Globular clusters contribute to the nuclear star clusters and galaxy centre γ-ray excess, moderated by galaxy assembly history

Yuan Gao,[1,2]* Hui Li [ORCID],[3,4]* Xiaojia Zhang,[2,5] Meng Su[1,2] and Stephen Chi Yung Ng[1,2]

[1]*Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China*
[2]*Laboratory for Space Research, The University of Hong Kong, Cyberport 4, Hong Kong SAR, China*
[3]*Department of Astronomy, Tsinghua University, Beijing 100084, China*
[4]*Department of Astronomy, Columbia University, Mail Code 5246, 538 West 120th Street, New York, USA*
[5]*Department of Earth Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China*

ABSTRACT
Two unresolved questions at galaxy centres, namely the formation of the nuclear star cluster (NSC) and the origin of the γ-ray excess in the Milky Way (MW) and Andromeda (M31), are both related to the formation and evolution of globular clusters (GCs). They migrate towards the galaxy centre due to dynamical friction, and get tidally disrupted to release the stellar mass content including millisecond pulsars (MSPs), which contribute to the NSC and γ-ray excess. In this study, we propose a semi-analytical model of GC formation and evolution that utilizes the Illustris cosmological simulation to accurately capture the formation epochs of GCs and simulate their subsequent evolution. Our analysis confirms that our GC properties at $z = 0$ are consistent with observations, and our model naturally explains the formation of a massive NSC in a galaxy similar to the MW and M31. We also find a remarkable similarity in our model prediction with the γ-ray excess signal in the MW. However, our predictions fall short by approximately an order of magnitude in M31, indicating distinct origins for the two γ-ray excesses. Meanwhile, we utilize the catalogue of Illustris haloes to investigate the influence of galaxy assembly history. We find that the earlier a galaxy is assembled, the heavier and spatially more concentrated its GC system behaves at $z = 0$. This results in a larger NSC mass and brighter γ-ray emission from deposited MSPs.

Key words: pulsars: general – Galaxy: centre – globular clusters: general – galaxies: evolution – gamma-rays: galaxies.

1 INTRODUCTION

A compact bright star cluster is commonly observed at the centres of galaxies of all types, known as the nuclear star cluster (NSC; e.g. Light, Danielson & Schwarzschild 1974; Kormendy 1985; Matthews & Gallagher 1997; Hughes et al. 2005). They take up the innermost a few to tens of parsecs at most, and have masses $10^5 \sim 10^8$ M$_\odot$ (Georgiev et al. 2016; Spengler et al. 2018), making them the densest known star clusters. Observationally, they are distinguished by prominently brighter luminosity on top of the disc or bulge component (e.g. Böker et al. 2002; Kim et al. 2004). Besides, some NSCs are also observed to co-exist with a supermassive black hole (SMBH) at the galaxy centre (e.g. Filippenko & Ho 2003; Nguyen et al. 2019).

NSCs consist of a mixed stellar population in terms of age, metallicity, etc (e.g. Böker et al. 2001; Kacharov et al. 2018). The complexity in stellar population has complicated the quest of NSC formation mechanisms. The young and metal-rich stars suggests an *in situ* formation scenario, where local star formation is triggered by the inflow of gas induced by various mechanisms, such as bar-driven infall (Shlosman, Begelman & Frank 1990) and the action of instabilities (Milosavljević 2004). The fact that this young stellar population is usually flattened, rotating, and concentrated at the centre of NSCs (Georgiev & Böker 2014; Carson et al. 2015) also favours the *in situ* formation scenario. On the other hand, the old and metal-poor population can naturally arise from massive globular clusters (GCs) which migrate into the galaxy centre due to dynamical friction. This was actually the first proposed NSC formation scenario (Tremaine, Ostriker & Spitzer 1975) one year after the ground-breaking observation of the NSC in M31 (Light, Danielson & Schwarzschild 1974). Observationally, this scenario is also supported by evidences such as the deficit of massive GCs (Capuzzo-Dolcetta & Mastrobuono-Battisti 2009) and the nucleation fraction tracing the fraction of galaxies that have GCs (Sánchez-Janssen et al. 2019).

Despite observational motivation for both NSC formation mechanisms, a comprehensive modelling is lacking. For *in situ* formation, direct simulation is challenging. There are limited studies focusing on different aspects of the process, such as gas inflow (Hopkins & Quataert 2010), momentum feedback and self-regulation (McLaughlin, King & Nayakshin 2006), and stellar two-body relaxation (Aharon & Perets 2015). *Ex situ* formation was more extensively studied (e.g. Tremaine, Ostriker & Spitzer 1975; Capuzzo-Dolcetta 1993; Lotz et al. 2001; Gnedin, Ostriker & Tremaine 2014; Leveque et al. 2022). In particular, Hartmann et al. (2011), Antonini et al. (2012), and Tsatsi et al. (2017) used direct *n*-body simulation to study the final parsec-scale evolution of spiraled-in GCs and resulting NSC morphological and kinematic properties. The general picture has

* E-mail: li.hui@columbia.edu (HL); u3502132@connect.hku.hk (YG)

been well-established, but previous studies adopted crude treatments of the initial conditions of GCs due to a lack of knowledge on their formation. Besides, the evolution of the old GC systems is closely correlated with the evolution of the host galaxy, which hasn't been taken good care of in previous studies. Up to today, it is still uncertain the roles of *in situ* and *ex situ* formation mechanisms (e.g. Guillard, Emsellem & Renaud 2016; Fahrion et al. 2022), with ongoing debates on topics such as a potential transition between the two mechanisms indicated by the NSC mass or galaxy stellar masses (Lyubenova & Tsatsi 2019).

In this study, we focus on the *ex situ* formation of NSCs, as the infall of GCs might contribute to another unsolved problem. A diffuse γ-ray excess has been observed at the centres of the Milky Way (MW) and Andromeda (M31) galaxies (Abazajian et al. 2014; Ackermann et al. 2017a). These excesses exhibit spherical symmetry and extend over a few parsecs, with a peak energy around a few GeV. Possible origins of these excesses are debated primarily over dark matter (DM; e.g. Calore et al. 2015a; McDaniel, Jeltema & Profumo 2018; Di Mauro et al. 2019; Cholis et al. 2022) and millisecond pulsars (MSPs; e.g. Brandt & Kocsis 2015; Haggard et al. 2017; Eckner et al. 2018; Feng et al. 2019; Fragione, Antonini & Gnedin 2019; Zimmer et al. 2022).

While the spatial distribution of DM is already expected to peak at galaxy centres (e.g. Navarro, Frenk & White 1997), MSPs are not commonly observed there due to difficulties in resolving individual γ-ray sources. However, MSPs are abundant in GCs, with much more observed number per unit mass compared to the galaxy field (Ransom 2007; Brandt & Kocsis 2015; Ye et al. 2019). This is due to the fact that MSPs are believed to originate from low-mass X-ray binaries (LMXBs), where the neutron star gets spun up through mass transfer from the companion star (Bhattacharya & van den Heuvel 1991). Thus, the high stellar density of GCs provide the desirable environment for both primordial binary formation and dynamical encounter. Galaxy centre MSPs can originate from GCs that have migrated in and tidally dissolved.

On the other hand, the central bulge region of galaxies is also a dense stellar environment, although less dense by an order of magnitude than most GCs (Voss & Gilfanov 2007). Thus, MSPs could in principle form *in situ* as well. However, recent studies using scaling relations to probe *in situ* MSP luminosity cast doubts on this mechanism as the sole origin of the γ-ray excess. The Galactic centre excess (GCE) has been examined by Cholis, Hooper & Linden (2015) and Haggard et al. (2017), who pointed out that LMXBs are too rarely observed in the MW bulge that *in situ* MSPs can account for less than a quarter of the excess luminosity. For M31, where $\sim10^4$ LMXBs are needed to explain the excess, <80 were detected within the inner 12 arcmin (~2.7 kpc; Voss & Gilfanov 2007). Consequently, it is crucial to explore the contribution of *ex situ* MSPs from GCs

From the previous analysis, we see that the evolution of GCs inevitably contributes to NSC formation and the γ-ray excess, and it is important to find out the extent of its contributions. However, the formation of GCs still remains highly uncertain. Many previous studies (e.g. Gnedin, Ostriker & Tremaine 2014; Fragione, Antonini & Gnedin 2019; Leveque et al. 2022) assumed that all GCs formed at a single redshift, which only serves as a primitive approximation. To achieve a more reliable modelling of GC evolution and mass deposition, a better formulation of GC formation is needed.

In this paper, we use a new semi-analytic model of GC formation and evolution to study its contribution to the NSC formation and galaxy centre γ-ray excess in galaxies similar to the MW and M31. The model adopts the GC formation scenario by Li & Gnedin (2014), Choksi, Gnedin & Li (2018), and Choksi & Gnedin (2019), where

GC formation was triggered by periods of rapid mass accretion on to the host galaxy across its assembly history, typically triggered by major galactic mergers. To obtain realistic galaxy merging histories, results from the Illustris cosmological simulation (Vogelsberger et al. 2014; Nelson et al. 2015) are used, and GCs are sampled at qualified simulation snapshots. After formation, the GC population is subject to orbital migration and mass loss depending on their mass and galactocentric distance. In addition, we also model an evolving background potential according to the galaxy assembly history. Through this new model, we hope to enhance our understanding of the connection between GCs, the NSC, and galaxy centre γ-ray excess.

We arrange this paper as follows. We introduce our modelling of GC formation and evolution in Section 2. In it, we also show how to account for the galaxy centre MSP luminosity at $z = 0$. The calculation of halo parameters from Illustris outputs is introduced in the Appendix A. In Section 3, we present our model predictions of GC properties, the NSC mass, and γ-ray emission by MSPs at $z = 0$. As we can retrieve from Illustris a collection of haloes of similar masses but with different assembly histories, we discuss the moderation effect of assembly history in Section 4. As our γ-ray luminosity prediction for the M31 falls short to observation, we also discuss alternative explanations. Caveats of our study are listed and discussed as well. Finally in Section 5, we summarize important results and suggest future work.

2 METHODS

In this section, we present our semi-analytical model for the formation and evolution of GCs in the framework of hierarchical structure formation. Then we describe the calculation of the γ-ray luminosity at $z = 0$ from deposited MSPs.

2.1 Formation of GCs in cosmological simulations

We introduce our modelling of GC formation in terms of formation times, initial masses, and spatial distribution.

In modelling GC formation times, we improve on the simple prescription by previous studies (e.g. Gnedin, Ostriker & Tremaine 2014; Fragione, Antonini & Gnedin 2019; Leveque et al. 2022) that GCs formed at a single redshift. While this assumption is partly justified due to the old ages of most GCs, it is recently recognized that GC formation covers a wide range of cosmic time with diverse formation histories (Forbes et al. 2018 and references therein). Therefore, a more physically motivated GC formation model is desired. Fortunately, over the past decades, our understanding of the origin of GCs in the framework of hierarchical structure formation has been revolutionized. Here, we adopt the GC formation model of Choksi, Gnedin & Li (2018, hereafter CGL), which was built upon earlier works by Muratov & Gnedin (2010) and Li & Gnedin (2014). The CGL model assumes that GC formation was triggered by periods of rapid mass accretion on to the host galaxy, typically during major mergers. This idea was motivated by multiple reasons such as more observed young massive clusters in interacting galaxies (e.g. Wilson et al. 2006; Portegies Zwart, McMillan & Gieles 2010), earlier formation times of GCs than the field stars and that galactic mergers were more frequent at high redshifts (Li & Gnedin 2014), and that galactic mergers are able to induce the high densities and pressures desired for cluster formation (e.g. Li et al. 2017; El-Badry et al. 2019).

In the CGL model, GC formation is painted on to the halo merger trees of the Illustris simulation, which captures the evolution of

halo properties from $z = 47$ to 0 (Vogelsberger et al. 2014; Nelson et al. 2015). GC formation is triggered when the specific halo mass accretion rate exceeds a threshold value, which is a tunable parameter. The total GC mass is mapped from the halo mass through observed stellar mass–halo mass relation (SMHM; Behroozi, Wechsler & Conroy 2013), the stellar mass–gas mass relation (Magdis et al. 2012; Lilly et al. 2013; Genzel et al. 2015; Tacconi et al. 2018), and GC mass fraction from total gas mass (the second tunable parameter). Once the total GC mass at the epochs of formation is fixed, individual GC masses are sampled from a power -2 mass function observed from young massive clusters (Portegies Zwart, McMillan & Gieles 2010; Schulz, Pflamm-Altenburg & Kroupa 2015). It is encouraging to note that this simple two-parameter model successfully reproduced several key observed properties of GC populations, such as metallicity bimodality, GC mass–halo mass relation, etc. This serves as a much improved model of GC formation for the study of NSCs in galaxies with different assembly histories.

As GCs form at multiple epochs across the galaxy assembly history, their initial spatial distribution varies and influences their subsequent orbital evolution. Thus, we need to carefully take care of this issue at each formation epoch. For GCs formed inside the galaxy (*in situ*), we assume that, similar to normal star formation, GCs follow similar spatial distribution of cold gas during their formation. As gas does not exhibit a bulge structure, we adopt the continuous spherical Sérsic density distribution, based on which GCs are assigned to specific galacto-centric radii. On the other hand, for GCs formed inside satellite galaxies and were brought in via galactic mergers (*ex situ*), we lack detailed knowledge of the stellar dynamics during galactic mergers. As a simplified treatment, we place these *ex situ* GCs at half the virial radius of the descendant halo. This shouldn't affect our GC distribution at the galaxy centre at $z = 0$, though, because these massive GCs carry large angular momenta relative to the descendant halo, which are highly unlikely to be sufficiently reduced by dynamical friction from the ambient stellar density.

The Sérsic spatial density distribution was proposed by Prugniel & Simien (1997) to match the well-established Sérsic surface brightness profile (Sérsic 1963). The spherical density distribution is given by:

$$\rho(r) = \rho_0 \left(\frac{r}{R_e} \right)^{-p} e^{-b(r/R_e)^{\frac{1}{N_s}}}, \tag{1}$$

here ρ_0 is a normalization, R_e is the effective radius of the galactic disc, N_s is the concentration index. The term b is a function of N_s to ensure half of the projected light is contained within R_e, and can be well approximated analytically by $b = 2N_s - 1/3 + 0.009876/N_s$ for $0.5 < N_s < 10$ (Prugniel & Simien 1997). The form of p is adopted from Márquez et al. (2000) as $p = 1.0 - 0.6097/N_s + 0.05563/N_s^2$ for $0.6 < N_s < 10$ and $10^2 \leq r/R_e \leq 10^3$ to match the Sérsic surface brightness profile (Terzić & Graham 2005).

The effective radius R_e is related to the virial radius of the halo, R_{vir}, and halo spin parameter, λ, by assuming a classical galactic disc formation model (e.g. Mo, Mao & White 1998):

$$R_e = \lambda R_{vir}/\sqrt{2}. \tag{2}$$

Since we do not have the information of the total energy of the DM halo to estimate the traditional spin parameter defined in Peebles (1980), we instead use an alternative definition $\lambda_B = j_{sp}/\sqrt{2} R_{vir} V_{vir}$ that only requires the specific angular momentum and virial velocity V_{vir} (e.g. Bullock et al. (2001)). Both M_{vir} and j_{sp} are provided at snapshots of the Illustris simulation, and v_{vir} and r_{vir} are calculated accordingly as explained in the Appendix.

Regarding the concentration index N_s, larger values were conventionally associated with higher concentration, but we found this
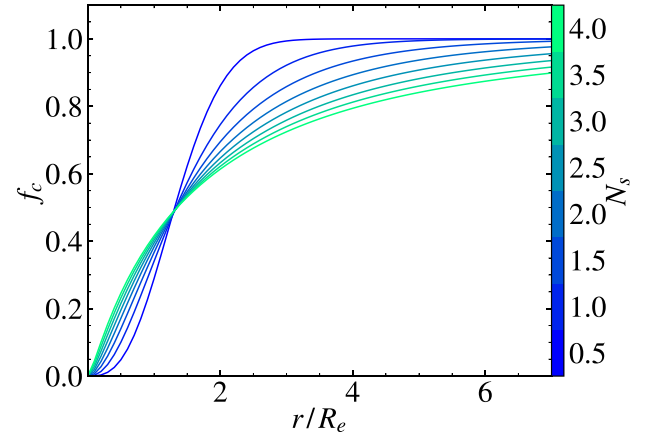
Figure 1. Distribution of the cumulative mass fraction (f_c) for the Sérsic volume density profile (equation 1), plotted against the radial distance r normalized by the effective radius R_e. Different concentration index N_s are plotted separately.

to be misleading. In Fig. 1, we present the cumulative mass fraction distribution versus normalized radial distance for different N_s. We see that larger N_s does exhibit higher concentration in the inner region. However, beyond the crossing point at $r \gtrsim R_e$ (which we refer to as Sérsic crossing hereafter), they reach saturation much slower than curves with smaller N_s. Consequently, while larger N_s values indicate a more peaked distribution, they also indicate a greater spread. In our subsequent studies, we investigate the influence of different N_s values on our results within the range of 0.5 to 4, with increments of 0.5. As a fiducial model, we select $N_s = 2$ following the work of Gnedin, Ostriker & Tremaine (2014).

To study the GC system in the MW and M31, we select Illustris haloes with similar masses at $z = 0$. For the MW, we adopt the findings by Deason et al. (2021), who estimated a virial mass of $1.01 \pm 0.24 \times 10^{12} \, M_\odot$. Among the Illustris haloes, 1099 fall within this range. For M31, there are significant uncertainties, therefore we adopt a wider range of 0.7–$2.5 \times 10^{12} \, M_\odot$ based on the review by Kafle et al. (2018). Accordingly, 2029 haloes were selected, including those selected for the MW.

2.2 Dynamical evolution of GCs

After birth, GCs migrate towards the galactic centre due to dynamical friction while depositing masses due to stellar evolution and tidal stripping along the way. In situations where they have migrated into the innermost region, the tidal force can be so strong that GCs get completely torn apart (Arca-Sedda & Gualandris 2018; Wang & Lin 2023). Thus, GCs that make all the way into the inner a few parsecs can build up NSC and contribute to the MSP populations there. We directly adopt the analytical prescriptions of Fragione, Antonini & Gnedin (2019) in modelling: (1) orbital migration, (2) tidal stripping, and (3) direct tidal disruption. These prescriptions include corrections to parameters originally proposed by Gnedin, Ostriker & Tremaine (2014) and Fragione, Antonini & Gnedin (2018). Below, we describe two improvements we made in this work.

The first improvement is an updated prescription of the mass loss due to stellar evolution, which is obtained from the stellar population synthesis model FSPS (Conroy & Gunn 2010). This allows us to account for the evolving nature of stars within GCs more accurately.

The second improvement is modelling the time-varying gravitational potential by extracting the time-evolution of the DM haloes from the Illustris simulation. For any cosmic time, we linearly interpolate the halo mass and spin between adjacent Illustris snapshots, and calculate galactic structural parameters accordingly. The overall gravitational potential comprises three components: the dark halo, the stellar component, and modelled GCs. The dark halo is described using the NFW distribution (Navarro, Frenk & White 1997), and the calculations for determining halo parameters are introduced in the Appendix A. The stellar component is modelled with a Sérsic distribution, as explained earlier. The mass of modelled GCs is also included in calculating the overall potential. For cosmological parameters, we adopt a flat lambda cold dark matter (ΛCDM) model with $h = 0.704$ and $\Omega_{m,0} = 0.2726$, consistent with the Illustris simulation.

To improve the efficiency of our simulation, we implement subcycling in the evolution of GCs. We divide the entire time span into 100 sections, each characterized by a constant background potential. To evolve individual GCs, we first calculate their evolution time-step dt_i being the smaller of the tidal and orbital evolution timescales multiplied by a fraction, ts_m and ts_r, respectively. Then, a GC with the smallest value of $t_i + dt_i$ means that it evolves the fastest, and exerts the strongest influence on other GCs. Thus, we find and evolve such GC at each step, until all GCs cross the current time span.

Meanwhile, we found that it is usually the first step of calculated dt_i that is overestimating. To efficiently address this, we introduce a maximum cutoff time-step dt_{max} on top of the ts factors. By testing single GCs with different masses and galactocentric positions, we optimize the values of dt_{max}, ts_m, and ts_r together. We find that $dt_{max} = 0.01$ Gyr and $ts_m = ts_r = 0.2$ strike a balance between efficiency and accuracy.

For calculating the change in GC mass and galactocentric distance in each step, we employ the Runge–Kutta 4th order method. This method offers a significant speed improvement of approximately 20 times compared to the 1st order Euler method.

2.3 MSPs from GCs

MSPs are deposited by GCs due to tidal stripping and disruption, thus they can contribute to the unresolved γ-ray excess. In calculating the γ-ray luminosity at $z = 0$ from MSPs, we follow Fragione, Antonini & Gnedin (2018, 2019) who set the total γ-ray luminosity of deposited MSPs equal to that of the debris of the GC, using luminosity–mass relation of GCs fitted to observations. To account for uncertainties in this fitting, they also used a constant luminosity–mass relation for comparison, and labeled it as the model 'C'. The model with fitted luminosity–mass relation was labeled 'EQ'.

Then, individual MSP luminosity was sampled according to the observed MSP luminosity function. To account for MSP spin-down due to the loss of rotational energy via magnetic dipole braking, Fragione, Antonini & Gnedin (2018, 2019) use two models of the spin-down time-scale: a Gaussian distribution model (GAU) and a log-normal distribution model (LON). In our study, we follow their prescribed methodology and calculate the γ-ray luminosity at redshift $z = 0$ using the four models, namely GAU-C, GAU-EQ, LON-C, and LON-EQ.

3 RESULTS

In this section, we present our results on the properties of GCs at $z = 0$ for MW and M31-like galaxies, and compare with observations. We then show the predicted mass of the NSC and γ-ray luminosity
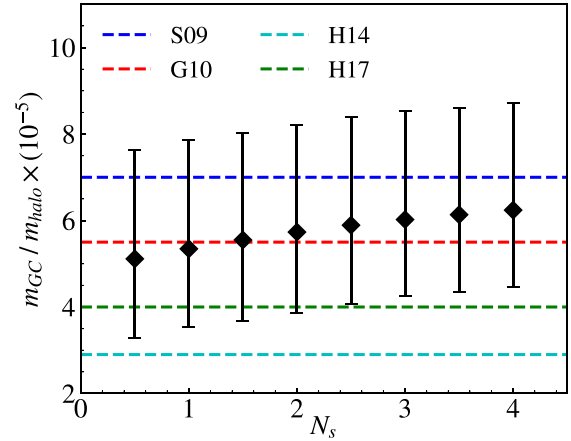
Figure 2. The ratio between the total GC system mass (m_{GC}) and the host halo mass (m_{halo}) at $z = 0$ for different Sérsic index N_s for all selected MW and M31-like haloes. The diamonds show the median ratio, while the error bars show the 25–75 interquartile range. For reference, we also overplot horizontal dashed lines for a compilation of recent observations from Spitler & Forbes (2009; S09), Georgiev et al. (2010; G10), Hudson, Harris & Harris (2014; H14), and Harris, Blakeslee & Harris (2017; H17).

distribution from MSPs. We also explore how the above quantities vary with different galaxy assembly histories.

3.1 GC-halo mass scaling relation

One of the most striking observations of GCs is a linear correlation between the mass of the GC system and its host halo (Spitler & Forbes 2009; Georgiev et al. 2010; Hudson, Harris & Harris 2014; Harris, Blakeslee & Harris 2017). The mass ratio is approximately 10^{-5} across 5 orders of magnitude (Harris, Harris & Hudson 2015). Therefore, it is crucial for our model to reproduces this scaling relation.

Fig. 2 shows our model prediction of the scaling relation for the MW and M31-like haloes combined for different N_s. We see that the runs with smaller N_s tend to produce lower mass in the GC system at $z = 0$. This is in line with our findings in Fig. 1, that a smaller N_s corresponds to a smaller spatial spread of formed GCs, which results in more mass loss due to stronger tidal effect. Nevertheless, the changes of total GC mass in different N_s is very small and all choices of N_s produce the GC mass in reasonably good agreement with observations. Consequently, our model successfully reproduces the GC-halo mass scaling relation.

3.2 Spatial distribution of GC number density

In this section, we closely examine the spatial distribution of GCs in their number density at $z = 0$. Fig. 3 shows the average trend for MW and M31 combined, as we have checked that they are barely distinguishable.

When we compare the initial and final GC distributions, we notice that the number density outside ≈ 3 kpc barely decreases. This can be attributed to the fact that beyond this position, most GCs with a mass of approximately $10^5 \, M_\odot$ have tidal and migration time-scales longer than a Hubble time. In addition, stellar evolution alone cannot exhaust a GC.

When we investigate the effect of different N_s values, it turns out that this made minimal difference for both the initial and final
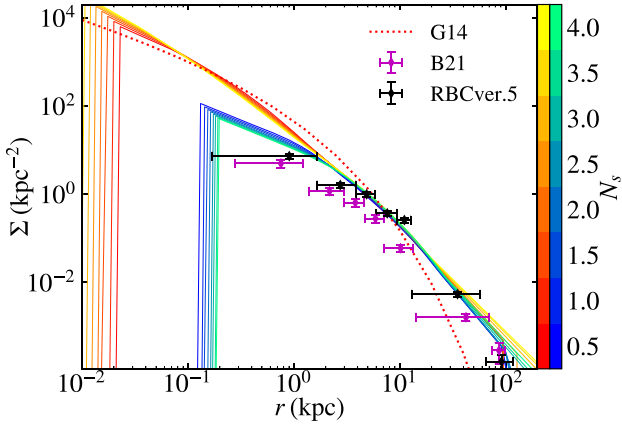
Figure 3. The surface density distribution of GC number with different N_s averaged for all selected haloes. We do not separate results for MW and M31-like haloes, because we have checked that they are barely distinguishable, and that the MW mass range is enclosed by that of the M31. In this plot, initial GCs correspond to all that formed at different redshift, and final GCs to those at $z = 0$. We show the initial GC distribution in Gnedin, Ostriker & Tremaine (2014) for comparison (G14). The error-bars (B21, RBC ver. 5) show the observed GC catalogue for the MW (Baumgardt et al. 2021) and M31 (Galleti et al. 2006). In B21, ωCen was manually removed as it was proved to be the stripped core of a disrupted dwarf galaxy (Noyola, Gebhardt & Bergmann 2008).

GC distributions, except for the innermost region. In the initial distribution, the typical Sérsic crossing is only observed ≈100 pc, since GC formation across cosmic times overlaps in the outer regions as the halo grows, erasing outer crossings. Our inclusion of *ex situ* GCs also contributes to this effect. In the final distribution, although the difference between N_s values remains minimal, the trend is opposite to that observed in the initial distribution. Larger N_s values within the Sérsic crossing lead to stronger tidal disruption, resulting in fewer surviving GCs and a smaller number of GCs in the final distribution.

Comparing our results with observations, we found a minor overshoot within ≈ 5 kpc for the MW. However, individual haloes exhibit a spread around the average distribution, and several candidate haloes demonstrate conformity to the observation. We present one such candidate halo in Fig. 4. Therefore, our model can be considered an acceptable fit to the observed spatial distribution of MW and M31 GCs. Additionally, we observed no preference for N_s, so we will keep $N_s = 2$ as the fiducial choice.

To further analyse the GC population, we utilize the orbital information in the Baumgardt et al. (2021) catalogue to distinguish between the *in situ* and *ex situ* sub-populations of MW GCS. In order to do so, we follow Massari, Koppelman & Helmi (2019), who analysed the dynamical differences of the two branches of GCs on the age-metallicity plot for 151 MW GCs. Based on their distinct features, they assigned 62 GCs as having formed *in situ*, which comprise of bulge GCs and disc GCs. The former are defined as having apocentre distance <3.5 kpc, while the latter as having orbital altitude <5 kpc and circularity >0.5. For our GC catalogue, we calculate the orbital and potential parameters using the PYTHON package GALPOT by Dehnen & Binney (1998) and McMillan (2017). Since our treatment of *ex situ* GCs is rudimentary, we focus on *in situ* GCs for now.

In Fig. 5, we plot the number density distributions of MW *in situ* GCs in a similar manner to Fig. 3. We can observe that most of the
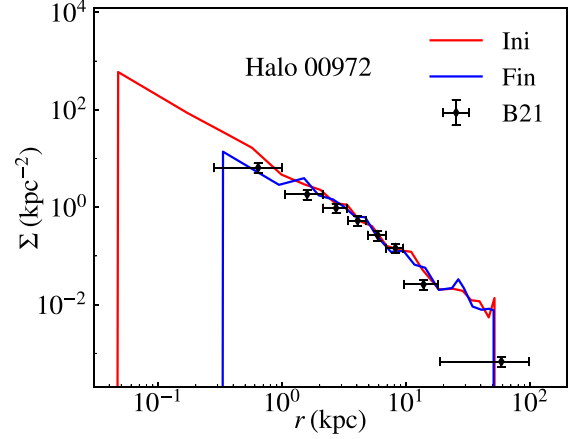
Figure 4. GC number density distribution for one candidate MW halo with Illustris subhalo ID 00972. Initial GCs and GCs at $z = 0$ are shown. The error bars are again the Baumgardt et al. (2021) observed GC catalogue.
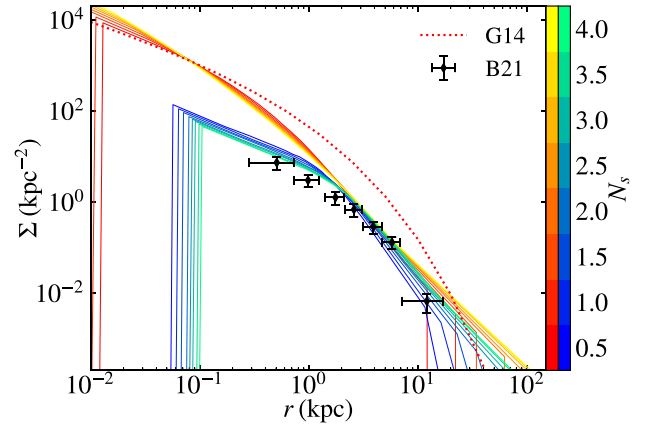

**Figure 5.** Same as Fig. 3, but only for MW *in situ* GCs.

observations made for the overall GC distribution also apply to the *in situ* GC distribution, except for two differences. First, *in situ* GCs exhibit a more concentrated distribution, with a significant drop in their numbers towards the outskirts. This suggests that *in situ* GCs preferentially populate the central regions of the MW-like haloes. Second, without the contribution from the *ex situ* population, the dispersion between different values of N_s becomes more pronounced in the outskirts, and we can observe the outer portion of the Sérsic crossing. Nevertheless, we can regard our model as satisfactorily reproducing the observed GC number density distribution, both overall and *in situ*, and $N_s = 2$ as a reasonable choice for the concentration parameter.

One advantage of utilizing the halo catalogue extracted from Illustris is the ability to investigate the influence of different assembly histories on haloes with similar masses. To parametrize the halo assembly histories, we employ the concept of half mass redshift, denoted as z_{hm}, which represents the redshift at which the halo acquired half of its present mass. In Fig. 6, we show the distribution of z_{hm} values. The histogram exhibits a log-normal shape centred around $z \sim 1.2$, corresponding to a look-back time of approximately 8 billion yr, with a tail extending towards higher redshifts.

We are interested in whether z_{hm} has any discernible effect on the spatial distribution of GCs. Fig. 7 illustrates the GC number density
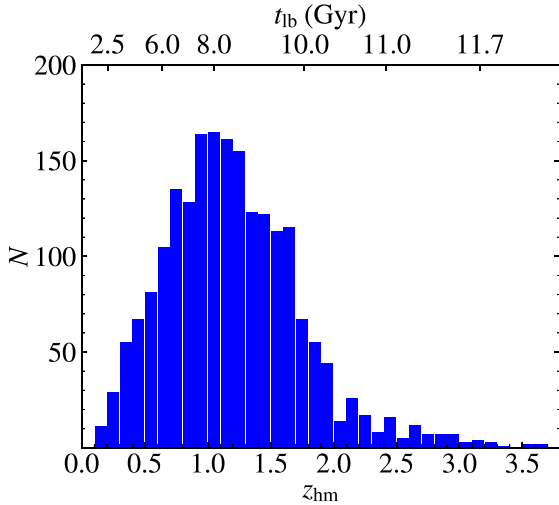
Figure 6. Histogram showing the half-mass redshift z_{hm} for all candidate haloes. Look-back time is shown on the upper horizontal axis as well.
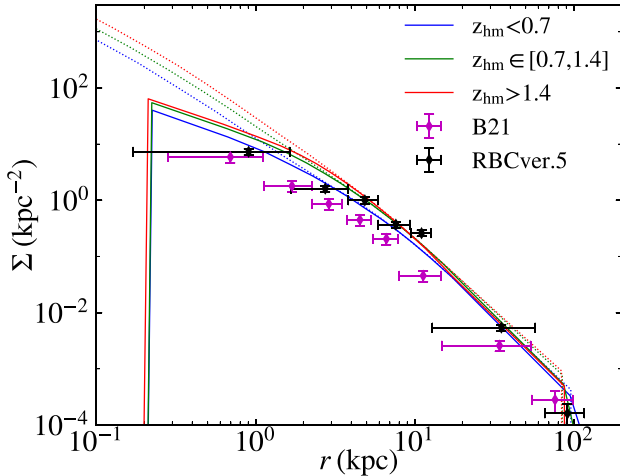
Figure 7. Number density distribution for all candidate haloes with different z_{hm}. For clarity purposes, we only show the average distribution for three z_{hm} ranges. Dashed lines correspond to initial GCs and solid line to final ones.

distribution with different z_{hm} for MW and M31-like galaxies. When we look at the influence of z_{hm} for initial GCs, galaxies that formed earlier (referred to as EFGs, earlier formed galaxies) tend to exhibit more concentrated distributions of GCs compared to haloes that formed later (referred to as LFGs, later formed galaxies). This trend can be attributed to the higher merger rate of haloes at larger redshifts, as indicated by previous studies (Fakhouri & Ma 2008; Fakhouri, Ma & Boylan-Kolchin 2010). Consequently, EFGs experience more GC formation events, starting from smaller halo sizes, while LFGs undergo fewer GC formation events, each resulting in significant growth of halo mass and size and leading to a more spread-out distribution of formed GCs.

In the case of final GCs, however, the impact of z_{hm} is less pronounced. This can be understood, as more concentrated GCs also experience stronger tidal disruption and dynamical friction, which reduces their numbers. Additionally, as these GCs formed earlier, these disruptive processes act over a longer period. Consequently,
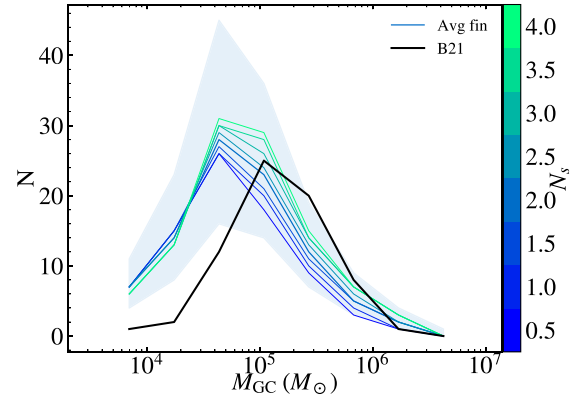
Figure 8. The mass function of *in situ* GCs at $z = 0$, showing the median of all MW-like haloes for different N_s. The shaded region corresponds to the 25–75 interquantile range for $N_s = 2$. The GC catalogue by Baumgardt et al. (2021) is plotted in thick solid line for comparison.

the final GC number density distribution shows minimal traces of the galaxy assembly history.

When comparing our results with observations, as the influence of z_{hm} on the final GC distribution is not significant, the results are similar to those of Fig. 3.

3.3 GC mass function

Besides the spatial distribution of GCs, their mass function at $z = 0$ is also an important property to compare with observations. Previous studies using the same GC formation model (Muratov & Gnedin 2010; Li & Gnedin 2014; Choksi, Gnedin & Li 2018; Choksi & Gnedin 2019; Chen & Gnedin 2022, 2023) have carried out comprehensive analyses, confirming a transformation from the initial power-law mass function to a log-normal shape at $z = 0$ that agrees with observations. Therefore, in this study we focus on the *in situ* GC population. In Fig. 8, we show the median mass function of candidate MW-like haloes for different N_s, and overplot observation results. We see that our model also shows a log-normal shape, although with its peak shifted towards smaller GC masses compared with observations. This can be due to the incompleteness in observed low mass GCs. Nevertheless, the shaded region corresponding to $N_s = 2$ can cover most of the observation trend. Indeed, candidate haloes can fit the observation quite well, as illustrated by Fig. 9. Thus, our model provides a good fit to the observed *in situ* GC mass function.

We also observe that N_s results in more noticeable discrepancies in the predicted mass function above $\sim 10^5 \, M_{\odot}$, while at lower masses, the influence of different N_s is flipped and less prominent. Larger N_s leads to slightly fewer light GCs but more heavier ones, and vice versa. This is due to the fact that those light GCs are prone to weak dynamical friction, thus they barely migrate inward. Their disruption is mostly determined by their position at birth. As we have observed in Fig. 1, larger N_s have peaked density at the galaxy centre, which leads to prominent disruption of lighter GCs that initially formed in the region. On the other hand, as the general distribution is more spreaded, heavier GCs have a longer journey to migrate to the galaxy centre and are subject to weaker tidal effect. Thus, the distribution leads to slightly stronger disruption of light GCs formed in the innermost region, but weaker disruption of heavier ones.
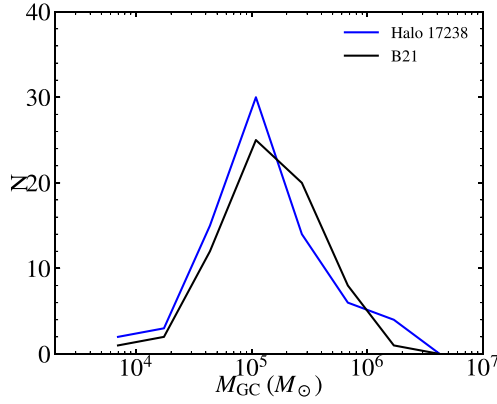
Figure 9. The same as Fig. 8, but for one candidate halo with Illustris subhalo ID 17238.
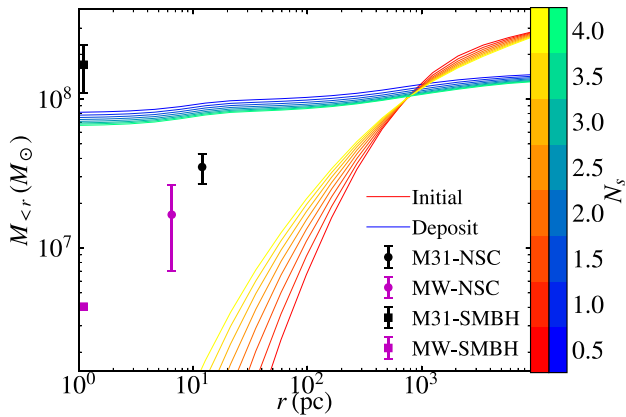
Figure 10. Spatial distribution of the cumulative masses of initial GCs (summed over all formation epochs) and their deposition at $z = 0$, averaged for MW and M31-like haloes with different N_s. The colour code is identical to Fig. 3. The MW SMBH and NSC observations are adopted from Do et al. (2019), Neumayer, Seth & Böker (2020), and M31 from Bender et al. (2005) and Leveque et al. (2022), respectively.

3.4 NSC mass

In this section, we examine the NSC mass contributed by GCs as they spiral in. On average, $\sim$80 and 100 GCs migrated into the NSC of the MW and M31 respectively, with smaller N_s corresponding to a few more GCs. The standard deviation is $\sim$30 and 40, respectively. In Fig. 10, we compare the deposited GC masses with the observed NSC mass at $z = 0$. Unlike the number density distribution, the cumulative mass distribution of initial GCs exhibits a prominent Sérsic crossing in the inner region, because the range of the y-axis is much smaller in this case. Regarding the deposited mass, we observe that GCs with smaller values of N_s contribute more to the total deposited mass, which aligns with the trend we observed in the GC-halo mass scaling relation discussed in Section 3.1. Regardless of the influence of N_s, the deposited mass is predominantly confined to the central regions. This contrasts with the results presented by Gnedin, Ostriker & Tremaine (2014) and Fragione, Antonini & Gnedin (2018), where a plateau is established starting from 4 pc outward and continues to rise prominently up to 10 kpc. The disparity arises due to our GC formation occurring throughout the entire assembly history of the halo, resulting in a more concentrated distribution of GCs, as shown
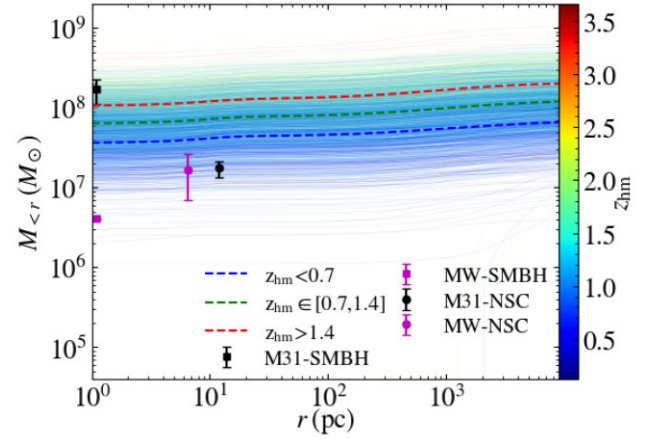
Figure 11. Cumulative spatial distribution of masses of the initial GCs and their deposition with different z_{hm} for all candidate haloes. Three dashed curves show the average for haloes falling under different z_{hm} spans. Observed positions and masses of the MW and M31 NSC, and SMBH are plotted with uncertainties.

in Fig. 3. Consequently, GCs are more susceptible to significant tidal effects and tend to deposit mass towards the galaxy centre.

While the average deposited mass exceeds the NSC mass, it is important to note that different galaxy assembly history yield substantial variations, as depicted in Fig. 11. As discussed in the preceding section, EFGs give rise to more GCs, which experience stronger and more prolonged tidal disruption. Consequently, haloes with larger z_{hm} have GCs depositing more mass towards the centre. Conversely, LFGs exhibit less deposited mass, implying that our MW NSC plausibly originates from a halo with $z_{hm} \sim 1$.
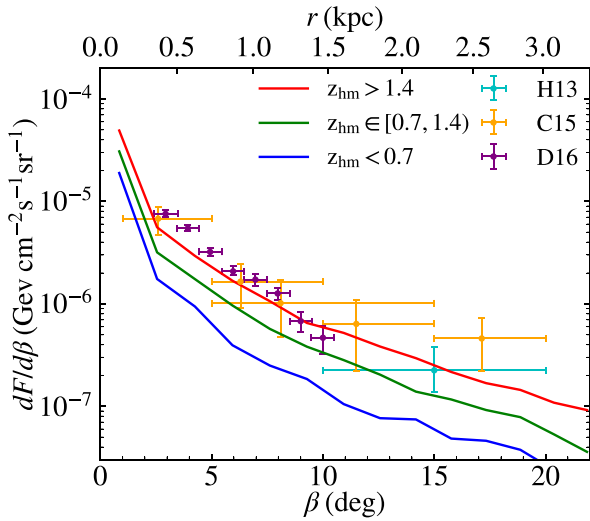
3.5 Gamma-ray luminosity

As our GC fittings turn out consistent with observations, we move forward with $N_s = 2$ to check the spatial distribution of the luminosity of deposited MSPs. As different authors analysed the γ-ray excess data in different methods, we present our results in both differential and cumulative distribution. Fig. 12 shows our model prediction of the differential flux distribution. Notably, different models yield similar results. And although they generally underestimate the excess flux, the overall shape is consistent. Consequently, we select the GAU-EQ model as the best choice and plot it for different z_{hm} intervals in Fig. 13. We observe that EFGs exhibit relatively higher flux emission, as their GCs deposit more MSPs. Haloes with $z_{hm} \gtrsim 0.7$ provide a good fit to the observations. Note that for clarity purpose, the colour code here does not precisely match the colourbar in Fig. 11. With that in mind, we notice that the results are compatible, indicating that a halo with $z_{hm} \gtrsim 0.7$ can successfully reproduce both the NSC mass and the spatial distribution of the γ-ray differential flux emission.

For the cumulative flux distribution, Fig. 14 illustrates the GAU-EQ model for different z_{hm}. We can observe that our model also shows a consistent shape with the observations, which appears better than the results by Fragione, Antonini & Gnedin (2018) as we exhibit more flux in the innermost region. This arise from our GC formation occurring throughout the entire assembly history of the halo. As a result, GCs started their evolution closer to the galaxy centre. Once again, haloes with $z_{hm} \gtrsim 0.7$ provide a good fit. And consequently, it is plausible that the GCE arose solely from deposited MSPs.
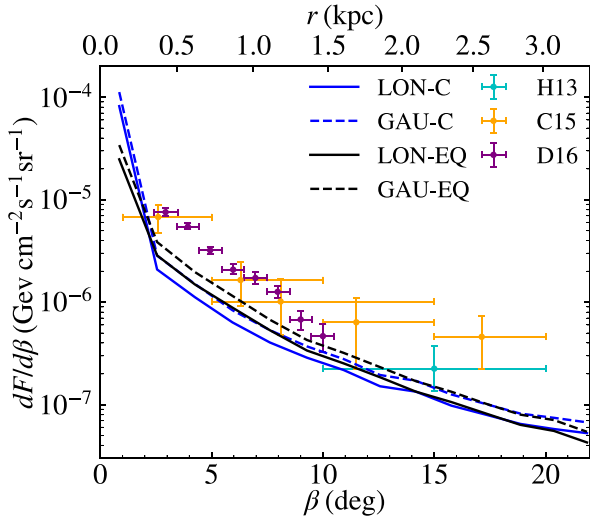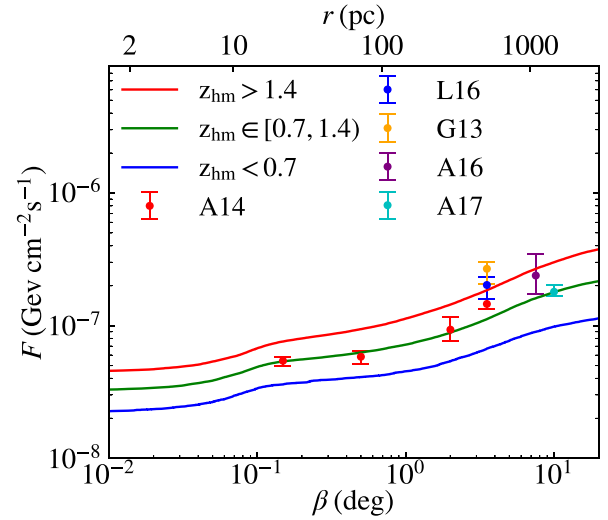
Figure 12. Differential flux distribution over angular distance from the MW centre (β) for the four MSP luminosity models. We overplotted observational constraints by Hooper & Slatyer (2013; H13), Calore, Cholis & Weniger (2015b; C15), and Daylan et al. (2016; D16).
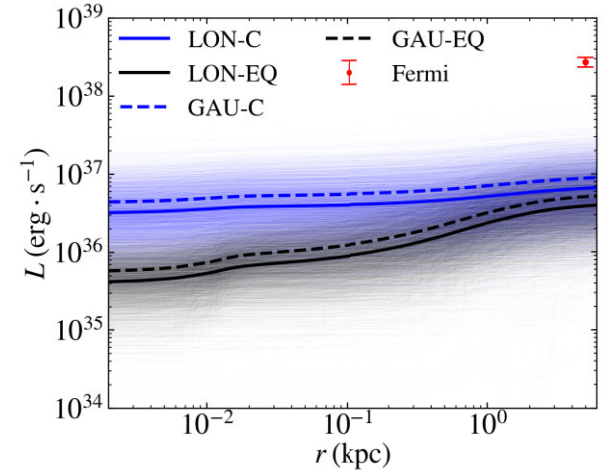

Figure 14. Cumulative flux distribution for the GAU-EQ model with three $z_{\rm hm}$ ranges. Observational constraints are from Abazajian et al. (2014; A14), Lacroix et al. (2016; L16), Gordon & Macías (2013; G13), Ajello et al. (2016; A16), and Ackermann et al. (2017b; A17).

Figure 13. Similar to Fig. 12, but for the GAU-EQ model with three $z_{\rm hm}$ ranges.

Figure 15. Cumulative luminosity distribution for the four MSP luminosity models for all 2029 M31-like haloes. The average distributions for each model are also plotted. The Fermi data were adopted from Fragione, Antonini & Gnedin (2019) in errorbar.

Having examined the GCE, we proceed to the M31 excess. Due to the lack of detailed analysis regarding the excess spatial distribution, we solely compare the cumulative luminosity at 6 kpc. Fig. 15 illustrates the cumulative luminosity distribution for the four MSP luminosity models. The models exhibit notable discrepancies, particularly in the innermost region, which diminish as we move towards the outskirts and nearly vanish at 6 kpc. This discrepancy stems from the fact that the EQ models employ a fitted relation, where $\log(L_\gamma/m_{\rm GC})$ is negatively proportional to $\log(m_{\rm GC})$. Consequently, heavier GCs are dimmer compared to the C models that utilize a constant luminosity–mass ratio. Since heavier GCs are more susceptible to dynamical friction and tidal disruption, they primarily contribute to the γ-ray emission in the innermost regions. As we move farther away from the galaxy centre, lighter GCs become more dominant, which reduces the discrepancy. Nevertheless, all

models converge and fall short of matching the excess signal at 6 kpc. Additionally, the highest luminosity among individual haloes only reaches approximately 7×10^{37} erg s^{-1}, which is less than one-third of the excess luminosity. Therefore, it is unlikely that MSPs alone account for the M31 excess. This suggests that the two galaxies have distinct origins of the excess emissions despite their similar masses.

4 DISCUSSION

4.1 Galaxy assembly history

We have observed in preceding sections that galaxy assembly history has clear influences on the GC and NSC properties. We shall take a systematic look in this section.
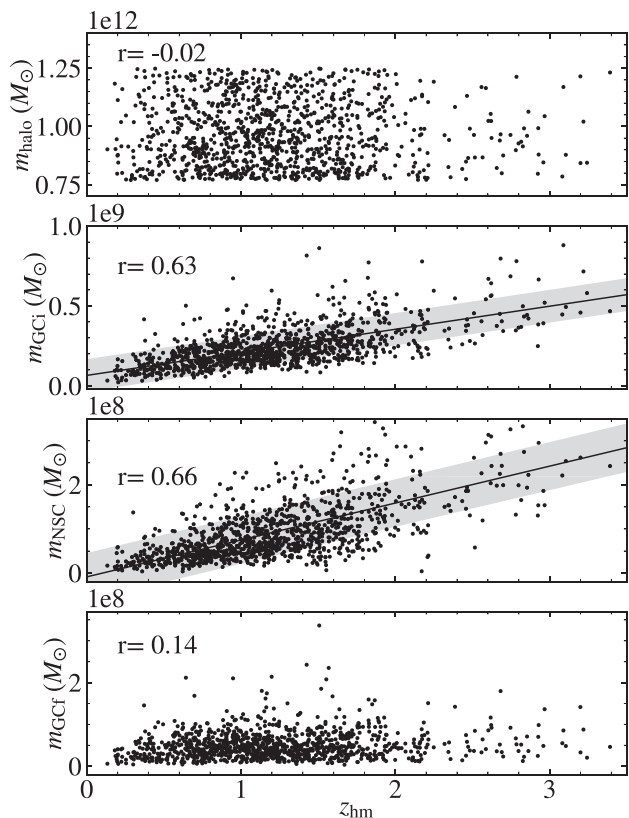
Figure 16. Correlations of z_{hm} with (from the top panel to the bottom) the mass of the halo, initial GCs, NSC, and final GCs for all MW-like haloes. Pearson correlation coefficients are shown at the upper left corner of each panel. Solid lines and shaded regions show the best-fitting trend and 1σ dispersion. The standard deviations of the four masses are 1.38×10^{11}, 1.30×10^8, 4.82×10^7, and $3.33 \times 10^7 \, M_\odot$, respectively, which corresponds to 0.14, 0.54, 0.76, and 0.66 of the respective mean.

In Fig. 16, we illustrate the correlations between z_{hm} and four important masses: the halo mass, NSC mass, total initial GC mass, and GC mass at $z = 0$. The halo mass does not correlate with z_{hm} since haloes of any mass can assemble early or late. However, the density of data points varies across z_{hm}, consistent with the log-normal distribution presented in Fig. 6. And we have demonstrated previously that EFGs give rise to more GCs, resulting in a positive linear trend between the initial GC masses and z_{hm}. We have also showed that EFGs have GCs distributed closer to the galaxy centre, making them susceptible to stronger and prolonged tidal disruption. Consequently, they contribute a greater amount of mass to the NSC, leading to a positive linear trend between the NSC mass and z_{hm}. Thus, the NSC mass serves as a good indicator in breaking the degeneracy to infer the galaxy assembly history. Larger NSC masses indicate relatively earlier accumulation of the halo mass, and vice versa. However, it is intriguing that z_{hm} appears to provide little information about the final GC mass. This seems to suggest a lack of correlation between initial and final GC masses either, contrary to what one might expect.

To verify this observation, we show in Fig. 17 the initial and final GC masses for all MW-like haloes, with colours denoting z_{hm}. A positive trend reasonably exist between initial and final GC masses, although there is relatively large dispersion, particularly at smaller final GC masses. If we observe z_{hm} across different initial GC masses,
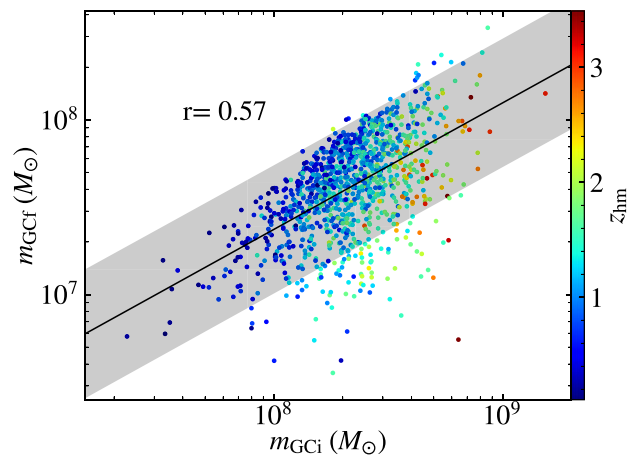
Figure 17. GC initial and final masses for all MW-like haloes, showing different z_{hm}. The best-fitting trend with 1σ band is plotted in solid line and shaded region.
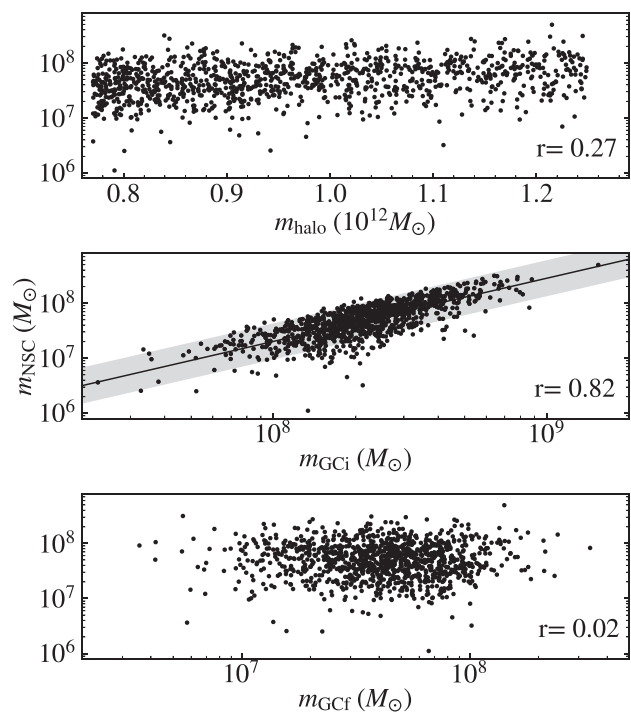
Figure 18. Correlations of the NSC mass with the halo mass (upper panel), total mass of initial GCs (middle panel) and that of final GCs (lower panel) for all MW-like haloes. Pearson correlation coefficients are shown at the lower right corner of each panel, and best-fitting trend and 1σ range was shown as well. Only the initial GC mass exhibits a positive trend with the NSC mass.

larger GC masses do correspond to larger z_{hm}. However, if we examine final GC masses, each value is associated with a large spread of z_{hm} values. This is due to the fact that the earlier-formed larger GC masses also suffer from stronger disruption, potentially resulting in smaller final masses at $z = 0$. Consequently, small final GC masses can arise from either small initial masses or earlier-formed large initial masses. This breaks the correlation between final GC masses and galaxy assembly history.

We are also interested in correlations associated with the NSC mass, a highly significant outcome of our GC model. Fig. 18 shows the NSC mass plotted against the halo mass, initial GC mass, and final GC mass. As previously demonstrated in Fig. 16, the NSC mass is positively correlated with z_{hm}, which has no relation to the halo mass. It is not surprising to observe that the NSC mass does not correlate with the halo mass. It is the earlier assembly of the halo that give rise to a heavier NSC, rather than the mass of the halo itself. However, it is worth noting that this observation may change when examining a broader range of halo masses, which will be investigated in our future studies.

In the middle panel of Fig. 18, a strong positive trend is observed between the mass of initial GCs and the NSC mass. Not only do GCs serve as the fuel for the build-up of the NSC, but larger initial GC masses also correlate with larger z_{hm} values, indicating earlier formation and more concentrated distribution. Collectively, these factors contribute to a larger NSC mass.

However, final GC masses exhibit no correlation with the NSC mass. This can be comprehended, as we have demonstrated in Fig. 16 that final GC masses do not correlate with z_{hm}. Each GC mass at $z = 0$ might arise from either early-formed large GC_i which builds up a large NSC, or later-formed lighter GC_i which contributes little to the NSC mass. Thus, the final GC mass does not serve as an indicator for the NSC mass.

In summary, we presented in this section various correlations associated with the galaxy assembly history. We found that z_{hm}, the total initial GC mass, and the NSC mass are correlated, due to the fact that EFGs give rise to an old, heavy, and concentrated GC system which contributes a larger amount of mass to the NSC. Thus, either value among z_{hm}, the initial GC mass and the NSC mass serves as an indicator of the other two. And one important observation is that we could utilize the NSC mass to infer knowledge on galaxy assembly history. Intriguingly, the final GC mass is not correlated with z_{hm} or the NSC mass. This stems from the degeneracy in the relation between initial and final GC masses.

4.2 Possible explanations of the M31 gamma-ray excess

While our results fall short of the M31 excess by an order of magnitude on average, candidate halo can reach approximately one third of the signal at its highest. Nevertheless, even with *in situ* MSPs combined, the MSP channel alone is unable to fully explain the M31 excess. Thus, it is evident that a contribution from DM is necessary.

Upon the first report on the detection of the M31 excess, Ackermann et al. (2017a) have brought up the possible explanation of DM. A primitive estimate inferred from a DM-origin GCE results in a flux deficit by five times, though the level of uncertainty was high. Subsequent investigations claimed to match the excess luminosity, but commonly identified tensions with observational constraints, such as the under-detection of DM emission in MW dwarf galaxies (Di Mauro et al. 2019) and a lack of DM radio emission for M31 (McDaniel, Jeltema & Profumo 2018). Additionally, McDaniel, Jeltema & Profumo (2018) found that the two preferred DM annihilation channels for M31, namely $b\bar{b}$ and an even mixture of $b\bar{b}/\tau^+\tau^-$, favour smaller DM masses compared to those suggested by the GCE. Consequently, it was suggested that DM alone does not explain the M31 excess. Combining these findings with our results, it becomes clear that a combination of MSPs and DM offers a promising and potentially inevitable way for explaining the M31 excess. However, the question remains as to why the MW and M31 have different origins for producing such excess emission.

4.3 Caveats and future works

In this section, we discuss the caveats in our model and improvements to be made in future works.

First, although we adopt the analytical expression of tidal disruption from Fragione, Antonini & Gnedin (2019), it was based on a static spherical galactic background with circular cluster orbits (Gieles & Baumgardt 2008). Our inclusion of the assembly history of galaxies, however, indicates a more complicated galactic background, especially at large redshifts when galactic mergers were more frequent. This complication is twofold. On one hand, the galactic background keeps varying, although in out treatment of linear interpolation, the variation is steady. Thus, the static approximation is not unreasonable. On the other hand, the process of galactic mergers, especially major mergers, perturbs the galactic environment and GC orbits. However, due to the vibrant nature and lack of knowledge on the process, we leave it to future research. The eccentricity of GC orbits also differs from the circularity approximation. However, to accurately capture realistic GC orbital evolution along galaxy assembly histories can be computationally expensive (e.g. Li et al. 2017; Li, Gnedin & Gnedin 2018; Chen & Gnedin 2022). We could only treat our method as a time-averaged approximation to real eccentric GC orbits. However, as very eccentric orbits usually happens for *ex situ* GCs, the approximation does not significantly affect our results on the properties of the NSC. With improving powers of *N*-body simulation on galactic mergers and dynamical friction, more knowledge will enable us to incorporate these effects into a more comprehensive model.

In investigating the formation of NSCs, we didn't take into account the *in situ* channel of young stars forming in nuclear regions, despite various observational evidence as mentioned in 1. Our model partially takes care of this channel, as the GC spatial distribution at formation can sometimes sample GCs at the centre of the host galaxies. Nevertheless, a more systematic investigation is warranted to obtain a thorough picture of NSC formation.

Besides fitting the NSC mass, future works will be carried out on investigating other NSC properties, such as the age/metallicity distribution and the internal mass profile of the NSC. Besides MW and M31-like galaxies, we will extend our investigation to a broader galaxy mass range and to different galaxy types.

There are also caveats in modelling the MSPs, as their distribution and evolution in GCs remains highly unknown to us. As stellar density increases towards the GC centre, LMXBs and MSPs are supposed to peak near the centre. There are evidences both from observations and simulation (see Ye et al. (2019) and references therein). Thus, the number of MSPs stripped and deposited to the ambient environment depends on the current size of the GC. Our treatment essentially assumes a uniform distribution of MSPs in GCs, which only serves as a lower limit to the MSP contribution to the galaxy centre γ-ray excess. On the other hand, following Ye et al. (2019) and Ye & Fragione (2022) studied the γ-ray excess problem by depositing MSPs only when the GC is fully disrupted. This methodology goes to another extreme by assuming all MSPs residing at the very centre. Further knowledge on the distribution of MSPs inside GCs will enable us to better evaluate their γ-ray contribution to the galaxy centre.

Furthermore, we follow Fragione, Antonini & Gnedin (2018, 2019) to assume the effect of new MSP formation cancelling MSP spin-down in GCs, which is somewhat arbitrary. More knowledge is required to properly evaluate these competing factors.

5 CONCLUSIONS

In this study, we have presented a comprehensive model of GC formation and evolution, based on the premise that they primarily form following periods of rapid halo mass accretion. Leveraging the results from the Illustris cosmological simulation, we sample GCs across the galaxy assembly history and simulate their subsequent evolution, accounting for the mass loss and radial migration within an evolving galactic background. Our model successfully reproduces key observations of the MW and M31 GC system at $z = 0$, including the mass scaling of the total GC system with the host halo, and the spatial distribution of GC number density. For the MW, we also reproduced the spatial distribution for the *in situ* sub-population.

With this model at hand, we investigate the spatial distribution of deposited masses of migrated GCs to study its link to the formation of the NSC and galaxy centre γ-ray excess. We find that both NSC masses of the MW and M31 can be reproduced. Detailed spatial distribution of the GCE can also come entirely from deposited MSPs. However, the M31 excess strength is three times as large as our most luminous candidate galaxy. Even factoring in *in situ* MSPs born at the galaxy centre, the MSP channel still cannot fully account for the excess emission. It becomes evident that DM must play a role in explaining the M31 excess, highlighting a fundamental astrophysical difference between the two galaxies. This constitutes another big difference between them, apart from their galaxy centre SMBHs differing in mass by about 50 times. Further investigations are demanded in figuring out the causes to and possible links between these differences.

Another intriguing aspect we discovered is the influence of galaxy assembly history on galaxy properties, which we investigated using halo half mass redshift z_{hm}. Interestingly, we found that it does not correlate with halo mass, but conveys valuable information about the GC system and NSC mass. Specifically, EFGs with large z_{hm} give rise to an old, heavy, and concentrated GC population as they formed, and vice versa. This results in more deposited mass from GCs and a heavier NSC, which in turn serves as an informative indicator of galaxy assembly history.

In conclusion, our comprehensive model of GC formation and evolution provides a robust framework for understanding the properties of the GC system, the NSC and galaxy centre γ-ray emissions. Additionally, our study unveils the significance of galaxy assembly history in shaping galaxy properties. However, the need to invoke DM to explain the M31 excess emphasizes the distinct astrophysical origins of these high energy emissions from the two similar galaxies. Further investigations are warranted to unravel the precise mechanisms that drive these differences and establish a comprehensive understanding of galaxy formation and evolution.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. The data sets were derived from the Illustris simulation results in the public domain: (The Illustris Collaboration, https://www.illustris-project.org/data/). The observation data of the MW and M31 are available via corresponding references in the article.

REFERENCES

Abazajian K. N., Canac N., Horiuchi S., Kaplinghat M., 2014, Phys. Rev. D, 90, 023526

Ackermann M. et al., 2017a, APJ, 836, 208

Ackermann M. et al., 2017b, ApJ, 840, 43

Aharon D., Perets H. B., 2015, ApJ, 799, 185

Ajello M. et al., 2016, ApJ, 819, 44

Antonini F., Capuzzo-Dolcetta R., Mastrobuono-Battisti A., Merritt D., 2012, ApJ, 750, 111

Arca-Sedda M., Gualandris A., 2018, MNRAS, 477, 4423

Baumgardt H., Sollima A., Hilker M., Bellini A., Vasiliev E., 2021, Fundamental Parameters of Galactic Gobular Clusters (as of May 2021). [online] Available at: <https://people.smp.uq.edu.au/HolgerBaumgardt/globular/)

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Bender R. et al., 2005, ApJ, 631, 280

Bhattacharya D., van den Heuvel E. P. J., 1991, Phys. Rep., 203, 1

Böker T., van der Marel R. P., Mazzuca L., Rix H.-W., Rudnick G., Ho L. C., Shields J. C., 2001, AJ, 121, 1473

Böker T., Laine S., van der Marel R. P., Sarzi M., Rix H.-W., Ho L. C., Shields J. C., 2002, AJ, 123, 1389

Brandt T. D., Kocsis B., 2015, ApJ, 812, 15

Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240

Calore F., Cholis I., McCabe C., Weniger C., 2015a, Phys. Rev. D, 91, 063003

Calore F., Cholis I., Weniger C., 2015b, J. Cosmology Astropart. Phys., 2015, 038

Capuzzo-Dolcetta R., 1993, ApJ, 415, 616

Capuzzo-Dolcetta R., Mastrobuono-Battisti A., 2009, A&A, 507, 183

Carson D. J., Barth A. J., Seth A. C., den Brok M., Cappellari M., Greene J. E., Ho L. C., Neumayer N., 2015, AJ, 149, 170

Chen Y., Gnedin O. Y., 2022, MNRAS, 514, 4736

Chen Y., Gnedin O. Y., 2023, MNRAS, 522, 5638

Choksi N., Gnedin O. Y., 2019, MNRAS, 486, 331

Choksi N., Gnedin O. Y., Li H., 2018, MNRAS, 480, 2343

Cholis I., Hooper D., Linden T., 2015, J. Cosmology Astropart. Phys., 2015, 043

Cholis I., Zhong Y.-M., McDermott S. D., Surdutovich J. P., 2022, Phys. Rev. D, 105, 103023

Conroy C., Gunn J. E., 2010, ApJ, 712, 833

Daylan T., Finkbeiner D. P., Hooper D., Linden T., Portillo S. K. N., Rodd N. L., Slatyer T. R., 2016, Phys. Dark Universe, 12, 1

Deason A. J. et al., 2021, MNRAS, 501, 5964

Dehnen W., Binney J., 1998, MNRAS, 294, 429

Di Mauro M., Hou X., Eckner C., Zaharijas G., Charles E., 2019, Phys. Rev. D, 99, 123027

Do T. et al., 2019, Science, 365, 664

Eckner C. et al., 2018, ApJ, 862, 79

El-Badry K., Quataert E., Weisz D. R., Choksi N., Boylan-Kolchin M., 2019, MNRAS, 482, 4528

Fahrion K., Leaman R., Lyubenova M., van de Ven G., 2022, A&A, 658, A172

Fakhouri O., Ma C.-P., 2008, MNRAS, 386, 577

Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267

Feng L., Li Z.-Y., Su M., Tam P.-H. T., Chen Y., 2019, Res. Astron. Astrophys., 19, 046

Filippenko A. V., Ho L. C., 2003, ApJ, 588, L13

Forbes D. A. et al., 2018, Proc. R. Soc. Lond., 474, 20170616

Fragione G., Antonini F., Gnedin O. Y., 2018, MNRAS, 475, 5313

Fragione G., Antonini F., Gnedin O. Y., 2019, ApJ, 871, L8

Galleti S., Federici L., Bellazzini M., Buzzoni A., Pecci F., 2006, A&A, 456, 985

Genzel R. et al., 2015, ApJ, 800, 20

Georgiev I. Y., Böker T., 2014, MNRAS, 441, 3570

Georgiev I. Y., Puzia T. H., Goudfrooij P., Hilker M., 2010, MNRAS, 406, 1967

Georgiev I. Y., Böker T., Leigh N., Lützgendorf N., Neumayer N., 2016, MNRAS, 457, 2122

Gieles M., Baumgardt H., 2008, MNRAS, 389, L28

Gnedin O. Y., Ostriker J. P., Tremaine S., 2014, ApJ, 785, 71

Gordon C., Macías O., 2013, Phys. Rev. D, 88, 083521

Guillard N., Emsellem E., Renaud F., 2016, MNRAS, 461, 3620

Haggard D., Heinke C., Hooper D., Linden T., 2017, J. Cosmology Astropart. Phys., 2017, 056

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

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

Hartmann M., Debattista V. P., Seth A., Cappellari M., Quinn T. R., 2011, MNRAS, 418, 2697

Hooper D., Slatyer T. R., 2013, Phys. Dark Universe, 2, 118

Hopkins P. F., Quataert E., 2010, MNRAS, 405, L41

Hudson M. J., Harris G. L., Harris W. E., 2014, ApJ, 787, L5

Hughes M. A. et al., 2005, AJ, 130, 73

Kacharov N., Neumayer N., Seth A. C., Cappellari M., McDermid R., Walcher C. J., Böker T., 2018, MNRAS, 480, 1973

Kafle P. R., Sharma S., Lewis G. F., Robotham A. S. G., Driver S. P., 2018, MNRAS, 475, 4043

Kim S. C., Sung H., Park H. S., Sung E.-C., 2004, Chinese J. Astron. Astrophys., 4, 299

Kormendy J., 1985, ApJ, 292, L9

Lacroix T., Macias O., Gordon C., Panci P., Bœhm C., Silk J., 2016, Phys. Rev. D, 93, 103004

Leveque A., Giersz M., Arca-Sedda M., Askar A., 2022, MNRAS, 514, 5751

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

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

Light E. S., Danielson R. E., Schwarzschild M., 1974, ApJ, 194, 257

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119

Lotz J. M., Telford R., Ferguson H. C., Miller B. W., Stiavelli M., Mack J., 2001, ApJ, 552, 572

Lyubenova M., Tsatsi A., 2019, A&A, 629, A44

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940

Magdis G. E. et al., 2012, ApJ, 758, L9

Márquez I., Lima Neto G. B., Capelato H., Durret F., Gerbal D., 2000, A&A, 353, 873

Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4

Matthews L. D., Gallagher John S. I., 1997, AJ, 114, 1899

McDaniel A., Jeltema T., Profumo S., 2018, prd, 97, 103021

McLaughlin D. E., King A. R., Nayakshin S., 2006, ApJ, 650, L37

McMillan P. J., 2017, MNRAS, 465, 76

Milosavljević M., 2004, ApJ, 605, L13

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution, Cambridge Univ. Press, Cambridge

Muratov A. L., Gnedin O. Y., 2010, ApJ, 718, 1266

Navarro J. F., Frenk C. S., White S. D. M., 1997, Astrophys. J., 490, 493

Nelson D. et al., 2015, Astron. Comput., 13, 12

Neumayer N., Seth A., Böker T., 2020, A&A Rev., 28, 4

Nguyen D. D. et al., 2019, ApJ, 872, 104

Noyola E., Gebhardt K., Bergmann M., 2008, ApJ, 676, 1008

Peebles P. J. E., 1980, The Large-scale Structure of the Universe. Princeton Univ. Press Princeton, NJ

Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, ARA&A, 48, 431

Prugniel P., Simien F., 1997, A&A, 321, 111

Ransom S. M., 2007, Pulsars in Globular Clusters, Proceedings of the International Astronomical Union. Cambridge University Press, Vol. 3(S246), p. 291

Sánchez-Janssen R. et al., 2019, ApJ, 878, 18

Schulz C., Pflamm-Altenburg J., Kroupa P., 2015, A&A, 582, A93

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Shlosman I., Begelman M. C., Frank J., 1990, Nature, 345, 679

Spengler C. et al., 2018, ApJ, 869, 85

Spitler L. R., Forbes D. A., 2009, MNRAS, 392, L1

Tacconi L. J. et al., 2018, ApJ, 853, 179

Terzić B., Graham A. W., 2005, MNRAS, 362, 197

Tremaine S. D., Ostriker J. P., Spitzer L. J., 1975, ApJ, 196, 407

Tsatsi A., Mastrobuono-Battisti A., van de Ven G., Perets H. B., Bianchini P., Neumayer N., 2017, MNRAS, 464, 3720

Vogelsberger M. et al., 2014, Nature, 509, 177

Voss R., Gilfanov M., 2007, MNRAS, 380, 1685

Wang L., Lin D. N. C., 2023, ApJ, 944, 140

Wilson C. D., Harris W. E., Longden R., Scoville N. Z., 2006, ApJ, 641, 763

Ye C. S., Fragione G., 2022, ApJ, 940, 162

Ye C. S., Kremer K., Chatterjee S., Rodriguez C. L., Rasio F. A., 2019, ApJ, 877, 122

Zimmer F., Macias O., Ando S., Crocker R. M., Horiuchi S., 2022, MNRAS

APPENDIX A: DETERMINING HALO PARAMETERS

The NFW profile is described in Navarro, Frenk & White (1997) as

$$\rho_{\mathrm{NFW}}(r) = \rho_0 \left(\frac{r}{R_{\mathrm{s}}} \right)^{-1} \left(1 + \frac{r}{R_{\mathrm{s}}} \right)^{-2}. \tag{A1}$$

The corresponding potential is given by

$$\Phi_{\mathrm{NFW}}(r) = -\frac{4\pi \rho_0 R_{\mathrm{s}}^3 G}{r} ln \left(1 + \frac{r}{R_{\mathrm{s}}} \right). \tag{A2}$$

Here ρ_0 is normalized by the virial mass $M_{\mathrm{vir}} = 4\pi \rho_0 R_{\mathrm{s}}^3 [ln(1 + c) - \frac{c}{1+c}]$, where the halo concentration $c = R_{\mathrm{vir}}/R_s$ is taken from Macciò, Dutton & van den Bosch (2008) with the analytic form

$$c = 9.354 \left(\frac{M_{\mathrm{vir}} h}{10^{12} \, \mathrm{M}_\odot} \right)^{-0.094}. \tag{A3}$$

The virial radius R_{vir} is mapped from M_{vir} and z via

$$R_{\mathrm{vir}} = \frac{163}{(1+z) h} \left(\frac{M_{\mathrm{vir}} h}{10^{12} \, \mathrm{M}_\odot} \right)^{\frac{1}{3}} \left(\frac{\Delta_{\mathrm{vir}}}{200} \right)^{-\frac{1}{3}} \Omega_{\mathrm{m},0}^{\frac{1}{3}} \,\, \mathrm{kpc}. \tag{A4}$$

Here Δ_{vir} is the average halo over-density at R_{vir}, which we take from the spherical collapse model as

$$\Delta_{\mathrm{vir}} = \frac{18\pi^2 + 82x - 39x^2}{x + 1}, \,\, x \equiv \Omega_{\mathrm{m}}(z) - 1. \tag{A5}$$

In our adopted cosmological model, the evolution of matter density parameter equals (Mo, van den Bosch & White 2010)

$$\Omega_{\mathrm{m}}(z) = \frac{\Omega_{\mathrm{m},0}(1+z)^3}{\Omega_{\Lambda,0} + \Omega_{\mathrm{m},0}(1+z)^3}. \tag{A6}$$

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