Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **513,** 4107–4129 (2022)
Advance Access publication 2022 April 29

https://doi.org/10.1093/mnras/stac1145

The chemo-dynamical groups of Galactic globular clusters

Thomas M. Callingham [1,2★] Marius Cautun [3] Alis J. Deason [1] Carlos S. Frenk [1] Robert J. J. Grand [4,5] and Federico Marinacci [6]

[1] *Institute of Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2] *Kapteyn Astronomical Institute, University of Groningen, Landeleven 12, NL-9747 AD Groningen, the Netherlands*
[3] *Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[4] *Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, E-38205 La Laguna, Tenerife, Spain*
[5] *Departamento de Astrofísica, Universidad de La Laguna, Av. del Astrofísico Francisco Sánchez s/n, E-38206 La Laguna, Tenerife, Spain*
[6] *Department of Physics & Astronomy 'Augusto Righi', University of Bologna, via Gobetti 93/2, I-40129 Bologna, Italy*

Accepted 2022 April 21. Received 2022 April 6; in original form 2022 February 22

ABSTRACT

We introduce a multicomponent chemo-dynamical method for splitting the Galactic population of globular clusters (GCs) into three distinct constituents: bulge, disc, and stellar halo. The latter is further decomposed into the individual large accretion events that built up the Galactic stellar halo: the *Gaia*–Enceladus–Sausage, Kraken and Sequoia structures, and the Sagittarius and Helmi streams. Our modelling is extensively tested using mock GC samples constructed from the AURIGA suite of hydrodynamical simulations of Milky Way (MW)-like galaxies. We find that, on average, a proportion of the accreted GCs cannot be associated with their true infall group and are left ungrouped, biasing our recovered population numbers to ~ 80 per cent of their true value. Furthermore, the identified groups have a completeness and a purity of only ~ 65 per cent. This reflects the difficulty of the problem, a result of the large degree of overlap in energy-action space of the debris from past accretion events. We apply the method to the Galactic data to infer, in a statistically robust and easily quantifiable way, the GCs associated with each MW accretion event. The resulting groups' population numbers of GCs, corrected for biases, are then used to infer the halo and stellar masses of the now defunct satellites that built up the halo of the MW.

Key words: methods: numerical – Galaxy: halo – galaxies: haloes – galaxies: kinematics and dynamics.

1 INTRODUCTION

Our stellar halo is a cosmic graveyard populated by the stars and globular clusters (GCs) that were once part of now destroyed dwarf galaxies. Halo assembly stems from hierarchical growth – the hallmark of the Λ Cold Dark Matter (ΛCDM) cosmological model (Davis et al. 1985) – whereby massive galaxies like the Milky Way (MW) evolve by devouring many lower mass galaxies, whose remains are mixed and spread into the stellar halo (e.g. Bullock & Johnston 2005; Cooper et al. 2010). Unravelling this galactic debris to reconstruct the assembly history of the MW is a difficult undertaking as ancient mergers have long since phase-mixed, effectively erasing information in physical space. However, simulation-based studies have shown that debris from the same progenitor remains localized, preserving structure in the space of the integrals of motion (e.g. Gómez et al. 2010). Combined with stellar age and chemistry information, which also persists over time, this raises the prospect that we may be able to reconstruct our Galaxy's past.

Of the accreted material in the stellar halo, GCs have long been recognized as sensitive probes of the accretion history of the MW (Searle & Zinn 1978). Several GCs are suspected of being the remnant nucleus of accreted dwarf galaxies (M54, M4, ω-Centauri, NGC 1851), directly showing where the cores of fallen progenitors came to rest. Furthermore, while major mergers dominate the stellar halo (Cooper et al. 2010; Deason, Mao & Wechsler 2016; Fattahi et al. 2020), it has been shown that GCs are generally associated with smaller accretion events in the MW's past (e.g. Harris, Harris & Hudson 2015; Amorisco 2019). When studying the origin of the MW's GC system, it is necessary to identify which of them were born natively in our Galaxy (*in situ* GCs) and which formed in dwarf galaxies and were later accreted.

On average, in the MW there is a rough trend for metal-poor GCs to be located at a larger radius, while metal-rich GCs are more centrally concentrated (Frenk & White 1980). However, this is not enough to distinguish populations by chemistry alone (Trujillo-Gomez et al. 2021). With precise age and metallicity data now available for many GCs, it has been shown that the MW GC's age–metallicity relation (AMR) contains two branches: a metal-poor one characterized by halo-like kinematics, and a metal-rich one whose GCs orbit the inner Galaxy, suggesting an *in situ* origin (Marín-Franch et al. 2009; Forbes & Bridges 2010; Leaman, VandenBerg & Mendel 2013). This behaviour can be understood using simple models, such as a leaky-box chemical enrichment model, in which the stellar birth environment in smaller dwarf galaxies is enriched more slowly than in larger galaxies such as our own.

The recent explosion of Galactic data, such as those from the *Gaia* mission (Gaia-Collaboration 2018), APOGEE (Majewski et al. 2017), the H3 survey (Conroy et al. 2019), and GALAH (Martell

★ E-mail: t.m.callingham@astro.rug.nl

et al. 2017), has revolutionized the field of Galactic astronomy. In particular, they have revealed evidence of an ancient major merger, *Gaia–Enceladus–Sausage* (GES) (Belokurov et al. 2018; Helmi et al. 2018). Combined with previous discoveries such as the stellar stream of the Sagittarius dwarf galaxy (Ibata, Gilmore & Irwin 1994) and the Helmi stream (Helmi et al. 1999), there is a wealth of known structures present in the Galactic stellar halo (Naidu et al. 2020). Characterizing the properties of the progenitors of these structures is challenging since their debris consists of extended, diffuse stellar distributions. One solution is to identify the GCs associated with these structures, since GCs are compact, bright objects whose properties and orbits can be measured accurately.

Arguably the easiest accreted group to identify is the set of GCs that belonged to the Sagittarius dwarf galaxy (Ibata et al. 1994) as it is currently being disrupted and its stars and GCs can be found as an identifiable stream (e.g. Law & Majewski 2010; Antoja et al. 2020; Bellazzini et al. 2020; Peñarrubia & Petersen 2021). Identifying members of other stellar halo structures remains a challenging problem. The works of Myeong et al. (2018a, c) have associated GCs with the GES debris, which is characterized by highly radial orbits. Likely members of the Helmi stream were identified by Koppelman et al. (2019a) from their proximity to selection cuts in the phase space of the stellar halo. The retrograde accretion event dubbed 'Sequoia' was, in part, born out of studies of notable retrograde GCs such as FSR1758 and ω-Centauri (Myeong et al. 2018b, 2019; Barba et al. 2019), with other GCs similarly associated.

The recent work by Massari, Koppelman & Helmi (2019), hereafter Massari19, was a significant development in this field. These authors used a sample of 160 Galactic GCs to identify the major GC groups. They did so by defining selection boxes in energy and angular momentum space that are based on 'known' accretion groups and expanding to include all likely GCs members. GCs leftover from this process without a clear accretion origin were divided into a high-energy group, which is likely a collection of smaller accretion events, and a lower energy group that was thought potentially to be a signature of an ancient accretion event. This GC grouping has been refined by Horta et al. (2020), hereafter Horta20, who have added APOGEE alpha element abundances for 46 inner GCs to make minor revisions.

The low-energy group of Massari19 is consistent with the Kraken event predicted by Kruijssen et al. (2019b, 2020) to be the MW's most ancient merger. This work identified the structure by comparing the observed distribution of MW GCs with the predictions of the EMOSAICs hydrodynamic simulations of GC formation and evolution (Pfeffer et al. 2018; Kruijssen et al. 2019a). This merger is likely the same as or significantly overlapping with the one that gave rise to the Koala structure of Forbes (2020), hereafter Forbes20, and the Inner Galaxy System (or later Heracles) of Horta et al. (2021). In this paper, we refer to this accretion event as Kraken.

Once the accretion groups of GCs have been identified, the number of GCs, and the age–metallicity and the dynamical distributions of the GCs can all provide information about the progenitor galaxy. The GC AMR relation provides clues to the formation time and the chemical enrichment of the progenitor dwarf (Forbes20), while groups of GCs with smaller apocentres indicate an ancient or massive merger (Pfeffer et al. 2020). Using these techniques, the Massari19 GC group memberships have been used in studies such as those by Forbes20 and Trujillo-Gomez et al. (2021) to reverse engineer the assembly history of the MW. Combined with insights from the EMOSAICs project, Kruijssen et al. (2020) used these groups to suggest that the MW has experienced two to three major mergers, and at least 15 smaller mergers contributing GCs in total.

The number of GCs in a progenitor galaxy is related to its mass. For LMC-mass and more massive galaxies, observations have revealed a linear relationship between the number of or total mass of GCs and the halo mass of the host galaxy (Forbes et al. 2018). Theoretical models reproduce this trend (e.g. Boylan-Kolchin 2017; Bastian et al. 2020; Burkert & Forbes 2020). However, it is unclear if this relation holds for dwarf galaxies with stellar masses below 10^9 $M_\odot$. Observationally, it is difficult to measure the halo mass of such systems, and theoretical predictions in this range often do not agree with one another. At lower masses, analytical models based on hierarchical clustering predict a continuation of the linear relation between GC mass and total halo mass (e.g. Boylan-Kolchin 2017), while the EMOSAICs project predicts a linear relation with stellar mass instead of halo mass (Bastian et al. 2020).

One limitation of the current GC groupings is that they are defined in a rather subjective way, mostly by eye. This methodology raises questions about whether the current groupings are statistically robust and physically relevant. Furthermore, subjective methods are very difficult to test using mock catalogues, but this represents an essential analysis step to trust the results (e.g. see Wu et al. 2021). Alternatively, recent work has seen the use of clustering algorithms to find structures in the halo (e.g. Helmi et al. 2017; Myeong et al. 2018b; Koppelman et al. 2019b; Necib et al. 2020; Ostdiek et al. 2020). These should give more objective, quantifiable results, but as noted in Naidu et al. (2020), it can be challenging to tune these clustering methods to the astrophysical problem of identifying groups of accreted material. A few studies have applied these sorts of techniques to GCs specifically. Examples include the use of a friends-of-friends clustering algorithm to associate GCs with the Sequoia merger (Myeong et al. 2018c) and the decomposition of GCs in the centre of our galaxy into bulge, disc, and halo components (Pérez-Villegas et al. 2020). However, we know of no studies that have yet been applied to the total Galactic population of GCs.

In this paper, we develop an objective methodology combining chemo-dynamical information to identify the likely progenitors of the full population of Galactic GCs. By fitting models to both the dynamical distribution in action space and the AMR of the accreted galaxy, we calculate membership probabilities for each GC and statistically link them to particular accretion events. We do so by modelling the GCs as a combination of bulge, disc, and halo components, the latter representing the focus of our study. The stellar halo is further decomposed into the massive merger events that built it, such as GES, Kraken, and Sagittarius, and an ungrouped component coming from lower mass mergers that did not contribute enough GCs to be robustly identified. This methodology is extensively tested and characterized using mock GC catalogues built from the AURIGA suite of hydrodynamical simulations (Grand et al. 2017). We apply the method to the Galactic GCs and fully account for observational errors to identify the most likely GCs associated with each merger event. Using these membership probabilities, properties of the progenitor galaxies, such as halo and stellar masses, are derived.

The structure of the paper is as follows. In Section 2, we describe our chemo-dynamical mixture model. Section 3 describes the construction of our mock globular GCs catalogues from AURIGA haloes. In Section 4, we apply our method to the mocks. In Section 5, we apply our method to the MW and discuss the individual cluster fits. We discuss the resulting implications for the MW's accretion history in Section 6. Finally, Section 7 summarizes and concludes the paper.

2 MULTICOMPONENT MODEL FOR THE GALACTIC GC POPULATION

We model the MW population of GCs as a combination of a bulge, disc, and stellar halo components. The latter is the main focus of our work and is further split into subgroups that correspond to all known major accretion events, such as GES and Kraken. The decomposition is performed using an expectation–maximization algorithm applied to chemo-dynamical data, that is combining age–metallicity information with orbital integrals of motions (i.e. action space). This section presents a detailed description of the decomposition method and its motivation.

For a general space, $\boldsymbol{X}$, which represents a combination of metallicity and action quantities, each GC component is modelled as a distribution, $F_{\rm c}(\mathbf{X}) \equiv F(\boldsymbol{X}|\boldsymbol{\theta}_c)$, specified in terms of a set of model parameters, $\boldsymbol{\theta}_c$, whose details will be given when discussing each model component. $F_{\rm c}(\mathbf{X})$ is normalized to integrate to 1 over the space $\mathbf{X}$. Then, the multicomponent model describing the overall population of GCs is written as the sum over each individual component:

$$F(\mathbf{X}) = \sum_{c}^{\rm Com} W_c F_c(\boldsymbol{X}) \,, \tag{1}$$

where W_c denotes the weight of component c and specifies the fraction of the GC population contributed by each component. The total distribution, $F(\mathbf{X})$, is normalized to unity over the space, which implies that

$$\sum_{c}^{\rm Com} W_c = 1 \,. \tag{2}$$

The probability that the i-th GC belongs to component c, which is often referred to as the 'responsibility' in multicomponent models, such as Gaussian mixture models, is given by

$$r_{ic} = \frac{W_c F_c(\boldsymbol{X}_i)}{\sum_{c'} W_{c'} F_{c'}(\boldsymbol{X}_i)} \equiv \frac{p_{ic}}{\sum_{c'} p_{ic'}} \,, \tag{3}$$

where $\boldsymbol{X}_i$ denotes the coordinates of the i-th GC in the chemo-dynamical space used to identify the different populations. For brevity, we also introduced the notation, $p_{ic} \equiv W_c F_c(\boldsymbol{X}_i)$, which gives the value of the F_c distribution at $\boldsymbol{X}_i$ multiplied by the weight of that component. The total log-likelihood, $\ln \mathcal{L}$, of the mixture model is given as

$$\ln \mathcal{L} = \sum_{i}^{\rm GCs} \ln F(\boldsymbol{X}_i) \equiv \sum_{i}^{\rm GCs} \ln \left(\sum_{c'}^{\rm Com} p_{ic'} \right) \,, \tag{4}$$

wherein the rightmost term the first sum is over all the GCs in the system and the second sum is over all components of the model. To find the maximum likelihood estimate, we need to find the maximum of $\mathcal{L}$ for the set of parameters $\{\boldsymbol{\theta}_c\} \equiv \{\boldsymbol{\theta}_{c=1}, \boldsymbol{\theta}_{c=2}, ..., \boldsymbol{\theta}_{c=K}\}$, where K is the number of components and each $\boldsymbol{\theta}_c$ is, in turn, a set of multiple parameters. For example, if we model a component as a Gaussian distribution, then $\boldsymbol{\theta}_c$ is the combination of peak position along each coordinate axis in $\boldsymbol{X}$-space and the corresponding covariance matrix. The maximization procedure is further complicated by the fact that the W_c weights that appear in the $p_{ic'}$ expression depends on the values of all the $\{\boldsymbol{\theta}_c\}$ parameters which makes for a very non-linear and multidimensional maximization procedure.

To solve this challenge, we use the expectation–maximization approach. This algorithm is often used to fit Gaussian mixture models efficiently. As explained below, our methodology is similar to this but adapted to include relevant astrophysics such as the AMR of the component. The algorithm corresponds to an iterative approach for finding the maximum likelihood and has the following steps:

(i) Initialization:
An initial guess is made for the responsibilities, r_{ic}. The outcome can be dependent on this initial choice. This dependence is tested and discussed in Section 4.

(ii) Maximization step:
In this step, we assume that the responsibilities, r_{ic}, are known, and we find the $\{\boldsymbol{\theta}_c\}$ parameters that maximize the log-likelihood, $\ln \mathcal{L}$, for fixed r_{ic} values. The advantage is that once the r_{ic} are known, maximizing $\ln \mathcal{L}$ reduces to a much simpler problem in which the parameters of one component are independent of the parameters of the remaining components. For component c, $\ln \mathcal{L}$ is maximal for the $\boldsymbol{\theta}_c$ values that maximize the expression

$$\sum_{i}^{\rm GCs} r_{ic} \ln F_c(\boldsymbol{X}_i) \,. \tag{5}$$

In the above equation, each data point contributes with a weight, r_{ic}, which is why r_{ic} is called the responsibility.

(iii) Expectation step:
The values of the responsibilities are updated using the $\{\boldsymbol{\theta}_c\}$ parameters found in the previous step.

(iv) Iteration:
Repeat the maximization and expectation steps until $\ln \mathcal{L}$ is converged. In practice, we assume convergence when $\ln \mathcal{L}$ changes between consecutive steps by less than 0.001 times the number of GCs.

The space $\mathbf{X}$ we use to identify the components of the GC population is a combination of orbital dynamical quantities, which we denote with $\mathbf{Y}$, and age–metallicity information, which we denote with $\mathbf{Z}$. We assume that the orbital quantities are uncorrelated with the chemistry of GCs, which implies that the distribution function of each component can be split into two independent distributions:

$$F_c(\mathbf{X}) = F_c^{\rm dyn}(\mathbf{Y})\, F_c^{\rm AMR}(\mathbf{Z}) \,. \tag{6}$$

In the following text, we describe how we model the distribution of dynamical quantities, $F_c^{\rm dyn}(\mathbf{Y})$, and of the AMR, $F_c^{\rm AMR}(\mathbf{Z})$, where we drop the superscripts for brevity. These functions are independent, and so can be fit by maximizing their respective likelihoods (with equation 5) independently.

2.1 Dynamical modelling

In this work, we primarily consider a four-dimensional dynamical space consisting of the orbital energy and the three orbital actions: the component of the angular momentum perpendicular on the disc plane, L_z, the radial action, J_R, and vertical action, J_z. The integrals of motion, $\boldsymbol{J}$, completely describe the orbit, which determines the orbital energy (for more information see Binney & Tremaine 2008). This means that in the $(E, \boldsymbol{J})$ four-dimensional space all orbits, including those of our GCs, lie on a three-dimensional surface. This suggests that the energy only contains redundant information about the orbits; however, tests on mock catalogues show that the combined $(E, \boldsymbol{J})$ space leads to a more accurate identification of GC populations than $(\mathbf{J})$ space, justifying our choice (more on this in Section 4).

2.1.1 Accreted GCs

The accreted components are modelled as multivariate Gaussian distributions in the $\boldsymbol{Y} = (E, L_z, J_R, J_z)$ space through

$$F_c(\mathbf{Y}) = N(\mathbf{Y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^{n_{\rm dim}}|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{Y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right), \quad (7)$$

where $n_{\rm dim} = 4$ is the number of dimensions of the space, $\mathbf{Y}$, $\boldsymbol{\mu}$ is the mean, and $\boldsymbol{\Sigma}$ is the covariance matrix. The values of these parameters that maximize the total model likelihood can be found analytically from equation (5) by calculating moments of the distribution.

In reality, not all the accreted material from a single merger event will necessarily be well represented by a Gaussian distribution. Typically, the bulk of the material is often centred around the orbit of the accreting galaxy, and can be well described by a single Gaussian component. However, some of the material can be in more complex substructures formed in accretion, such as leading or trailing stream arms of a stream, which can have a different dynamical distribution. It should be noted that it is likely that material very near, or on, the boundaries of our chosen dynamical space $(E, \mathbf{J})$ (such as the maximally circular orbits) can be poorly described. However, we find that due to the relatively small number of GCs, alternative 'assumption free' distributions, such as density kernels, do not work effectively, and it is necessary to assume a form for the distribution.

If the number of points to which an unconstrained multivariate Gaussian distribution is fit, $N_{\rm points}$, is equal to or less than $n_{\rm dim}$ then the covariance matrix becomes degenerate with some eigenvalues equalling zero (or infinitesimal). The corresponding principle axes then have infinitesimal width, which can give unrealistically large probability values. For example, in two-dimensional space, two points will be fit as a line, with the fit and grouping unable to develop further. To prevent this, we fix the value of the smallest principal axis using the procedure described in Appendix B.

2.1.2 Ungrouped GCs

Some GCs cannot be attributed to any known accretion event, such as the High Energy group in the Massari19 analysis. This could be because they fell in as small groups that do not contain enough information to be robustly identified. Alternatively, the GC's orbit could have evolved such that it no longer resembles those of the rest of the group. Our model accounts for such GCs which are classified as the 'ungrouped' component.

The ungrouped component is modelled as a uniform background distribution, normalized to integrate to one over the convex hull volume, V, of the dynamical space filled by all of the GCs. That is,

$$F_{\rm Ung} = \frac{1}{V}, \quad (8)$$

where V is calculated using SCIPY's convex hull module (Virtanen et al. 2020).

2.1.3 In situ components

In the MW, we cannot be certain if the GCs are accreted or have an *in situ* origin. Therefore, we need to include models of the bulge and disc components. The dynamics of these components are not well described by Gaussians, and instead, we model them as distribution functions in action space using the implementations in AGAMA (Vasiliev 2019).

When modelling the bulge and disc components in $(E, \mathbf{J})$ space, we assume that the energy distribution can be separated from the action distribution, that is,

$$F(E, \mathbf{J}) = F(E)\, F(\mathbf{J}). \quad (9)$$

The energy distribution is calculated numerically from the prescribed action distribution of the components.

We originally modelled the action distribution of the bulge as a double power law with a cut-off as introduced in Posti et al. (2015). In practice, we found that the fitting converges on values consistent with the simpler exponential fit:

$$F_{\rm Bulge}(\mathbf{J}) = \frac{4}{J_{\rm Cut}^3\sqrt{3\pi^3}} \exp\left[-(J_{\rm Tot}/J_{\rm cut})^2\right], \quad (10)$$

where $J_{\rm Tot} = J_R + |L_z| + J_z$ and $J_{\rm cut}$ is a free parameter that controls the steepness of the cut-off.

The disc is modelled using the quasi-isothermal disc, first described in Binney (2010). This is also used to model GCs in Posti & Helmi (2019), whose assumptions we follow. The action distribution is given as

$$F_{\rm Disc}(\mathbf{J}) = \frac{\Sigma \nu \Omega}{2\pi^2 \kappa \sigma_R^2 \sigma_z^2} f_{\pm,d} \exp\left(-\frac{\kappa J_{\rm R}}{\sigma_R^2} - \frac{\nu J_z}{\sigma_z^2}\right)$$
$$\Sigma = \exp\left[-R_c(L_z)/R_d\right]$$
$$f_{\pm,d} = \begin{cases} 1 & L_z \geq 0 \\ \exp\left(2\Omega L_z/\sigma_R^2\right) & L_z < 0 \end{cases}, \quad (11)$$

where Σ describes the disc surface density and $f_{\pm,d}$ controls the rotation of the disc. The circular, radial, and vertical epicycle frequencies are denoted by Ω, κ, and ν, respectively, and are evaluated at the radius of the circular orbit, $R_c = R_c(J_{\rm Tot})$, with angular momentum $J_{\rm Tot} = J_R + |L_z| + J_z$. The radial velocity dispersion is given as $\sigma_R = \sigma_{R0}\exp(-R_c/R_\sigma)$, and the vertical velocity dispersion is fixed at a constant scale height, $\sigma_z = \sqrt{2}h_d\nu$. The disc is chosen to match the thick disc of Piffl et al. (2014) with $R_\sigma = 13$ kpc and $h_d = 0.2R_d$. This leaves two free parameters: the disc scale length, R_d, and the central radial dispersion, $\sigma_{R,0}$.

2.2 Age–metallicity relation

We use the leaky-box chemical evolution model to describe the AMR for GCs as given in Forbes (2020):

$$[{\rm Fe/H}] = -p_{\rm yield} \log\left(\frac{t}{t_f}\right), \quad (12)$$

where $p_{\rm yield}$ is a measure of how quickly the system enriches and t_f is the formation time of the system. Larger galaxies enrich in metallicity faster, giving a higher $p_{\rm yield}$ and steeper evolutionary track. Note that this is equivalent to equation (4) of Kruijssen et al. (2019a), with rearranged and renamed constants, and is similar to the relation of Massari19 (equation 1).

We proceed by fitting equation (12) to the GCs associated with each component taking into account the weights, i.e. the responsibilities, associated with each object. The fitted relation can be inverted to obtain the expected age as a function of metallicity, which we denote as $t_{\rm fit}([{\rm Fe/H}])$. The probability of the GCs observed age being part of the modelled relation is then given by a normal distribution, centred on the expected age with dispersion equal to the error in age, σ_t, i.e.

$$F_c^{\rm AMR}([t, [{\rm Fe/H}]]) = N(t|\mu = t_{\rm fit}([{\rm Fe/H}]), \sigma = \sigma_t). \quad (13)$$

For the GCs that do not have age–metallicity data, we assume that they have a constant probability to be assigned to the component in

the age–metallicity space. This is taken to be the inverse of the range of ages of the GCs (i.e. maximum age–minimum age), similar to the uniform probability of the ungrouped component in dynamical space. For the ungrouped component, we do not expect all group members to be from a single accretion event or follow the same AMR. The probability is then taken as a constant value as if there were no age–metallicity data.

2.3 Observational errors

To model the MW effectively it is necessary to include the statistical uncertainty from observational errors. For this, we use the Monte Carlo method described in Section 5.1 that samples the uncertainties in the measured velocity and position of GCs. The Monte Carlo samples of a single cluster are treated as independent points, with their own responsibilities and are fit independently. When the model has converged, the final probabilities of cluster i is given as

$$p_{ic} = W_c \sum_{j}^{\mathrm{MC}} F_c \left(\boldsymbol{X}_i^j \right), \tag{14}$$

where $\mathbf{X}_i^j$ is j-th Monte Carlo realization of the i-th GC and the sum is over all the Monte Carlo samples of the GC. These probabilities are then used to calculate the responsibilities of the final results, according to equation (3).

3 MOCK CATALOGUES OF GCS

We now describe our construction of mock GC catalogues from the AURIGA hydrodynamical simulations. The AURIGA project consists of a suite of high-resolution cosmological zoom-in simulations of individual MW-like haloes (Grand et al. 2017) with halo masses in the range 1–2 $\times$ $10^{12} M_\odot$. The haloes were selected from the 100^3 Mpc^3 periodic cube of the EAGLE project, a ΛCDM cosmological hydrodynamical simulation (Crain et al. 2015; Schaye et al. 2015) adopting Planck1 (Planck Collaboration I 2014) cosmological parameters. Using the N-body and moving mesh magnetohydrodynamic AREPO code (Springel 2011), these haloes were resimulated to produce a zoom-in simulation of each halo. We selected these simulations because they have been shown to reproduce many properties of the MW and other MW-mass galaxies, such as the satellite luminosity function (Shao et al. 2018; Simpson et al. 2018), stellar bulge and disc structures (Gómez et al. 2017; Grand et al. 2017), and stellar halo (Fattahi et al. 2019; Grand et al. 2019; Monachesi et al. 2019; Deason et al. 2021). We use the level 4 resolution sample, with a DM particle mass of $\sim 3 \times 10^5 M_\odot$ and an initial gas resolution element of mass $\sim 5 \times 10^4 M_\odot$. This sample contains 30 haloes which we label Au1 to Au30.

Of the 30 level 4 AURIGA haloes, 13 are unrelaxed at the present day according to the criteria of Neto et al. (2007). These unrelaxed haloes are poorly modelled by static axisymmetric potentials. This is typically because they are currently, or recently, undergoing a disruptive transient event such as a merger. We therefore restrict our analysis to the 17 relaxed haloes. There is some debate whether the presence of the LMC would cause the MW to be classified as unrelaxed according to the same criteria (Cautun et al. 2019; Erkal, Belokurov & Parkin 2020; Erkal et al. 2021) and is, in fact, poorly modelled by a static axisymmetric potential. We leave the effects of a time-dependent potential to future work.

The AURIGA simulations do not 'natively' contain GCs. To represent groups of accreted GCs we select old accreted stars in the stellar halo (c. f. Halbesma et al. 2020). For each accretion event we identify the accreted stars and randomly assigned GCs to a subsample of them based on the properties of the progenitor galaxy. To assign GCs, we select only accreted halo stars older than 10 Gyr and require them to be within R_{200} of the host galaxy at the present day. This is motivated by age estimates of the MW GCs which are, with a few exceptions, older than 10 Gyr. To determine the origin of the stars, we use the accretion catalogue of stars as Fattahi et al. (2019).

The birthplace of the star is defined as the subhalo in which it resides at the first simulation snapshot (as defined by the SUBFIND algorithm of Springel, Yoshida & White 2001) after its formation. If the star is born in the main halo, it is defined as an *in situ* star. If the star is born outside of the main halo, its origin is defined to be the last subhalo to which it belonged before it fell into the main halo. This prescription identifies the accreted stars that are associated with the accretion event that brought them into the main halo. The few stars that formed from the gas of infalling satellites in the main halo are classified as *in situ*.

To create the *in situ* GCs we generate test particles, with positions and velocities randomly drawn from an action distribution of the *in situ* components using AGAMA (see Section 2). We use the bulge and disc action distributions described in Section 2, fit to the GCs identified in the Massari19 groupings. These action distributions are scaled appropriately by the mass of the AURIGA galaxy (M_{200}^{Au}), such that $F_{\mathrm{Au}}(\mathbf{J}) = F_{\mathrm{MW}}(\lambda \mathbf{J})$, where $\lambda = \left(M_{200}^{\mathrm{MW}} / M_{200}^{\mathrm{Au}} \right)^{2/3}$. We take $M_{200}^{\mathrm{MW}} = 1.17 \times 10^{12}$ M$_\odot$ from Callingham et al. (2019), which also contains further discussion of this mass scaling technique. We create 1000 mock catalogues of accreted and *in situ* clusters for every relaxed AURIGA halo.

3.1 The GC populations

To generate the mocks, we must choose the size of the membership of each GC group. For the *in situ* component, we assume fixed populations of 40 GCs for the bulge and 20 for the disc, motivated by previous groupings in the literature. For the accreted groups, we adopt the Burkert & Forbes (2020) model in which the number of GCs is proportional to the total mass of the host. The mean expected number of GCs, N_{GC}, for an accretion event of mass, M_{Host}, is given by

$$N_{\mathrm{GC}} = \frac{M_{\mathrm{Host}}}{5 \times 10^9 M_\odot} . \tag{15}$$

From this mass–number of GC relation ($M_H - N_{GC}$) we generate 1000 GC mocks for each accreted satellite. To keep the analysis as clear as possible, each random realization has an equal number of GCs given by the mean expectation, rounded to the nearest integer. In principle, we could include scatter on this relation (as given in Burkert & Forbes 2020). However, in these tests, we are principally interested in the changes caused by the sampling of dynamics of the accretion events, not those caused by random variance in the population numbers.

While the expected number of GCs from a single small accretion event (objects of mass less than $5 \times 10^9 M_\odot$) is less than one, we estimate that on average the expected *total* number of GCs from small accretion events is typically approximately five. This population of small accretion events bring in individual, ungrouped GCs. To include them, we assign individual GCs starting from the largest 'small' accretion event until the expected population is accounted for.

The resulting population of accreted GCs in our mocks is compared to the observed Galactic GCs in Figs 1 and 2. For the MW data, we take the total mass estimate from Callingham et al. (2019) and the

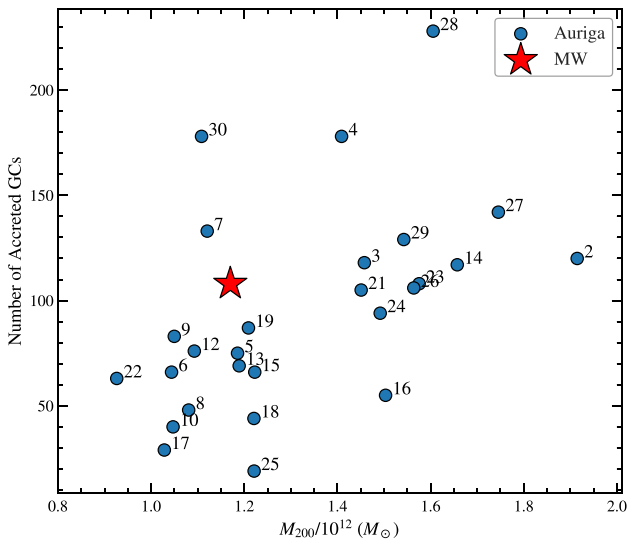

Figure 1. The relation between total mass, M_{200}, and the number of accreted GCs for our AURIGA mock catalogues (blue symbols) and for the MW (red star).

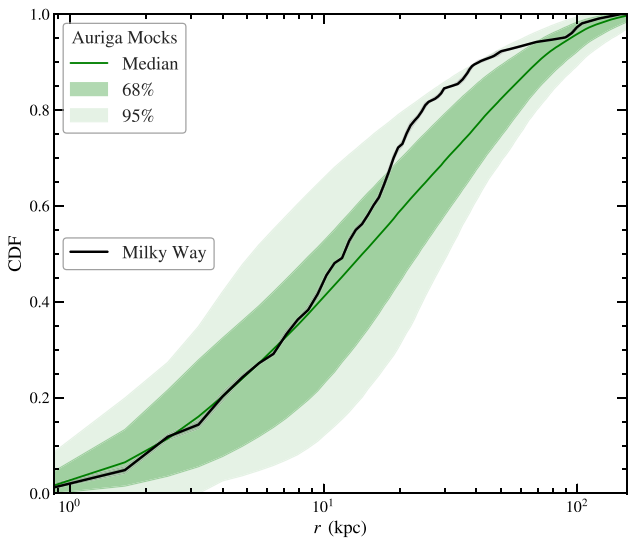

Figure 2. The cumulative radial distribution of accreted GCs in our AURIGA mock catalogues and in the MW (black line). The green solid line shows the median in the mocks, and the shaded regions give the 68 and 95 percentile regions.

number of accreted GCs that we find in Section 5. The number of GCs in the mocks increases with the mass of the host galaxy, as expected from observations and theoretical models (see discussion in Section 1). Fig. 1 shows that the number of accreted GCs in our mocks is consistent with the MW estimates. The AURIGA mocks with a total mass of $\sim 1.2 \times 10^{12}$ $M_\odot$ have slightly fewer GCs than the MW, but the scatter is rather large and there are at least two systems with more GCs.

Fig. 2 compares the radial distribution of GCs, where the distance of the GCs in the AURIGA mocks was scaled by $R_{200}^{\rm MW}/R_{200}^{\rm AURIGA}$ to account for the different sizes of the AURIGA systems. For this, we assumed $R_{200}^{\rm MW} = 222$ kpc from Callingham et al. (2019). The radial distribution of GCs in our mocks is similar to the observed one, although the MW is slightly more centrally concentrated in the 20–30 kpc region than most of the AURIGA sample. This could potentially reflect that the Galactic stellar halo was mostly built from a few massive early accretion events (e.g. Kruijssen et al. 2019b) whose remains are primarily found in the inner region of the MW. Alternatively, it has been suggested that the limited resolution of a simulation can cause accreting satellites to disrupt before reaching the galaxies centre, reducing the concentration of accreted stellar material (e.g. Springel 2005; Grand et al. 2021).

The orbital dynamics of the GCs (including the energy, pericentres, apocentres, actions, angles, and frequencies) for all stars in the main AURIGA halo at the present day are calculated using the AGAMA package (Vasiliev 2019). The potential is modelled from the $z = 0$ simulation snapshot, representing the contribution of the hot gas and DM as a spherical harmonic expansion and the contribution of the stars and cold gas as an azimuthal harmonic expansion (using AGAMA).

3.2 The AMR

Hydrodynamical simulations generally have difficulties reproducing the metallicity of dwarf galaxies and their GCs (e.g. Halbesma et al. 2020), which is potentially due to uncertainties in stellar yields. To mimic the observed AMR of GCs we assign metallicity values to our mock GCs using the relation given in equation (12). For each accretion event, we first choose an AMR, setting a formation time that is equal to the oldest star in that galaxy and a yield determined by the yield–stellar mass relation described in Forbes (2020). The AURIGA galaxies have a somewhat high stellar mass for their halo mass, and to mitigate this we recalculate the progenitor stellar masses using the halo mass at infall and the stellar mass–halo mass (SMHM) relation of Behroozi et al. (2019). This gives yields more comparable with those predicted for the MW than if we had used the original stellar mass of AURIGA. To mimic observational uncertainties, we add normally distributed errors with a mean of 1 Gyr to the age estimates, which corresponds to the average errors for the MW GCs. Note that these uncertainties are applied after determining the appropriate metallicity values so that the final age–metallicity data do not lie exactly on the AMR relation.

For the *in situ* clusters, we randomly assign ages between 12 and 14 Gyr, motivated by the age distribution of MW *in situ* clusters. We then follow the same procedure used for generating the accreted metallicity values, using the AMR fit to the MW *in situ* clusters. This process generates an age–metallicity distribution that is comparable with the MW's own, with a distinctive steeper *in situ* branch and a shallower, wider accreted branch.

4 MOCK TESTS OF THE MIXTURE MODEL

We proceed by testing our multicomponent model for the GC population using our mock catalogues. These tests will help select the optimal dynamical quantities to identify GC groups and characterize the extent to which our modelling approach recovers the true GC groups predicted by the cosmological simulations.

First, we illustrate an example of a chemo-dynamical fit for a mock catalogue. The fit is obtained by following the steps described in detail in Section 2. Fig. 3 shows the $(E, \boldsymbol{J})$ distribution, and Fig. 4 the AMR fits, of the six most massive accretion events for the AURIGA 5 halo. These events are numbered from 0 to 5 in descending order of their total mass at accretion and are the systems containing four or more GCs. The remaining GCs, i.e. those from accretion events that brought three or fewer objects, are labelled as 'ungrouped'. In this fit, our model agrees with the groupings of 110 out of the 145 GCs,

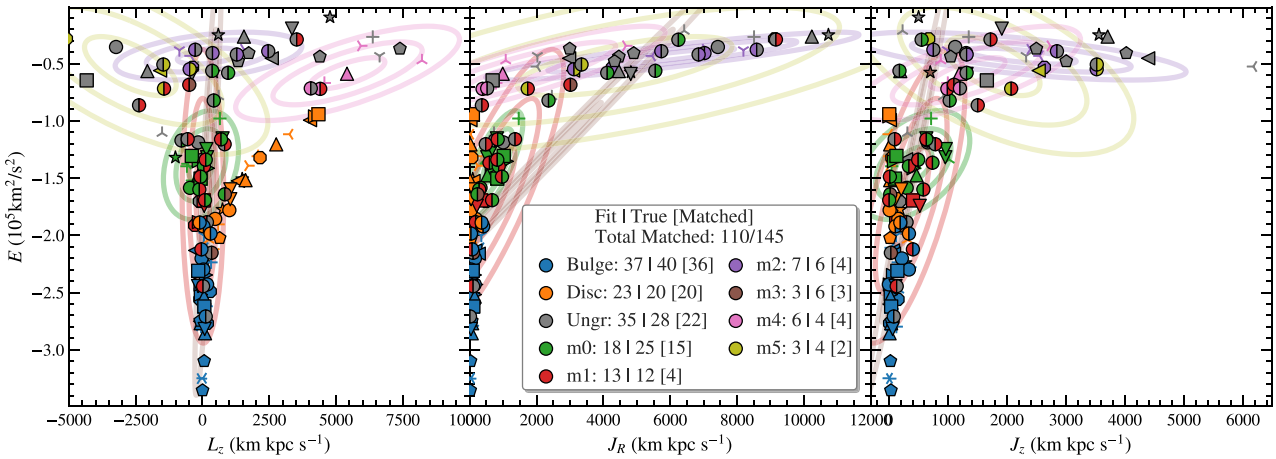

Figure 3. The chemo-dynamical model of Section 2 fit to a mock GC sample from the AURIGA 5 halo. The axes show the dynamical component of the fit in energy (E) and action space (angular momentum L_z, radial action J_R, vertical action J_z). The accompanying age–metallicity fit is shown in Fig. 4. GCs are represented by symbols, which are consistent for each object between the three panels, to help identify individual GCs. Different colours represent the different accretion groups, labelled m0 to m5 by decreasing accretion mass. The groups are modelled as Gaussian distributions, with the contours giving the 1σ, 2σ, and 3σ regions. The 'ungrouped' group contains GCs that do not fit well in other groups or contain less than four members in the true groups. A GC is attributed to the group for which it has the highest probability of belonging. The solid symbol colours indicate where the model agrees with the true grouping. Symbols that are split in colour show the original true grouping on the right-hand side and the assigned group on the left-hand side.

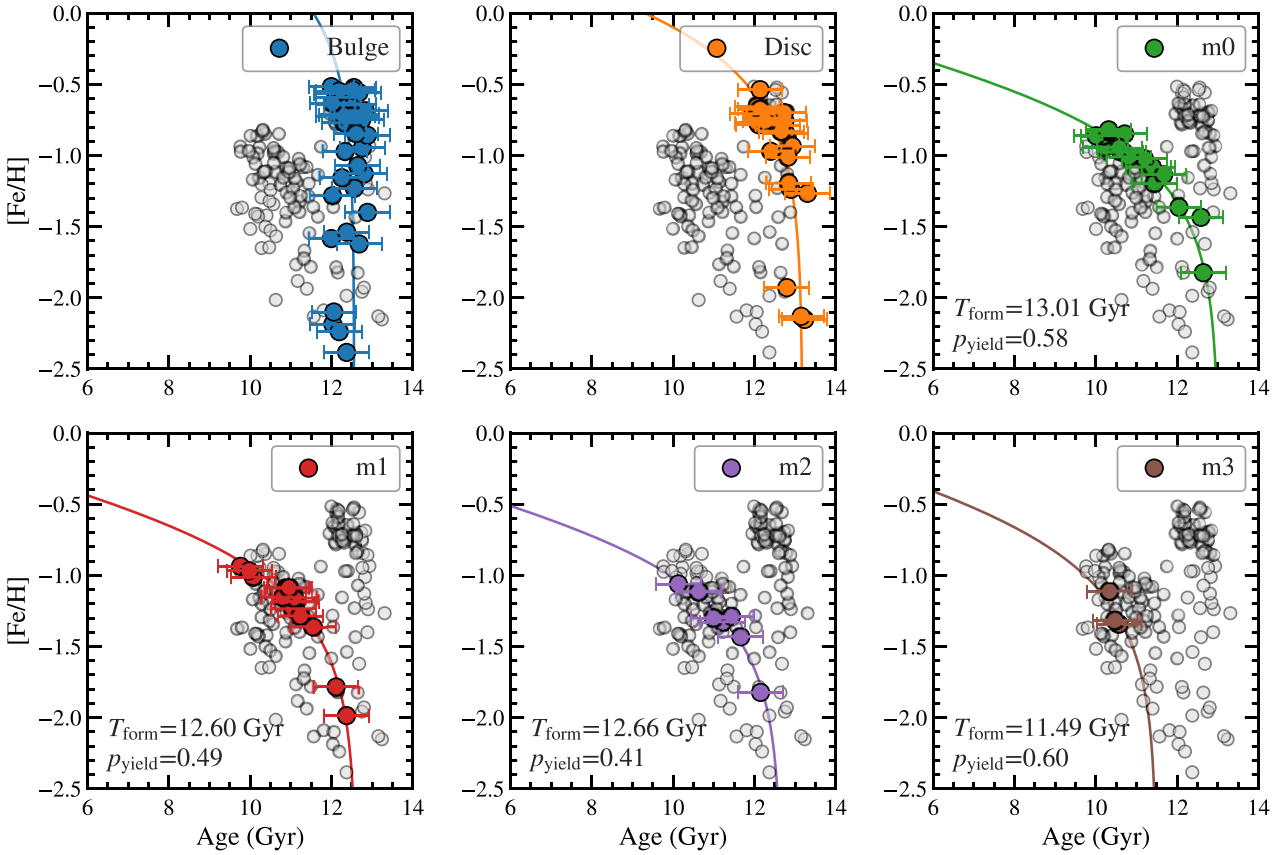

Figure 4. The AMR for a mock GC sample from the AURIGA 5 halo. The different panels show the position of the different GC groups as identified by our method and the AMR fit to them (see the main text for details). The *in situ* components (bulge and disc) can be seen to be steeper branches than the *in situ* components.

where the GCs are associated with the group for which they have the highest probability of being a member.

We can see that in the dynamical space, the accreted distributions overlap in all three panels, with some distributions very widely spread. Some individual GCs of a group can be far from the rest of the group's GCs and the fitted distribution, and have little chance of being correctly identified. The accretion groups of these mock catalogues are undoubtedly more complex than our current picture

of the MW, highlighting the need for realistic testing to understand the feasibility of identifying these groupings.

The *in situ* clusters are reasonably well-identified in both dynamical and age–metallicity space. The accreted components seem to be more distinct and easy to identify at higher energy, where the smaller groups can remain as compact distributions. For the larger groups at lower energy, we see significant overlap with other groups, making them difficult to identify confidently.

4.1 Initial groups

To apply our algorithm to the GCs, we must first make a choice of starting groups. Through testing, we have found that with different initializations it is possible to generate different final groupings as the algorithm converges to different local maxima. To overcome this, we apply our algorithm to many starting configurations. The log-likelihood of the fits can then be compared, with the largest chosen as the best-fitting grouping.

However, it is not feasible to try all possible starting groups. We have experimented extensively with different methodologies to generate the initial groups, including using other clustering algorithms and seeding the groups with random GCs and overdensities. However, none of these alternatives returned satisfactory results, reflecting the difficulty of the problem. Instead, we choose the 'sensible' starting configurations described next and apply both a bootstrapping-based approach and hand-selected variations to test.

In the mock catalogues, we know the true accretion groups of the GCs, and so we use them as the sensible starting point. We tested the robustness of this initialization step by reassigning a fraction of the GCs to plausible alternative groups. We find that the outcome is generally robust to such changes as long as the reassigned fraction is $\lesssim 35$ per cent, with the smaller groups being the most affected. This is likely due to the average position and spread of the distribution describing the groups remaining similar until a large fraction of the group members are lost. From these distributions, the group can recover its members. Further details of our tests on initial groupings can be found in Appendix C.

For the MW case discussed in Section 5 we use selections from the literature as our starting points. There is already a rough decomposition of GCs into accretion events with the main limitation being that the boundary between these groups is rather subjectively defined (e.g. Massari19). When analysing the MW sample, we find all the known groups with more than nine members, which suggests that having a modest fraction of mislabelled GCs does not strongly impact the outcome. This is discussed in Section 5.

4.2 Testing on mock samples

We now apply our method to all the mock catalogues, testing each of the 1000 sets of mock GCs for each relaxed AURIGA galaxy.

Arguably the most important quantity for inferring the properties of the progenitor galaxies is our ability to estimate the population of each of the accreted groups. In a first step, we study how the population sizes of our recovered groups compares to the truth. This test does not fully characterize our method, since it does not indicate whether the individual GCs have their correct groups identified. To further quantify how well the method recovers each GC group, we define the purity, P, and completeness, C, as

$$P = N_{\cap}/N_{\rm fit} \tag{16}$$

$$C = N_{\cap}/N_{\rm true}, \tag{17}$$

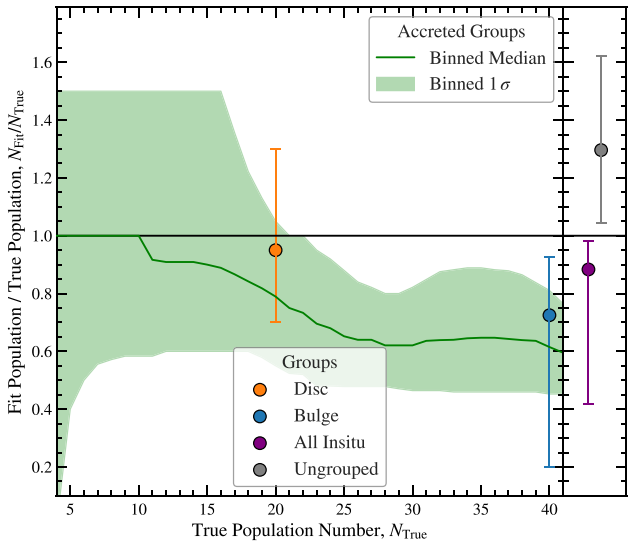

Figure 5. The ratio, $N_{\rm fit}/N_{\rm true}$, between the group population size recovered by our method and the true value as a function of $N_{\rm true}$ for our mock GC catalogues. We bin the accreted groups (not including the ungrouped component) in $N_{\rm true}$, and, using a Gaussian smoothing kernel of $\sigma = 2$, show the median (solid green line) and 16 per cent–84 per cent range (shaded green region) for the distribution of $N_{\rm fit}/N_{\rm true}$. The symbols with error bars show the individual *in situ* components, the total *in situ* sample, and the ungrouped components (see legend). Note that the disc and bulge components of the mocks have populations of 20 and 40 by construction, whereas the population of the ungrouped component depends upon the accretion history of the halo.

where $N_{\rm True}$ is the true GC population of the group, $N_{\rm Fit}$ is the number of members identified by our fitting procedure, and $N_{\cap}$ is the intersection of the fit and true groups.

First, we study how the recovered number of GCs in each group compare against the true number of members. This is shown in Fig. 5. Across the total sample, we recover the accreted population numbers with an underbias of ~ 10 per cent. This is the average value per component and, since there are more small groups than large ones, is biased towards small groups. There are clear trends if we look at the results as a function of the group richness. The smallest groups ($\lesssim$10) are recovered without an underbias, but as the true size of the group increases so does a trend to systematically underestimate the population. For smaller groups there is a significant fractional scatter (~ 50 per cent), as the changing membership of a single cluster corresponds to a larger proportion of the group. This scatter decreases as the true population grows. Furthermore, the smallest groups can be seen to have a small chance of going extinct (their population dropping to zero), which can happen when the group is too spread out to be reliably identified.

Similar to Fig. 5, we now consider the purity and completeness of the groups as a function of the true group richness (see Fig. 6). We find that, on average, for all groups we achieve a purity and a completeness of ~ 67 per cent. For the accreted groups, we find a purity of ~ 64 per cent and a completeness of ~ 55 per cent. There are no clear trends in the purity against the true population number. However, there is a clear decrease in the completeness of the groups as the richness increases, matching the systematic underbias in the fit population numbers seen in Fig. 5.

The dependence of the completeness on $N_{\rm True}$ seen in Fig. 6 is likely a reflection of the characteristics of large versus small groups. Large groups have a greater spread in phase space due to their higher internal velocity dispersion before accretion and tend to exist at

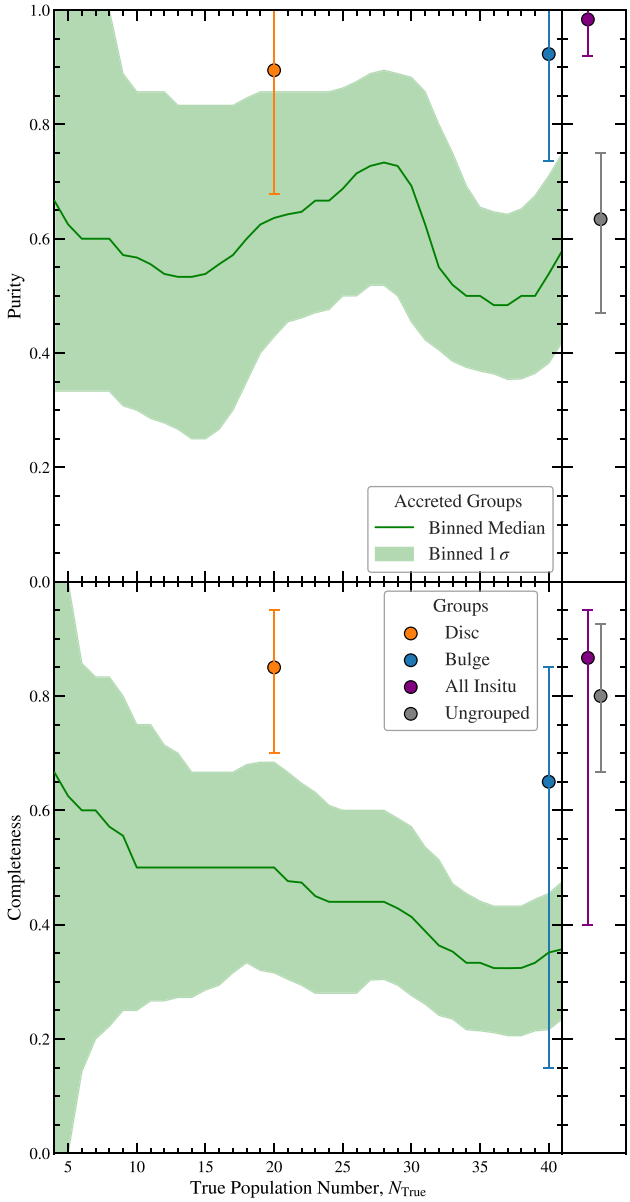

Figure 6. The true group population, $N_{\rm true}$, against the purity, P (top panel), and completeness, C (bottom panel), of our fitted mock GC sample (see the main text for definitions). We bin the results for the accreted groups in $N_{\rm true}$, using a Gaussian smoothing kernel of $\sigma = 2$, and show the median and 16 per cent–84 per cent range with the shaded regions for the distribution of $N_{\rm fit}$. The symbols with error bars show the individual bulge and disc components, total *in situ* sample, and ungrouped component.

lower energies because they experience greater dynamical friction. These factors directly impact our ability to recover these groups. Due to the wider spread in dynamical space, and the crowded nature of the lower energy regions, more of the GCs are misattributed to other groups. The smaller groups tend to be more compact in phase space, and typically exist at higher energies (unless accreted at early times). Providing the group itself has enough members to be reliably identified; they are recovered with greater confidence.

The total *in situ* population is, in general, well recovered with very high purity. Rarely does the methodology misidentify an accreted GC as an *in situ* one in our mock tests, with a median purity of 98 per cent. Compared to most of the accreted components, the *in situ* components occupy distinct positions in the chemo-dynamical space. Groups that do overlap with the *in situ* populations tend to be older and more massive mergers that can bring their material to the heart of the galaxy. The purity and completeness of the bulge and disc components is marginally worse than for the *in situ* population as a whole, and the decrease is due to our method shuffling GCs between the disc and bulge groups.

We find that the richness of the ungrouped component is systematically overestimated by ~ 30 per cent. This is driven by the inclusion of GCs that could not be identified with their true groups (effectively, the missing clusters that cause the average ~ 10 per cent underbias in the fit groups). These are typically separated from the rest of their accreted group in phase space, where they are difficult to identify, and so they fall into the ungrouped component. This is also reflected in the low purity of this component, while the completeness is marginally better, suggesting that ungrouped GCs are not normally being misattributed to other structures.

4.3 Unbiased population estimates

From the results of our mock tests, we find that our fit population numbers, $N_{\rm Fit}$, is a biased estimator of the true population numbers, $N_{\rm True}$. This can be seen explicitly in the top panel of Fig. 7, where we consider the $N_{\rm Fit}/N_{\rm True}$ ratio as a function of $N_{\rm Fit}$. Both the median bias (denoted $\beta(N_{\rm Fit})$) and the 16 to 84 percentiles change as a function of $N_{\rm Fit}$, with an underbias that increases for richer groups (as suggested by Fig. 5). Using these results we can correct our estimate to obtain an unbiased estimate, $N_{\rm Est}$, defined as

$$\log_{10}(N_{\rm Est}) = \log_{10}(N_{\rm Fit}) - \beta(N_{\rm Fit}), \tag{18}$$

as shown in the bottom panel of Fig. 7. The group-to-group scatter, whose 16–84 percentiles are shown as a shaded region, can then be used to give quantifiable uncertainty in $N_{\rm Est}$. We use this bias-corrected estimate and uncertainties, alongside the uncorrected estimates, in our analysis of the MW's infall groups in Section 6.

4.4 Choices of chemo-dynamical spaces

We also have used the mocks to investigate which dynamical spaces best recover the true GC groups, which we define as the space that returns the highest purity and completeness. When fitting in various dynamical spaces, we consider only the accreted components, as our *in situ* fitting scheme applies only in action-based space. We found that the $(E, \boldsymbol{J})$ space is best at recovering the true groupings performing better than $\boldsymbol{J}$ space alone (completeness of ~ 50 per cent), or combinations between E and angular momentum $\boldsymbol{L}$, components. These include (E, L_z, L_p), where L_p is the $\boldsymbol{L}$ component in the disc plane (completeness of ~ 50 per cent), which has been used by Massari19, and the two-dimensional space (E, L) (completeness of ~ 43 per cent).

When fitting without the AMR, the purity and completeness of the accreted groups decrease by ~ 10 per cent in $(E, \mathbf{J})$ space. This trend is similar across the other spaces tested. Without the AMR relation, our ability to identify the *in situ* components is significantly reduced. The average total purity of the group decreases to ~ 70 per cent. In areas where the dynamical distributions of the groups overlap, it is this additional information that allows the memberships to be identified. It should be noted that for our real sample of Galactic GCs, only 96 of the 170 GCs have age–metallicity data, likely hindering our ability to confidently identify the groupings.

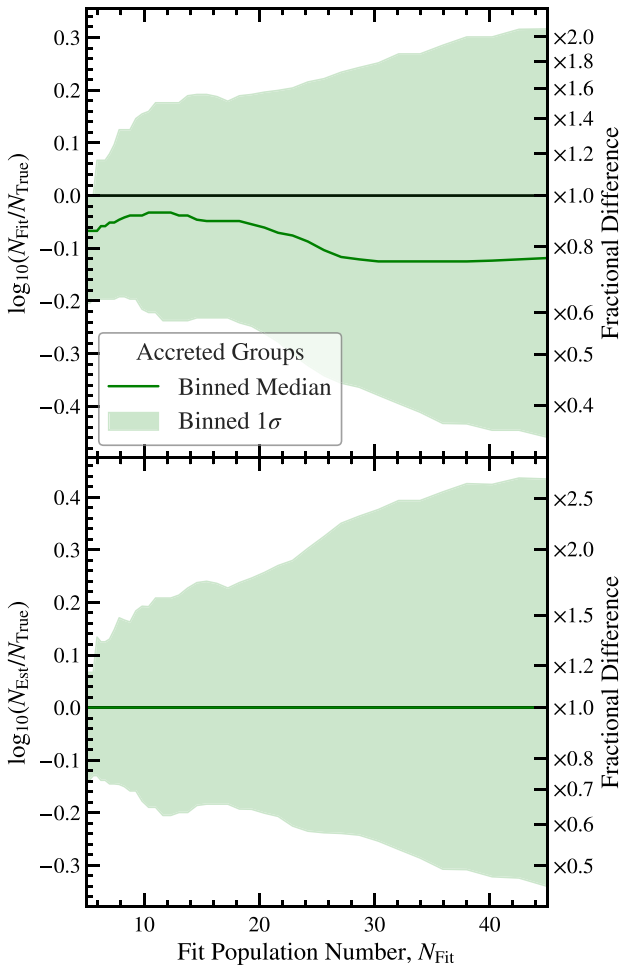

Figure 7. Top panel: the ratio of the fit to true GC number, $N_{\rm Fit}/N_{\rm True}$, for our mock catalogues. We bin the accreted groups in $N_{\rm Fit}$, and, using a Gaussian smoothing kernel in log space of $\sigma = \log_{10}1.1$, show the median (solid green line) and 16 per cent–84 per cent range with shaded regions for the distribution of $N_{\rm true}$. These results are used to obtain an unbiased estimate, $N_{\rm Est}$, of the likely number of GCs in each group. Bottom panel: the ratio of the unbiased estimate to the true GC number, $N_{\rm Est}/N_{\rm True}$, following the format of the top panel.

5 FITTING THE GALACTIC GCS

We now proceed to apply our multicomponent model to the Galactic GC data.

5.1 Observational data

We use the largest Galactic GC sample to date, which consists of the 170 GCs studied by Vasiliev & Baumgardt (2021) that have 6D phase space (i.e. position and velocity) data. The GCs proper motions are based on the *Gaia* Early Data Release 3 (EDR3) and represent an improvement in precision by roughly a factor of 2 compared to the previous *Gaia* Data Release 2 (DR2) measurements (Gaia Collaboration 2021). Where available, we updated the Vasiliev & Baumgardt GC distances with those from Baumgardt & Vasiliev (2021), which are based on the mean values of a combination of *Gaia* EDR3, *Hubble Space Telescope*, and literature data.

To transform the observations to a Galactocentric reference frame we assume the following: a Local Standard of Rest of LSR = 232.8 km s^{-1} (McMillan 2017), a solar radius of $R_\odot = 8.2$ kpc, a solar height of $z = 0$ pc (assumed negligible), and a local solar motion of $(U, V, W) = (11.1, 12.24, 7.25)$ km s^{-1} (Schönrich, Binney & Dehnen 2010).

To calculate the dynamics of the GCs, we use the AGAMA package (Vasiliev 2019) and assume the McMillan (2017) potential of the MW, as implemented in AGAMA. We have tried other potentials, such as that of Cautun et al. (2020), and we find that while the energy of the GCs shifts by an approximately constant value, the individual groupings experience only minor changes. We calculate a range of dynamical quantities, including the energy, actions, and angular momentum of the GCs' orbits.

To account for measurement errors in the positions and velocities of GCs, we create a Monte Carlo sample of 1000 points in observed space (i.e. radial distance and velocity, and celestial proper motions) using the quoted measurement errors which we model as Gaussians for each measured quantity. These are then transformed into positions and velocities with respect to the Galactic Centre, and fed into AGAMA to generate a Monte Carlo sample of dynamical quantities. The precision of these phase-space coordinates is typically limited by distance uncertainties.

The age and chemistry data are taken from a compilation of literature data by Kruijssen et al. (2019b), which provides ages and values of [Fe/H] for 96 GCs. These are averaged from values derived by Forbes & Bridges (2010), VandenBerg et al. (2013), Dotter et al. (2010), and Dotter, Sarajedini & Anderson (2011). We neglect measurement uncertainties in metallicity since these are considerably smaller than the errors in the age.

5.2 Fitting the MW

We now apply our model to the MW. The first step is to initialize our expectation–maximization algorithm by postulating a set of starting groups. We experimented with different initial groupings taken from the literature, primarily from Massari19, Horta20, and Forbes20. We also tried a bootstrap-inspired approach, relabelling one GC at a time as 'ungrouped' and refitting the model to check for a higher likelihood. In general, we find little dependence of the final groups on these small changes.

The results we present below are for the maximum likelihood model over all these variations in the initialization of the expectation–maximization algorithm. The final fit is shown in Figs 8 and 9, and the fit parameters are listed in the Appendix. The derived properties of the groups can be found in Table 2 (Section 6), and group memberships are discussed in the next subsection. In general, we find good agreement with previous work, with all groups being distinct in either chemical or dynamical space.

At the centre of our Galaxy, we find significant overlap in dynamical space between the two *in situ* components (bulge and disc) and part of the Kraken group. This is where we see the most change from previous literature groupings, with a substantial increase in GCs identified as Kraken. To separate these groups with confidence, we rely on the age–metallicity space for the GCs, where these data are available. However, the *in situ* and accreted tracks overlap for old, low-metallicity GCs and cannot be distinguished. In this region, we find that there is not enough information to separate all the GCs confidently into distinct groups.

Within the Kraken and *in situ* groups, some GCs clusters can be identified with high membership probability. We can be confident that there is an accreted group at low energy, from both the distributions in the age–metallicity space and the dynamics. It is the exact extent

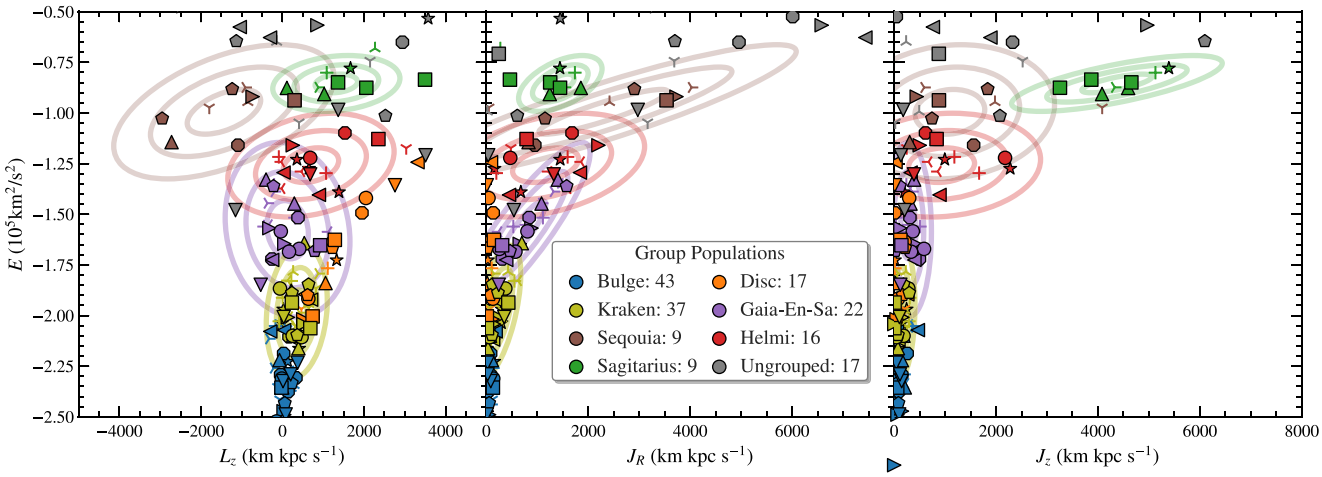

Figure 8. Dynamical groups in energy-action space as inferred by our chemo-dynamical model of the Galactic GC population. The companion age–metallicity modelling can be found in Fig. 9. Symbols are the observed GCs and are consistent across panels to help identify individual GCs. Different colours indicate different groups, with each GC coloured by its most likely group. Accreted components are modelled as Gaussian distributions, with the 1σ, 2σ, and 3σ intervals given by the contours. *In situ* components, bulge and disc, are in the inner regions of the galaxy and typically contain the most bound GCs. The ungrouped component is modelled as a uniform distribution and contains objects accreted in small groups that cannot be reliably identified. Note that a few high-energy GCs are beyond the axes limits and are not shown.

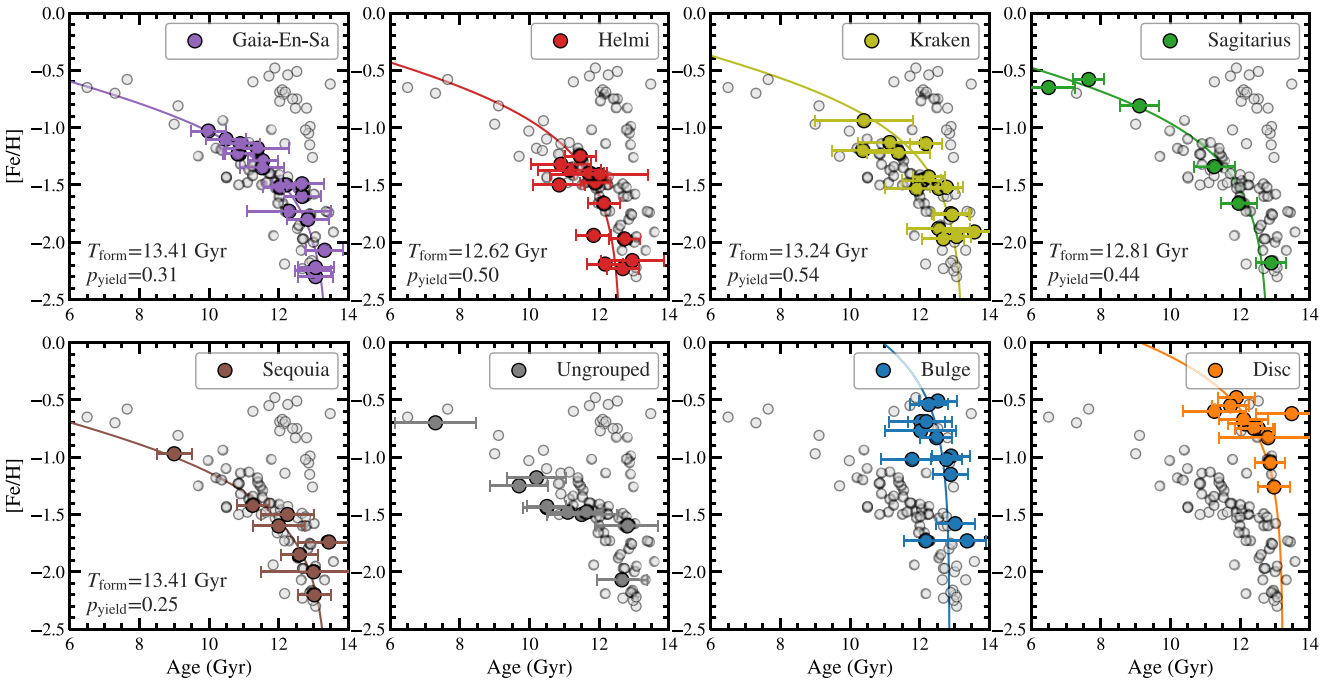

Figure 9. The AMR for the Galactic GCs split according to the component with which they are associated. The solid lines show the AMR fit to the GCs associated with each component (see the main text for details).

of the group that we find difficult to confidently ascertain, and we caution against taking our proposed Kraken membership without considering these factors.

We find that Sequoia and GES cannot be convincingly fit by a single group. While there is no clear difference in the age–metallicity space, the dynamics of the two groups seem to be distinct. The possibility that Kraken is the core of GES was briefly discussed by Horta et al. (2021). We agree with their conclusions that Kraken and GES are unlikely to have the same origin. The dynamics of the two groups seem to be distinct, and Kraken has a steeper metallicity–age relation (higher $p_{\rm yield}$) than GES.

We find no convincing evidence for additional subgroups, such as the LMS-1/Wukong structure suggested to potentially contain GCs including ESO280, NGC 5024, NGC 5053, and Pal 5 (Naidu et al. 2020; Yuan et al. 2020; Malhan et al. 2021). When we model these GCs as separate groups we find that the group becomes extinct as the GCs are absorbed into the GES group. However, we note that this group is in the regime where the number of points is less than the dimensions of the space, and thus the groups are poorly modelled (see Section 2.1.1 for details). We can therefore not rule out the possibility of this substructure, or other small groups.

Table 1. The GC members of the accretion groups of the MW. Note that these are the most probable memberships. To see the membership probability of individual GCs, see Table 2.

Component	Membership
Bulge	Djorg2 (ESO456), Terzan6 (HP5), Terzan2 (HP3), NGC 6380 (Ton1), NGC 6440, Liller1, NGC 6642, NGC 6388, NGC 6535, NGC 6401, Terzan5 (11), NGC 6638, NGC 6528, 1636-283 (ESO452), Terzan9, NGC 6624, NGC 6558, Terzan4 (HP4), HP1 (BH229), NGC 6325, NGC 6453, NGC 6626 (M28), NGC 6304, NGC 6522, Terzan1 (HP2), Pal6, NGC 6652, NGC 6266 (M62), NGC 6342, NGC 6637 (M69), NGC 6355, NGC 6540 (Djorg), NGC 6717 (Pal9), NGC 6293, NGC 6256, NGC 6517, NGC 6144, VVVCL001, NGC 6723, VVVCL002, NGC 6171 (M107), NGC 6093 (M80), Gran1,
Disc	NGC 6838 (M71), NGC 5927, NGC 104 (47Tuc), NGC 6496, ESO93, NGC 6362, NGC 6366, NGC 6352, BH176, Pal10, E3, NGC 6218 (M12), NGC 6441, Pal11, IC1276 (Pal7), Lynga7 (BH184), Pfleiderer,
Gaia-En-Sa	NGC 6205 (M13), NGC 362, NGC 6779 (M56), NGC 7089 (M2), NGC 2298, NGC 1851, NGC 2808, NGC 7099 (M30), NGC 6341 (M92), NGC 5286, NGC 1261, ESO-SC06 (ESO280), NGC 288, NGC 5139 (oCen), NGC 6864 (M75), NGC 5897, NGC 6235, Ryu879 (RLGC2), BH140, NGC 6656 (M22), NGC 7078 (M15), IC1257,
Helmi	NGC 5904 (M5), NGC 4147, NGC 5634, NGC 5272 (M3), NGC 5053, Pal5, NGC 7492, NGC 5024 (M53), NGC 6229, NGC 4590 (M68), NGC 6981 (M72), Rup106, NGC 6584, Bliss1, NGC 6426, NGC 1904 (M79),
Kraken	NGC 6254 (M10), NGC 6712, NGC 6544, NGC 5946, NGC 6121 (M4), NGC 6809 (M55), NGC 4833, NGC 6681 (M70), NGC 6287, NGC 5986, NGC 6541, Terzan10, NGC 6752, NGC 6749, NGC 6760, UKS1, NGC 6284, Mercer5, NGC 6397, Terzan3, FSR1716, FSR1735, NGC 6539, Ton2 (Pismis26), Terzan12, NGC 6402 (M14), Pal8, NGC 6139, Djorg1, NGC 6553, NGC 6316, NGC 4372, NGC 6273 (M19), BH261 (AL3), NGC 6569, NGC 6333 (M9), NGC 6356,
Sagitarius	Arp2, NGC 6715 (M54), Terzan8, Terzan7, Pal12, Whiting1, Munoz1, Kim3, Ko1,
Seqouia	NGC 5466, NGC 6101, NGC 7006, NGC 3201, IC4499, Pal13, NGC 5694, Pal15, AM4,
Ungrouped	Ryu059 (RLGC1), Ko2, Pal3, NGC 6934, Crater, Pyxis, Segue3, Pal14, AM1, Eridanus, Pal4, Pal1, NGC 5824, NGC 2419, Laevens3, Pal2, FSR1758,

5.3 Component fits and membership

We now discuss the groups individually. The group membership can be found in Table 1, while the individual membership probabilities are compiled in Table A1 in Appendix A. To find the number of GCs associated with each accretion event, including uncertainties, we use the GC membership probabilities. We generate a Monte Carlo sample by drawing from the membership probability of each GC. From this sample, we find the expected membership and the 68 per cent confidence interval. It should be stressed that the mock tests of our methodology demonstrate that it is very difficult to correctly identify the membership of each individual cluster (although, on average, the population of accreted groups can be approximately recovered). Therefore, we caution against placing undue emphasis on single GC memberships.

Note that when discussing the expected populations of the components in this section we refer to those returned by our fitted model, not the bias-corrected estimates, as we are referring to the membership of the individual GCs. The bias-corrected estimates are used in the following section (Section 6), where we discuss the implications of the GC memberships for the properties of the accretion events of the MW.

5.3.1 In situ

We find that ∼60 of our GCs are likely to have an *in situ* origin, with the bulge group containing an expected 42^{+2}_{-1} GCs, and the disc group an expected 17^{+2}_{-2} GCs. This is comparable with the numbers of Massari19 who find 62 *in situ* GCs and Kruijssen et al. (2019b) who predict 67 out of their 157 to have an *in situ* origin. The slightly lower total number is again likely the result of our larger Kraken component. Individually, our bulge group is larger and our disc smaller than Massari19's 36 bulge and 26 disc GCs. We find that there is little information to distinguish the disc GCs at low radius and energy from the bulge component.

The bulge GCs typically have energies below $-2 \times 10^5\,\mathrm{km^2\,s^{-2}}$ and apocentres below 5 kpc. This component does not have any significant rotation and has an AMR track that is slightly steeper than the disc. The disc extends to higher energy, but all the GCs have $z_{\max} < 6$ kpc, eccentricity $e < 0.6$, and circularity >0.5. In the very centre of the Galaxy, the disc overlaps with the bulge, leaving a hole in the middle of the radial distribution, with no disc GCs having an apocentre < 3 kpc.

We find that VVVCL001, VVVCL002, and Gran1, previously uncategorized by Massari19, are likely to be bulge members, but could also plausibly fit into the Kraken group. This is in agreement with the work studying the individual GCs. Gran et al. (2021), find Gran1 either as *in situ* or an ancient merger such as Kraken. Minniti et al. (2021a) find VVVCL002 as the GC closest to the centre of our Galaxy, strongly suggesting that likely it is of *in situ* origin. However, Fernández-Trincado et al. (2021) find VVVCL001 to be very metal-poor GC on an eccentric orbit, and tentatively suggest an accretion origin, likely Sequoia or GES.

Other noteworthy observations of *in situ* GCs include:

(i) We find that 10 GCs previously associated with the disc are, instead, probable Kraken members. All these GCs lack age/metallicity data or have [Fe/H] < −1.5. This highlights the overlapping nature of the *in situ* and lower energy components.

(ii) Liller1 and NGC 6388 are very likely part of the bulge, in agreement with Horta20.

(iii) ESO93, previously uncategorized, is almost certainly a member of the disc.

(iv) E3 (ESO37-1) has previously been associated with the Helmi streams (Koppelman et al. 2019a). We find that it has over 94 per cent probability of being a disc member, reflecting its position on the *in situ* AMR track (in agreement with Kruijssen et al. 2020). It does however reach the highest height above the plane of any disc cluster, $z_{\max} \approx 5.5$ kpc.

5.3.2 Kraken

We expect 37^{+2}_{-1} GCs in the Kraken group, a substantial increase from 25 in Massari19, with the difference consisting of a net contribution of nine from the disc and four from the GES component.

We find that two GCs previously unclassified by Massari19, UKS1 and Mercer5, are highly likely to be Kraken members. However, previous works studying the individual clusters believed that they were likely members of the bulge. Notably, UKS1, as an old but metal-poor GC, was suggested to belong to the bulge in Fernández-Trincado et al. (2020), but the result was highly dependent on its then very uncertain distance. With our more recent distance estimates of Baumgardt & Vasiliev (2021), we find that it is a likely Kraken member.

On its discovery in Longmore et al. (2011), Mercer5 was believed to be a typical bulge GC due to its position in the inner galaxy ($\sim$ 5.5 kpc) and in subsequent extensive chemical follow-up (Peñaloza et al. 2015). With the chemo-dynamical information used in our methodology, the GC is classified as likely part of the Kraken group.

Our model predicts that the energy distribution of the Kraken group is approximately normally distributed with a mean of $-2 \times 10^5\,\mathrm{km^2\,s^{-2}}$ and dispersion values of $0.1 \times 10^5\,\mathrm{km^2\,s^{-2}}$. This is higher Energy distribution than other selections in the literature, which typically give values of $E < -2 \times 10^5\,\mathrm{km^2\,s^{-2}}$ (such as Massari19, Horta et al. 2021). Notably, our Kraken group seems to have bridged the gap seen in stars by Horta et al. (2021) at energies $-2 < E/10^5\,\mathrm{km^2\,s^{-2}} < -1.85$. Furthermore, unlike previous results, our Kraken group has net prograde motion, with angular momentum L_z distributed with a mean of $\sim 350\,\mathrm{kpc\,km\,s^{-1}}$ and a dispersion of $\sim 250\,\mathrm{kpc\,km\,s^{-1}}$. This prograde bias suggests that perhaps some disc GCs have been included in the group.

For GCs on lower energy orbits, and without age–metallicity information, or for those that have low metallicity where the *in situ* and accreted branches overlap, we have found that distinguishing between membership of Kraken or the *in situ* groups is difficult. In future, further chemistry information may allow us to distinguish better between the accreted and *in situ* components at low energy.

5.3.3 Sagittarius

Our Sagittarius group contains an expected population of 9^{+1}_{-0} GCs. Due to recent accretion and tidal stripping, much of its material is in an easily identifiable stream. This allows seven GCs to be identified with a high degree of certainty as being associated with the Sagittarius dwarf: Terzan7, Arp2, Terzan8, Pal12, Whiting1, and M54 (NGC 6715), which is believed to possibly be the nucleus of Sagittarius (Law & Majewski 2010; Antoja et al. 2020; Bellazzini et al. 2020; Peñarrubia & Petersen 2021). We find that these GCs have a near-certain membership. We also find that two uncategorized GCs, Munoz1 and Kim3, also have over 90 per cent probability of membership, and Koposov1 has $\sim$ 70 per cent probability. Koposov1 has been previously noted to lie close to a distant branch of the Sagittarius stream (Koposov et al. 2007; Paust, Wilson & van Belle 2014). These high probabilities are driven by the Sagittarius group's high group density in dynamical space and a distinct age–metallicity branch.

Several other GCs have been tentatively linked to Sagittarius in the literature, but we find no other likely members. Compared to the literature, we find:

(i) Pal2 has been proposed to lie on the trailing arm of the stream (Law & Majewski 2010; Bellazzini et al. 2020). However, we find that it has a 81 per cent probability of being ungrouped and a 14 per cent probability of being associated with Sequoia.

(ii) NGC 2419 and NGC 5824 are commonly linked to Sagittarius (Antoja et al. 2020; Bellazzini et al. 2020; Peñarrubia & Petersen 2021), but we find them almost certainly to be ungrouped. The orbit of NGC 2419 is more radial than the average, more vertical Sagittarius orbit, and NGC 5824 is at lower energy than the other GCs.

(iii) NGC 5634 and NGC 5053 have been proposed as lying on ancient wraps of the stream (Bellazzini et al. 2020). We find that they are not likely members (in agreement with Law & Majewski 2010); they are near-certain members of the Helmi group.

(iv) AM4 was attributed to Sagittarius by Forbes20 based on chemistry since, at the time, AM4 did not have *Gaia* kinematics (for this reason Massari19 did not assign the cluster to a group). We find that, as a prograde cluster, its orbit is incompatible with the Sagittarius orbit. Instead, we find that it is a likely member of Sequoia, but has a 17 per cent chance of being ungrouped.

(v) Before Koposov 1 and 2 (Ko1, Ko2) had measured radial velocities, Paust et al. (2014) suggested that they could plausibly lie on the Sagittarius stream. Improved observations by Vasiliev & Baumgardt (2021) have placed Ko1 as a likely member, but Ko2 is almost certainly ungrouped.

5.3.4 Gaia–Enceladus–Sausage

Our analysis gives 23^{+2}_{-1} GCs in the GES structure, in good agreement with Massari19 (25) and Forbes20 (28). The GES group is consistent with having no net rotation and has an energy distribution with a mean of $-1.58 \times 10^5\,\mathrm{km^2\,s^{-2}}$ and a dispersion of $0.15 \times 10^5\,\mathrm{km^2\,s^{-2}}$ This approximately agrees with previous literature selections: $-1.75 < E/10^5\,\mathrm{km^2\,s^{-2}} < -1.3$ in Horta et al. (2021), and $-1.86 < E/10^5\,\mathrm{km^2\,s^{-2}} < -0.9$ in Massari19. The latter also notes that the apocentres are mostly less than 25 kpc, in good agreement with Deason et al. (2018). We also find that our GES GC apocentres lie between 10 and 20 kpc.

(i) The previously unclassified clusters, Ryu879 (RLGC2) and BH140, are likely members of GES.

(ii) In contrast to Myeong et al. (2019), we infer that NGC 4147 and NGC 6981 (M72) are part of the Helmi Streams and NGC 7006, Pal15, and NGC 5694 are associated with Sequoia.

(iii) We find that four GCs that have been previously associated with GES are now associated with the Kraken structure (NGC 4833, NGC 6284, Djorg1, and Terzan10).

(iv) Pal2 is likely to be an ungrouped GC, despite being linked to GES by Massari19 and Forbes20. We find it is at higher energy ($-1.1 \times 10^5\,\mathrm{km^2\,s^{-2}}$) than the rest of the GES group.

(v) Our method classifies ω-Centauri as an almost certainly GES member, in agreement with a tentative classification by Massari19. This cluster has been claimed to be the nucleus of Sequoia by Myeong et al. (2019) and is discussed more in the following section.

5.3.5 Sequoia

We predict 9^{+1}_{-0} members in the Sequoia group, comparable to seven attributed by Myeong et al. (2019) and Massari19, and nine by Forbes20.

The Sequoia group has a narrow energy distribution, with a mean of $-1 \times 10^5\,\mathrm{km^2\,s^{-2}}$ and a dispersion of $0.1 \times 10^5\,\mathrm{km^2\,s^{-2}}$. However, the angular momentum has a wide distribution, with a mean of $-1400\,\mathrm{kpc\,km\,s^{-1}}$ and a dispersion of $900\,\mathrm{kpc\,km\,s^{-1}}$. This is noticeably smaller than other selections in the literature, such as that by Myeong et al. (2019) (and Massari19) of $-1.5 < E/10^5\,\mathrm{km^2\,s^{-2}} < -0.7$ and $-3700 < L_z\,\mathrm{kpc\,km\,s^{-1}} < -850$. The distribution in J_z and J_R is very broad, stretching across the space.

(i) FSR1758 was characterized by Barba et al. (2019), where its unusually large size (for a GC) lead to it being dubbed as 'a Sequoia in the garden'. Later, Myeong et al. (2019) suggested that it is the nucleus of a retrograde accretion event, with the entire accretion event being named Sequoia after the first paper. However, other works, such as Romero-Colmenares et al. (2021), have found the links of this GC with the main structure tenuous, making it more likely that it belongs to a different group such as GES. We find that FSR1758 has a 28 per cent chance of being associated with Sequoia and a 62 per cent chance of being ungrouped. This is primarily driven by its position at lower energy than the rest of Sequoia. It should be noted that, as the heart of an accretion event, it is plausible that FSR1758 has suffered more dynamical friction and fallen to lower energy than other accreted material.

(ii) ω-Centauri is another cluster that has been previously classified as a key Sequoia member. Due to its peculiar chemistry, ω-Centauri has long been suspected by some to be the nucleus of a dwarf galaxy (Bekki & Freeman 2003), thought to have a mass of $\sim 10^{10} M_{\odot}$ (Valcarce & Catelan 2011). Forbes20 and Myeong et al. (2019) believed this to be Sequoia, based on its retrograde orbit. As noted in Myeong et al. (2018b), ω-Centauri could have sunk to lower energy by dynamical friction. We find it has a near-certain GES membership.

(iii) As briefly discussed in the Sagittarius Section, the AM4 cluster has been tentatively linked to Sagittarius before, but we find it is likely a Sequoia member. If true, it is the youngest Sequoia member, with an age of $\sim$9 Gyr, approximately 2 Gyr younger than the rest of the group. Its position in energy-action space is also unusual for Sequoia; while at the centre of the angular momentum distribution, it has negligible radial action and is primarily on a vertical orbit, in agreement with Sagittarius. We flag this cluster as a potential outlier.

5.3.6 Helmi streams

Koppelman et al. (2019a) identified seven GC members, (NGC 4590, NGC 5272, NGC 5904, NGC 5024, NGC 5053, NGC 5634, NGC 6981). This group was tentatively expanded by Massari19 to include three additional members, suggesting a total of 10 members. We find that these are indeed very likely members, inferring 15 ± 1 GCs in the Helmi streams group. Our Helmi stream structure has a narrow energy distribution, with a mean of $\sim -1.25 \times 10^5\,\mathrm{km^2\,s^{-2}}$ and a dispersion of $0.08 \times 10^5\,\mathrm{km^2\,s^{-2}}$. The distribution in angular momentum is broader, with a mean of $\sim 700\,\mathrm{kpc\,km\,s^{-1}}$ and a dispersion of $700\,\mathrm{kpc\,km\,s^{-1}}$. This is in approximate agreement with values in the literature (Koppelman et al. 2019a; Massari et al. 2019; Naidu et al. 2020).

(i) Our Helmi group includes four probable members that were previously associated with GES, NGC 4147, NGC 7492, NGC 6229, and NGC 1904 (Forbes & Bridges 2010; Myeong et al. 2019).

(ii) We find that the previously unclassified cluster, Bliss1, is likely a Helmi member.

(iii) NGC 6441 is almost certainly a member of the disc and not of a Helmi Stream or Kraken, as suggested by Massari19.

5.3.7 Ungrouped

We find 17 ± 1 GCs that are ungrouped or do not fall into any of the other accretion groups. This is likely to be a collection of GCs from different low-mass dwarfs that have otherwise left no significant stellar material to be identified. In their equivalent high-energy group, Massari19 (and Forbes20) identified 11 members.

Pal1 has previously been linked with the disc (Massari19) and GES (Forbes20), but instead we find that it has a very high probability of being ungrouped (in agreement with Kruijssen et al. 2020). It is on a circular orbit compatible with the outskirts of the disc, but it is young and has high [Fe/H] similar to the young Sagittarius GCs Whiting1 and Terzan7. Other hints from its chemistry support this view (Sakari et al. 2011). Naidu et al. (2020) associated Pal1 with a newly identified Aleph structure due to chemo-dynamical similarities.

Clusters NGC 5824 and NGC 2419 have been previously associated with Sagittarius, but we find it is highly likely that they have a different accretion origin; we associate them with the ungrouped component. NGC 5824 has also been associated with the Cetus stream (Yuan et al. 2019; Chang et al. 2020). As the only associated GC with the structure, it would be correct to categorize it as Ungrouped.

Pal2 and NGC 6934 lie close together in (E, $\mathbf{J}$) space, at an energy just below that of the Sagittarius group. These clusters have previously been associated with GES (Massari19), but we find that they are at higher energy than other GES clusters. The fit of the GES group is improved by their removal.

The rest of the group members are all at high energy ($E > -0.75 \times 10^5\,\mathrm{km^2\,s^{-2}}$). We find four GCs uncategorized by Massari19 (Ryu059, Ko2, Segue3, and Laevens3) that are highly likely to be ungrouped. We find no obvious subgroups in these GCs.

5.4 Completeness of *in situ* sample

We expect the GC members of the *in situ* components to be phase mixed and consistent with having axisymmetric distributions. This can be tested by checking that ϕ, the angle in the plane of the disc, is uniformly distributed with a Kuiper test. Similar to the more commonly used KS test, the Kuiper test can be used to quantify if the cumulative of two distributions are statistically compatible, but it is particularly suited to test distributions of modular variables as the statistic is invariant under cyclic transformations of the random variable (Kuiper 1960).

We find that the *in situ* components are not consistent with being axisymmetric, with p values of 0.039 for the bulge, 0.154 for the disc, and 0.03 for the combined sample. This is consistent with an overabundance of GCs on the near side of the Galactic Centre. Binning the GCs into quarter slices with the Sun at $\phi = 0$, we find that our 60 *in situ* clusters are distributed in angle as 24 in $-\pi/4 < \phi \leq \pi/4$, 11 in $\pi/4 < \phi \leq 3\pi/4$, 14 in $-3\pi/4 < \phi \leq -\pi/4$, 11 in $\phi \leq -3\pi/4$, $3\pi/4 < \phi$. In the grouping of Massari19, out of 62 *in situ* clusters, 30 fall in $-\pi/4 < \phi < \pi/4$, with the rest evenly distributed. These results suggest either incomplete observations, with on the order 30 missing *in situ* clusters, or an otherwise undiscovered structure in the GCs identified as *in situ*.

The Kuiper test can also be applied to the angles of the orbital actions of the GCs, which should also be uniformly distributed between 0 and 2π if the group is phase mixed. We find that the accreted components are all consistent with being phase mixed, apart from Sagittarius (with a combined p value of less than 10^{-4}).

6 INFERRING THE PROPERTIES OF ACCRETED GALAXIES

From the fit to the MW data, we have found the likely population numbers of each of the GC groups that have been accreted on to the MW. Using the results from the mock tests, we correct these fit populations to give an unbiased estimate of the true population

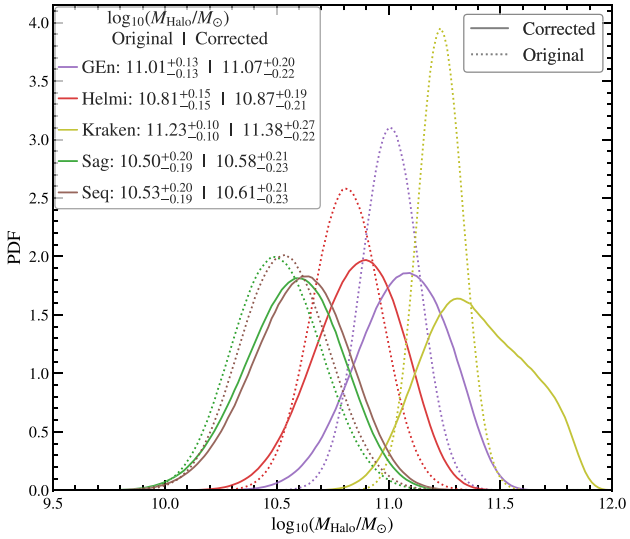

Figure 10. The halo mass likelihood for accreted satellites. This is derived from their probable populations of GCs using the relation between halo mass and number of GCs of Burkert & Forbes (2020). This includes uncertainties from grouping the clusters and theoretical uncertainties from the relation. The dotted lines show the estimates using the original group population numbers as found from the fitting method. The solid lines show the estimates when correcting for bias and including the group-to-group scatter in recovering the true number of GCs (see the main text for details).

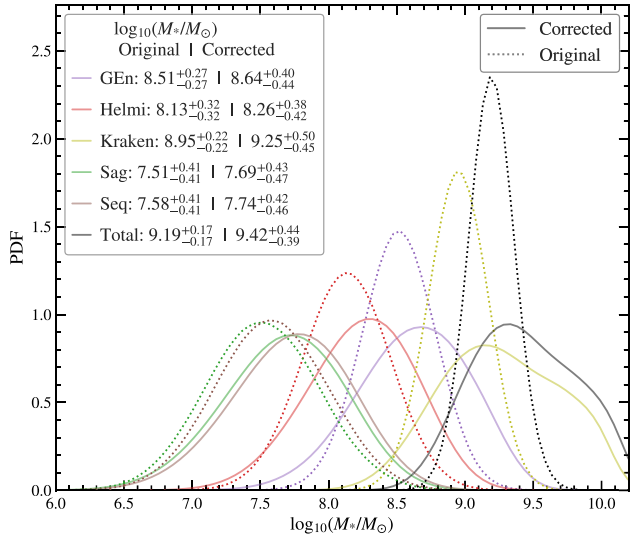

Figure 11. The stellar mass likelihoods of accreted galaxies. These were calculated by assuming the stellar mass–halo mass relation of Behroozi et al. (2019) (assuming $z = 0$) to transform the halo mass PDFs of Fig. 10. The effect of accounting for the redshift of accretion is considered in Appendix D. The dotted lines show the estimates using the original group population numbers as found from the fitting method, and the solid lines show the corrected unbiased estimates (see the main text for details).

numbers with estimates of the uncertainty (see Section 4.3). We now use the M_H–N_{GC} relation (equation 15) to estimate the mass of the progenitor dwarf galaxies. We also include the theoretical uncertainties in this relation (normally distributed as $\sigma_{\rm Ngc}$), given in fig. 2 of Burkert & Forbes (2020) as $\sigma_{\rm Ngc}/N_{\rm GC} \approx (N_{\rm GC}/2)^{-1/2}$. This relationship then gives a probability density function (PDF):

$$p\left(N_{GC}|M\right) = N\left(\mu = \frac{M}{5\times10^9\ \mathrm{M}_\odot}, \sigma = \sqrt{\frac{M}{2.5\times10^9\ \mathrm{M}_\odot}}\right), \tag{19}$$

where N is the normal distribution.

Using Bayes theorem, this relationship can be inverted:

$$p\left(M|N_{GC}\right) = \frac{p\left(M\right)}{p\left(N_{GC}\right)} p\left(N_{GC}|M\right). \tag{20}$$

For the mass prior, $p(M)$, we adopt the mass function of Boylan-Kolchin et al. (2010, equations 7, 8; fig. 5). This relation gives the distribution in terms of the ratio of the mass of the accreted satellite to the virial mass of the host galaxy at present day. For this we take the mass of the MW as $1.17 \times 10^{12} M_\odot$ (Callingham et al. 2019). The N_{GC} prior, $p(N_{GC})$, is effectively the normalization factor.

To account for the probabilistic nature of our GC memberships, we randomly draw population samples of the accretion groups from the membership probabilities. Each drawn population, N_{GC}, is then used to derive a PDF, $p(M|N_{GC})$. The total mass PDF is then the sum over this sample. These results are given as PDFs in Fig. 10. We note that this methodology has significant limitations, as discussed by Kruijssen et al. (2019b). We do not include any redshift dependence in the M_H–N_{GC} relation, assuming that this is sufficiently flat. The errors assumed in this relation are theoretical, and they, as well as the underlying relation, are the subject of extensive debate in the literature.

The estimated halo masses (including uncertainties) can be combined with an SMHM relation to infer the likely stellar masses of the accretion events. We use the SMHM relation of Behroozi et al. (2019), including the given uncertainties. This relation has a non-negligible dependence on redshift; here, we assume the $z = 0$ relation. The resulting stellar-mass PDFs are presented in Fig. 11; the median and 68 per cent confidence limits of these results are summarized in Table 2. Alternatively, one could assume that star formation in a galaxy stops approximately around the redshift of accretion. This assumption has the effect of lowering the stellar masses, particularly of the older accretion events such as Kraken. Further discussion of this effect can be found in Appendix D and Fig. D1.

In general, we find good agreement with results in the literature, particularly for GES, the Helmi streams, and Sequoia. We have more GCs than earlier work, and so we find slightly higher halo and stellar masses, but none the less consistent within the uncertainty interval. The greatest difference between our results and those in the literature is the higher mass for Kraken that we infer. This reflects the considerable increase in $N_{\rm GC}$ that we attribute to the Kraken event, but we note the difficulty in distinguishing between Kraken and the *in situ* component. As a result, we find better agreement with the higher mass estimates, our stellar mass of $\sim10^9$ $\mathrm{M}_\odot$ being closer to that of Horta et al. (2021), who estimated a log stellar mass of 8.7, approximately twice the stellar mass of GES.

From dynamical arguments, the total mass of Sagittarius is thought to be greater than 6×10^{10} $\mathrm{M}_\odot$ (Laporte et al. 2019, 2018), around twice our estimate of its halo mass, $\sim3.4 \times 10^{10}$ $\mathrm{M}_\odot$. However, recent work has claimed to find up to an additional 20 plausible GCs in the body of Sagittarius (Minniti et al. 2021b, 2021c). In our analysis, this would suggest a log halo mass of $\sim$11.16. While this revised estimate is high compared to the majority of the literature, it agrees with recent work by Bland-Hawthorn & Tepper-García (2021), who suggested that the infall mass of Sagittarius has been underestimated because

Table 2. Properties of the Galactic GC accretion groups, as derived in this work and other works in the literature. The first section gives our results; the second column gives the expected number of GCs (N_{GC}), including 68 per cent confidence interval, as inferred from our chemo-dynamical model. From this, using the halo mass-number of GCs relation of Burkert & Forbes (2020), we find the halo mass of the accretion event. The halo mass is used to further infer the stellar mass (fourth column) from the SMHM relation of Behroozi et al. (2019). The halo mass and stellar masses of the ungrouped (and *in situ* components) cannot be estimated in the same way as for the accreted components. The tests on mock catalogues have showed a bias in our method. The fifth and sixth columns give the halo and stellar masses, $\hat{\mathrm{M}}_{\mathrm{halo}}$ and $\hat{\mathrm{M}}_{\star}$, respectively, corrected for this bias and also including the considerable group-to-group dispersions in recovering the true number of GCs (details in the main text). The second section gives relevant values from the literature, with references given below. Our summed total stellar mass is the amount accreted by the named groups, which are assumed to make up the bulk of the contribution to the stellar halo.

	This work					Literature		
Acc. event	N_{GC}	$\log_{10}\mathrm{M}_{\mathrm{halo}}$	$\log_{10}\mathrm{M}_{\star}$	$\log_{10}\hat{\mathrm{M}}_{\mathrm{halo}}$	$\log_{10}\hat{\mathrm{M}}_{\star}$	N_{GC}	$\log_{10}\mathrm{M}_{\mathrm{halo}}$	$\log_{10}\mathrm{M}_{\star}$
Gaia–En–Sa	23^{+2}_{-1}	$11.00^{+0.13}_{-0.13}$	$8.51^{+0.27}_{-0.27}$	$11.07^{+0.20}_{-0.22}$	$8.64^{+0.44}_{-0.40}$	20, 28	10.98 ± 0.08, (11–11.7)	8.43 ± 0.15, (8.7–9.7)
Helmi	15^{+1}_{-1}	$10.80^{+0.16}_{-0.15}$	$8.13^{+0.32}_{-0.32}$	$10.87^{+0.21}_{-0.19}$	$8.26^{+0.38}_{-0.42}$	5, 10	10.74 ± 0.1	7.96 ± 0.18, 8
Kraken	37^{+2}_{-1}	$11.22^{+0.10}_{-0.10}$	$8.95^{+0.22}_{-0.22}$	$11.38^{+0.27}_{-0.22}$	$9.25^{+0.45}_{-0.50}$	13, 25	10.92 ± 0.1	8.28 ± 0.18, 8.7
Sagittarius	9^{+0}_{-1}	$10.49^{+0.20}_{-0.20}$	$7.51^{+0.41}_{-0.41}$	$10.58^{+0.21}_{-0.23}$	$7.69^{+0.43}_{-0.47}$	7, 8	10.94 ± 0.1, >10.8	8.44 ± 0.22
Sequoia	9^{+1}_{-0}	$10.52^{+0.20}_{-0.20}$	$7.58^{+0.41}_{-0.41}$	$10.61^{+0.21}_{-0.22}$	$7.74^{+0.42}_{-0.46}$	3, 7	10.70 ± 0.06, (10–10.7)	7.90 ± 0.11, (6.7–6.9)
Ungrouped	17^{+1}_{-1}	–	–	–	–	11	–	–
Bulge	42^{+2}_{-1}	–	–	–	–	36	–	–
Disc	17^{+2}_{-2}	–	–	–	–	26	–	–
TOTAL	170	–	$9.19^{+0.17}_{-0.17}$	–	$9.42^{+0.44}_{-0.39}$	151	–	$9.15^{+0.11}_{-0.15}$, $8.95^{+0.09}_{-0.05}$

Notes. References: Kruijssen et al. (2020), Massari et al. (2019), Myeong et al. (2019), Horta et al. (2021), Laporte et al. (2019), Laporte et al. (2018), Koppelman et al. (2019a), Deason, Belokurov & Sanders (2019) (total), Mackereth & Bovy (2020) (accreted).

of rapid tidal stripping. This could make its halo mass comparable to the LMC (with halo mass $\sim 10^{11} M_{\odot}$).

We estimate the total stellar mass accreted to be $\sim 2.6^{+4.6}_{-1.5} \times 10^9$ $\mathrm{M}_{\odot}$ (or $\sim 1.5^{+0.7}_{-0.4} \times 10^9$ $\mathrm{M}_{\odot}$ with the uncorrected mass estimates), obtained by summing the stellar masses of the fitted named groups. This does not include estimates of the stellar mass of the ungrouped component, which we assume to be subdominant to the larger accretion events. In comparison with results in the literature, our estimate is higher than the results of Deason et al. (2019) who estimate the total stellar mass in the halo as $1.4 \pm 0.4 \times 10^9$ $\mathrm{M}_{\odot}$, but within 1σ. Similarly, Mackereth & Bovy (2020) estimate a total stellar mass of $1.3^{+0.3}_{-0.2} \times 10^9$ $\mathrm{M}_{\odot}$, but then conclude that only ~ 70 per cent ($0.9^{+0.2}_{-0.1} \times 10^9$ $\mathrm{M}_{\odot}$) has been accreted. We note that including a redshift dependence in the stellar mass–halo mass relation reduces the individual and total stellar masses (see Appendix D). However, the systematic uncertainties on both the redshift dependence of the SMHM relation and the uncertainties on the accretion time make the extent of this effect unclear.

7 CONCLUSIONS

We have introduced a multicomponent model for the GCs in the MW that splits the population into three individual constituents: bulge, disc, and stellar halo. The latter is further decomposed into the individual large accretion events that built up the Galactic stellar halo. The identification of the components has been performed in a chemo-dynamical space for GCs that combines information on the AMR with the orbital energy, E, and the action, $\boldsymbol{J}$. Our study has aimed to obtain an objective and statistically robust identification of accreted GCs groups. These have been modelled as multivariate Gaussian distributions in (E, $\boldsymbol{J}$) space that follow the AMR proposed by Forbes (2020).

We have extensively tested our methodology using mock GC catalogues built from the AURIGA suite of zoom-in simulations of MW-like galaxies. The mocks roughly reproduce the number, radial distribution, and, by construction, the AMR of GCs in our galaxy. These show that the best space for our modelling is the combined energy-action space; including the age–metallicity information improves our results by ~ 10 per cent.

Our approach recovers, on average, the population numbers of the GCs associated with each merger event in the simulations with an underbias of $\sim$ 10–20 per cent, but with considerable group-to-group scatter. We find that a proportion of 'missing' clusters causing the underbias cannot be associated with their true groups and are identified as ungrouped. Using our mock test results, we create an unbiased estimate for the true population of the fitted groups with realistic uncertainties of our methodology. However, the grouped GCs are not always associated with the actual objects brought in by a particular merger event; the fit accretion groups have an average purity and completeness of only $\sim$60 per cent. These relatively low values reflect the large overlap between various accretion events, which makes it difficult unequivocally to associate many GCs with a single accretion group.

We then have applied this methodology to the Galactic GC data, accounting for measurement errors. The result is a decomposition of the GC population into bulge, disc, and the following halo components: GES, Kraken, Sagittarius, Sequoia, and Helmi groups and an ungrouped component. This ungrouped component contains 17 'left over' GCs that have a uniform background distribution and are likely associated with many small accretion events that do not contain enough members to be robustly identified.

We find that it is difficult to separate some of the low-energy GCs into a contribution from Kraken and the *in situ* components. Where these groups overlap in dynamical space, age–metallicity information is needed to help identify the groups. However, in the region where the AMRs of the *in situ* and Kraken components overlap (high age and low metallicity), or for GCs without age–metallicity information, there is not enough information to identify them confidently. This likely leads to an overestimate of the membership of the Kraken component. This is supported by our Krakens slight net rotation, possibly indicating the inclusion of disc GCs.

Combining the resulting groupings of GCs with the relation between halo mass and GC number of Burkert & Forbes (2020),

we have inferred the halo mass of the progenitor of each accretion event. Combining these halo masses with the SMHM relation of Behroozi et al. (2019), we then inferred the progenitor stellar masses. We find the Kraken group to be the most massive accreted galaxy ($M_{\rm halo} \sim 2.4^{+2.0}_{-1.0} \times 10^{11}$ M$_\odot$), likely slightly larger than GES ($M_{\rm halo} \sim 1.2^{+0.7}_{-0.4} \times 10^{11}$ M$_\odot$).

We find evidence in the phase distribution that the sample of *in situ* MW GCs are probably incomplete, with on the order of 20–30 GCs 'missing' from the far side of the Galaxy. These are likely to be obscured by the Galactic centre and disc.

There are two relatively straightforward possible improvements of our study:

(i) Our tests with mock catalogues indicate that increasing the sample of GCs with age–metallicity data would improve our ability to identify groups, particularly for the *in situ* component. Furthermore, the age-chemistry modelling used in our method could, in principle, be refined by including more detailed models of chemical evolution. The obvious choice for this would be to include the α abundances which are currently available only for a small subset of GCs (Horta et al. 2020). This could be very useful for disentangling the *in situ* and Kraken groups.

(ii) Our methodology could be extended to include stellar halo stars. The automated and statistical nature of our method makes it straightforward to handle large samples, as well as the much larger observational errors of stellar samples. Increasing the number of dynamic tracers by several orders of magnitude would allow a much more accurate inference of the MW's accretion history.

This work has developed a GC grouping methodology that combines dynamical and chemical data in a statistically robust manner. A crucial part of our analysis has been the tests using mocks which have highlighted the difficulties inherent in this kind of study. In the face of considerable uncertainties due to the messy nature of accretion, we believe that this philosophy represents an improvement on previous work. In the future, with further development and more data, our method should allow a stronger inference of the MW accretion history.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their comments, which have improved our paper. TC and CSF were supported by the Science and Technology Facilities Council (STFC) (grant numbers ST/F001166/1, ST/I00162X/1, ST/P000541/1) MC also acknowledges support by the EU Horizon 2020 research and innovation programme under a Marie Skłodowska-Curie grant agreement 794474 (DancingGalaxies). AD is supported by a Royal Society University Research Fellowship. CSF acknowledges European Research Council (ERC) Advanced Investigator grant DMIDAS (GA 786910). This work used the DiRAC Data Centric system at Durham University, operated by the ICC on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. RG acknowledges financial support from the Spanish Ministry of Science and Innovation (MICINN) through the Spanish State Research Agency, under the Severo Ochoa Program 2020-2023 (CEX2019-000920-S). FM acknowledges support through the program 'Rita Levi Montalcini' of the Italian MUR.

DATA AVAILABILITY

The used and data produced in this paper are available upon reasonable request to the corresponding author. The membership probabilities of the GCs and the chemo-dynamical fits of the groups are available at https://github.com/TomCallingham/Callingham22_MW_GCs.

REFERENCES

Amorisco N. C., 2019, MNRAS, 482, 2978
Antoja T., Ramos P., Mateu C., Helmi A., Anders F., Jordi C., Carballo-Bello J. A., 2020, A&A, 635, L3
Barba R. H., Minniti D., Geisler D., Alonso-Garcia J., Hempel M., Monachesi A., Arias J. I., Gomez F. A., 2019, ApJ, 870, L24
Bastian N., Pfeffer J., Kruijssen J. M. D., Crain R. A., Trujillo-Gomez S., Reina-Campos M., 2020, MNRAS, 498, 1050
Baumgardt H., Vasiliev E., 2021, MNRAS, 505, 5957
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Bekki K., Freeman K. C., 2003, MNRAS, 346, L11
Bellazzini M., Ibata R., Malhan K., Martin N., Famaey B., Thomas G., 2020, A&A, 636, A107
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Binney J., 2010, MNRAS, 401, 2318
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton
Bland-Hawthorn J., Tepper-García T., 2021, MNRAS, 504, 3168
Boylan-Kolchin M., 2017, MNRAS, 472, 3120
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., 2010, MNRAS, 406, 896
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931
Burkert A., Forbes D. A., 2020, AJ, 159, 56
Callingham T. M. et al., 2019, MNRAS, 484, 5453
Cautun M., Deason A. J., Frenk C. S., McAlpine S., 2019, MNRAS, 483, 2185
Cautun M. et al., 2020, MNRAS, 494, 4291
Chang J., Yuan Z., Xue X.-X., Simion I. T., Kang X., Li T. S., Zhao J.-K., Zhao G., 2020, ApJ, 905, 100
Conroy C. et al., 2019, ApJ, 883, 107
Cooper A. P. et al., 2010, MNRAS, 406, 744
Crain R. A. et al., 2015, MNRAS, 450, 1937
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Deason A. J., Mao Y.-Y., Wechsler R. H., 2016, ApJ, 821, 5
Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, ApJ, 862, L1
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Deason A. J. et al., 2021, MNRAS, 501, 5964
Dotter A. et al., 2010, ApJ, 708, 698
Dotter A., Sarajedini A., Anderson J., 2011, ApJ, 738, 74
Erkal D., Belokurov V. A., Parkin D. L., 2020, MNRAS, 498, 5574
Erkal D. et al., 2021, MNRAS, 506, 2677
Fattahi A. et al., 2019, MNRAS, 484, 4471
Fattahi A. et al., 2020, MNRAS, 497, 4459
Fernández-Trincado J. G. et al., 2020, A&A, 643, A145
Fernández-Trincado J. G. et al., 2021, ApJ, 908, L42
Forbes D. A., 2020, MNRAS, 493, 847
Forbes D. A., Bridges T., 2010, MNRAS, 404, 1203
Forbes D. A. et al., 2018, Proc. R. Soc., 474, 20170616
Frenk C. S., White S. D. M., 1980, MNRAS, 193, 295
Gaia-Collaboration, 2018, A&A, 616, A1
Gaia Collaboration, 2021, A&A, 649, A8
Gómez F. A., Helmi A., Brown A. G. A., Li Y.-S., 2010, MNRAS, 408, 935
Gómez F. A., White S. D. M., Grand R. J. J., Marinacci F., Springel V., Pakmor R., 2017, MNRAS, 465, 3446
Gran F. et al., 2022, MNRAS, 509, 4962

Grand R. J. J. et al., 2017, MNRAS, 467, 179
Grand R. J. J., Deason A. J., White S. D. M., Simpson C. M., Gómez F. A., Marinacci F., Pakmor R., 2019, MNRAS, 487, L72
Grand R. J. J. et al., 2021, MNRAS, 507, 4953
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
Harris W. E., Harris G. L., Hudson M. J., 2015, ApJ, 806, 36
Helmi A., White S. D. M., de Zeeuw P. T., Zhao H., 1999, Nature, 402, 53
Helmi A., Veljanoski J., Breddels M. A., Tian H., Sales L. V., 2017, A&A, 598, A58
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Horta D. et al., 2020, MNRAS, 493, 3363
Horta D. et al., 2021, MNRAS, 500, 1385
Ibata R. A., Gilmore G., Irwin M. J., 1994, Nature, 370, 194
Koposov S. et al., 2007, ApJ, 669, 337
Koppelman H. H., Helmi A., Massari D., Roelenga S., Bastian U., 2019a, A&A, 625, A5
Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019b, A&A, 631, L9
Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019a, MNRAS, 486, 3134
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019b, MNRAS, 486, 3180
Kruijssen J. M. D. et al., 2020, MNRAS, 498, 2472
Kuiper N. H., 1960, Indag. Math., 63, 38
Laporte C. F. P., Johnston K. V., Gómez F. A., Garavito-Camargo N., Besla G., 2018, MNRAS, 481, 286
Laporte C. F. P., Minchev I., Johnston K. V., Gómez F. A., 2019, MNRAS, 485, 3134
Law D. R., Majewski S. R., 2010, ApJ, 718, 1128
Leaman R., VandenBerg D. A., Mendel J. T., 2013, MNRAS, 436, 122
Longmore A. J. et al., 2011, MNRAS, 416, 465
Mackereth J. T., Bovy J., 2020, MNRAS, 492, 3631
Majewski S. R. et al., 2017, AJ, 154, 94
Malhan K., Yuan Z., Ibata R., Arentsen A., Bellazzini M., Martin N. F., 2021, ApJ, 920, 51
Marín-Franch A. et al., 2009, ApJ, 694, 1498
Martell S. L. et al., 2017, MNRAS, 465, 3203
Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4
McMillan P. J., 2017, MNRAS, 465, 76
Minniti D., Fernández-Trincado J. G., Smith L. C., Lucas P. W., Gómez M., Pullen J. B., 2021a, A&A, 648, A86
Minniti D., Fernández-Trincado J. G., Gómez M., Smith L. C., Lucas P. W., Ramos R. C., 2021b, A&A, 650, L11
Minniti D., Gómez M., Alonso-García J., Saito R. K., Garro E. R., 2021c, A&A, 650, L12
Monachesi A. et al., 2019, MNRAS, 485, 2589
Myeong G. C., Evans N. W., Belokurov V., Amorisco N. C., Koposov S. E., 2018a, MNRAS, 475, 1537
Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018b, MNRAS, 478, 5449
Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018c, ApJ, 863, L28
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, ApJ, 901, 48
Necib L., Ostdiek B., Lisanti M., Cohen T., Freytsis M., Garrison-Kimmel S., 2020, ApJ, 903, 25
Neto A. F. et al., 2007, MNRAS, 381, 1450
Ostdiek B. et al., 2020, A&A, 636, A75
Paust N., Wilson D., van Belle G., 2014, AJ, 148, 19
Peñaloza F., Pessev P., Vasquez S., Borissova J., Kurtev R., Zoccali M., 2015, PASP, 127, 329
Peñarrubia J., Petersen M. S., 2021, MNRAS, 508, L26
Pérez-Villegas A., Barbuy B., Kerber L., Ortolani S., Souza S. O., Bica E., 2020, MNRAS, 491, 3251
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
Piffl T. et al., 2014, A&A, 562, A91
Planck Collaboration I, 2014, A&A, 571, A1
Posti L., Helmi A., 2019, A&A, 621, A56
Posti L., Binney J., Nipoti C., Ciotti L., 2015, MNRAS, 447, 3060
Romero-Colmenares M. et al., 2021, A&A, 652, A158
Sakari C. M., Venn K. A., Irwin M., Aoki W., Arimoto N., Dotter A., 2011, ApJ, 740, 106
Schaye J. et al., 2015, MNRAS, 446, 521
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Searle L., Zinn R., 1978, ApJ, 225, 357
Shao S., Cautun M., Frenk C. S., Grand R. J. J., Gómez F. A., Marinacci F., Simpson C. M., 2018, MNRAS, 476, 1796
Simpson C. M., Grand R. J. J., Gómez F. A., Marinacci F., Pakmor R., Springel V., Campbell D. J. R., Frenk C. S., 2018, MNRAS, 478, 548
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2011, Proc. Int. Astron. Union, 270, 203
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
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
Valcarce A. a. R., Catelan M., 2011, A&A, 533, A120
VandenBerg D. A., Brogaard K., Leaman R., Casagrande L., 2013, ApJ, 775, 134
Vasiliev E., 2019, MNRAS, 482, 1525
Vasiliev E., Baumgardt H., 2021, MNRAS, 505, 5978
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Wu Y., Valluri M., Panithanpaisal N., Sanderson R. E., Freese K., Wetzel A., Sharma S., 2021, MNRAS, 509, 5882
Yuan Z., Smith M. C., Xue X.-X., Li J., Liu C., Wang Y., Li L., Chang J., 2019, ApJ, 881, 164
Yuan Z., Chang J., Beers T. C., Huang Y., 2020, ApJ, 898, L37

APPENDIX A: MEMBERSHIP PROBABILITIES

Table A1. The membership probability of individual GCs, as found by our chemo-dynamical model. We give the most likely group and probability of each cluster, and the second most probable alternate group. We also give the groupings from the literature where possible: M19 corresponds to Massari et al. (2019), F20 corresponds to Forbes (2020), and H20 corresponds to Horta et al. (2020).

Name	Alternative	Main group	Prob	Alt group	Alt prob	M19	F20	H20
Djorg2	ESO456	Bulge	1.00	–	–	Bulge	-	–
Terzan6	HP5	Bulge	1.00	–	–	Bulge	–	–
Terzan2	HP3	Bulge	1.00	–	–	Bulge	–	Bulge
NGC 6380	Ton1	Bulge	1.00	–	–	Bulge	–	Bulge
NGC 6440	–	Bulge	1.00	–	–	Bulge	–	–
Liller1	–	Bulge	1.00	–	–	–	–	Ungr
NGC 6642	–	Bulge	1.00	–	–	Bulge	–	–
NGC 6388	–	Bulge	1.00	–	–	Bulge	–	Seq/Bulge
NGC 6535	–	Bulge	1.00	–	–	Kraken/Seq	Seq	–
NGC 6401	–	Bulge	1.00	–	–	Kraken	Kraken	–
Terzan5	11	Bulge	1.00	–	–	Bulge	–	–
NGC 6638	–	Bulge	1.00	–	–	Bulge	–	–
NGC 6528	–	Bulge	1.00	–	–	Bulge	–	–
1636-283	ESO452	Bulge	1.00	–	–	Bulge	–	–
Terzan9	–	Bulge	1.00	–	–	Bulge	–	–
NGC 6624	–	Bulge	1.00	–	–	Bulge	–	–
NGC 6558	–	Bulge	1.00	–	–	Bulge	–	–
Terzan4	HP4	Bulge	1.00	–	–	Bulge	–	–
HP1	BH229	Bulge	1.00	–	–	Bulge	–	Bulge
NGC 6325	–	Bulge	1.00	–	–	Bulge	–	–
NGC 6453	–	Bulge	0.99	–	–	Kraken	Kraken	–
NGC 6626	M28	Bulge	0.99	Disc	0.01	Bulge	–	–
NGC 6304	–	Bulge	0.99	Disc	0.01	Bulge	–	–
NGC 6522	–	Bulge	0.99	Disc	0.01	Bulge	–	Bulge
Terzan1	HP2	Bulge	0.99	Disc	0.01	Bulge	–	–
Pal6	–	Bulge	0.99	Kraken	0.01	Kraken	–	Kraken
NGC 6652	–	Bulge	0.99	Disc	0.01	Bulge	–	–
NGC 6266	M62	Bulge	0.99	Disc	0.01	Bulge	–	–
NGC 6342	–	Bulge	0.99	Disc	0.01	Bulge	–	–
NGC 6637	M69	Bulge	0.98	Disc	0.02	Bulge	–	–
NGC 6355	–	Bulge	0.98	Kraken	0.02	Bulge	–	–
NGC 6540	Djorg	Bulge	0.97	Disc	0.03	Bulge	–	Bulge
NGC 6717	Pal9	Bulge	0.97	Disc	0.03	Bulge	–	–
NGC 6293	–	Bulge	0.95	Kraken	0.05	Bulge	–	–
NGC 6256	–	Bulge	0.94	Disc	0.06	Kraken	Kraken	–
NGC 6517	–	Bulge	0.93	Kraken	0.07	Kraken	Kraken	–
NGC 6144	–	Bulge	0.89	Disc	0.11	Kraken	Kraken	–
VVVCL001	–	Bulge	0.89	Kraken	0.11	–	–	–
NGC 6723	–	Bulge	0.83	Disc	0.17	Bulge	–	Bulge
VVVCL002	–	Bulge	0.80	Kraken	0.19	–	–	–
NGC 6171	M107	Bulge	0.75	Disc	0.25	Bulge	–	Bulge
NGC 6093	M80	Bulge	0.70	Disc	0.16	Kraken	Kraken	–
Gran1	–	Bulge	0.66	Kraken	0.33	–	–	–
NGC 6838	M71	Disc	1.00	–	–	Disc	–	Disc
NGC 5927	–	Disc	1.00	–	–	Disc	–	–
NGC 104	47Tuc	Disc	1.00	–	–	Disc	–	Disc
NGC 6496	–	Disc	1.00	–	–	Disc	–	–
ESO93	–	Disc	1.00	–	–	–	–	–
NGC 6362	–	Disc	1.00	–	–	Disc	–	–
NGC 6366	–	Disc	1.00	–	–	Disc	–	–
NGC 6352	–	Disc	1.00	–	–	Disc	–	–
BH176	–	Disc	1.00	–	–	Disc	–	–
Pal10	–	Disc	0.99	–	–	Disc	–	Disc
E3	–	Disc	0.99	Ungr	0.01	Helmi/?	–	–
NGC 6218	M12	Disc	0.98	Bulge	0.01	Disc	–	Disc
NGC 6441	–	Disc	0.97	Bulge	0.03	Kraken	–	Kraken
Pal11	–	Disc	0.71	GEn	0.28	Disc	–	–
IC1276	Pal7	Disc	0.68	Kraken	0.28	Disc	–	–
Lynga7	BH184	Disc	0.62	Bulge	0.37	Disc	–	–
Pfleiderer	–	Disc	0.38	GEn	0.32	–	–	–
NGC 6205	M13	GEn	1.00	–	–	GEn	GEn	GEn
NGC 362	–	GEn	1.00	–	–	GEn	GEn	GEn

Table A1 *– continued*

Name	Alternative	Main group	Prob	Alt group	Alt prob	M19	F20	H20
NGC 6779	M56	GEn	1.00	–	–	GEn	GEn	–
NGC 7089	M2	GEn	1.00	–	–	GEn	GEn	GEn
NGC 2298	–	GEn	1.00	–	–	GEn	GEn	–
NGC 1851	–	GEn	1.00	–	–	GEn	GEn	GEn
NGC 2808	–	GEn	1.00	–	–	GEn	GEn	GEn
NGC 7099	M30	GEn	1.00	–	–	GEn	GEn	–
NGC 6341	M92	GEn	1.00	–	–	GEn	GEn	GEn
NGC 5286	–	GEn	1.00	–	–	GEn	GEn	–
NGC 1261	–	GEn	1.00	–	–	GEn	GEn	–
ESO-SC06	ESO280	GEn	1.00	–	–	GEn	–	–
NGC 288	–	GEn	1.00	–	–	GEn	GEn	GEn
NGC 5139	oCen	GEn	1.00	–	–	GEn/Seq	Seq	–
NGC 6864	M75	GEn	1.00	–	–	GEn	GEn	–
NGC 5897	–	GEn	0.99	Disc	0.01	GEn	GEn	–
NGC 6235	–	GEn	0.95	Kraken	0.05	GEn	GEn	–
Ryu879	RLGC2	GEn	0.90	Kraken	0.09	–	–	–
BH140	–	GEn	0.90	Disc	0.06	–	–	–
NGC 6656	M22	GEn	0.85	Disc	0.15	Disc	–	Disc
NGC 7078	M15	GEn	0.83	Disc	0.17	Disc	–	Disc
IC1257	–	GEn	0.66	Helmi	0.34	GEn	GEn	–
NGC 5904	M5	Helmi	1.00	–	–	Helmi/GEn	Helmi	GEn/Helmi
NGC 4147	–	Helmi	1.00	–	–	GEn	GEn	–
NGC 5634	–	Helmi	1.00	–	–	Helmi/GEn	Helmi	–
NGC 5272	M3	Helmi	1.00	–	–	Helmi	Helmi	Helmi
NGC 5053	–	Helmi	1.00	–	–	Helmi	Helmi	Helmi
Pal5	–	Helmi	1.00	–	–	Helmi/?	Helmi	Helmi
NGC 7492	–	Helmi	1.00	–	–	GEn	GEn	–
NGC 5024	M53	Helmi	1.00	–	–	Helmi	Helmi	Helmi
NGC 6229	–	Helmi	1.00	–	–	GEn	GEn	GEn
NGC 4590	M68	Helmi	1.00	–	–	Helmi	Helmi	Helmi
NGC 6981	M72	Helmi	0.99	GEn	0.01	Helmi	Helmi	–
Rup106	–	Helmi	0.93	Ungr	0.07	Helmi/?	Helmi	–
NGC 6584	–	Helmi	0.92	GEn	0.08	Ungr	Ungr	–
Bliss1	–	Helmi	0.88	Ungr	0.12	–	–	–
NGC 6426	–	Helmi	0.73	GEn	0.27	Ungr	Ungr	–
NGC 1904	M79	Helmi	0.59	GEn	0.40	GEn	GEn	GEn
NGC 6254	M10	Kraken	1.00	–	–	Kraken	Kraken	Kraken
NGC 6712	–	Kraken	1.00	–	–	Kraken	Kraken	–
NGC 6544	–	Kraken	1.00	–	–	Kraken	Kraken	Kraken
NGC 5946	–	Kraken	1.00	–	–	Kraken	Kraken	–
NGC 6121	M4	Kraken	1.00	–	–	Kraken	–	Kraken
NGC 6809	M55	Kraken	1.00	–	–	Kraken	Kraken	Kraken
NGC 4833	–	Kraken	1.00	–	–	GEn	GEn	–
NGC 6681	M70	Kraken	1.00	–	–	Kraken	Kraken	–
NGC 6287	–	Kraken	1.00	–	–	Kraken	Kraken	–
NGC 5986	–	Kraken	1.00	–	–	Kraken	Kraken	–
NGC 6541	–	Kraken	0.99	–	–	Kraken	Kraken	–
Terzan10	–	Kraken	0.99	–	–	GEn	GEn	–
NGC 6752	–	Kraken	0.99	GEn	0.01	Disc	–	Disc
NGC 6749	–	Kraken	0.99	Disc	0.01	Disc	–	–
NGC 6760	–	Kraken	0.99	Disc	0.01	Disc	–	Disc
UKS1	–	Kraken	0.99	GEn	0.01	–	–	–
NGC 6284	–	Kraken	0.99	GEn	0.01	GEn	GEn	–
Mercer5	–	Kraken	0.98	Disc	0.02	–	–	–
NGC 6397	–	Kraken	0.98	Disc	0.01	Disc	–	Disc
Terzan3	–	Kraken	0.98	Disc	0.02	Disc	–	–
FSR1716	–	Kraken	0.97	Disc	0.03	Disc	–	–
FSR1735	–	Kraken	0.97	Bulge	0.02	Kraken	Kraken	–
NGC 6539	–	Kraken	0.97	Disc	0.02	Bulge	–	Bulge
Ton2	Pismis26	Kraken	0.96	Bulge	0.03	Kraken	Kraken	–
Terzan12	–	Kraken	0.96	Bulge	0.03	Disc	–	–
NGC 6402	M14	Kraken	0.94	Bulge	0.05	Kraken	Kraken	–
Pal8	–	Kraken	0.94	Bulge	0.05	Disc	–	–
NGC 6139	–	Kraken	0.94	Bulge	0.05	Kraken	Kraken	–

Table A1 – *continued*

Name	Alternative	Main group	Prob	Alt group	Alt prob	M19	F20	H20
Djorg1	–	Kraken	0.94	GEn	0.05	GEn	GEn	–
NGC 6553	–	Kraken	0.93	Disc	0.07	Bulge	–	Bulge
NGC 6316	–	Kraken	0.92	Bulge	0.07	Bulge	–	–
NGC 4372	–	Kraken	0.92	Disc	0.04	Disc	–	–
NGC 6273	M19	Kraken	0.90	Bulge	0.10	Kraken	Kraken	–
BH261	AL3	Kraken	0.85	Bulge	0.12	Bulge	–	–
NGC 6569	–	Kraken	0.84	Bulge	0.12	Bulge	–	–
NGC 6333	M9	Kraken	0.84	GEn	0.16	Kraken	Kraken	–
NGC 6356	–	Kraken	0.83	GEn	0.15	Disc	–	–
Arp2	–	Sag	1.00	–	–	Sag	Sag	–
NGC 6715	M54	Sag	1.00	–	–	Sag	Sag	–
Terzan8	–	Sag	1.00	–	–	Sag	Sag	–
Terzan7	–	Sag	1.00	–	–	Sag	Sag	–
Pal12	–	Sag	1.00	–	–	Sag	Sag	–
Whiting1	–	Sag	1.00	–	–	Sag	Sag	–
Munoz1	–	Sag	0.99	Ungr	0.01	–	–	–
Kim3	–	Sag	0.97	Ungr	0.03	–	–	–
Ko1	–	Sag	0.71	Ungr	0.29	–	–	–
NGC 5466	–	Seq	1.00	–	–	Seq	Seq	Seq
NGC 6101	–	Seq	1.00	–	–	Seq/GEn	Seq	–
NGC 7006	–	Seq	1.00	–	–	Seq	Seq	–
NGC 3201	–	Seq	1.00	–	–	Seq/GEn	Seq	Seq
IC4499	–	Seq	1.00	–	–	Seq	Seq	–
Pal13	–	Seq	0.99	Ungr	0.01	Seq	Seq	–
NGC 5694	–	Seq	0.98	Ungr	0.02	Ungr	Ungr	–
Pal15	–	Seq	0.96	Ungr	0.04	GEn/?	GEn	–
AM4	–	Seq	0.83	Ungr	0.15	–	Sag	–
Ryu059	RLGC1	Ungr	1.00	–	–	–	–	–
Ko2	–	Ungr	1.00	–	–	–	–	–
Pal3	–	Ungr	1.00	–	–	Ungr	Ungr	–
NGC 6934	–	Ungr	1.00	–	–	Ungr	Ungr	–
Crater	–	Ungr	1.00	–	–	Ungr	Ungr	–
Pyxis	–	Ungr	1.00	–	–	Ungr	Ungr	–
Segue3	–	Ungr	1.00	–	–	–	–	–
Pal14	–	Ungr	1.00	–	–	Ungr	Ungr	–
AM1	–	Ungr	1.00	–	–	Ungr	Ungr	–
Eridanus	–	Ungr	1.00	–	–	Ungr	Ungr	–
Pal4	–	Ungr	1.00	–	–	Ungr	Ungr	–
Pal1	–	Ungr	1.00	–	–	Disc	GEn	–
NGC 5824	–	Ungr	0.99	Helmi	0.01	Sag	Sag	–
NGC 2419	–	Ungr	0.98	Seq	0.02	Sag	Sag	–
Laevens3	–	Ungr	0.94	Seq	0.06	–	–	–
Pal2	–	Ungr	0.81	Seq	0.14	GEn	GEn	–
FSR1758	–	Ungr	0.62	Seq	0.28	Seq	Seq	–

APPENDIX B: GAUSSIAN FITTING OF SMALL GROUPS

In our model, in principle, all data points contribute to each component, although some points can have very low responsibilities. On average, each component fits $N_{\rm points} = W_c \times N_{\rm total}$ points, where $N_{\rm total}$ is the total number of GCs. If the weight of the component is such that $N_{\rm points} < n_{\rm dim}$, where $n_{\rm dim}$ is the number of dimensions of the space, then Σ tends to become degenerate within machine precision. This causes the responsibility of the GCs to tend to one and the fit is unable to improve. To prevent this, after calculating the covariance matrix, we change the $\lfloor n_{\rm dim} - N_{\rm points} \rfloor$ smallest eigenvalues to half of the smallest non-degenerate value. If $N_{\rm points}$ drops below 1.5, we set the eigenvalue to be 0.05 (note that internally the space is scaled by the 25–75 per cent range to be dimensionless). If $N_{\rm points}$ drops below 0.5, the cluster is then considered extinct, and the normalization weight is set to zero. We note that the weights of the MW groups generally do not decrease sufficiently when performing the multicomponent fit to cause this issue. This affects only a few groups from the mock samples, but is nevertheless important to include to accurately fit the groups.

APPENDIX C: MISCLASSIFYING THE INITIAL GROUPS

Here, we study how sensitive our GC grouping algorithm is on the initial groupings used as the starting point of our iterative method. We have explored this by selecting a subset of GCs and by changing their label to another group. When misclassifying GCs in observations, they are not assigned to a random group, but actually to a neighbouring group. To identify in a simple way the closest incorrect neighbour for each GC we proceed by calculating the best-fitting distributions in a non-iterative way. This corresponds to applying the maximization

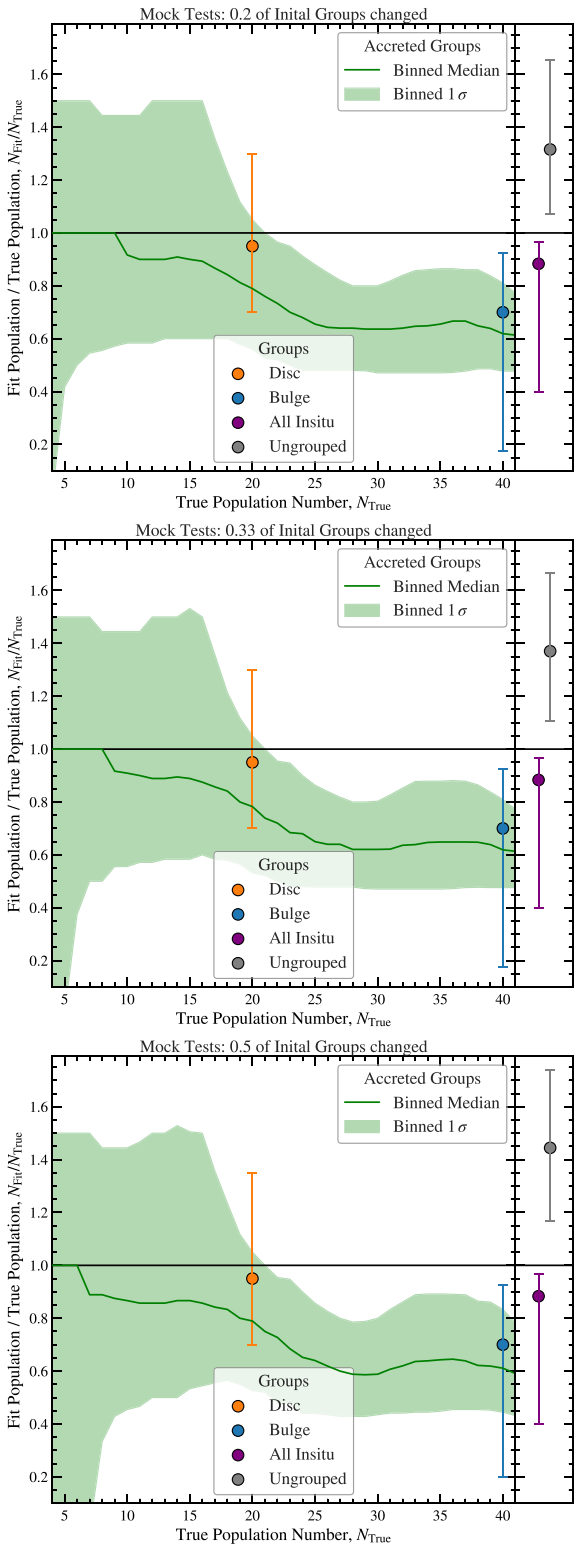

Figure C1. Like Fig. 5, but now showing how the ratio $N_{\rm fit}/N_{\rm true}$ as a function of $N_{\rm true}$ depends on the fraction of mislabelled GCs present in the groupings used to initialize our iterative clustering algorithm. The tests are done on the AURIGA mock GC catalogues. In each panel, from top to bottom, we reassign a fraction of, respectively, 0.2, 0.33, and 0.5 of the GCs from their true group to the next closest group.

step of our algorithm, where the responsibilities are calculated using the true GC labels, and then updating the responsibilities using these new best-fitting distributions. Then, we assign to the misclassified GCs the label associated with the most likely group that was not their true group. We study what is the impact of such an initial mislabelling of GC groups as a function of the misclassified fraction. In particular, we are interested in testing the effect of starting with 33 per cent of the GCs incorrectly identified, approximately in line with the average final purity and completeness of our methodology.

We redo most of the analysis presented in Section 4.2 but now starting from initial labels that have various fractions of misclassified GCs. To start with, we study how the ratio, $N_{\rm fit}/N_{\rm true}$, of the recovered to true group richness changes when mislabelled 20 per cent, 33 per cent, and 50 per cent of the GCs. This is shown in Fig. C1. We can see that increasing the fraction of GCs with incorrect initial groups increases the error spread gradually. The greatest effects can be seen in the smallest groups. This makes intuitive sense: changing a single GC in a poor group represents a more significant change than for a richer group. Furthermore, this large change can lead to the group no longer being well modelled in dynamical space, potentially leading the group to go extinct as the method iterates. The larger groups are, in general, robust to even large changes, as enough of the true members remain for the average position of the group to be found and the majority of the group recovered.

However, it is not until the change reaches the 50 per cent level that the median trend is changed, and then only for small groups. For a 33 per cent misclassification fraction, roughly the uncertainty resulting from our clustering process, there is little change in the distribution of $N_{\rm fit}/N_{\rm true}$ compared to the case of the true starting groups. While not shown, the trends in recovered group purity and completeness are similarly robust to initially mislabelled GCs, especially for a 33 per cent or lower misclassification fraction.

To conclude, this shows that our methodology is insensitive to potential misclassifications as large as 33 per cent (and even 50 per cent) of the GC groups used to initialize our clustering algorithm. While the test performed here is not a direct equivalent to any mislabelling present in the literature groupings used to initialize our method when applied to the MW data, we believe that it indicates that our methodology and results are robust in application to the MW.

APPENDIX D: REDSHIFT DEPENDENCE OF THE SMHM

In the main paper, when using the SMHM relation of Behroozi et al. (2019), we assumed the present day, $z = 0$, relation (see Fig. 11). Alternatively, it is reasonable to assume that star formation in the accreted satellites approximately stops upon accretion. We consider the approximate infall time estimates from Kruijssen et al. (2020): Kraken, $z_{\rm Acc} = 2.26$; Helmi Streams, $z_{\rm Acc} = 1.75$; Sequoia, $z_{\rm Acc} = 1.46$; GES, $z_{\rm Acc} = 1.35$; and Sagittarius, $z_{\rm Acc} = 0.76$. The resulting PDFs, and original $z = 0$ estimates, can be seen in Fig. C1.

The effect of truncating the star formation histories in accreted galaxies lowers the inferred stellar masses. This truncation has the greatest effect on older mergers, such as Kraken (with approximately a three times decrease in stellar mass). The total mass is approximately reduced by a factor of 2.5. This effect is considerable but depends on the highly uncertain accretion times and the considerable systematic uncertainty of the SMHM relation at high redshift.

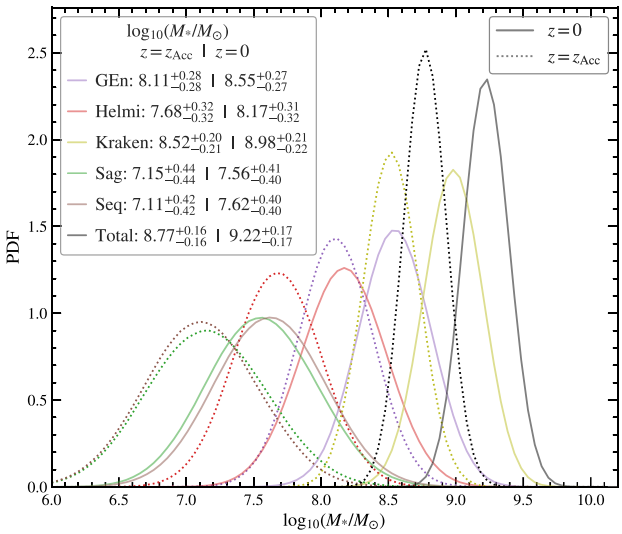

Figure D1. The likelihoods of the stellar mass of accreted galaxies, calculated by assuming the SMHM relation of Behroozi et al. (2019) to transform the halo mass PDFs of Fig. 10. This relationship is dependent on redshift. In this figure, we compare the results found by assuming the present day, $z = 0$, relationship (as shown in Fig. 11). We use estimates of the approximate accretion time of the groups found in the literature (see the main text). Due to the large systematic uncertainties on these accretion times, we consider these results as demonstrative of the effect of redshift dependence.

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