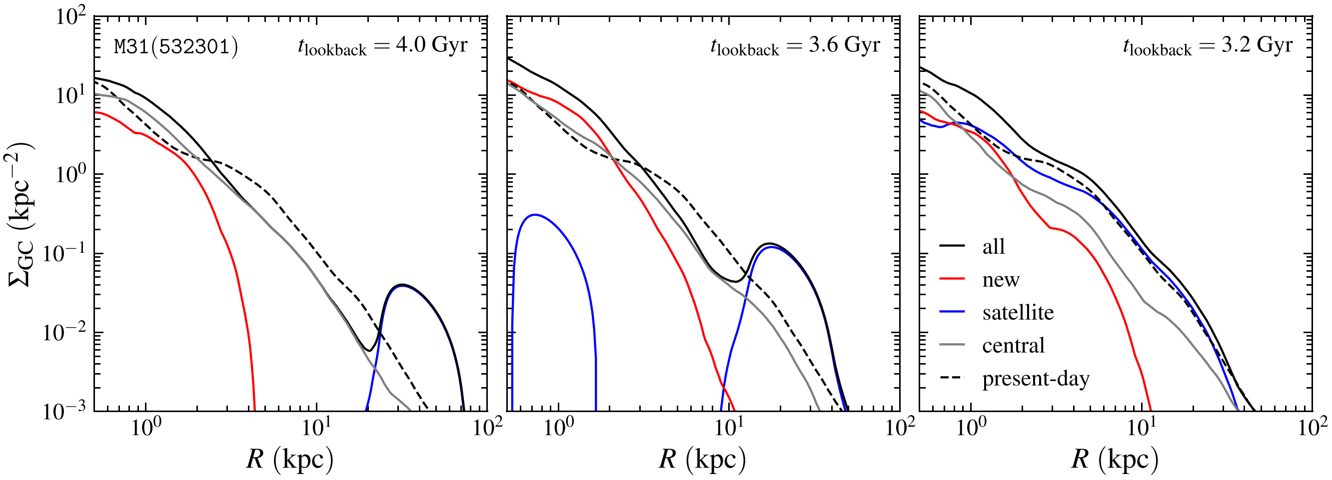

Figure 4. Evolution of the GC number density profile during the recent major merger of `532301`. From left to right, the three panels correspond to the epochs before the merger, at the satellite apocentre after the first passage, and at the end of the merger, respectively. We plot the profiles of all surviving clusters at each epoch as the solid black curve, with individual contributions of clusters brought by the satellite as blue, clusters formed in the central galaxy or accreted prior to this merger as grey, and new clusters formed during this merger as red. For comparison, we plot the present-day profile as the dashed curve.

are still bound to the satellite at this stage, we do not analyse this outlier further.

As the satellite approaches the main galaxy around $t_{\rm lookback} = 3.6$ Gyr, it does not largely alter the spatial distribution of central clusters $\lesssim$10 kpc (approximately the pericentre distance of the satellite's first passage). In contrast, it pulls the clusters in the outer region $\gtrsim$10 kpc outwards, as the satellite is massive enough that it perturbs the potential of the central galaxy significantly in this region.

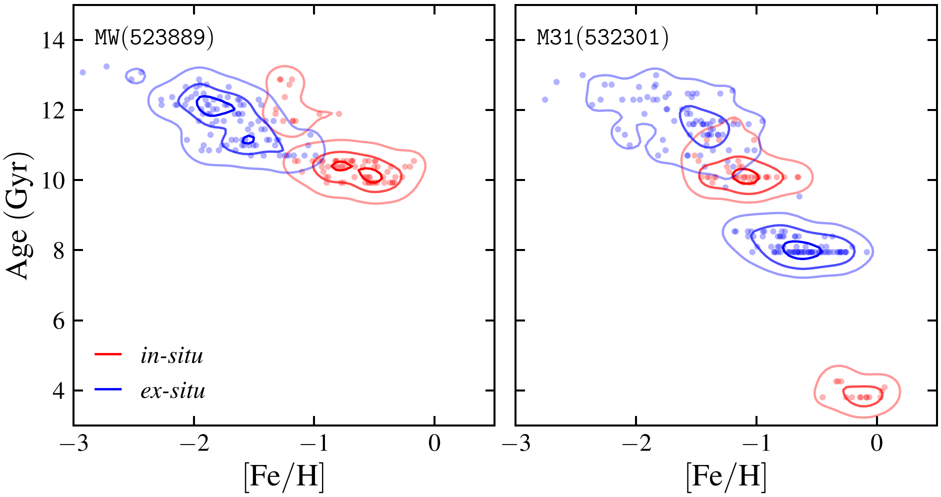

Figure 5. Cluster age–metallicity relation (AMR) for a typical MW analogue (`523889`, *left panel*) and a typical M31 analogue (`532301`, *right panel*). Each dot represents a surviving cluster either formed *in-situ* (red) or *ex-situ* (blue). We calculate the contours using Gaussian KDE with bandwidth =0.1 dex for horizontal axis, and 0.3 Gyr for vertical axis. The contours from dark to light enclose 10, 50, and 90 per cent of total number, respectively.

The formation of about 110 new clusters increases the number density within $\lesssim 3$ kpc. However, strong tidal disruption quickly brings the numbers down by the end of the merger ($t_{\rm lookback} = 3.2$ Gyr). In contrast to such dramatic early disruption, the number of new clusters only drops gradually in the next 3.2 Gyr until the present. Similarly, the early tidal disruption removes around 40 central clusters during the merger, whereas the subsequent evolution only disrupts less than 30 central clusters in the next 3.2 Gyr, leaving ∼90 central clusters surviving until the present.

Despite the change in the normalization, the radial profile of the central clusters does not change significantly during the merger. The merger only raises the effective radius of the central clusters from 2.5 to 3.7 kpc. However, the satellite clusters have a radial profile significantly more outspread than the centrals: the effective radius of the satellite clusters is 6.1 kpc at the end of the merger, raising the overall effective radius to 4.6 kpc.

In the subsequent 3.2 Gyr, the radial profile does not change much, except for the inner part being lowered by tidal disruption, slightly increasing the final effective radius to 5.2 kpc.

By adding a large number of *ex-situ* clusters with distinct radial distribution from the central cluster population, the recent major merger can significantly enlarge the GC effective radius by a factor of 2 in less than 1 Gyr. However, the merger does not significantly alter the spatial distribution of central clusters $\lesssim 10$ kpc. Also, the merger triggers the formation of a young *in-situ* GC population within $\lesssim 3$ kpc. Combining these two factors may explain why many M31 clusters are still located around the disc without being stripped away by the merger. But as we show below in Section 5.1, the last major merger can still heat up the *in-situ* clusters to higher energy.

5 DISCUSSION

5.1 Investigating galaxy assembly with the catalogue

The main goal of constructing this catalogue is to investigate connections between the galaxy assembly history and the properties of GCs. For example, we show the AMRs of surviving clusters for a typical MW analogue (`523889`) and a typical M31 analogue (`532301`) in Fig. 5. GCs in the MW analogue are in general older than the M31 counterparts because of the early assembly history of MW. This leads to almost no GC formation after $t_{\rm lookback} = 10$ Gyr. In contrast, GCs in M31 span a wider range of ages. For `532301` specifically, there are four bursts of GC formation: (1) ∼40 *ex-situ* GCs formed at $t_{\rm lookback} \approx 12$ Gyr, corresponding to active cluster formation in the early Universe when galaxy mergers are frequent; (2) subsequent 'bottom-up' assembly of satellites contributed a significant amount of mass to the main galaxy, leading to the formation of majority of *in-situ* GCs at ≈10 Gyr; (3) similar growth of the largest companion galaxy at ≈8 Gyr, forming ∼130 GCs; (4) the major merger between the main galaxy and its companion boosted *in-situ* GC formation at ≈4 Gyr (as illustrated in Section 4.3). Other M31 analogues also have more extended and discrete GC formation histories, but the timing and order for each GC population may differ among galaxies.

Another noticeable difference is the larger metallicity separation between *in-situ* and *ex-situ* clusters in the MW analogue. Because the assembly of MW is less hierarchical than that of M31, the main progenitor galaxy of MW is more dominant in the GC populations. This enlarges the gap between the metallicity of clusters formed in the main galaxy and the satellites. In contrast, most of *ex-situ* clusters in the M31 analogue formed in the largest companion galaxy which had mass and metallicity comparable to the main galaxy. The metallicity of these *ex-situ* clusters is thus similar to the *in-situ* clusters formed before the merger.

In addition to the AMR, galaxy assembly histories shape the orbits of GCs. For example, Fig. 6 shows the normalized IOM space for the MW and M31 analogues (`523889` and `532301`). The circularity parameter ε is defined as $L_z/L_{\rm circ}(E)$, where $L_{\rm circ}(E)$ is the angular momentum of an in-plane circular orbit with the same energy E as the cluster. The normalized energy e is defined as $E/|E_0|$, where E_0 is the gravitational potential at the galaxy centre. We have $\varepsilon \in [-1, 1]$ and $e \in [-1, 0)$. Perfectly circular and in-plane orbits have $\varepsilon = 1$ (prograde) or $\varepsilon = -1$ (retrograde), while purely radial or polar orbits have $\varepsilon = 0$.

We find that the *ex-situ* populations in both MW and M31 analogues have similar distribution around $(\varepsilon, e) = (0, -0.3)$. However, the *in-situ* populations have quite different distributions: in the MW *in-situ* GCs are located near $(\varepsilon, e) = (0.5, -0.6)$, indicating significantly lower energy and more discy orbits. The *in-situ* GCs

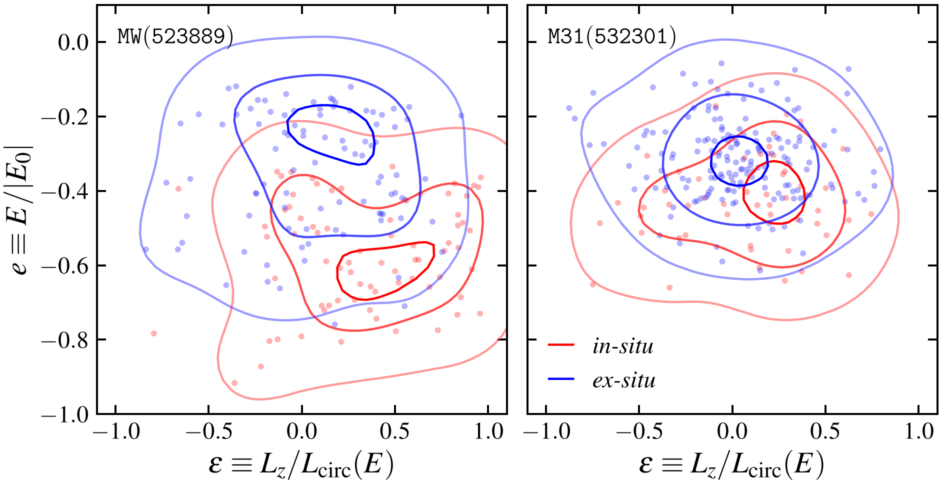

Figure 6. Normalized IOM space for clusters in a typical MW analogue (`523889`, *left panel*) and a typical M31 analogue (`532301`, *right panel*). Each dot represents a surviving cluster either formed *in-situ* (red) or *ex-situ* (blue). We calculate the contours using Gaussian KDE with bandwidth =0.2 for horizontal axis, and 0.1 for vertical axis. The contours from dark to light enclose 10, 50, and 90 per cent of total number, respectively.

of the M31, however, have a more similar distribution to the *ex-situ* population and centre around $(\varepsilon, e) = (0.2, -0.4)$. Considering the significant scatter $\Delta\varepsilon = 0.3$–0.4 and $\Delta e = 0.1$–0.15 in each population, the distributions of *in-situ* and *ex-situ* clusters in the M31 analogue largely overlap in the normalized IOM space. This is likely due to the recent major merger which heated up the *in-situ* clusters to higher energy and moved some of them away from the disc.

Since GCs formed in different progenitor galaxies are located in distinct regions within the chemical and kinematic property space, numerous attempts (e.g. Massari, Koppelman & Helmi 2019; Malhan et al. 2022; Belokurov & Kravtsov 2023) have been made to identify GCs of various origins by classifying isolated clusters/groups in the AMR, IOM, and chemistry spaces. However, these studies often yield conflicting results. Our catalogue can serve as a tool to assess the accuracy of the classification algorithms employed in these studies since our model provides 'true' labels from the simulations (as indicated by the `host_id_from` and `host_id_accrete` entries in Table 1). Furthermore, our catalogue offers a unique opportunity to investigate the spread of properties among GCs originating from the same progenitor and to explore how much GCs retain their original kinematic signatures during galaxy assembly. Finding the answers to these questions is crucial for characterizing the accuracy of current classification schemes.

5.2 Why are TNG50 GC systems more extended?

As shown in Section 4.1, the GC systems in TNG50 galaxies tend to have more radially extended distribution than the observations: the half-number radii of the two best-matching MW analogues in TNG50 are around 1.7 times the observational value ∼5 kpc. In fact, all seven TNG50 galaxies in the MW sample have greater half-number radii, with a median value ∼12 kpc. We find a similar trend for the M31 sample. In contrast, all three MW-like galaxies in the LG simulations can match the radial profile of MW GCs with half-number radii ∼5 kpc.

These extended distributions of the TNG50 systems may be a general outcome of the large-scale environment probed by the TNG50 simulation. For example, Semenov et al. (2023) showed that approximately 90 per cent of TNG50 galaxies developed their disc later than the MW. This difference arises from TNG50 lacking a quiescent, merger-free environment, which is necessary for the survival of an early-formed disc. While we limit our MW sample galaxies to those without major mergers with a mass ratio less than 4:1 in the last 10 Gyr, we do not impose constraints on higher merger ratios (∼10:1) which can still play a significant role in structure formation.

Additional evidence indicating that MW may be less merger-dominated than typical TNG50 galaxies lies in the high *in-situ* fraction of MW GCs. Observational classification studies suggest that the *in-situ* fraction may range from about 40 per cent (Massari, Koppelman & Helmi 2019) to 56 per cent (Malhan et al. 2022) and up to 66 per cent (Belokurov & Kravtsov 2023). In contrast, the TNG50 galaxies in the MW sample typically have *in-situ* fractions between 20 per cent and 50 per cent, once again suggesting that even the selected MW sample galaxies may be too merger-dominated compared to the real MW.

The relatively low fraction of *in-situ* clusters in TNG50 leads to two discrepancies between the model and observed data. First, the low *in-situ* fraction accounts for the too extended distribution of TNG GCs, as shown in the lower-left panel of Fig. 2. This is because the median radius of *ex-situ* clusters is approximately four times larger than that of the *in-situ* clusters (see CG22). Secondly, the low *in-situ* fraction in TNG50 also explains why the AMR of TNG GCs lacks a sequence of old, metal-poor *in-situ* clusters (as shown in the left panel of Fig. 5), which is a robust feature of the MW assembly history. To align the simulations more closely with observed data, a more realistic simulation of the MW assembly is needed, particularly to match the observed AMR of the *in-situ* clusters.

6 SUMMARY

We introduce a catalogue of model star clusters in the MW and M31 analogues, drawn from the TNG50 and LG simulations. By applying criteria based on galaxy virial mass, circular velocity profile, and defining assembly events, we select a sample comprising 10 MW analogues and 14 M31 analogues. We then apply an analytical model of GC formation and evolution to the merger trees and particle outputs of these simulated galaxies. This enables us to obtain key observables such as the mass, age, metallicity, positions, and velocities of the surviving cluster population. We calibrate the model parameters to optimize agreement with the observed total cluster mass, mass function, metallicity distribution, radial profile, and velocity dispersion. From these model results, we select the three best-matching GC systems each for MW and M31. One of the best-matching MW analogues is from the LG simulations; the other five are from TNG50 (see Fig. 1 for the mass growth histories of the six galaxies).

The GC systems in the three best-matching MW analogues successfully reproduce the observed mass function and metallicity distribution (Fig. 2) after model calibration. The galaxy from the LG simulations also matches the observed radial distribution, yielding a KS p-value > 0.1. However, the radial profiles of the other two GC systems from TNG50 are more extended, with effective radii approximately 1.7 times the observed value. Likewise, the three best-matching M31 analogues have mass and metallicity distributions that agree with observational data. However, only one M31 analogue reproduces the observed radial distribution, while the remaining two exhibit more extended profiles, with effective radii $\sim$1.5 times the observed value. This discrepancy may arise from the higher rate of mergers in TNG50 compared to the LG environment. Such mergers typically introduce a significant population of *ex-situ* clusters, which tend to be located $\sim$4 times father away from the galaxy centre than their *in-situ* counterparts.

An in-depth study of the last major merger in M31 analogues also supports the above conclusion. We find that even a 1:1 merger has a limited impact on the spatial distribution of *in-situ* clusters within $\lesssim$10 kpc (Fig. 4). However, such a merger can notably enlarge the overall GC effective radius by a factor of 2 in a short period $\lesssim$1 Gyr, primarily by bringing in a large number of *ex-situ* clusters. Moreover, the merger triggers the formation of a new population of young *in-situ* clusters near the galactic centre. This population, in addition to older clusters brought in by the merger, forms distinct groups in the AMR (Fig. 5). These groups have lower age and higher metallicity compared to the central clusters formed prior to the merger, resulting in an AMR spanning a wide range in both age and metallicity.

Furthermore, we show that the galaxy assembly history significantly influences GC kinematics. The MW analogues have a quiescent formation history over the past 10 Gyr. Such a long period preserves the original location of clusters in the IOM space (Fig. 6). In this space, *in-situ* clusters tend to have lower energy and angular momentum L_z closer to $L_{\rm circ}(E)$, whereas *ex-situ* clusters generally have higher energy and $L_z \approx 0$. This distinction is particularly helpful when using the GC properties to decode the merger history of their host galaxy. Conversely, the last major merger of M31 elevates the *in-situ* populations to higher energy, making the *in-situ* and *ex-situ* populations less distinguishable in the IOM space.

We make our catalogue available in a machine-readable format. Along with all the variables directly output by the model, the catalogue also includes several orbital parameters: the gravitational potential, actions, and pericentre/apocentre radii. A comprehensive list of keys and descriptions for these variables is provided in Table 1. The catalogue can be accessed at www.github.com/ognedin/gc_model_mw.

ACKNOWLEDGEMENTS

We thank Eric Bell, Leandro Beraldo e Silva, Behzad Tahmasebzadeh, Monica Valluri, Gillen Brown, and Xi Meng for insightful discussions. This research was mainly conducted with the PYTHON programming language, employing the following packages: NUMPY (Harris et al. 2020), MATPLOTLIB (Hunter 2007), SCIPY (Virtanen et al. 2020), AGAMA (Vasiliev 2019), YT (Turk et al. 2011), ILLUSTRIS_PYTHON (Nelson et al. 2021), and PRJ_PLOTTER (Chen 2023). OYG and YC were supported in part by the U.S. National Science Foundation through grant AST-1909063 and by National Aeronautics and Space Administration through contract NAS5-26555 for Space Telescope Science Institute programme HST-AR-16614.

DATA AVAILABILITY

The catalogues of model clusters in the MW and M31 galaxies are available at www.github.com/ognedin/gc_model_mw. Other data that support the findings of this study are available from the corresponding author, upon reasonable request. The TNG50 simulation data are publicly available at www.tng-project.org/data.

REFERENCES

Asplund M., Amarsi A. M., Grevesse N., 2021, A&A, 653, A141
Baumgardt H., Sollima A., Hilker M., 2020, Publ. Astron. Soc. Aust., 37, e046
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57
Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689
Belokurov V., Kravtsov A., 2023, preprint (arXiv:astro-ph/2309.15902)
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Birkin J. E. et al., 2023, preprint (arXiv:astro-ph/2307.10412)
Boyett K. et al., 2023, preprint (arXiv:astro-ph/2303.00306)
Chen Y., 2023, ybillchen/prj_plotter: v0.1.2 (v0.1.2). Zenodo. Available at: https://doi.org/10.5281/zenodo.8165883
Chen Y., Gnedin O. Y., 2022, MNRAS, 514, 4736 (CG22)
Chen Y., Gnedin O. Y., 2023, MNRAS, 522, 5638 (CG23)
Chen B. et al., 2016, AJ, 152, 45
Choksi N., Gnedin O. Y., 2019a, MNRAS, 486, 331
Choksi N., Gnedin O. Y., 2019b, MNRAS, 488, 5409
Choksi N., Gnedin O. Y., Li H., 2018, MNRAS, 480, 2343
Crain R. A. et al., 2015, MNRAS, 450, 1937
Creasey P., Sales L. V., Peng E. W., Sameie O., 2019, MNRAS, 482, 219
Curti M., Mannucci F., Cresci G., Maiolino R., 2020, MNRAS, 491, 944
Curti M. et al., 2022, MNRAS, 518, 425
Curti M. et al., 2023, preprint (arXiv:astro-ph/2304.08516)
Deason A. J., Belokurov V., Weisz D. R., 2015, MNRAS, 448, L77
Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, ApJ, 862, L1
D'Souza R., Bell E. F., 2018, Nat. Astron., 2, 737
Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, ApJ, 871, 120
Epanechnikov V. A., 1969, Theory Probab. Appl., 14, 153
Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 646, 107
Faisst A. L. et al., 2016, ApJ, 822, 29
Gaia Collaboration, 2016, A&A, 595, A2
Gaia Collaboration, 2018, A&A, 616, A1
Gaia Collaboration, 2023, A&A, 674, A1
Galleti S., Federici L., Bellazzini M., Fusi Pecci F., Macrina S., 2004, A&A, 416, 917
Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., Lee K., 2014, MNRAS, 438, 2578
Genzel R. et al., 2015, ApJ, 800, 20
Gieles M., Baumgardt H., 2008, MNRAS, 389, L28
Gieles M., Gnedin O. Y., 2023, MNRAS, 522, 5340
Halbesma T. L. R., Grand R. J. J., Gómez F. A., Marinacci F., Pakmor R., Trick W. H., Busch P., White S. D. M., 2020, MNRAS, 496, 638
Hammer F., Puech M., Chemin L., Flores H., Lehnert M. D., 2007, ApJ, 662, 322
Harris W. E., 1996, AJ, 112, 1487
Harris C. R. et al., 2020, Nature, 585, 357
Heintz K. E. et al., 2023, Nat. Astron., in press

Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Hilker M., Baumgardt H., Sollima A., Bellini A., 2019, Proc. IAU Symp. 372, The Era of Multi-Messenger Solar Physics. Cambridge Univ. Press, Cambridge, p. 451
Hsiao T. Y.-Y. et al., 2023, preprint (arXiv:astro-ph/2305.03042)
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jung I. et al., 2023, preprint (arXiv:astro-ph/2304.05385)
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kravtsov A. V., Gnedin O. Y., 2005, ApJ, 623, 650
Kravtsov A., Manwadkar V., 2022, MNRAS, 514, 2667
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Kruijssen J. M. D., 2009, A&A, 507, 1409
Kruijssen J. M. D., Lamers H. J. G. L. M., 2008, A&A, 490, 151
Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Icke V., 2011, MNRAS, 414, 1339
Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019, MNRAS, 486, 3134
Leitner S. N., 2012, ApJ, 745, 149
Lewis Z. J. et al., 2023, preprint (arXiv:astro-ph/2304.12343)
Li H., Gnedin O. Y., 2014, ApJ, 796, 10
Li H., Gnedin O. Y., 2019, MNRAS, 486, 4030
Li H., Gnedin O. Y., Gnedin N. Y., Meng X., Semenov V. A., Kravtsov A. V., 2017, ApJ, 834, 69
Li H., Gnedin O. Y., Gnedin N. Y., 2018, ApJ, 861, 107
Li M. et al., 2023, ApJ, 955, L18
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Mackey D. et al., 2019, Nature, 574, 69
McMillan P. J., 2017, MNRAS, 465, 76
Maiolino R. et al., 2008, A&A, 488, 463
Majewski S. R. et al., 2017, AJ, 154, 94
Malhan K. et al., 2022, ApJ, 926, 107
Mannucci F. et al., 2009, MNRAS, 398, 1915
Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4
Matthee J., Mackenzie R., Simcoe R. A., Kashino D., Lilly S. J., Bordoloi R., Eilers A.-C., 2023, ApJ, 950, 67
Muratov A. L., Gnedin O. Y., 2010, ApJ, 718, 1266
Mészáros S. et al., 2020, MNRAS, 492, 1641
Mészáros S. et al., 2021, MNRAS, 505, 1645
Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, preprint (arXiv:astro-ph/2301.12825)
Nelson D. et al., 2019, MNRAS, 490, 3234
Nelson D. et al., 2021, preprint (arXiv:astro-ph/1812.05609)
Pancino E. et al., 2017, A&A, 601, A112
Peng E. W. et al., 2006, ApJ, 639, 95
Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, MNRAS, 475, 4309
Pfeffer J. L., Trujillo-Gomez S., Kruijssen J. M. D., Crain R. A., Hughes M. E., Reina-Campos M., Bastian N., 2020, MNRAS, 499, 4863
Phipps F., Khochfar S., Varri A. L., Dalla Vecchia C., 2020, A&A, 641, A132
Pillepich A. et al., 2018, MNRAS, 473, 4077
Pillepich A. et al., 2019, MNRAS, 490, 3196
Pillepich A. et al., 2023, preprint (arXiv:astro-ph/2303.16217)
Planck Collaboration XIII, 2016, A&A, 594, A13
Reina-Campos M., Hughes M. E., Kruijssen J. M. D., Pfeffer J. L., Bastian N., Crain R. A., Koch A., Grebel E. K., 2020, MNRAS, 493, 3422
Reina-Campos M., Keller B. W., Kruijssen J. M. D., Gensior J., Trujillo-Gomez S., Jeffreson S. M. R., Pfeffer J. L., Sills A., 2022, MNRAS, 517, 3144
Renaud F., Gieles M., Boily C. M., 2011, MNRAS, 418, 759
Renaud F., Agertz O., Gieles M., 2017, MNRAS, 465, 3622
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Sanders R. L. et al., 2021, ApJ, 914, 19
Sanders R. L. et al., 2023, ApJ, 942, 24
Savaglio S. et al., 2005, ApJ, 635, 260
Schaye J. et al., 2015, MNRAS, 446, 521
Schechter P., 1976, ApJ, 203, 297
Semenov V. A., Conroy C., Chandra V., Hernquist L., Nelson D., 2023, preprint (arXiv:astro-ph/2306.09398)
Simien F., Athanassoula E., Pellet A., Monnet G., Maucherat A., Courtès G., 1978, A&A, 67, 73
Springel V., 2010, MNRAS, 401, 791
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Strom A. L., Rudie G. C., Steidel C. C., Trainor R. F., 2022, ApJ, 925, 116
Tacconi L. J. et al., 2018, ApJ, 853, 179
Topping M. W. et al., 2021, MNRAS, 506, 1237
Tremonti C. A. et al., 2004, ApJ, 613, 898
Trujillo-Gomez S., Kruijssen J. M. D, Reina-Campos M., Pfeffer J. L., Keller B. W., Crain R. A., Bastian N., Hughes M. E., 2021, MNRAS, 503, 31
Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, ApJS, 192, 9
Valenzuela L. M., Moster B. P., Remus R.-S., O'Leary J. A., Burkert A., 2021, MNRAS, 505, 5815
Valenzuela L. M., Remus R.-S., McKenzie M., Forbes D. A., 2023, preprint (arXiv:astro-ph/2309.11545)
Vasiliev E., 2019, MNRAS, 482, 1525
Vasiliev E., Belokurov V., Erkal D., 2021, MNRAS, 501, 2279
Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Wang T.-M. et al., 2022, A&A, 660, A142
Williams H. et al., 2023, Science, 380, 416
Yu S. et al., 2023, MNRAS, 523, 6220
Zahid H. J., Kewley L. J., Bresolin F., 2011, ApJ, 730, 137
Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, ApJ, 791, 130

APPENDIX: THE REDSHIFT-DEPENDENT MASS–METALLICITY RELATION

We calculate the metallicity of GCs using a redshift-dependent MZR. Before introducing the MZR in depth, we first introduce the definition of metallicity. Traditionally, metallicity is expressed in the logarithmic abundance ratio log(X/Y), which is short for $\log(N_{\rm X}/N_{\rm Y})$. A common practice is to normalize this value with a fixed offset of 12 (e.g. $\log\epsilon_{\rm X}\equiv 12+\log({\rm X/H})$), or with the solar value: $[{\rm X/Y}]\equiv\log({\rm X/Y})-\log({\rm X/Y})_\odot$. The former is often used to describe gas-phase metallicity, while the latter is more frequently used for stellar metallicity. In this work, we use the solar reference values from Asplund, Amarsi & Grevesse (2021):

$$\begin{aligned}12+\log({\rm O/H})_\odot &= 8.69\\ 12+\log({\rm Fe/H})_\odot &= 7.46.\end{aligned} \tag{A1}$$

The MZR used to be poorly constrained for dwarf galaxies ($M_\star \lesssim 10^8\,{\rm M}_\odot$) and in the early Universe ($z\gtrsim 6$). The launch of *JWST* has significantly extended the observed galaxy metallicity in both directions, enabling us to recalibrate the MZR. We use the following factorized power-law relation:

$$[{\rm Fe/H}] = \alpha_M \log\frac{M_\star}{M_0} - \alpha_z\log(1+z) + [{\rm Fe/H}]_0, \tag{A2}$$

where α_M and α_z are the logarithmic slopes of the stellar mass and redshift dependence, respectively. $[{\rm Fe/H}]_0$ is the metallicity at some characteristic mass scale M_0 at $z=0$. We take $M_0=10^9\,{\rm M}_\odot$ to represent the typical mass of host galaxies at the formation time of the majority of surviving clusters. Note that such a simple power-law scaling is not valid for most massive galaxies ($M_\star\gtrsim 10^{11}\,{\rm M}_\odot$), where metallicity saturates at a slight supersolar value (see e.g. Tremonti et al. 2004). Therefore, we follow Choksi, Gnedin & Li (2018) to

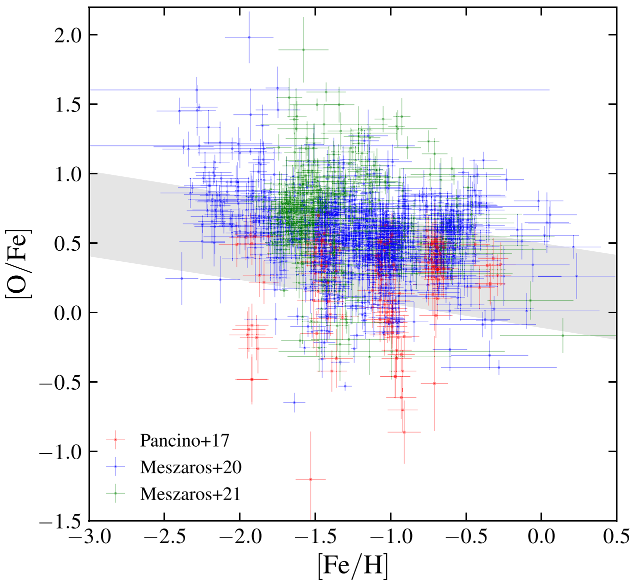

Figure A1. [O/Fe]–[Fe/H] relation of stars in nearby GCs. We plot the GC stars from the Gaia-ESO survey as red symbols (Pancino et al. 2017) and the SDSS-IV APOGEE-2 survey as blue (excluding ω Cen, Mészáros et al. 2020) and green symbols (ω Cen only, Mészáros et al. 2021) with errorbars representing the measurement uncertainty. We show the linear fit of the combined data as the grey shaded region with a 0.3 dex scatter.

cap the maximum value $[\mathrm{Fe/H}]_{\mathrm{max}} = +0.3$ dex to avoid unreasonably large [Fe/H] of GCs in massive galaxies.

We assume that the metallicity of GCs is directly inherited from the surrounding gas at formation. Since most measurements of the gas-phase metallicity in high-redshift galaxies provide the oxygen abundance rather than the iron abundance, we need to convert the oxygen abundance to [Fe/H] using the observed [O/Fe]–[Fe/H] relation. Combining the observational data of nearby GCs from the Gaia-ESO survey (Pancino et al. 2017) and the Sloan Digital Sky Survey (SDSS-IV) APOGEE-2 survey (Mészáros et al. 2020, 2021), we find an anticorrelation between [O/Fe] and [Fe/H], see Fig. A1. This relation can be quantified by a linear relation: [O/Fe] = 0.37–0.17[Fe/H], with a 0.3 dex scatter, in the range of $-2 \lesssim$ [Fe/H] $\lesssim 0$. We emphasize that this relation is only valid for stars in GCs as these works only in analysed stellar abundances in nearby GCs.

Given log(O/Fe), we can now convert the measured oxygen abundance to the gas-phase iron abundance:

$$\log(\mathrm{Fe/H}) = \log(\mathrm{O/H}) - \log(\mathrm{O/Fe}). \tag{A3}$$

We then rewrite this equation in terms of [Fe/H] and [O/Fe]:

$$[\mathrm{Fe/H}] = \log(\mathrm{O/H}) - [\mathrm{O/Fe}] - \log(\mathrm{O/H})_{\odot}. \tag{A4}$$

Combining equations (A1) and (A4), and the linear anticorrelation between [O/Fe] and [Fe/H], we obtain

$$[\mathrm{Fe/H}] = 1.20\,(12 + \log(\mathrm{O/H})) - 10.92. \tag{A5}$$

Although the MZR is well constrained for nearby ($z \sim 0$) galaxies across a wide range of stellar mass, we must calibrate the parameters of equation (A2) specifically at earlier times when GC formation is most active. Therefore, we incorporate observational gas-phase MZR from $z \sim 0.7$ up to $z \sim 12$ from multiple surveys as described below.

Maiolino et al. (2008) combined MZR measurements from earlier works (Tremonti et al. 2004; Savaglio et al. 2005; Erb et al. 2006; Kewley & Ellison 2008) at $z \sim 0$–3 and the Assessing the Mass-Abundance redshift[-Z] Evolution programme at $z \sim 3.5$. These authors used several emission line ratios as metallicity diagnostics to derive the gas-phase metallicity. Note that their data at $z \sim 3.5$ are primarily star forming Lyman-break galaxies, which may introduce a systematic bias in the derived MZR.

Zahid et al. (2014) put together observational data from the SDSS, Smithsonian Hectospec Lensing Survey, Deep Extragalactic Evolutionary Probe 2, and the Fiber Multi-Object Spectrograph-Cosmic Evolution Survey (FMOS-COSMOS) spanning a redshift range $0 < z < 1.7$. They obtained the gas-phase oxygen abundance via metallicity–line ratio scaling.

Lewis et al. (2023) measured oxygen abundances for 145 galaxies at $z \sim 0.7$. They estimated the metallicity via a Bayesian framework using emission lines from the Large Early Galaxy Astrophysics Census survey.

Strom et al. (2022) calculated the gas-phase metallicity of a subsample of the Keck Baryonic Structure Survey galaxies at $z \sim 2$–3. They tested various methods to determine the oxygen abundance and found noticeable discrepancy. Here, we apply the metallicity indicated by their photoionization model.

Li et al. (2023) measured the metallicity of 55 galaxies in the Abell 2744 (by the GLASS *JWST* Early Release Science programme) and SMACS J0723−3732 (by the *JWST* Early Release Observations programme) galaxy cluster fields. These galaxies span a redshift range $z = 2$–3. They also employed metallicity–line ratio scaling to derive the gas-phase metallicity.

Sanders et al. (2021) analysed the metallicity for about 300 galaxies at $z \sim 2.3$ and 150 galaxies at $z \sim 3.3$ in the MOSFIRE Deep Evolution Field (MOSDEF) survey. They obtained the oxygen abundance first by fitting various line ratios as a function of log(O/H), and then searching for the best-fitting oxygen abundance using the χ^2 minimization technique.

Sanders et al. (2023) incorporated the earlier data (Zahid, Kewley & Bresolin 2011; Curti et al. 2020; Sanders et al. 2021; Topping et al. 2021) and fitted a functional form to match the MZR from $z = 0.08$ to 3.3.

Curti et al. (2023) calculated the gas-phase metallicity of 66 galaxies at $z = 3$–10. These galaxies are observed with the *JWST*/NIRSpec as part of the *JWST* Advanced Deep Extragalactic Survey. These authors also used line ratio fitting to obtain metallicity via the minimal likelihood technique. They presented their results in combination with Nakajima et al. (2023).

Nakajima et al. (2023) focused on 135 galaxies at $z = 4$–10 from the *JWST*/NIRSpec data by the ERO, GLASS, and Cosmic Evolution Early Release Science (CEERS) programmes. They used a combination of direct T_{e} (electron temperature) method and metallicity–line ratio scaling to estimate log(O/H).

Faisst et al. (2016) measured the metallicity of 224 galaxies at $z \sim 5$ from COSMOS. They estimated the gas-phase oxygen abundance using the ultraviolet equivalent width–metallicity correlation. Their sample includes Ly α emitting galaxies and galaxies without Ly α emission. Here, we only consider the average MZR of both categories.

Matthee et al. (2023) analysed 117 galaxies at $z = 5.33$–6.93 from the Emission-line galaxies and the Intergalactic Gas in the Epoch of Reionization (EIGER) programme using *JWST*/NIRSpec. These authors employed photoionization modelling for spectral energy distribution fitting to calculate the gas-phase metallicity along with the stellar mass and other model parameters.

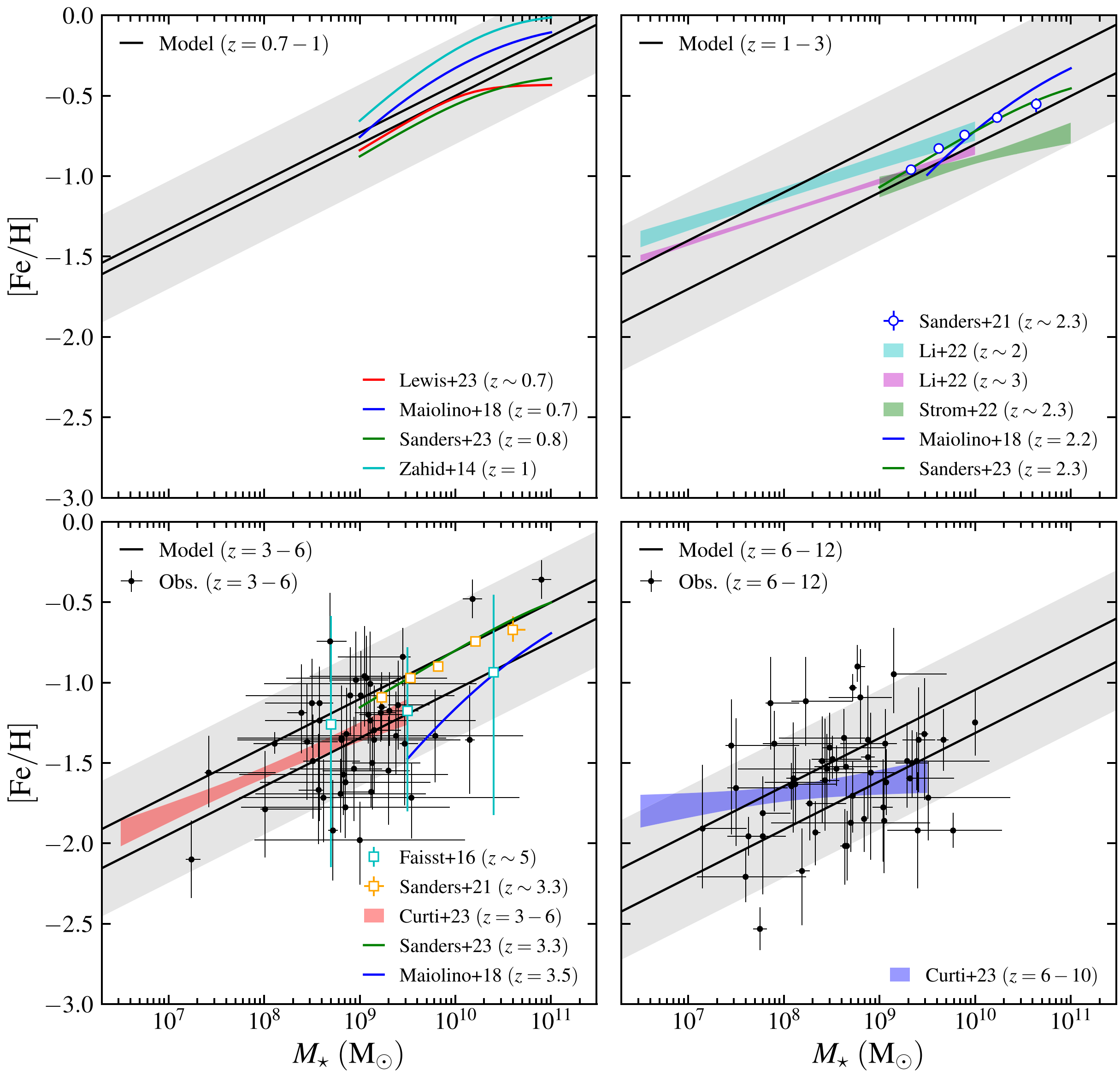

Figure A2. Galaxy stellar mass–gas-phase metallicity relations in different redshift ranges: $z = 0.7$–1 (*upper left*), 1–3 (*upper right*), 3–6 (*lower left*), and 6–12 (*lower right*). Points with black errorbars show data for individual galaxies (Curti et al. 2022; Boyett et al. 2023; Hsiao et al. 2023; Jung et al. 2023; Williams et al. 2023). The observational compilations that fit the data with functional forms are shown as coloured curves or bands (Maiolino et al. 2008; Zahid et al. 2014; Faisst et al. 2016; Sanders et al. 2021; Strom et al. 2022; Curti et al. 2023; Heintz et al. 2023; Lewis et al. 2023; Li et al. 2023; Matthee et al. 2023; Nakajima et al. 2023). We convert the observed oxygen abundance 12 + log(O/H) to iron abundance [Fe/H] using equation (A5). The mean MZR adopted in this paper is represented by two black lines in each panel, evaluated at the lower/upper redshift bounds of the panel. The grey shaded ranges show the scatter $\sigma_g = 0.3$ dex around the bounding mean relations.

Heintz et al. (2023) measured the metallicity of 16 galaxies at $z = 7$–10 with *JWST*/NIRSpec from the gravitational lensing clusters Abell 2744 and RXJ-2129, and the CEERS survey. They also determined the metallicity using direct T_e method and metallicity–line ratio scaling.

In addition to the above surveys, we also include *JWST* measurements from Birkin et al. (2023, two galaxies at $z \sim 4$), Jung et al. (2023, three galaxies at $z = 7.47$–7.75), Curti et al. (2022, three galaxies at $z \sim 8$), Boyett et al. (2023, a galaxy at $z = 9.31$), Williams et al. (2023, a galaxy at $z = 9.51$), and Hsiao et al. (2023, a galaxy at $z = 11.7$).

In Fig. A2, we plot these observational data grouped by redshift bins. Galaxies at different redshifts and from different surveys show slopes of the [Fe/H]–$M_\star$ relation between 0.2 and 0.4. We find a single value ∼0.3 can match the observational data at $1 \lesssim z \lesssim 6$ when our model forms majority of surviving clusters. We therefore set $\alpha_M = 0.3$ for the model MZR.

In addition, we find a dependence on redshift from $z \sim 0.7$ to $z \sim 12$, which can be characterized by $\alpha_z = 1.0$. This value is slightly greater than what was employed in the previous version of the model, $\alpha_z = 0.9$, which accounts for the approximately 0.6 dex drop of metallicity from $z = 0$ to $z = 4$ at a fixed mass (Mannucci et al. 2009).

Finally, we normalize the model MZR by selecting $[\mathrm{Fe/H}]_0$ that brings equation (A2) to the correct scale. The resulting functional form for the shape and evolution of MZR is

$$[\mathrm{Fe/H}] = 0.3 \log \frac{M_\star}{10^9\,\mathrm{M}_\odot} - 1.0 \log(1+z) - 0.5. \tag{A6}$$

We plot the final relation in Fig. A2 for comparison with the data. Our MZR can match the observations across a broad mass range ($M_\star = 10^7$–$10^{11}\,\mathrm{M}_\odot$) and redshift range ($z =$ 0.7–12). The standard deviation of the difference between the observed metallicity of individual galaxies and the predicted metallicity by our MZR is around 0.3 dex. We apply this value as the scatter of mean galaxy metallicity, $\sigma_{\mathrm{g}} = 0.3$ dex.

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