Exploring the diversity and similarity of radially anisotropic Milky Way-like stellar haloes: implications for disrupted dwarf galaxy searches

Matthew D. A. Orkney [1,2]★ Chervin F. P. Laporte [1,2] Robert J. J. Grand [3,4,5] Facundo A. Gómez [6] Freeke van de Voort [7] Azadeh Fattahi [8] Federico Marinacci [9] Rüdiger Pakmor [10] Francesca Fragkoudi[8] and Volker Springel [10]

[1]*Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain*
[2]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[3]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[4]*Instituto de Astrofísica de Canarias, Calle Via Lactea 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]*Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile*
[7]*Cardiff Hub for Astrophysics Research and Technology, School of Physics and Astronomy, Cardiff University, Queen's Buildings, The Parade, Cardiff CF24 3AA, UK*
[8]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[9]*Department of Physics & Astronomy 'Augusto Righi', University of Bologna, via Gobetti 93/2, I-40129 Bologna, Italy*
[10]*Max Planck Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching, Germany*

ABSTRACT

We investigate the properties of mergers comparable to the Gaia–Sausage–Enceladus (GSE) using cosmological hydrodynamical simulations of Milky Way-like galaxies. The merger progenitors span an order of magnitude in their peak stellar mass ($3 \times 10^8 < M_\star/M_\odot < 4 \times 10^9$) and include both rotation and pressure-supported galaxies ($0.10 < D/T < 0.77$). In a minority of cases, the GSE-like debris is comprised of stars from more than one merger progenitor. However, there is a close similarity in their chemodynamical properties and the triaxial shapes of their debris, and so it is not always possible to distinguish them. The merger progenitors host a variety of luminous satellites (0 and 8 with $M_\star > 10^6\,M_\odot$), but most of these do not follow the merger to low orbital energies. Between 0 and 1 of these satellites may survive to $z = 0$, but with no clear signatures of their past association. We show that the fraction of stars originating from GSE-like mergers is reduced for lower metallicities (reaching a minimum around [Fe/H] $= -2$), and also within 5 kpc of the Galactic Centre. Whilst these central regions are dominated by *in-situ* stars, the *ex-situ* fraction trends towards a 100 per cent asymptote when considering the most metal-poor stars ([Fe/H] $\ll -2.5$). Considering this, its near proximity, and its small volume on the sky, the Galactic Centre lends itself as a prime environment in the search for the stars from the earliest galaxies, whilst avoiding contamination from GSE stars.

Key words: methods: numerical – Galaxy: abundances – Galaxy: centre – Galaxy: evolution – Galaxy: kinematics and dynamics – Galaxy: structure.

1 INTRODUCTION

In the Lambda cold dark matter (ΛCDM) paradigm, present-day galaxies are formed through the hierarchical assembly of numerous mergers (Press & Schechter 1974; Searle & Zinn 1978; White & Rees 1978; Blumenthal et al. 1984; White & Frenk 1991; Kauffmann, White & Guiderdoni 1993), and their final properties are a product of this intricate assembly process. Galaxy mergers often produce visible substructure within the stellar haloes of the host galaxy, such as surviving remnant objects (e.g. the Sagittarius dwarf in the MW Ibata, Gilmore & Irwin 1994; Majewski et al. 2003), shells (e.g. Schweizer 1986; Martínez-Delgado et al. 2008; Dey et al. 2023), or streams (e.g. Lynden-Bell & Lynden-Bell 1995; Belokurov et al. 2006, 2007b;

Martínez-Delgado et al. 2015). However, many of the most ancient accreting galaxies will have gradually dissolved due to cumulative perturbations and dynamical relaxation (see Ibata et al. 2002; Mayer et al. 2002; Peñarrubia et al. 2006; Gómez et al. 2013; Buist & Helmi 2015; Maffione et al. 2015; Erkal, Sanders & Belokurov 2016; Ngan et al. 2016; Price-Whelan et al. 2016; Maffione et al. 2018), and eventually become smoothly integrated within the stellar halo. In these cases, there may yet be evidence of their existence in the form of lingering features in the chemical and dynamical planes.

Observations infer that ancient Milky Way (MW) type galaxies underwent a phase of rapid merging in the early Universe (Unavane, Wyse & Gilmore 1996; Bell et al. 2008), and this is reinforced by the predictions of cosmological simulations (Bullock & Johnston 2005; Bell et al. 2008; Cooper et al. 2010; Kruijssen et al. 2019; Renaud et al. 2021a). Fortunately, the progenitor galaxies of these mergers can be distinguished by their unique chemical

★ E-mail: m.d.a.orkney@gmail.com

evolution, which is itself tied to the details of their formation history (e.g. Freeman & Bland-Hawthorn 2002; Venn et al. 2004; Gallart, Zoccali & Aparicio 2005; Robertson et al. 2005; Tolstoy, Hill & Tosi 2009). Furthermore, their debris is linked by shared integrals of motion that relate to their infall and internal kinematics (e.g. Helmi & de Zeeuw 2000; Font et al. 2011; Amarante et al. 2022). Given the long dynamical times within the stellar halo of MW-mass galaxies ($\mathcal{O}$(Gyrs)), chemodynamic signatures of ancient mergers may persist to the present day (e.g. Eggen, Lynden-Bell & Sandage 1962; Freeman & Bland-Hawthorn 2002; Helmi, White & Springel 2003; Nissen & Schuster 2010; Hayes et al. 2018; Vera-Casanova et al. 2022), albeit with some degree of phase-mixing. Investigating these signatures, and using them to reconstruct the history of our own Galaxy, are a core goal in the field of Galactic archaeology.

Early exploration of old metal-poor halo stars around the solar neighbourhood revealed a characteristic radially anisotropic velocity distribution (Chiba & Beers 2000). This result was interpreted in the context of preexisting formation models of the inner galaxy that this population of stars formed from the monolithic collapse of radially infalling gasses (Eggen et al. 1962), and/or they were inherited from radially infalling dwarf galaxy mergers (Searle & Zinn 1978; Helmi et al. 1999), or even that they were thick-disc stars which were dynamically heated onto radial orbits following a disruptive merger event (e.g. Quinn, Hernquist & Fullagar 1993; Villalobos & Helmi 2008). Chemodynamical simulations favoured the satellite accretion scenario (Brook et al. 2003), with some citing observed α-abundance versus metallicity trends as evidence for a relation to the hypothetical progenitor galaxy of the ω Cen globular cluster (Meza et al. 2005; Nissen & Schuster 2010; Limberg et al. 2022).

Astrometric and photometric data from the *Gaia* (Gaia Collaboration et al. 2016) and SDSS (Kollmeier et al. 2017) missions have supported more extensive explorations of the solar neighbourhood than ever before. This has revealed that the velocities of MW halo stars are highly radially anisotropic at sub-solar metallicities (Belokurov et al. 2018). The velocity anisotropy can be parametrized following Binney (1980) as

$$\beta(r) = 1 - \frac{\langle v_t^2 \rangle}{2 \langle v_r^2 \rangle}, \qquad (1)$$

where $\langle v_t^2 \rangle$ and $\langle v_r^2 \rangle$ are the squared velocity dispersions in the tangential and radial directions. Here, $\beta = 1$ describes a velocity distribution that is radially anisotropic, $\beta = 0$ is isotropic, and $\beta = -\infty$ is tangentially anisotropic. Following this definition, the anisotropy parameter of MW halo stars peaks at $\beta \sim 0.9$ for metallicities [Fe/H] > −1.7 (Belokurov et al. 2018). This extreme value cannot easily be explained by an accretion of numerous dwarf galaxies, because the orbital variation between each progenitor galaxy would yield a more isotropic velocity distribution in their debris. Instead, Belokurov et al. (2018) invoked a single massive dwarf accretion at a redshift between $1 < z < 3$ with virial mass $> 10^{10}\,\mathrm{M}_\odot$. This scenario was corroborated elsewhere due to discoveries of breaks in the stellar halo density profile (Deason et al. 2013; Amorisco 2017), evidence for a two-component halo in colour–magnitude diagrams (Gaia Collaboration et al. 2018; Haywood et al. 2018; Helmi et al. 2018), stellar kinematics (Koppelman, Helmi & Veljanoski 2018), and stellar ages (Gallart et al. 2019a). In particular, Helmi et al. (2018) showed that the α-abundance versus metallicity trends of these halo stars were more consistent with having formed in a separate galaxy than in the MW thick disc. This possible accretion event

has since become known as the Gaia–Sausage–Enceladus (GSE), and evidence favouring its existence has continued to accumulate (e.g. Bignone, Helmi & Tissera 2019; Mackereth et al. 2019; Naidu et al. 2020; Myeong et al. 2022). However, constraining various progenitor properties, including the absolute number of progenitors, may be more difficult than anticipated (i.e. see Rey et al. 2023).

As 6D phase-space and chemical observations of the stellar halo have improved, it has become possible to identify substructures related to the GSE debris (Simion, Belokurov & Koposov 2019; Perottoni et al. 2022), as well as a myriad of other possibly distinct debris features (e.g. *Sequoia:* Myeong et al. 2018; Barbá et al. 2019; Myeong et al. 2019, *Kraken*[1]: Kruijssen et al. 2019; Horta et al. 2021; Naidu et al. 2022, *Thamnos:* Koppelman et al. 2019, *Wukong:* Naidu et al. 2020; Yuan et al. 2020b, *Icarus:* Re Fiorentin et al. 2021, *Pontus:* Malhan et al. 2022, etc.). Whilst evidence in favour of the GSE massive merger scenario has continued to grow, the consensus around the origins of these other debris groups is more uncertain. The very retrograde high-energy halo debris may be the aggregate of three independent accretion events (Myeong et al. 2019; Naidu et al. 2020) or stars from the outskirts of the GSE progenitor (Koppelman, Bos & Helmi 2020; Naidu et al. 2021; Amarante et al. 2022). There are arguments that *Kraken* could be a population of low-metallicity *in-situ* stars (Belokurov & Kravtsov 2022; Myeong et al. 2022; Orkney et al. 2022; Rix et al. 2022), though statistical methods in Horta et al. (2022) suggest that *Kraken* is distinguishable from this *in-situ* population by its lower α-abundance. Some other objects are difficult to differentiate from the pervasive GSE debris, and may represent overdensities in a fragmented GSE debris footprint (e.g. Amarante et al. 2022), or possibly satellite galaxies of other major progenitors. Furthermore, the tendency for different merger debris distributions to overlap and fragment can make it impossible to correctly identify their origins (Jean-Baptiste et al. 2017). A more rigorous consideration of the GSE debris, its extents in various chemical and dynamical properties, and its contribution fraction compared to other debris groups, would help to substantiate or relieve these concerns.

In this paper, we analyse a selection of MW-type galaxies from the AURIGA simulation suite. Fattahi et al. (2019, hereafter F19) show that these galaxies are host to radially anisotropic inner halo debris features with properties broadly comparable to the GSE, and further identify the main mergers that contribute to these debris features. We explore the composition of these debris features, finding in some cases there are two or three separate mergers that contribute non-negligible mass fractions. We then investigate the properties of the merger that contributes the largest fraction of the GSE-like debris, both in terms of their pre-infall galaxies and the properties of their resulting debris at $z = 0$. These mergers span a wide range of properties, but these are not necessarily a good predictor of the properties in their debris.

We describe the AURIGA suite, our simulation sample and our post-processing choices in Section 2. Our results are shown in Section 3: we illustrate the contributions of each merger to their respective GSE-like features in Section 3.1, then we investigate a selection of different properties for both the progenitor mergers and their debris in Section 3.2, their populations of luminous satellites in Section 3.3, and overall trends in the halo debris in Section 3.4. We discuss the implication of these results in Section 4, and conclude in Section 5.

---

[1]Also known as *Heracles*.

2 METHOD

2.1 Simulation suite

The simulations analysed in this paper are taken from the AURIGA project (Grand et al. 2017). AURIGA includes thirty magnetohydrodynamic simulations of isolated MW-mass galaxies within a full cosmological context. These galaxies represent a wide diversity of accretion histories, and were not specifically designed to match the assembly history of the MW.

The target galaxies were originally selected from the dark matter (DM) only version of the Ref-L100N1504 cosmological volume in the EAGLE project (Schaye et al. 2015), and then resimulated using the Tree-PM moving-mesh code AREPO (Springel 2010) using the 'zoom' approach (Katz & White 1993; Frenk et al. 1996). The initial conditions were generated with the Gaussian white-noise realization PANPHASIA (Jenkins 2013), within a periodic cosmological box of side length 100 cMpc. The suite uses cosmological parameters from Planck Collaboration XVI (2014), which are $\Omega_\mathrm{m} = 0.307$, $\Omega_\mathrm{b} = 0.04825$, $\Omega_\Lambda = 0.693$, and a Hubble constant of $H_0 = 100h\,\mathrm{km^{-1}\,Mpc^{-1}}$, where $h = 0.6777$.

AURIGA includes physical models for a spatially uniform photoionizing UV background, primordial and metal line cooling, star formation, stellar evolution and supernovae feedback, supermassive black hole growth and feedback, and magnetic fields. See Grand et al. (2017) for a description of these sub-grid physics models, which are able to reproduce a range of expected galaxy properties in cosmological contexts (Vogelsberger et al. 2013; Genel et al. 2014; Marinacci, Pakmor & Springel 2014). The AURIGA galaxies have realistic properties that are generally compatible with MW-type galaxies in terms of their halo mass–metallicity relations (Monachesi et al. 2019), disc oscillations (Gómez et al. 2017a), rotation curves and star formation rates (Grand et al. 2017), thin/thick disc dichotomy (Grand et al. 2018), and the chemodynamical properties of their innermost regions (Fragkoudi et al. 2020).

In this work, we utilize the 'level-4' resolution versions of each AURIGA simulation, where the target high-resolution region is resolved with a DM particle mass of $\sim 3 \times 10^5\,\mathrm{M_\odot}$ and baryonic mass of $\sim 5 \times 10^4\,\mathrm{M_\odot}$. Each simulation includes 128 snapshots over the range $127 \geq z \geq 0$. Hereafter, we refer to the AURIGA simulations as Au-i, where i indicates the particular halo.

2.2 Post-processing

Virial properties and a varied assortment of group and subhalo properties are calculated using the SUBFIND halo finder (Springel et al. 2001), and are derived using a sphere of mean density 200 times the critical density of the universe. Haloes and subhaloes are linked across time-steps using the LHALOTREE merger tree algorithm (Springel et al. 2005).

AURIGA tracks a selection of different chemical abundances, including iron and α-process elements. We normalize the abundance ratios to solar values following Asplund et al. (2009), and apply further systematic shifts of -0.4 to the [Fe/H] ratios (and equivalent for other abundance ratios). Similar corrections were also used in F19 and Grand et al. (2020), and are calibrated based on observations of the MW abundance ratios.

We designate a star particle as '*in situ*' if it was bound to the potential well of the main progenitor halo at its formation time (as in Cooper et al. 2015), and '*ex situ*' if it formed within the potential well of satellite haloes or their progenitors. This choice means that stars forming within recently stripped gas are counted as *in situ*, although

realistically their identification could be more ambiguous. Therefore, we expect this method to represent an upper-bound on the fraction of *in-situ* stars.

All galaxies are reoriented to align on the angular momentum of the innermost *in-situ* star particles ($R_\mathrm{G} < 0.1 \times R_{200}$, where R_G is the galactocentric radius), such that any co-rotating discs are viewed face-on in the *x–y* plane. The gravitational potential energies of each star particle are recovered directly from the simulation data, and are normalized such that the spherically symmetric potential profile at the R_{200} virial radius is 0.

It is occasionally necessary to distinguish the co-rotating stellar disc from the stellar halo. We achieve this by calculating the circularity parameter for each star particle (as in Abadi et al. 2003; Grand et al. 2017; Gómez et al. 2017b), defined as

$$\epsilon = \frac{L_z}{L_z^\mathrm{max}(E)}, \tag{2}$$

where L_z is the *z*-component of the angular momentum, and $L_z^\mathrm{max}(E)$ is the maximum angular momentum that is allowed for the given specific orbital energy. Here, $\epsilon = 1$ corresponds to prograde circular orbits in the plane of the disc, $\epsilon = -1$ corresponds to retrograde orbits, and $\epsilon = 0$ corresponds to orbits with low angular momentum in the *z* direction.

We follow the assumption that each galaxy consists of a stellar spheroid with net zero rotational velocity and a symmetric distribution centred on $\epsilon = 0$ (the stellar halo), and a component that is co-rotating with $\epsilon > 0$ (the stellar disc/bar). The spheroid is estimated by mirroring the retrograde ϵ distribution about 0. The probability of a star particle being in the halo (p_halo) can be estimated by dividing the total ϵ distribution by the spheroidal ϵ distribution. Each star particle is then assigned to the disc or halo by drawing a random number n in the range 0–1, where $n > p_\mathrm{halo}$ corresponds to a disc star.

This method does not discriminate between different kinds of co-rotating distributions, such as stellar discs, bars, or clumpy features. The probabilistic determination used in this technique means there is always a chance that a disc particle is erroneously assigned to the high-ϵ portion of the halo, and vice versa. The disc and halo populations should typically have unique chemical abundance distributions, with disc stars tending to be more metal rich than halo stars, and these distributions will become smeared together in OUR selection. Therefore, we include the additional requirement that the retrograde and prograde sides of the stellar halo should share the same [Fe/H] distribution function. This assumption is incorporated into our calculation of p_halo, but we note that none of our reported results are dependent on it.

2.3 GSE-like features in AURIGA

F19 investigated the velocity anisotropy of inner halo stars for 28 AURIGA simulations. They identified MW-like galaxies in 10 simulations which possessed prominent radially anisotropic debris features with $\beta > 0.8$ and contributions to the stellar halo greater than 50 per cent, which were considered to be comparable to the GSE. These simulations were Au-5, 9, 10, 15, 17, 18, 22, 24, 26, and 27.

F19 also identified the progenitor merger that contributed the greatest mass fraction of this debris feature within specific metallicity, galactic height, and velocity ranges. In many cases, these mergers are responsible for the majority of the mass fraction, but in some other cases they are responsible for only a plurality. We focus on only the main progenitor mergers for the purposes of this paper, and ignore any other mergers that also contribute to the radially anisotropic

Table 1. The properties of the GSE-type mergers considered are in this work. Here, z_{infall} is the time at which the merger passes the coeval R_{200} virial radius of the central AURIGA galaxy. Pre-infall masses are determined for all member particles at the snapshot prior to z_{infall}, and are therefore minimally impacted by gravitational tides. The merger mass fraction ($M_{200}/M_{200,\mathrm{Host}}$) is found at the last snapshot where each object is a distinct group according to SUBFIND. Gas fractions are calculated as $M_{\mathrm{gas}}/(M_{\mathrm{gas}} + M_\star)$, for all material within twice the stellar half-mass radius. We also include the peak stellar masses using all bound stars found by SUBFIND, and corresponding stellar half-mass radii. The pre-infall disc to total mass ratio, as defined in Section 2.2, is given as *D/T*.

Run	GSE-type merger	z_{infall}	M_{200}(pre-infall) $[10^{10}\,\mathrm{M}_\odot]$ (pre-infall)	$M_{200}/M_{200,\,\mathrm{Host}}$ (pre-infall)	M_{gas} (pre-infall) $[10^{10}\,\mathrm{M}_\odot]$ (pre-infall)	f_{gas}	$M_\star$(pre-infall/peak) $[10^{10}\,\mathrm{M}_\odot]$ [kpc]	$R_{\mathrm{half},\,\star}$(pre-infall/peak) (pre-infall)	*D/T*
Au-5	Au-5-M	0.90	9.59	0.24	1.73	0.53	0.32/0.38	2.51/2.63	0.10
Au-9	Au-9-M	1.91	7.02	0.68	1.43	0.84	0.15/0.19	3.66/5.24	0.30
Au-10	Au-10-M1	0.90	3.62	0.08	0.88	0.71	0.09/0.10	4.40/4.41	0.77
	Au-10-M2	0.75	3.33	0.06	0.73	0.63	0.08/0.09	2.24/2.43	0.18
Au-15	Au-15-M	0.82	10.92	0.41	2.08	0.44	0.22/0.25	4.49/3.84	0.57
Au-17	Au-17-M	2.47	3.12	0.11	0.70	0.77	0.03/0.04	1.86/2.21	0.20
Au-18	Au-18-M	1.39	3.82	0.16	0.87	0.64	0.13/0.14	3.10/3.16	0.10
Au-22	Au-22-M	2.85	2.54	0.20	0.62	0.83	0.03/0.03	2.29/2.27	0.21
Au-24	Au-24-M	1.35	9.13	0.15	1.65	0.42	0.22/0.26	3.72/3.56	0.48
Au-27	Au-27-M	1.59	10.82	0.25	1.94	0.64	0.39/0.41	3.89/3.98	0.57

feature. We make an exception for Au-10, for which there are two mergers that contribute almost equivalent mass fractions at the solar radius, and in this case, we investigate both objects to see if they can be distinguished. We exclude Au-26 from our analysis because the main progenitor merger is so massive ($M_\star > 10^{10}\,\mathrm{M}_\odot$) that it is difficult to reconcile with the expected properties of the GSE. We list these mergers, alongside a selection of key properties, in Table 1.

Throughout this paper, we use the nomenclature 'GSE-type' to refer to the merger identified as contributing the most stars to the radially anisotropic debris feature. Similarly, 'GSE-like debris' refers only to the debris originating from this particular merger event. We refer to the MW-mass galaxy as the 'central AURIGA' galaxy.

3 RESULTS

3.1 Radially anisotropic debris

3.1.1 *ex-situ contributions*

The *ex-situ* stellar density for each selected AURIGA simulation is shown at $z = 0$ in Fig. 1 as a series of grey-scale histograms. We perform a cut on stars with galactocentric radii $R_{\mathrm{G}} < 30$ kpc, metallicities $-2 < [\mathrm{Fe/H}] < -0.5$, and height above the disc plane of $|Z| > 3$ kpc, which is intended to generously bracket the expected properties of GSE-type debris.

We define a 'radially anisotropic debris' region to be all *ex-situ* stars with polar velocity coordinates of $|v_\phi| < 50\,\mathrm{km\,s^{-1}}$ and $100 < |v_{\mathrm{r}}|/\mathrm{km\,s^{-1}} < 400$, as indicated by the dashed black rectangles in each panel. These are similar to the kinematic cuts employed in F19, and are designed to eliminate most of the isotropic halo stars. We include a pie chart in the lower right-hand corner of each panel, which displays the mass contribution fractions within these velocity cuts. The stars from the four most major contributors are shown with unique colours (black, dark blue, blue, and light blue), and the debris from all remaining progenitor galaxies is shown in combination (red).

We take the four mergers that contribute the greatest radially anisotropic mass fractions, and draw contours around the perimeters of their debris in order to convey the shapes of their full velocity distributions. This is based upon an isodensity contour of a smoothed Gaussian kernel density estimate that encircles 80 per cent of the stellar mass, or else a mass-weighted area density of $10\,\mathrm{M}_\odot\,\mathrm{km^{-2}\,s^2}$ if the star particles are sparsely distributed. The contour line width reflects the total mass contribution without the velocity cuts, and is

therefore complementary to the pie chart. The debris from the GSE-type merger is in all cases radially anisotropic, with $\beta > 0.6$ over the property ranges used here.

The debris from other significant mergers adopt a range of contribution fractions and anisotropy. In Au-9, there is a subdominant merger that is even more radially biased ($\beta = 0.92$) than the GSE-type merger ($\beta = 0.63$), although it contributes only 11 per cent of the anisotropic mass fraction. In Au-10 and Au-22, there are subdominant mergers which contribute high-mass fractions and high anisotropy ($\beta > 0.7$). In particular, the two greatest contributors in Au-10 are near-equal over a wide range of radii. On the other hand, there are subdominant mergers such as that in Au-24 which has a relatively high mass fraction of 26 per cent but a significantly lower anisotropy ($\beta = 0.44$).

We show an example merger tree visualization in Appendix A, with the evolutionary tracks of the four largest contributors from Fig. 1 in the same colours. We take this opportunity to highlight a problem in the LHALOTREE algorithm used to generate these merger trees. The subhalo identification occasionally stops tracking a merger remnant as it nears a pericentre passage. This behaviour is normally corrected by allowing a halo descendant to skip a snapshot (see Springel et al. 2005), but in some cases, the remnant is re-identified after emerging on the other side of its pericentre as an entirely independent merger event. We have found two cases where this situation affects the mergers identified in Fig. 1, and we have corrected this by summing their debris together. As a result of this change, the most massive contributor in Au-22 is a different object to the one found in F19.

3.1.2 *Progenitor mass*

The progenitor masses shown in Table 1 bracket almost an order of magnitude, and there is a wide variation in their merger mass ratios (from <1: 10 to 1: 4). None the less, all subsequent mergers have a lower merger mass ratio than these GSE-type mergers. Furthermore, there are only two examples (Au-15, Au-24) in which there is a subsequent merger with a mass ratio >1: 10, and only one of these examples (Au-24) contributes a significant stellar mass fraction to the solar neighbourhood. Therefore, these GSE-type mergers can be considered the last 'significant' merger in their respective host galaxies, similar to what is expected for the MW.

Belokurov et al. (2018) originally estimated that the GSE progenitor would need a virial mass of $> 10^{10}\,\mathrm{M}_\odot$, which is consistent with the GSE-type mergers in Table 1. Using the redshift-dependent stellar
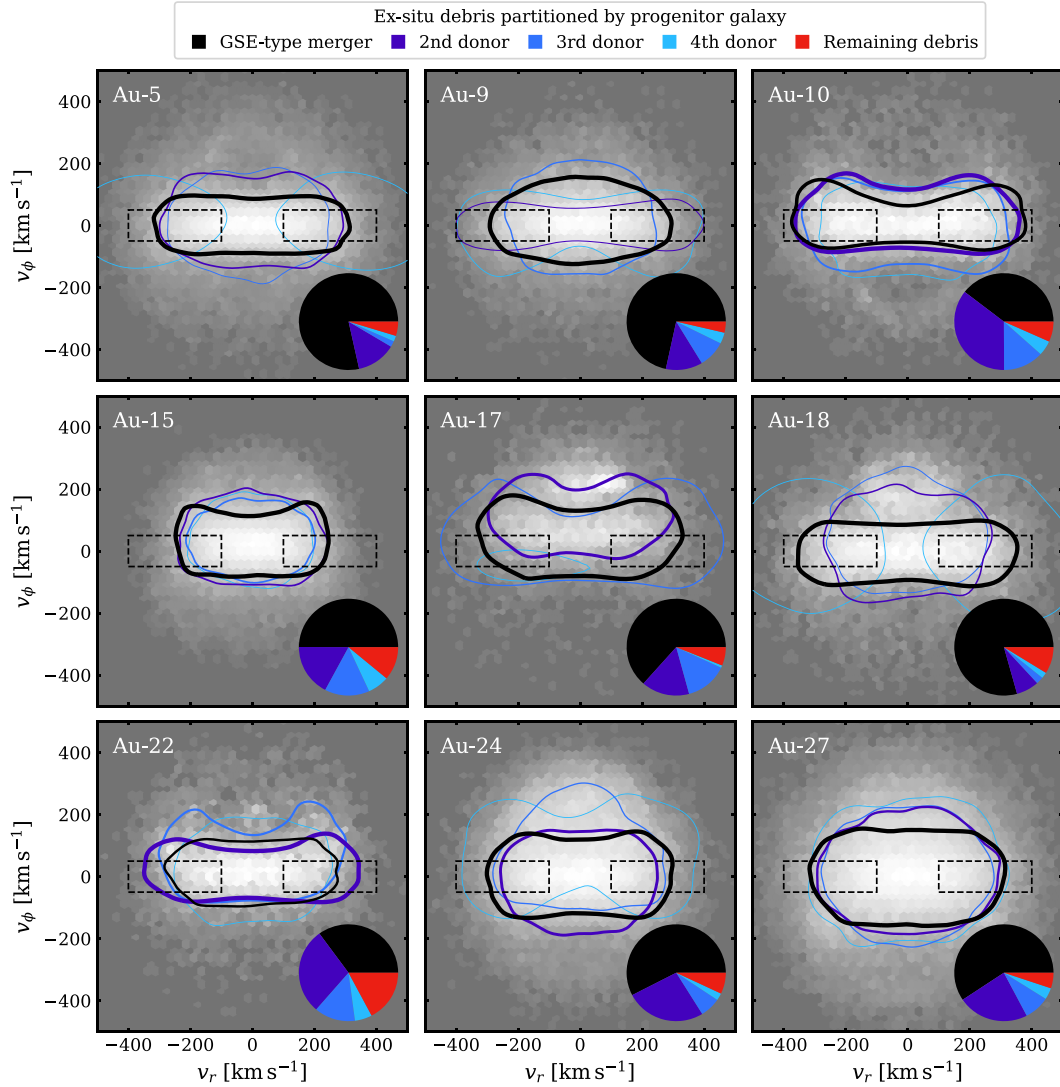
Figure 1. Each panel shows a histogram of *ex situ* stellar mass in v_ϕ/v_r spherical polar coordinates, with the following parameter cuts: $R_G < 30$ kpc; $-2 < [\text{Fe/H}] < -0.5$; $|Z| > 3$ kpc. The pie charts in the lower right-hand corners show the mass contribution fractions of stars within the radially anisotropic black-dashed rectangle regions: $|v_\phi| < 50$ km s^{-1} and $100 < |v_r|/\text{km s}^{-1} < 400$. Each pie segment represents a different progenitor galaxy as indicated in the legend. The stellar debris from each of these progenitors is indicated with a coloured outline that encircles 80 per cent of their mass, produced using an isodensity contour around a smoothed Gaussian kernel. Line width indicates the total debris mass. These provide a visualization of the radial velocity elongation. Whilst there is typically one dominant radially anisotropic merger, some realizations have significant contributions from other mergers.

to halo mass relation of Moster, Naab & White (2013), this would suggest a progenitor stellar mass of $> 4 \times 10^6\,\text{M}_\odot$. There have been many more estimates of the GSE progenitor stellar mass, including $5-6 \times 10^8\,\text{M}_\odot$ (Helmi et al. 2018), $2-4 \times 10^8\,\text{M}_\odot$ (Kruijssen et al. 2020), $3 \times 10^8\,\text{M}_\odot$-$10^9\,\text{M}_\odot$ (Mackereth et al. 2019), $5-6 \times 10^8\,\text{M}_\odot$ (Fernández-Alvar et al. 2018; Vincenzo et al. 2019), $2-5 \times 10^8\,\text{M}_\odot$ (Mackereth & Bovy 2020), $7 \times 10^8\,\text{M}_\odot-7 \times 10^9\,\text{M}_\odot$ (Feuillet et al. 2020), $4-7 \times 10^8\,\text{M}_\odot$ (Naidu et al. 2020), $6-8 \times 10^8\,\text{M}_\odot$ (Han et al. 2022a), $1-2 \times 10^8\,\text{M}_\odot$ (Lane, Bovy & Mackereth 2023), etc. Das, Hawkins & Jofré (2020) identify a sample of purely accreted stars which they associate with the GSE progenitor. They argue that the kinematic properties of this sample support a larger progenitor mass of $\sim 3.4 \times 10^{11}\,\text{M}_\odot$, from which they infer a stellar mass of $\sim 3 \times 10^9\,\text{M}_\odot$. To summarize, there is general agreement that the progenitor stellar mass of the GSE was between that of the Small and Large Magellanic Clouds, but with

a large scatter between different estimates of roughly an order of magnitude.

In comparison to these literature estimates, the stellar masses of the GSE-type mergers in AURIGA are rather high (see the stellar masses in Table 1). Grand et al. (2017) compare the stellar to halo mass relations for the central AURIGA galaxies against the semi-empirical relations of Moster et al. (2013). They find that most of the MW-mass galaxies are within the 1σ scatter of the semi-empirical model at $z = 0$, but lie predominantly above the scatter at $z = 3$. This was ruled to be a consequence of the sub-grid physics model employed in AURIGA, which lacks sufficient feedback to regulate star formation at earlier times. The discrepancy is even greater for lower mass galaxies, with dwarfs in the range $10^8 < M_\star/\text{M}_\odot < 10^9$ lying above the semi-empirical model by almost 3σ at $z = 0$.

As a result, we expect that all simulated galaxies presented here have a greater stellar mass than would be expected from the relations
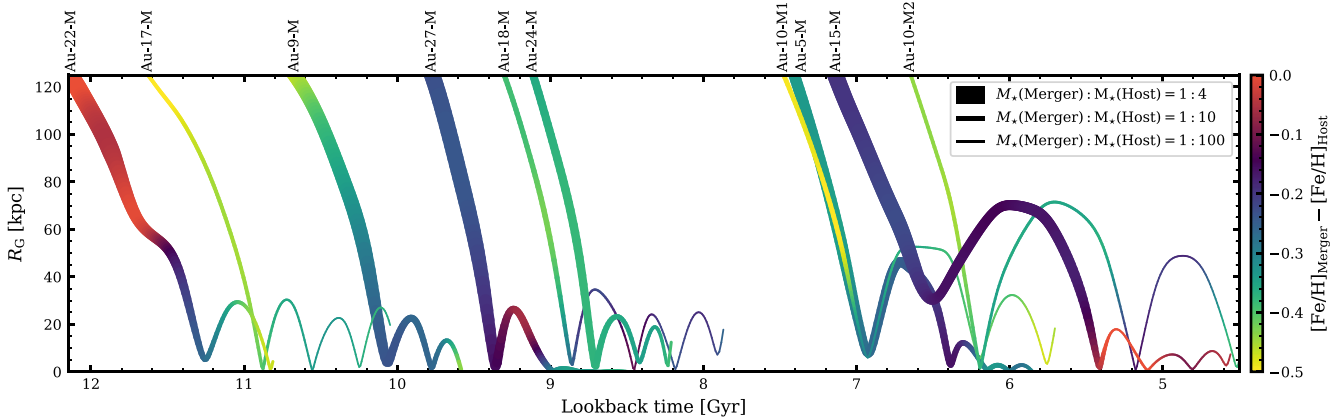
Figure 2. The infalls of the GSE-type mergers are listed in Table 1, as marked on the top-axis. The stellar mass ratio of the merger is represented by the line width, and the difference in mean [Fe/H] between the merging galaxy and central AURIGA galaxy is represented by the line colour (as calculated for stars within the stellar half-mass radius). The paths have been fitted with a cubic spline to improve the time resolution, and to better resolve the orbits. The infalls are highly radial, and in most cases the first or second apocentres are within 30 kpc.

in Moster et al. (2013), and this should be considered carefully before making any direct comparisons to the GSE. However, we note that the mass-to-light fractions remain very high in these galaxies despite their raised stellar masses, and so their dissolution within the host MW-like galaxies is likely to be unaffected. Furthermore, there is already a high scatter in the literature predictions for the GSE progenitor mass, and a scatter in the stellar to halo mass relations from Moster et al. (2013) (especially in this mass regime), and so these simulations remain useful for the investigative purposes here. Therefore, we do not expect this to change the qualitative nature of our results, but it will systematically shift properties like the stellar mass and chemical enrichment.

3.1.3 Merger infalls

Many of the literature works cited in Section 3.1.2 also include an estimate for the original accretion time of the GSE, ranging from $2 > z > 1$, which translates to look back times of approximately $10.5 > t_{lookback}/\text{Gyr} > 7.9$. This wide range of accretion times may in part reflect the prolonged interaction period between the GSE and the MW (as reasoned in Naidu et al. 2021).

We illustrate the infall of each GSE-type merger in Fig. 2. The line thickness indicates the stellar mass ratio between the GSE-type merger and the central AURIGA galaxy. The line colour indicates the metallicity difference within twice the stellar half-mass radius of each object.

All mergers infall on highly radial trajectories, with orbital circularities (see equation 2) of approximately 0. Each merger, except for Au-15-M, falls to within 10 kpc on the first pericentre passage. These steep infalls are responsible for the high radial velocity anisotropy of the debris seen in Fig. 1, and are a characteristic property of GSE-type mergers in general.[2] The first apocentres range from between 25 and 70 kpc, whereas breaks in the MW halo and anisotropy profiles suggest that the GSE apocentre occurred at between 25 and 30 kpc (e.g. Deason et al. 2018; Lancaster et al. 2019). The merger remnants are rapidly disrupted by the host potential, with the stellar mass ratios dropping below 1:100 within around 2 Gyr. Most of the merger

---

[2]Radial GSE-type mergers occur across much of the AURIGA simulation suite, but in other cases, there are other major mergers which either overwhelm or directly disturb the radially anisotropic debris footprints.

progenitors are less iron enriched and have between 4 and 10 times less stellar mass than the corresponding central AURIGA galaxies.

The star formation in each merging galaxy is rapidly quenched upon infall, whereas the central AURIGA galaxies experience an excitation in their star formation rates. Even so, in some cases, the metallicity of the merger appears to grow at a faster rate than the central AURIGA galaxy. In fact, this is an illusion caused by the tidal dissolution of the merger: The metal-poor outskirts of the merging galaxy are preferentially stripped, leaving the metal-rich core intact.

3.2 GSE-type merger properties

We now investigate the diversity in the main GSE-type mergers, in terms of their pre-infall progenitor galaxies and their resulting debris.

3.2.1 L_z distributions

Using controlled idealized simulations of GSE-type merger events, Amarante et al. (2022) show that the angular momentum of merger debris is related to the strength of the stellar feedback. Stronger feedback inhibits the recovery of the central density after pericentre passages, hastening the dissolution of the merger and therefore affecting the distribution of its debris in L_z. This leads to debris with a wider and more asymmetric spread in L_z, and offset from $L_z = 0$. Whether the merger is pressure-supported or rotationally supported may also influence the final distribution (e.g. Koppelman et al. 2020).

Here, we investigate whether these merger properties have a meaningful influence on the final distribution of their angular momenta. We estimate the disc fraction in each GSE-type merger at the snapshot before z_{infall} following the methodology described in Section 2.2, at which time we assume there is minimal tidal disruption to the inner galaxy. The fraction of disc to total stellar mass (D/T) is listed for each GSE-type merger in Table 1. We calculate the central density of the GSE-type merger at this same time, evaluated for all matter within the stellar half-mass radius. There is a wide range in both of these properties.

We show the pre-infall central density (upper panel) and D/T values (middle panel) versus the angular momentum distribution (L_z) of the merger debris at $z = 0$ in Fig. 3. The thin and thick error bars represent the $\pm 1\sigma$ and $\pm 2\sigma$ percentiles of the L_z
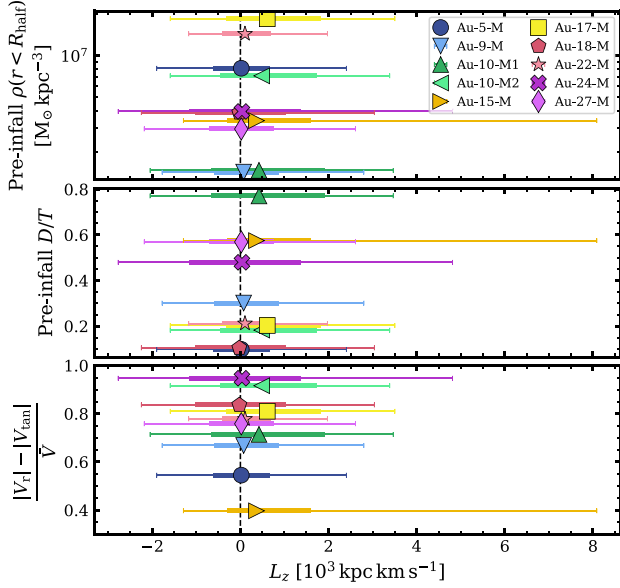
Figure 3. Upper panel: the pre-infall total mass density of each GSE-type merger within the stellar half-mass radius, versus the median z-component of the angular momentum of the debris at $z = 0$. Middle panel: the pre-infall disc to total stellar mass ratio of each GSE-type merger, versus the median z-component of the angular momentum of the debris at $z = 0$. Lower panel: the normalized radial orbital velocity minus the tangential orbital velocity, recorded at the moment when the merger progenitor crosses the R_{200} radius of the central AURIGA galaxy. A value of $+1$ corresponds to radial orbits, and -1 to circular orbits. The error bars correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ percentiles of the distribution, whereas coloured markers correspond to the medians. There are no overwhelming relationships between pre-infall density or disc fraction and the angular momentum distribution in the debris.

distribution, thereby indicating both the spread and asymmetry. The coloured markers represent the median of the distribution. There is no clear relationship between the pre-infall central density and the L_z distribution, nor between the pre-infall disc fraction and the L_z distribution. Normalizing the results by the mass or density of the central AURIGA galaxy does not affect this outcome.

The simulations in Amarante et al. (2022) are idealized and non-cosmological, whereas these AURIGA realizations represent a wide diversity in both merger and host properties. Therefore, trends relating L_z to the central density and/or disc fraction of the progenitor galaxies may be overwhelmed by the inherent stochasticity in their other properties. This shows that, at least in a varied cosmological setting, the L_z distribution of merger debris is not highly dependent on the central density or disc fraction.

The L_z distribution of merger debris can also be affected by the properties of the merger infall. Idealized simulations in Naidu et al. (2021) highlight a correlation with the orbital circularity of the merger: increasingly circular orbits shift the debris towards more retrograde (positive) L_z. The mergers in our sample are on highly radial infall trajectories, although Au-15-M is an outlier with by far the least radial infall (see Fig. 2). We evaluate this in terms of the velocity of merger progenitor at infall, and show the result in the lower panel of Fig. 3. The debris from Au-15-M is notably shifted into the retrograde side of L_z, as would be expected from Naidu et al. (2021), but the rest of our sample exhibits no clear relationship.

The stellar velocities of the Au-15-M debris are shown in Fig. 1, which reveals that it is the least radially extended of all the GSE-type mergers shown here. This suggests that AURIGA requires mergers

with lower orbital circularity than in Naidu et al. (2021) in order to produce debris with radial anisotropy comparable to the GSE.

3.2.2 Metallicity gradients

The hierarchical formation of galaxies generally results in a radial metallicity gradient, with the most evolved stars forming in the Galactic Centre and less evolved stars accreting at higher radii. As merging galaxies are tidally stripped during infall, the most tightly bound stars are shielded until the merger remnant has decayed to lower orbital energies. Therefore, a pre-infall metallicity gradient should give rise to a corresponding gradient within the merger debris.

We show the [Fe/H] (upper panels) and [Mg/Fe] (lower panels) metallicity gradients for each GSE-type merger in Fig. 4. The first column shows the radial gradients in the GSE-type merger progenitor at the snapshot before z_{infall}. The second column shows the resulting radial gradient of the merger debris at $z = 0$. The third column shows the gradient of the merger debris once more, but in terms of the angular momentum L_z. In each panel, a single black error bar is used to indicate the typical standard deviation of each stellar metallicity distribution.

There is a clear metallicity gradient within all the GSE-type mergers progenitors. The slopes of these gradients are steepest within ~ 5 kpc (which is typically comparable to the stellar half-mass radius), spanning the range $-0.2 < \Delta[\mathrm{Fe/H}]/\Delta R_{\mathrm{Merger}}(\mathrm{kpc}) < -0.06$ and $0.002 < \Delta[\mathrm{Mg/Fe}]/\Delta R_{\mathrm{Merger}}(\mathrm{kpc}) < 0.008$. The slopes of the [Fe/H] gradients are only slightly shallower outside of ~ 5 kpc, whereas the [Mg/Fe] gradients are flattened in most cases. There is no clear relationship between these slopes and the stellar mass of the merger.

The slope of the $z = 0$ debris gradients are shallower than before infall, and steepest within the inner ~ 20 kpc, spanning the range $-0.04 < \Delta[\mathrm{Fe/H}]/\Delta R_{\mathrm{Host}}(\mathrm{kpc}) < 0.0$ and $0 < \Delta[\mathrm{Mg/Fe}]/\Delta R_{\mathrm{Host}}(\mathrm{kpc}) < 0.002$. For some of the realizations, the [Mg/Fe] gradients are almost completely flat over all radii.

This change in slope is due to three effects: (i) the stars are spread over a much wider radial range; (ii) the merger is not necessarily dissolved within a single infall, leading to a smearing of the pre-infall gradient over multiple pericentre passages; (iii) the distribution of stellar orbits is more radially biased than they were in the pre-infall merger galaxy, meaning some stars that were originally liberated at high-radii are seen at the low-radii pericentres of their new orbits. These effects are not consistent between each GSE-type merger, and the metallicity gradients in some examples are flattened more than others. A similar investigation and discussion are performed for AURIGA galaxies in Monachesi et al. (2019).

As highlighted in Naidu et al. (2021), there is evidence for a weak metallicity gradient in the L_z angular momenta of GSE stars. In the third column, we show the metallicity gradients in the L_z angular momentum distributions for each GSE-type merger debris. These gradients are roughly symmetric about $L_z = 0$ for $|L_z| < 2$, but with some deviations outside of this regime. The (absolute) gradient slopes span the range $0.05 < \Delta[\mathrm{Fe/H}]/\Delta L_z(10^3 \, \mathrm{kpc\,km\,s^{-1}}) < 0.22$ and $0.0 < \Delta[\mathrm{Mg/Fe}]/\Delta L_z(10^3 \, \mathrm{kpc\,km\,s^{-1}}) < 0.01$.

3.2.3 The shape of GSE-type merger debris

The ellipsoidal shape of purely collisionless merger debris is known to relate to the infall of the progenitor galaxy (Moore et al. 2004; Cooper et al. 2010), with massive radial mergers creating aspherical and prolate shapes that are aligned with the direction of infall.
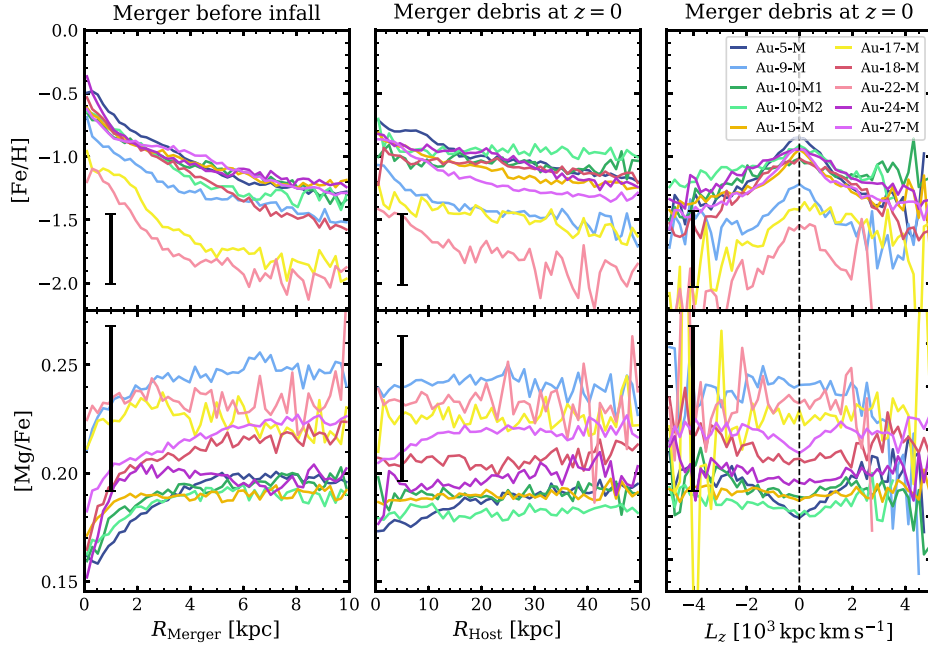
Figure 4. Left-hand panels: the pre-infall median [Fe/H] (upper panel) and [Mg/Fe] (lower panel) radial gradients shown out to a maximum radius of 10 kpc. A single black error bar is included to represent the typical standard deviation within 5 kpc. Middle panels: the radial metallicity gradient of the merger debris at $z = 0$, out to a maximum radius of 50 kpc. Here, the error bar represents the typical standard deviation within 20 kpc. Right-hand panels: the metallicity gradient of the merger debris but in terms of the L_z angular momentum. There is not necessarily a consistent mapping between the abundance gradients before infall and in the debris at $z = 0$.

Therefore, it is possible that the stellar debris from GSE-type mergers has a unique shape that can be distinguished from the rest of the *ex-situ* stellar halo.

To consider this possibility, we estimate the shape of stellar debris following the methods described in Dubinski & Carlberg (1991), Katz (1991), and Warren et al. (1992). This involves solving the moment of inertia tensor for stars within an initially spherical shell:

$$S_{ij} = \frac{\sum_k m_k \boldsymbol{r}_{k,i} \boldsymbol{r}_{k,j}}{\sum_k m_k}, \tag{3}$$

where the right-hand side refers to the elements of S in terms of a stellar particle k with mass m and galactocentric position vector r. The axial ratios of each volume can be derived from the eigenvalues of S, and the orientation of the volume can be derived from its eigenvectors. We iterate this procedure, each time updating the initial ellipsoidal shell with the shape from the previous iteration, until a convergence criterion is met. In this case, we define convergence as when the axial ratios between iterations vary by less than 0.1 per cent. We apply a bootstrap method to this algorithm, where we resample the stellar distribution 100 times using the RESAMPLE function in the SKLEARN PYTHON package (Pedregosa et al. 2011). We use the argument 'replace = True', which replaces a subset of the data array with a random sample of the data. We estimate the best value and $\pm 1\sigma$ uncertainties using percentiles.

The resulting axial ratios for each AURIGA simulation at $z = 0$ are shown in Fig. 5. We consider three distinct galactic components: the stellar debris from the main GSE-type merger (black), the remaining *ex-situ* stellar halo (blue) and the *in-situ* disc (red). Disc stars are kinematically selected following the methodology described in Section 2.2. The figure is then divided into three panels corresponding to the inner galaxy ($0 < a/\mathrm{kpc} < 5$), the solar neighbourhood ($5 < a/\mathrm{kpc} < 10$), and the outer galaxy ($10 < a/\mathrm{kpc} < 50$), where

a is the major axis length. The GSE-type merger debris tails off towards and beyond 50 kpc, so realistically the outer galaxy shape calculation will be dominated by stars in the $10-30$ kpc range. Grey-dashed lines mark constant T, where T is the triaxiality parameter defined in Franx, Illingworth & de Zeeuw (1991) as

$$T = \frac{1 - b^2/a^2}{1 - c^2/a^2}, \tag{4}$$

where $T > 2/3$ is prolate (cigar-shaped) and $T < 1/3$ is oblate (pancake-shaped). These lines converge at $c/a = b/a = 1$, where the shape is maximally spheroidal.

Inner galaxy (left-hand panel): Most of the disc stars are prolate due to the presence of rotating stellar bars (see Blázquez-Calero et al. 2020; Fragkoudi et al. 2020). Au-15 is a solitary outlier, and this is the single AURIGA galaxy from our sample that does not possess a stellar bar. The *ex-situ* stellar haloes are more spheroidal than the disc, and on the whole do not strongly favour prolate or oblate forms. The GSE-like debris has a similar shape to the *ex-situ* halo, though with large uncertainties due to the low number of stellar particles compared to the other components.

Solar neighbourhood (middle panel): The disc components have shifted closer to an oblate shape, marking the transition from bar dominated to disc dominated. The *ex-situ* haloes and GSE-type merger debris are, in most cases, highly oblate.

Outer galaxy (right-hand panel): The few remaining disc stars are converged at a highly oblate form. The *ex-situ* halo and GSE-type merger debris have similar overall shapes as they did at the solar radius, but are less strongly converged at highly oblate configurations.

Over all three radial regimes, there is no clear separation between *ex-situ* haloes and GSE-type merger debris across our simulation selection. However, there are some notable differences between the components in *individual* realizations.
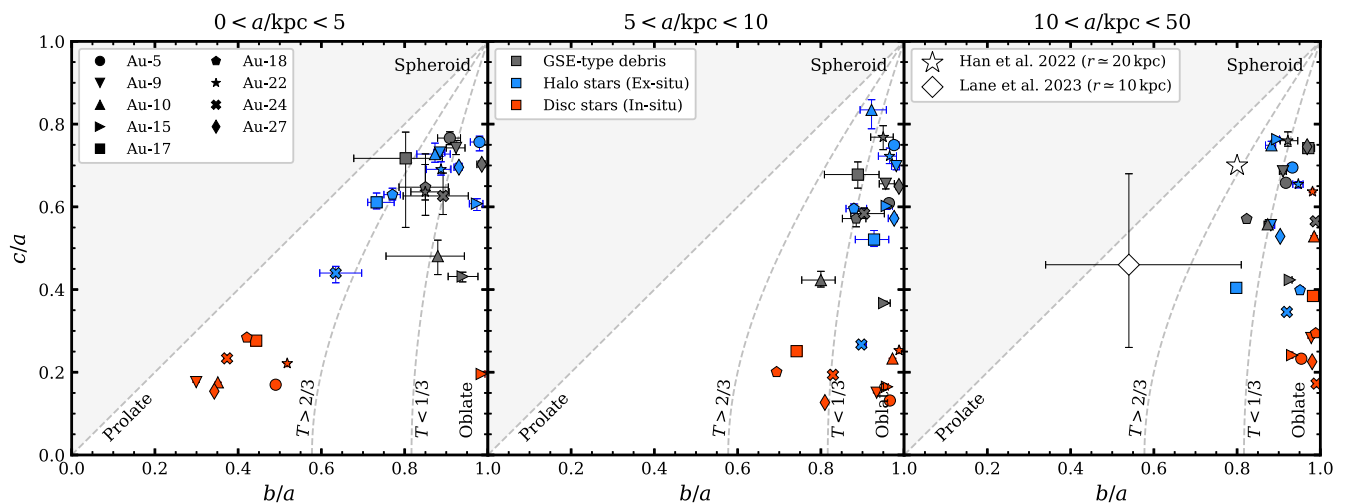
Figure 5. The shape of different stellar components in each AURIGA galaxy at $z = 0$, as measured by the axial ratios b/a and c/a. Black markers indicate the shape of stellar debris from GSE-type merger events. For the purposes of this plot, the double-mergers in Au-10 are considered together. Red markers indicate the shape of kinematically selected disc stars. Blue markers indicate the shape of kinematically selected halo stars from an *ex-situ* origin, excluding stars that originated from the GSE-type merger. The shape is calculated in three radial bins, corresponding to the three figure panels, which represent the inner galaxy ($0 < a/\text{kpc} < 5$), the solar neighbourhood ($5 < a/\text{kpc} < 10$), and the outer galaxy ($10 < a/\text{kpc} < 50$). Error bars denote the $\pm 1\sigma$ uncertainty derived from a bootstrap method. Grey-dashed lines depict the regions that are prolate ($T > 2/3$) and oblate ($T < 1/3$). The white star in the right-hand panel is based on observed GSE stars from Han et al. (2022a), and the white diamond on observed GSE stars from Lane et al. (2023). The shape of GSE-stars is oblate in most cases, except for within the inner 5 kpc.

The *ex-situ* halo in Au-24 has shape approaching that of the disc stars across all three radial regimes. This halo is investigated in detail in Gómez et al. (2017b), where it is shown that a substantial *ex-situ* disc forms from the accretion of mergers on tangentially biased orbits that are preferentially aligned with the disc plane. The GSE-type merger also infalls along the disc plane, but does so on an extremely radial trajectory. The *ex-situ* accretions are more isotropically distributed in the other realizations, leading to more spheroidal *ex-situ* haloes.

The GSE-like debris in Au-15 is more oblate and less spheroidal than in other realizations. This is because the infall trajectory of Au-15-M is the most tangential of all GSE-type mergers (see Fig. 2), leading to the debris adopting a pancake shape.

The GSE-like debris in Au-10 is less spheroidal than the *ex-situ* halo ($\Delta(c/a) > 0.2$), and also less spheroidal than the GSE-type mergers from most other simulations. This is unexpected because the majority of the GSE-like debris in Au-10 is comprised almost equally of stars from two separate merger events, and the superposition of two shapes should favour spheroidal symmetry. However, the two GSE-type mergers infall along nearly polar opposite directions, and their combined angular momenta are complementary.

We show the time evolution of the GSE-type debris shape for the 'solar neighbourhood' regime in Fig. 6. This demonstrates that the debris from every GSE-type merger event is initially elongated and prolate, exactly as would be expected given their highly radial infall trajectories. Time intervals of 2 Gyr are marked with black rings, which shows that the shape progresses towards rounder and increasingly oblate shapes. Whilst not shown here, the evolution of the DM debris shape follows very similar paths. There are several mechanisms that could be contributing to this shape change:

(i) Torques from the galactic disc/bar, as well as the underlying DM halo, will encourage the debris to both align with and adopt the shape of the disc/bar and DM halo. The growth of the stellar disc/bar can itself impact the shape and alignment of surrounding debris (e.g. Berentzen & Shlosman 2006; DeBuhr, Ma & White 2012).
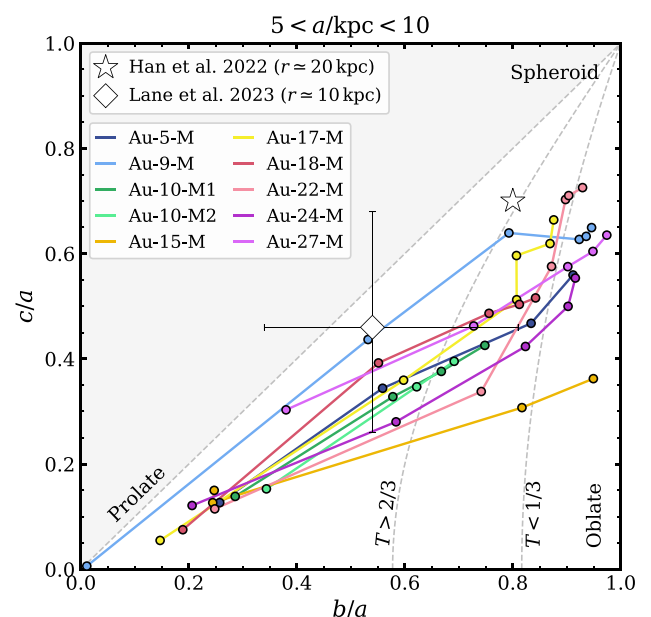

Figure 6. The shape evolution of the GSE-type stellar debris in the range $5 < a/\text{kpc} < 10$, as measured by the axial ratios b/a and c/a. These axial ratios are averaged over bins in look-back time with a 2 Gyr width which progress from the left- to right-hand panels, where each interval is marked by a black circle. The distribution is initially highly prolate in all cases.

(ii) The transition between centrally DM dominated and baryon dominated can turn prolate shapes into rounder and more oblate shapes (Tomassetti et al. 2016).

(iii) A dynamic gravitational potential can perturb the orbits of particles, which transforms prolate shapes into rounder and more oblate shapes. Such shape transformations are widely reported in the DM haloes of galaxies (Tissera & Dominguez-Tenreiro 1998; Kazantzidis, Abadi & Navarro 2010; Zhu et al. 2017), including in

the AURIGA simulation suite (Prada et al. 2019), and the stellar halo will be similarly affected. This process is most efficient in regions of high gas density, although the halo shapes in the 'inner galaxy' are not notably more spheroidal than elsewhere.

The GSE-type merger debris appears to be influenced by these background galactic processes. In all cases, the stellar disc undergoes rapid growth during the time of the GSE-type merger accretion. The early discs in the central AURIGA galaxies are overmassive because the AURIGA galaxies lie above the typical stellar mass halo mass relation at early times (as highlighted in Section 3.1.2). As such, the torques from the disc component may have been stronger than in the MW.

In the 'outer galaxy' panel, we include the axial ratios as derived for GSE stars in Han et al. (2022a) as a white star. These stars are selected with chemical cuts and a cut on orbital eccentricity of >0.7, and the shape is evaluated at a 'flattened' radius of 20 kpc. They find a spheroidal shape which tends towards a prolate configuration. A similar shape is also determined in Iorio & Belokurov (2019).

We also include the axial ratios as derived for a high-purity sample of GSE stars in Lane et al. (2023) as a white diamond. These are selected with chemical cuts and a selection in action space, and the sample is centred around radii of ~ 10 kpc. They find a more elongated shape with axial ratios of 1:0.55:0.45, but with greater uncertainties due to the smaller size of their sample.

It is curious that none of the GSE-type debris in our AURIGA selection is consistent with any of these measurements, though we emphasize that the absolute differences in the axial ratios are small when compared to Han et al. (2022a). To investigate this further, we reproduce the shape fits on the GSE-type debris using the same selection criteria as in Han et al. (2022a). The resulting shapes at $0 < a/\text{kpc} < 5$ and $5 < a/\text{kpc} < 10$ are more spheroidal, with raised c/a ratios (by $\Delta(c/a) \simeq 0.2$ in the most extreme case) but similar b/a ratios. The difference in the c/a ratios can be attributed to the eccentricity cut, which favourably removes merger stars that have become captured by the disc. However, the shape of the debris at $10 < a/\text{kpc} < 50$ is insensitive to the selection criteria. Overall, the resulting shape fits remain inconsistent.

3.2.4 The alignment of GSE-type debris with the in-situ disc

In addition to estimating the shape of the observed GSE debris, Iorio & Belokurov (2019), Han et al. (2022a), and Lane et al. (2023) also estimate the angular tilt with respect to the Galactic disc plane. They find angles of $20°$, $25°$ and $16°$, respectively.

One possibility is that this tilt reflects the infall trajectory of the GSE, with the major axis pointing in the direction that the merger originated. In support of this scenario, Chandra et al. (2022) detect overdensities in the outer 60–90 kpc of the MW halo, which they link to apocentric shells created during a retrograde and high-inclination GSE merger event. The Virgo Overdensity (Vivas et al. 2001) and Hercules-Aquila Cloud (Belokurov et al. 2007a) may also be formed by stars shed during the GSE infall. Conversely, *N*-body simulations in Naidu et al. (2021) suggest that the infall trajectory is not *necessarily* preserved in the orientation of the merger debris, finding that an infall inclination of $15°$ above the disc yields a debris distribution that is inclined at $35°$.

Gómez et al. (2017b) investigated the discs in the AURIGA simulations, and showed that the trajectory of massive mergers which contribute to co-rotating stars in the disc plane become aligned with the disc over time-scales of ~ 2 Gyr. This is due both to the merger aligning with the disc, and to the disc aligning with the merger,
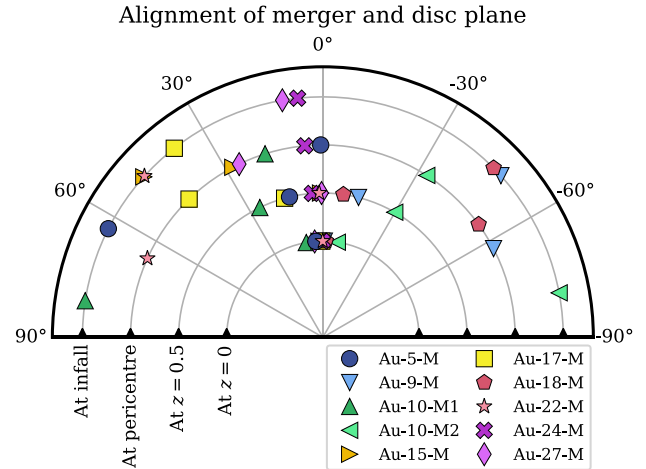

Figure 7. The alignment angle between the angular momentum vector of the *in-situ* disc and the GSE-type merger, defined such that an angle of $0°$ corresponds to an exact alignment. We show this alignment at four different times, as indicated by the axis labels. At infall, the alignment of the GSE-type merger is defined by its orbital velocity vector. Otherwise, the alignment is based on the minor-axis of the merger debris, derived using a shape fit to all debris over the radial range $5 < a/\text{kpc} < 50$. Regardless of the alignment at earlier times, all GSE-type debris has become aligned with the disc by $z = 0$.

caused by the transfer of angular momentum and dynamical friction (i.e. Huang & Carlberg 1997; Read et al. 2008; Earp et al. 2019).

We observe a similar effect for the shapes of GSE-type mergers here, continuing long after the merger is disrupted. In Fig. 7, we show the angular alignment between the *in-situ* disc and the GSE-type merger, defined such that an angle of $0°$ corresponds to alignment and ± 90 to perpendicular misalignment. By $z = 0$, all GSE-type stellar debris is aligned with the disc to within $\pm 15°$. The evolution towards alignment progresses from the inside-out, with debris at low radii aligning most quickly. At higher radii, beyond the influence of the disc, the debris preferentially aligns with the underlying DM halo instead. Over time, this underlying DM halo is itself brought into alignment with the disc.

This evolution towards closer alignment could be due to the response of the merger debris to the growth of the stellar disc, an effect which is known to be the most efficient for stars on radial orbits such as these (Binney & May 1986, and see also Dubinski & Kuijken 1995). Additionally, the gas donated by the GSE-type mergers may encourage a tilting of the disc into closer alignment with the merger (as in Debattista et al. 2015; Earp et al. 2019). Similar effects are also seen in the ARTEMIS cosmological simulations (Dillamore et al. 2022, and see also Dodge et al. 2023).

Of the GSE-type mergers presented here, Au-10-M1 and Au-10-M2 infall at an almost perpendicular angle to the disc. Nevertheless, by $z = 0$ the disc and the GSE-type debris are closely aligned. These results are not consistent with the larger tilt angles found for the GSE in Iorio & Belokurov (2019) and Han et al. (2022a), but may still indicate that the modern GSE-disc alignment is not necessarily a reliable reflection of the alignment at infall, or even of the original infall trajectory with respect to cosmic filaments. Whilst Naidu et al. (2021) also find a large tilting angle for their best-fitting model, their simulations were pure *N*-body and did not model the growth of the stellar disc. If baryonic physics were included, then their merger debris may have become more closely aligned with the disc.

If torques between the disc and merger debris were operating at this same level in the MW, it may suggest that the GSE accreted later than expected and has not yet become fully aligned with the disc. This is, however, inconsistent with the ancient accretion times inferred by direct measurements of the GSE age (e.g. Belokurov et al. 2020; Bonaca et al. 2020). As already stated, the alignment time-scales in AURIGA may be artificially fast due to the high stellar mass at earlier times, which would help to mitigate this problem.

3.2.5 Consideration of the DM halo

Han et al. (2022b) show that both the shape and tilt of stellar debris can be preserved for many Gyrs when immersed in a DM distribution that shares the same shape and tilt. This may be a natural outcome if the merger contributes a meaningful fraction of the local DM. In Iorio & Belokurov (2019), it is estimated that the GSE could contribute as much as 50 per cent of the total DM budget within 30 kpc. Then again, other works instead favour a minor-merger scenario in which the GSE would contribute a far lower fraction (e.g. Lane et al. 2023).

The shape and alignments of the DM haloes in AURIGA have previously been investigated in Prada et al. (2019) and Gómez et al. (2017a). Prada et al. (2019) find typically oblate DM haloes that favour alignment with the stellar disc. In particular, Gómez et al. (2017a) show that this alignment is closest for the DM within 10 kpc, whereas DM at higher radii is increasingly tilted and time-variable in some realizations. This tilting is, in most cases, due to interactions from massive merger encounters in the last few Gyrs.

We also investigate the shape and alignment of DM that originated from the GSE-type mergers. Whilst initially prolate, the shape gradually grows increasingly oblate with time, similar to what is seen for the stellar debris. Interestingly, the alignment of this DM is rapidly coupled to the alignment of the DM in the host galaxy, which in turn is usually aligned with the stellar disc.

To summarize, both the stellar debris and DM halo have grown into alignment with the disc by $z = 0$, but this time-scale may be unrealistically fast due to the high disc mass at early times. The tilting reflects the coupling of the merger debris with the orientation of the host halo, and may not reflect the infall trajectory of the merger itself.

3.2.6 Distinguishing the debris from two progenitors

As shown in Section 3.1.1, the radially anisotropic debris features in AURIGA contain the debris from several merger events, but typically have one dominant originator. Au-10 experiences two mergers that contribute an almost equivalent mass fraction of radially anisotropic stellar debris. Whilst their debris forms similar ellipsoids in v_ϕ/v_r coordinates, they may have distinct distributions in their other properties. Here, we investigate whether it is possible to distinguish their debris at $z = 0$.

We compare a selection of properties for Au-10-M1 and Au-10-M2 in Fig. 8. These include the galactocentric radius of the debris, the [Fe/H] and [Mg/Fe] metallicity distributions, the triaxiality and alignment of the shape with respect to the *in-situ* stellar disc, and the formation times of member stars. We base these comparisons upon the stars present within each merger prior to infall, and over the final radial range $0 < R_G/\mathrm{kpc} < 50$. For each property, we calculate the percentiles of the distribution and display the results in the form of
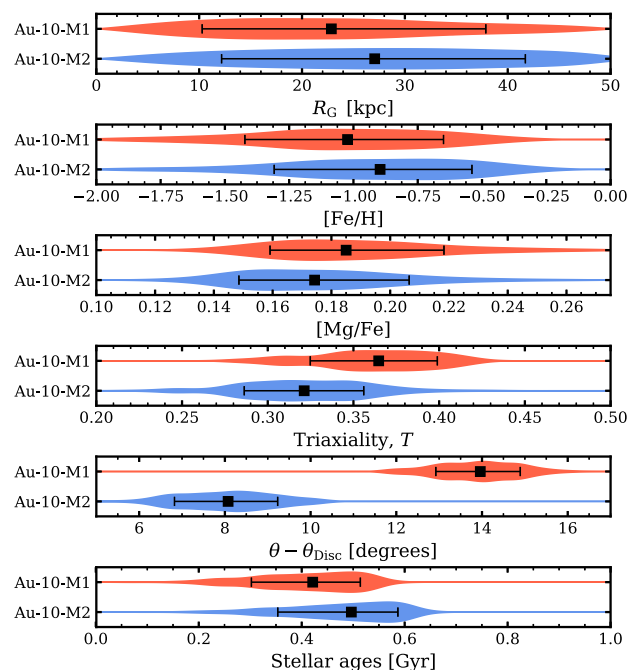

Figure 8. Violin plots for various properties of the GSE-type debris in Au-10. Black squares indicate the medians of each distribution, and error bars the $\pm 1\sigma$ standard deviation. The coloured bands represent a normalized Gaussian kernel density estimate fit to the simulation data. There are substantial overlaps in the distributions of most properties, which owe to the similar properties of each progenitor galaxy.

a violin plot. The median of each property is included as a black square, with error bars representing the $\pm 1\sigma$ percentiles.

The medians of the radial, chemical, and shape distributions all overlap within their $\pm 1\sigma$ percentile limits. The alignment of the debris shape appears to be distinct due to the narrow percentile distributions, but the absolute difference of $\sim 6°$ would be exceedingly difficult to detect in practice. The stellar ages are the most effective in differentiating each merger. However, it would remain difficult to discern this difference when considering the superposition of both distributions, especially once observational uncertainties of $\mathcal{O}(\mathrm{Gyr})$ are taken into account (see Miglio et al. 2017; Gallart et al. 2019b; Miglio et al. 2021a). These results are comparable to Rey et al. (2023), where it is shown that the debris from several distinct merger events can be difficult to parse.

If the merger progenitors are massive and gas rich, then the gas shocks resulting from their accretion could trigger bursts of *in-situ* star formation (see Mihos & Hernquist 1994; Tissera et al. 2002; Gargiulo et al. 2019; Gallart et al. 2019b; Ruiz-Lara et al. 2020; Massana et al. 2022; Orkney et al. 2022). Indeed, there are excitations in the *in-situ* star formation rate of up to half an order of magnitude during the first pericentre passage of all GSE-type mergers presented in this work. In some cases, there are further excitations resulting from subsequent pericentre passages, but the duration and magnitude of these bursts diminishes rapidly. A multiple-peaked *in-situ* star formation history could be a signature of multiple merger events, but would not necessarily mean that those merger events were GSE like (e.g. Orkney et al. 2022). Furthermore, it would remain difficult to distinguish these peaks if the mergers were accreted within a short time of each other, as is the case here.
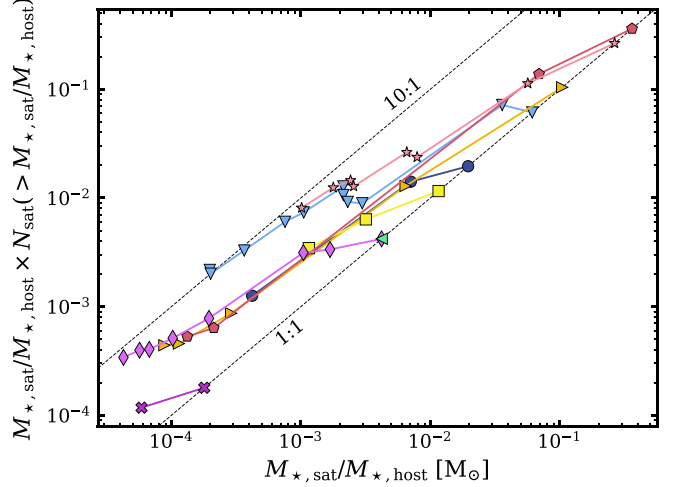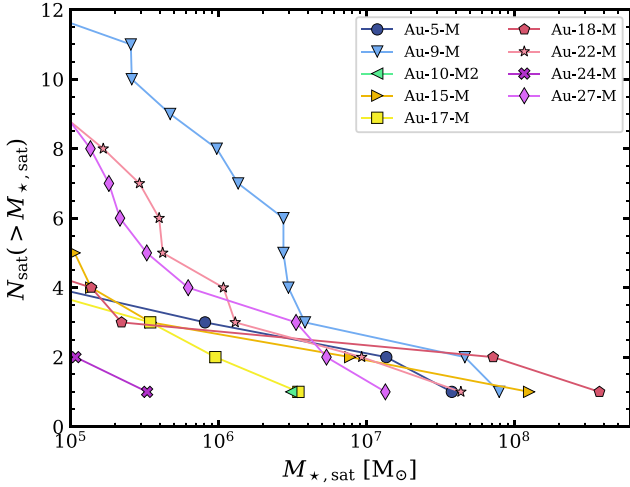
Figure 9. Left-hand panel: The satellite mass distribution function for each of the GSE-type mergers (not including the GSE-type merger itself), down to a lower stellar mass limit of 10^5 M$_\odot$. The satellite mass is given by the bound stellar mass identified by SUBFIND. Right-hand panel: The same distribution function, but normalized with respect to the pre-infall stellar mass of the central galaxy in the GSE-type merger system. If the distribution functions were proportional to the stellar mass of the host, then these normalized distributions would overlap. This shows that there is a wide diversity in the satellite mass distribution functions, and at least some of that diversity is independent of the host mass.

3.3 Satellites of GSE-type mergers

3.3.1 Satellite mass function

Investigations of the LMC have suggested that it brought several smaller galaxy companions into the MW (e.g. Jethwa, Erkal & Belokurov 2016; Sales et al. 2017). Assuming that the GSE progenitor had a stellar mass similar to the LMC as suggested by Das et al. (2020) and Evans et al. (2022), then it too may have been accompanied by a number of luminous satellites. If the debris or remnants of these satellites can be identified, then they may help to constrain the properties of the GSE itself. Alternatively, the debris may be mistaken for entirely independent accretion events.

Many of the GSE-type mergers presented here host a population of luminous satellites. We identify this population as satellites within the Friends-of-Friends (FoF) group of the pre-infall merger galaxy before it itself falls into the central AURIGA galaxy. Then, we exclude any satellites for which their orbital kinetic energy exceeds the gravitational potential energy binding them to the merger system, and perform a visual check that the satellites remain associated with the merger system until z_infall. Many of the remaining bound satellites are accreted only a few 100 Myr before z_infall, though there is a selection bias because some satellites that accrete earlier have already dissolved before this time.

We show the satellite stellar mass distribution function for each GSE-type merger in the left-hand panel of Fig. 9. There is a wide variation in the total number and masses of these satellites, with 0–8 satellites at $M_\star > 10^6$ M$_\odot$. This may simply reflect the wide variation in the mass of the host GSE-type mergers themselves (see Table 1). We account for this in the right-hand panel, where we normalize the distribution functions with respect to the coeval mass of the GSE-type merger. If the original distribution functions are entirely proportional to the stellar mass of the host, then these normalized distribution functions would overlap. Instead, a moderate degree of scatter remains, with variation in the normalized satellite mass of $\approx$4 orders of magnitude, and variation in the normalized number of $\approx$0.5 orders of magnitude. See Sales et al. (2013) for an investigation into the observed satellite stellar mass distribution functions, which
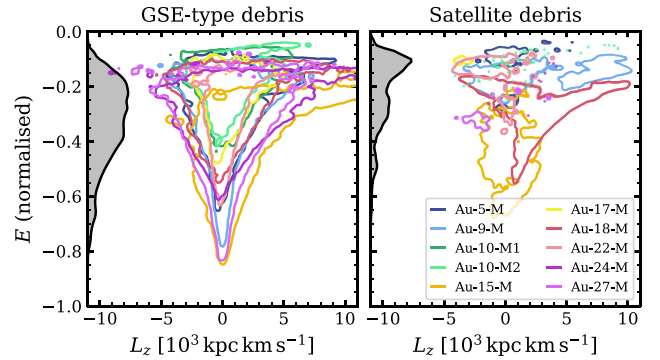

Figure 10. Left-hand panel: outer contours of the total specific orbital energy versus the specific angular momentum for stars that came from the main galaxy in each GSE-type merger system. Right-hand panel: the same, but for stars that came from satellite galaxies of the GSE-type merger system. A grey histogram is included on the y-axis which shows the summed PDF across all realizations. Satellite debris is favourably deposited at higher energies, and with a less symmetric distribution in L_z.

finds it to be relatively independent of the host mass for hosts of $M_\star < 10^{10}$ M$_\odot$.

3.3.2 Satellite debris

We have established that GSE-type mergers in AURIGA can host a variable number of luminous satellites, and now we show where those stars are deposited in the central AURIGA galaxies at $z = 0$.

We show the total specific orbital energy versus the specific angular momentum distributions for each GSE-type merger in the left-hand panel of Fig. 10. The energy has been normalized in the range $0 > E > -1$ to aid comparison, where $E = 0$ corresponds to the R_{200} radius, and $E = -1$ is the potential of the host MW-type galaxy at an arbitrarily chosen inner radius ($R_\mathrm{G} = 30$ pc). The energy distributions of stars that belonged to satellites of the GSE-type mergers are shown in the right-hand panel. We discard any satellite stars that coalesced with the main GSE-type merger prior to
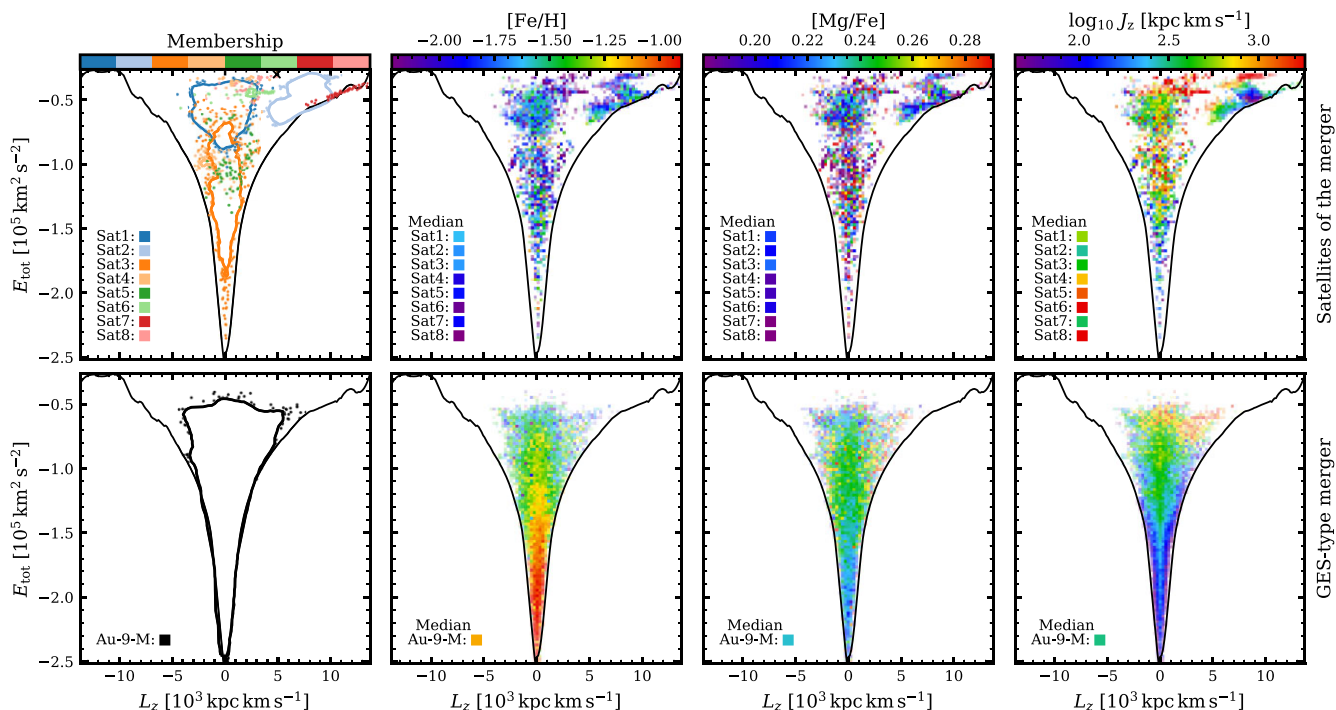
Figure 11. The total specific orbital energy versus the z-component of the specific angular momentum for the Au-9-M GSE-type merger. A black line marks the smoothed edge of the total stellar distribution in Au-9. Upper panels: stars that were exclusively associated with $M_\star > 10^6\,M_\odot$ satellite galaxies of the GSE-type merger prior to infall. Lower panels: Stars that were associated with the main GSE-type merger at the time of its peak stellar mass. The left-most column groups the stars by their pre-infall galaxy (contours and points). The presence of a surviving satellite remnant (Sat8) is indicated with a black cross. The following three columns include stacked, mass-weighted histograms coloured by the metallicity ([Fe/H]), the α-metallicity ([Mg/Fe]), and the z-action (J_z). The histogram opacity is based on the mass within each bin. The legends list the median property value for each distinct group. Whilst there is a large gap between the median properties of the satellites and the main merger, these differences are minimized in the energy-space regions where the populations overlap.

it achieving peak stellar mass, and any stars within surviving satellite remnants.

The debris from each GSE-type merger spans a wide range of energies that extends from near the virial radius to the solar radius and below. The debris from some GSE-type mergers reaches far lower energies, and this can be understood from their merger mass ratios in Fig. 2. The remnants of higher-mass ratio mergers are shielded against tidal disruption for a longer time, and also infall more rapidly due to a greater dynamical friction force.[3]

In most cases, the satellite debris is preferentially deposited at higher average energies. This is because the satellites are among the first stars to be stripped from the merging system. Once the satellites are independent of the GSE-type system, the dynamical friction they experience is greatly reduced and they start dissolving due to gravitational tides. This leads to the deposition of their debris at similar energies to where they were originally stripped. Despite this, the absolute density of the satellite debris is often overwhelmed by the GSE-type merger debris at the same energies.

Next, we perform a more detailed investigation into one example AURIGA simulation. We choose to focus on Au-9-M, because this object has the greatest number of satellite galaxies over a wide range of stellar masses. We consider only the satellites that had a stellar mass $> 10^6\,M_\odot$ prior to infall, yielding eight objects. We show merger debris in the plane of the total specific orbital energy against the z-component of specific angular momentum in

Fig. 11. The upper panels show stars that were within satellites of Au-9-M prior to its infall, excluding any which became bound to Au-9-M at the time of its peak stellar mass (but not any which became bound after this time). The lower panels show all stars that were within the GSE-type merger at the time of its peak stellar mass.

The median chemical abundances of Au-9-M and each of its satellites span approximately 1 dex in [Fe/H] and 0.05 dex in [Mg/Fe], in line with expectations from their differing stellar masses and evolutionary stages (see the second and third columns). However, due to the post-infall metallicity gradients discussed in Section 3.2.2, the high-energy tail of the Au-9-M debris is comprised of the least chemically evolved stars. Consequently, these metallicity differences are minimized in the regions where the debris of Au-9-M and its satellites overlap. This is a natural result, given that the metal-poor outskirts in these GSE-type mergers incorporate the debris of recently disrupted satellites (see the merger tree visualizations in Appendix A).

In the fourth panel, we colour the debris by the vertical action (J_z). This can be interpreted as the vertical excursion of particle orbits with respect to the plane of the galaxy (see Binney & Tremaine 2008, for a mathematical description). Actions are commonly used to assist in the identification of debris from various merger events (e.g. Yuan et al. 2020a; Limberg et al. 2021; Malhan et al. 2022). We calculate J_z using AGAMA (Vasiliev 2019) within a static axisymmetric potential fit to the mass of the central AURIGA galaxy at $z = 0$. The debris from both Au-9-M and its satellites span a wide range in J_z, with this variation being mostly independent of the boundaries between each debris group. An underlying pattern emerges, with

---

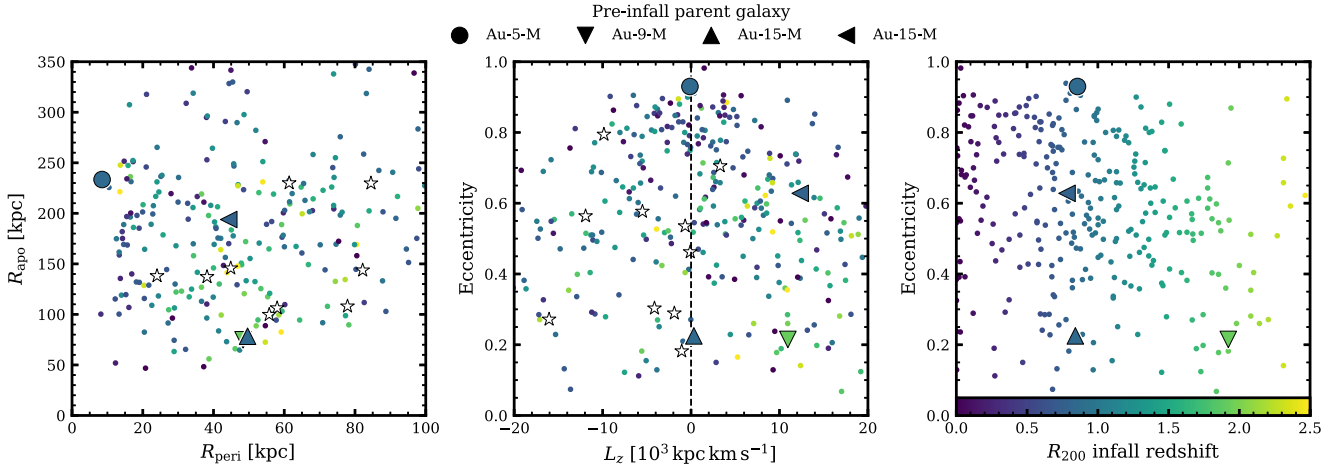[3]The dynamical friction time-scale is proportional to the merger mass ratio (Binney & Tremaine 1987).

Figure 12. Various orbital properties for surviving satellites in the dwarf-mass regime (defined here as $10^5 < M_\star/\mathrm{M}_\odot < 10^7$). Left-hand panel: The apocentre and pericentre for all satellites. Middle panel: The orbital eccentricity versus the z-component of the specific angular momentum. Right-hand panel: The orbital eccentricity versus the infall redshift (simulation data only). White stars represent observed MW dwarf spheroidal galaxies from Pace et al. (2022). Small circles represent surviving satellites across the full AURIGA suite, and are coloured by their infall time. The satellites which were associated with GSE-type mergers are shown with unique markers. These satellites are not exceptional in any obvious way, though this cannot be said with high confidence given that there are only four data points.

J_z at higher L_z adopting the lowest values ($J_z < 10^2 \, \mathrm{kpc \, km \, s^{-1}}$ for regions close to the perimeters of the L_z distribution), and J_z at higher energies adopting the highest values ($J_z > 10^3 \, \mathrm{kpc \, km \, s^{-1}}$ for energies $> -0.7 \times 10^5 \, \mathrm{km^2 \, s^{-2}}$). This same pattern appears across the rest of the AURIGA selection. The wide range of J_z values between each debris group, and even internal to larger debris groups, would make it difficult to draw connections between each debris. Some regions of the GSE-type debris may even appear to be unrelated to one another (as in Amarante et al. 2022).

3.3.3 Surviving satellite remnants

Only a small fraction of GSE-type satellites survive until $z = 0$, and only four of those are well resolved ($M_\star > 10^5 \, \mathrm{M}_\odot$). However, the survival of these four satellites raises the possibility that one or more of the MW dwarf spheroidal galaxies may have originated from a GSE merger event. In this section, we ask whether it would be possible to distinguish these galaxies from the rest of the satellite population.

We show a selection of orbital properties for surviving satellites in the stellar mass range $10^5 < M_\star/\mathrm{M}_\odot < 10^7$ at $z = 0$ in Fig. 12. Satellites that are unaffiliated with GSE-type mergers are represented by points, where the colour corresponds to the infall redshift. The five satellites that were associated with GSE-type merger events are represented by enlarged, unique markers. In order to provide some context, we also include the properties of MW dwarf spheroidal galaxies from Pace, Erkal & Li (2022) as white stars, using the same stellar mass cuts.

For the simulation data, we select all satellites within the R_{200} virial radius of the central AURIGA galaxy at $z = 0$. The infall redshift is defined as the last time the satellite crossed into this R_{200} virial radius. The instantaneous orbital properties are then determined by integration within AGAMA (Vasiliev 2019), assuming a static and axisymmetric host density profile that was fit to the total mass of the central AURIGA galaxy. For the observational data from Pace et al. (2022), we use the 'edr3' values for the proper motion in RA and Dec. These are then converted into galactocentric coordinates

using SKYCOORD from the ASTROPY package (Astropy Collaboration et al. 2022). We estimate the stellar mass from the V-band luminosity with a stellar mass-to-light ratio of $M/L_V = 2$, which is a reasonable assumption for the older metal-poor stars in galaxies of this mass scale (McLaughlin & van der Marel 2005).

There is no strong signature feature in the orbital properties of the surviving GSE-type satellites in Fig. 12, and neither are there tight correlations in the full satellite populations. There is only a weak relationship between orbital eccentricity and infall redshift, as shown in the right-hand panel, whereby an earlier infall leads to less eccentric orbits. This trend is impacted by survivorship bias; satellites on eccentric infalls are more likely to be disrupted by strong gravitational tides at the Galactic Centre. However, it is worth considering that the long-term survivability of these satellites is limited by the simulation resolution (see Grand et al. 2021).

3.4 Implications of GSE-type mergers on the search for ancient disrupted relic galaxies

Debris linked to the GSE merger dominates the mass fraction of the stellar halo around the solar neighbourhood (e.g. Lancaster et al. 2019; Myeong et al. 2022). The same is true for many of the GSE-type mergers in AURIGA. The overwhelming contribution from the GSE-type merger could be saturating the chemodynamical parameter space, thereby obscuring the signatures of other less massive mergers.

3.4.1 Debris in the solar neighbourhood and beyond

Here, we investigate the relative fraction of GSE-type debris in one example simulation. We select Au-5, because its debris features are especially clear and the radially anisotropic debris feature at $z = 0$ is heavily dominated by a single merger progenitor (see Section 3.1.1). Furthermore, this realization was shown to be the most comparable to the GSE in F19. None the less, we note that the trends reported in this example are general across our AURIGA selection, and we show this later in Section 3.4.4. We also include the full plots for
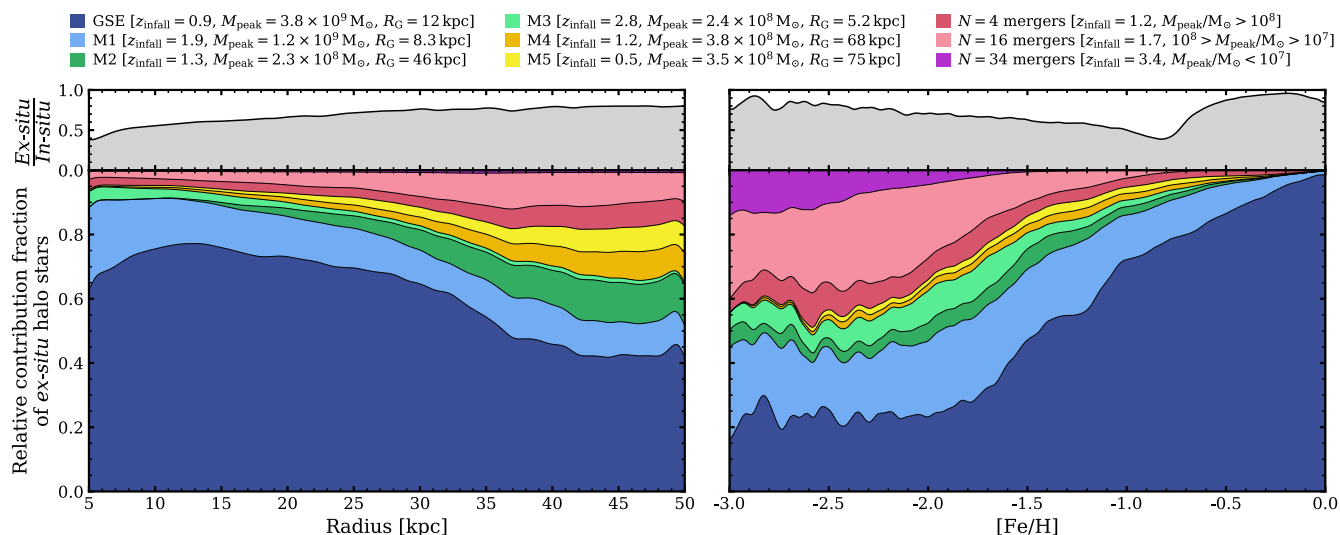
Figure 13. The radial distributions (left-hand panels) and metallicity distributions (right-hand panels) of *ex-situ* stars in the radial range $5 < R_G/\text{kpc} < 50$ in Au-5. '*Ex-situ*' refers to stars that were originally unbound to the central AURIGA galaxy. Upper panels: A histogram of the *ex-situ* to *in-situ* stellar halo mass fraction, where disc stars have first been removed following the kinematic decomposition described in Section 2.2. Lower panels: Stacked histograms representing the contribution of different merger events, as a fraction of all *ex-situ* stars. The GSE-type merger is labelled 'Au-5-M', and then the five next most dominant mergers in this radial range are labelled 'M1-5'. All other debris is binned according to the peak stellar mass of their progenitor galaxies. The histograms are constructed with a Gaussian KDE using Silverman's rule. In the case of individual mergers, the infall redshift, peak stellar mass, and the median orbital radius of the debris are included in the legend. In the case of mergers binned by their progenitor mass, the total number of distinct galaxies within this radial range are included in the legend, along with their median infall redshift. The relative fraction of GSE-type stars is greater than 50 per cent over most radii, but this fraction is reduced for increasingly low-metallicity stars. Similar plots for the other AURIGA realizations are included at the following link, or in the supplementary material.

each realization at the following link, and see also the supplementary material.

Inspired by the halo decomposition exercise performed in Naidu et al. (2020), we dissect the *ex-situ* halo of Au-5 in Fig. 13. The stacked coloured histograms in the lower panels represent the contributions from different merger events to the fraction of all *ex-situ* halo stars at $z = 0$. We have excluded stars that remain bound to substructure. The grey histogram in the upper panels represents the fraction of *ex-situ* to *in-situ* halo stars. In all cases, co-rotating disc stars have been kinematically selected as in Section 2.2 and then removed. This cut includes *ex-situ* stars that have adopted disc-like orbits.

The first coloured band, labelled 'Au-5-M', is the stellar debris originating from the GSE-type merger. The bands labelled 'M1-5' represent the next five mergers that contribute the largest fractions of the *ex-situ* stellar mass over the radial range $5 < R_G/\text{kpc} < 50$. The final three histograms represent the remaining *ex-situ* stars, binned with respect to the peak mass of their progenitor galaxies.

In the left-hand panels, the debris is plotted with respect to its galactocentric radius over the range $5 < R_G/\text{kpc} < 50$. The debris from the GSE-type merger dominates that of all other mergers, and contributes in excess of 50 per cent of all *ex-situ* stars over the radii considered here. Stars from mergers with a progenitor stellar mass $M_\star < 1 \times 10^8 \, M_\odot$ (the pink and purple bands) contribute a negligible fraction over all radii.

In the right-hand panels, the debris is plotted with respect to its [Fe/H] metallicity over the range $-3 > [\text{Fe/H}] > 0$. We choose these limits because they effectively bracket the simulation data, whilst ensuring there are still a statistically meaningful number of star particles across the entire range. The fraction of GSE-type stars is almost 100 per cent around solar metallicities ([Fe/H] = 0), but the relative abundance declines for lower metallicities ($\sim$25 per cent for

[Fe/H] < -2). Stars from mergers with a progenitor mass $M_\star < 1 \times 10^8 \, M_\odot$ now contribute nearly 40 per cent of *ex-situ* stars at the lowest metallicities. Furthermore, the *ex-situ* fraction in the stellar halo remains relatively high at these same low metallicities ($\sim$90 per cent).

This result can be explained by considering the histories of the mergers that contribute to the stellar halo at these radii. The GSE-type merger infalls at a relatively late epoch ($z = 0.90$), and so it was free to evolve towards higher metallicities unimpeded. In contrast, the mergers marked 'M1-5' infall at $z = 3.33$-1.60. Their mass and chemical evolution was quenched prematurely, with no or few stars attaining solar metallicities. This is a natural consequence of GSE-type features typically owing to the last major merger that contributed to the inner galaxy.

All galaxies forming in relative isolation will begin their evolution from a low metallicity, and so there will be a fraction of early low-metallicity stars within the debris of every merger. The more massive mergers, such as GSE-type mergers, rapidly self-enrich and form a larger quantity of stars with raised metallicity. At the lowest metallicities, however, the relative contribution fraction is divided more equitably across all accreting satellites.

3.4.2 Debris in the Galactic Centre

With the continued improvements in observational data and analytical techniques, there is a growing interest in investigating the centre of the MW (i.e. Ness et al. 2013; Howes et al. 2017; Lucey et al. 2019; Arentsen et al. 2020; Rix et al. 2022) – a region that is often avoided due to dust contamination and crowding effects. The Galactic Centre is likely to have been constructed in part by accretion events in the very early Universe, and is an ideal environment to search for ancient and low-metallicity merger debris (Schlaufman & Casey 2014; El-Badry et al. 2018; Rix et al. 2022). However, the short relaxation
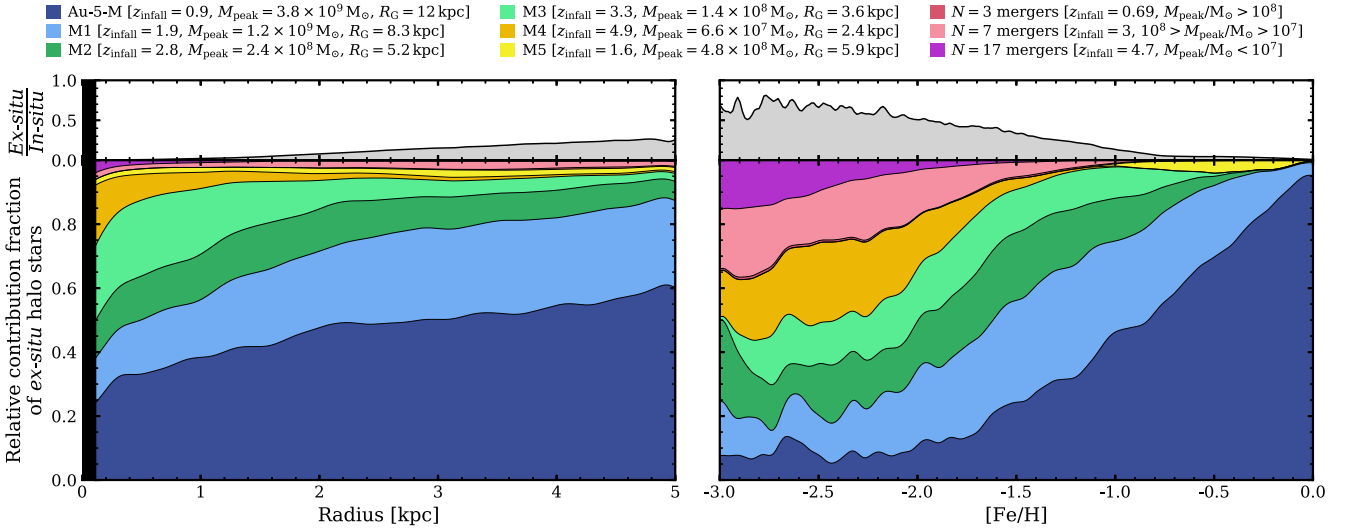
Figure 14. The same as Fig. 13, but for stars in an inner radial range of $0 > R_G/\mathrm{kpc} > 5$. Regions with insufficient particle counts are filled in black. The stellar disc and bulge dominate within this regime ($R_d = 3.58\,\mathrm{kpc}$, $R_{\mathrm{eff}} = 0.84\,\mathrm{kpc}$, Grand et al. 2017). The five most dominant mergers in this inner radial range are not necessarily the same as in Fig. 13, and typically sample older accretion events. As compared to Fig. 13, the relative contribution fraction of the GSE-type merger is reduced. Similar plots for the other AURIGA realizations are included at the following link, or in the supplementary material.

times and the phase-space bottleneck in the inner Galaxy rapidly erodes the rich kinematic landscape that is seen at the solar radius and beyond (e.g. Brauer et al. 2022). This makes it far more challenging to identify and distinguish different accreted populations.

We reproduce our stacked histogram figures for the inner 5 kpc of Au-5 in Fig. 14. The contribution fraction of the GSE-type merger now descends below 50 per cent within the inner 2.5 kpc, and the contributions of more ancient mergers are enhanced. When examining the metallicity distribution, a similar trend is seen in these inner regions as in the outer regions. However, the fractional contribution of low-mass mergers is proportionally greater.

The fraction of *ex-situ* stars falls rapidly towards the Galactic Centre due to the presence of the *in-situ* bulge and protohalo (Gargiulo et al. 2019; Fragkoudi et al. 2020; Grand et al. 2020), but most of these *in-situ* stars are formed with higher metallicities. The *ex-situ* fraction rises to around 60 per cent for metallicities of [Fe/H] < −2, and further to around 70 per cent for metallicities of [Fe/H] < −2.5. We anticipate these trends will continue to even lower metallicities, but we avoid commenting on this regime due to the reduced number of stellar particles and because the sub-grid physics models do not model the first generations of stars.

These results are encouraging in that even the most rudimentary parameter cuts can eliminate the majority of *in-situ* and GSE-type contaminants.

3.4.3 Debris in the galactic outskirts

The outskirts of the MW stellar halo are sparsely populated with few confirmed stars. Whilst it is currently prohibitive to detect more than the brightest populations [e.g. blue horizontal branch (Deason et al. 2012), M-giant stars (Bochanski et al. 2014) and RR Lyrae (Cohen et al. 2017)], these outskirts will be revealed in ever-increasing detail by future observational surveys (e.g. *James Webb Space Telescope*, *JWST*, Gardner et al. 2006; LSST, LSST Science Collaboration et al. 2009; WFIRST, Spergel et al. 2015).

We reproduce our stacked histogram once again, this time for the range $50 < R_G/\mathrm{kpc} < 200$ in Fig. 15. We do not extend this figure out to the full virial radius, because the number of star particles is too low for statistically meaningful analysis. Predictably, there are very few *in-situ* stars present.

The contribution fraction of the GSE-type debris is far less significant at these higher radii, and this is because there is a greater contribution from an assortment of other merger debris. The mergers that contribute to this radial regime tend to have been accreted more recently than those at lower radii, and there is a greater proportion of stars from low-mass mergers (pink and purple bands) which were rapidly disrupted upon infall. This is entirely expected and in accordance with previous works (e.g. Bullock & Johnston 2005; Fattahi et al. 2020).

The metal-rich core of the GSE-type merger is shielded against tidal disruption until it has fallen to within ∼ 50 kpc of the central AURIGA galaxy. Consequently, it contributes almost 0 per cent of stars approaching solar metallicities. Instead, the contribution fraction becomes dominated by metal-rich mergers that were disrupted at higher radii. This behaviour is not seen in Au-10 or Au-18, because those galaxies do not undergo such mergers.

3.4.4 Overall trends

We illustrate trends across the rest of our AURIGA selection in Fig. 16. The upper panels represent the *ex-situ* to *in-situ* stellar halo fractions, and the lower panels represent the fraction of *ex-situ* stars that originated from GSE-type mergers. The GSE-type debris in Au-10 is shown for each merger individually, though they could arguably be considered as a single debris population. The left column represent stars in the range $0 < R_G/\mathrm{kpc} < 5$ as in Fig. 14, the middle column represent stars in the range $5 < R_G/\mathrm{kpc} < 50$ as in Fig. 13, and the right-hand column represent stars in the range $50 < R_G/\mathrm{kpc} < 200$ as in Fig. 15. We perform cuts on the metallicity in order to illustrate the difference between a wide range of metallicities (−3 < [Fe/H] < 0) and metal-poor stars (−3 < [Fe/H] < −2.5).

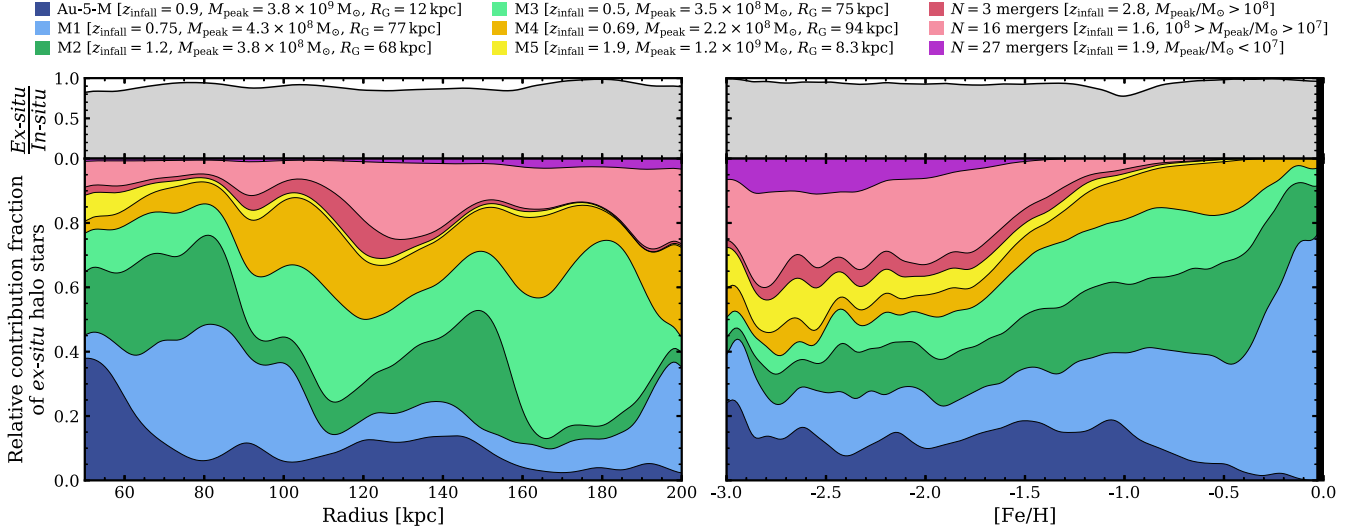**Figure 15.** The same as Fig. 13, but for stars in an outer radial range of $50 > R_G/\text{kpc} > 200$. Regions with insufficient particle counts are filled in black. This regime is dominated by *ex-situ* stars, originating from a wide variety of different accretion events. Again, the five most dominant mergers are not necessarily the same as in Fig. 13. The debris from more recent mergers is favoured. Similar plots for the other AURIGA realizations are included at the following link, or in the supplementary material.
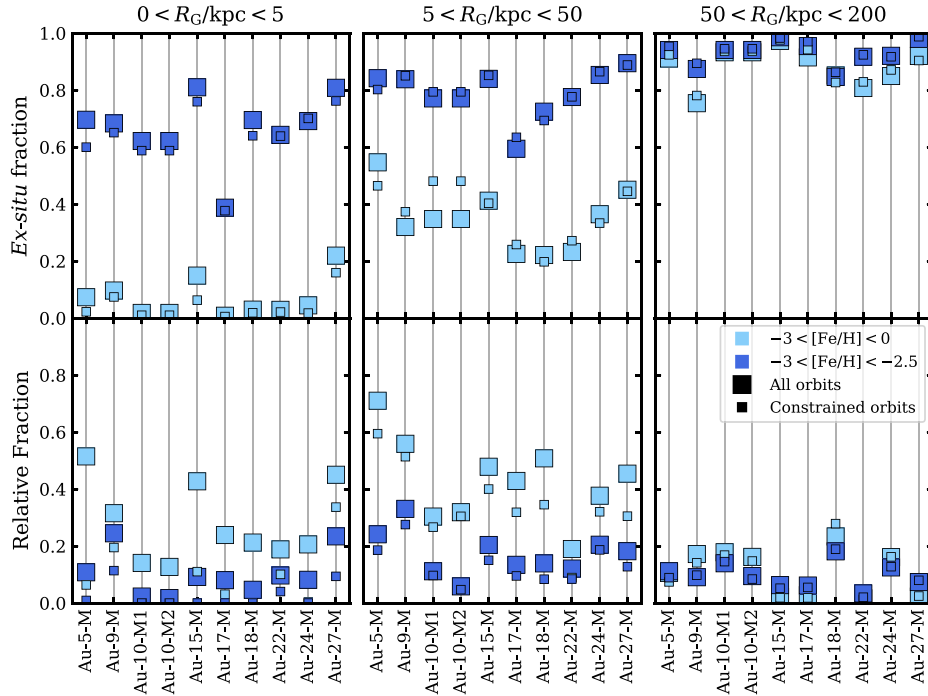
Figure 16. Upper panels: The fraction of *ex-situ* to *in-situ* halo stars (excluding the stellar disc). Lower panels: The fraction of *ex-situ* stars that originated from a GSE-type merger event. The panel titles indicate the radial limits considered. The marker colour represents different metallicity cuts, with all stars shown in light blue and metal-poor stars in dark blue. The marker size represents different orbital cuts, where all stars are shown with large markers and orbits constrained within the radial limits are shown with small markers. In these cases, all stellar pericentres are greater than the minimum radius and all stellar apocentres are less than the maximum radius.

Our basic radial cuts do not account for stars that are on orbits which take them outside of the defined radial limits. We estimate the pericentre and apocentre for each star particle by integrating its full orbit in an axisymmetric potential using AGAMA (Vasiliev 2019). We show all stars using large markers, and then stars that remain within the stated radial range over their entire orbits using small markers.

The figure shows that the *ex-situ* fraction is far lower in the inner radial range, and this is due to the presence of dense *in-situ* stellar bulges within the inner few kpc. However, the *ex-situ* fraction is increased when metal-rich stars are excluded.

In almost all cases, the relative fraction of GSE-type debris is reduced once more metal-rich stars are excluded. As mentioned in

Section 3.4.1, this is because the bulk of the stars in these GSE-type mergers are more enriched than the rest of the *ex-situ* population, which is a direct consequence of the GSE-type mergers typically being among the most recent and massive mergers that contributed to the inner galaxy.

The effect of the orbital cuts is subdominant compared to the metallicity cuts, and with predictable outcomes. The *ex-situ* fractions over the radii $5 < R_G/\mathrm{kpc} < 50$ are, on average, increased after the orbital cuts. This is because the cuts remove high eccentricity *in-situ* stars from the dense bulge region, which have pericentres that take them below the 5 kpc lower limit. In contrast, the *ex-situ* fraction in at $0 < R_G/\mathrm{kpc} < 5$ are slightly reduced after the orbital cuts. This is because there are many *ex-situ* stars on highly radial orbits, such as those from the GSE-type mergers, which have apocentres greater than 5 kpc. It is for this same reason that the relative fraction of GSE-type stars is, on average, reduced after the orbital cut.

There is little difference between different parameter cuts for radii $50 < R_G/\mathrm{kpc} < 200$, with uniformly low fractions of stellar debris from GSE-type mergers.

In summary, these results indicate that the contamination from GSE-type debris is reduced at lower metallicities (e.g. [Fe/H] < −2.5), towards the Galactic Centre (e.g. $R_G < 5\,\mathrm{kpc}$) and outskirts (e.g. $R_G > 50\,\mathrm{kpc}$), and excluding stars with higher orbital apocentres (e.g. $r_{\mathrm{apo}} > 5\,\mathrm{kpc}$). More specifically, both the fraction of GSE-type debris and the fraction of *in-situ* stars are reduced for increasingly low metallicity. It is possible that this trend continues even below [Fe/H] = −3, although there are few star particles below this limit and so we avoid making any firm interpretations. We also note that the fraction of *in-situ* stars can only ever be overestimated due to the manner in which they are identified, meaning the *ex-situ* fractions reported here should be considered as a lower bound.

In Appendix B, we test whether these relationships fail for stars that occupy similar E/L_z as the GSE-type debris. We find that it makes little difference.

4 DISCUSSION

4.1 Diversity and similarities of GSE-type mergers

The GSE-type mergers presented in Table 1 span a wide range of progenitor properties, many of which have been discussed in this work. In some cases, there is more than one single merger which contributes a large mass fraction to the radially anisotropic stellar debris near the solar radius (∼ 8 kpc).

Despite this inherent diversity, there is a great deal of degeneracy in the final debris footprints. As described in F19, the resulting GSE-type features contain a dominant component with velocity anisotropy of $\beta > 0.8$, and a contribution in excess of 50 per cent to the stellar halo. Furthermore, the overall form of the debris in energy–space and configuration–space converge upon similar qualitative properties.

There is a great stochasticity in both the number and mass of pre-infall luminous satellites associated with GSE-type mergers, with little dependence on the stellar mass of the host. Furthermore, the majority of these satellites are separated from their parent GSE-type merger soon after infall, and do not descend to the low specific orbital energies that is typical for GSE-type debris. Only a marginal fraction of these satellites survive as remnant objects at $z = 0$, though we note that their survival may be impacted by resolution effects (see Grand et al. 2021). These few survivors do not have exceptional orbital properties that might indicate that they were once related to the GSE-type merger event.

We show that the debris from the superposition of two separate mergers may be almost entirely degenerate with one another. This is a similar problem to that raised in Jean-Baptiste et al. (2017), where it is found that the debris from various merger events can overlap and fragment in ways that make it difficult to determine their origins. Whilst ages and chemical abundances can alleviate those challenges, they are less helpful in the case of two near-mass mergers that merge at a similar time. This could make it nearly impossible to confirm whether a GSE-like debris feature was sourced from one or two progenitors, especially since such a scenario would only change the progenitor mass estimate by a factor of two – which is already well within the range of current estimates for the GSE. Similar results were found in Rey et al. (2023), where it is shown that a ΛCDM cosmology naturally leads to several mergers contributing radially anisotropic halo debris, and that the properties of this debris are relatively insensitive to its assembly. There are also arguments for this 'multiple radial merger' scenario based on observational data of halo stars (Donlon Thomas et al. 2022; Donlon & Newberg 2023).

In essence, it is exceedingly difficult to accurately constrain many of the progenitor properties from the chemodynamics of merger debris at $z = 0$, even when the privileges of simulation data are readily available.

4.2 Revealing the centre of the MW

In the ΛCDM cosmology and hierarchical galaxy formation more generally, the proto-MW is assembled from a spectrum of high-mass ratio major mergers at early times (e.g. Renaud et al. 2021a, b). Much of the debris from these ancient mergers will remain locked in the low-energy potential well, concealed amongst the stellar bulge, bar, and disc (e.g. Bullock & Johnston 2005; Wetzel 2011; Rocha, Peter & Bullock 2012; van den Bosch et al. 2016; Starkenburg et al. 2017).

There has been growing evidence for this scenario in the MW. Analysis of the age and chemical distribution of globular cluster populations suggest they were either born *in situ* or donated by known satellite accretions (Kruijssen et al. 2019; Massari, Koppelman & Helmi 2019; Forbes 2020). Some 16 per cent, however, are associated with an unknown group at low binding energies. This led to the prediction of a high-mass ratio ancient merger, accreting around $z = 1$ with a progenitor stellar mass in the region $\sim 2 \times 10^9\,\mathrm{M_\odot}$ (Kruijssen et al. 2019). This merger was named *Kraken*, and possible evidence of its debris have since been uncovered near the bottom of the MW potential well in *Gaia* and APOGEE data (Horta et al. 2021; Naidu et al. 2022).

Further examination has suggested that *Kraken* may not be a genuine merger at all, but rather a low-energy extension to the GSE debris or even a misidentified *in-situ* population. Lane, Bovy & Mackereth (2022) highlighted that the apparently distinct energy level of *Kraken* debris was aligned with an energy gap in the survey selection function of APOGEE data. However, the distinct chemistry of *Kraken* stars cannot easily be explained if they were exclusively an extension of the low-energy GSE tail (i.e. Amarante et al. 2022). Around the same time, Belokurov & Kravtsov (2022) identified a population of old, high-α and *in-situ* stars toward the MW bulge which they named *Aurora*. The chemistry of this population overlaps with the *Kraken* debris, and the high-apocentre tail of *Aurora* stars could feasibly travel out of the bulge and masquerade as *Kraken* stars. Further decomposition of the stellar halo has supported this scenario (Myeong et al. 2022; Rix et al. 2022), and Orkney et al. (2022) showed that debris from a *Kraken*-style merger may be difficult to

distinguish from contemporaneous *in-situ* populations. On the other hand, Horta et al. (2022) are able to separate the *Aurora* and *Kraken* populations by their distinct α-abundance at low metallicity using a promising statistical method. It would, therefore, be incredibly valuable to assess the significance of this result using mock data.

It is clear that the pervasiveness of the GSE debris, and contamination from other stellar populations, are a real danger when investigating the inner MW. None the less, the origins of those remaining globular cluster populations remains to be conclusively decided, which could suggest yet undiscovered high-redshift merger debris.

We have shown that, across our simulation sample, debris from the main GSE-type progenitor is the single largest contributor to the stellar halo around the solar neighbourhood, making up roughly 50 per cent of all *ex-situ* halo stars. However, this relative contribution is greatly reduced within the inner 5 kpc and for metallicities [Fe/H] < -2.5. The contribution can be decreased even further by excluding stars with apocentres greater than 5 kpc, in some cases to less than 5 per cent. The *ex-situ* halo mass fraction is exceedingly low in the inner few kpc (typically less than 20 per cent, with much of the contamination arising from kinematically heated disc stars). However, this fraction is raised dramatically when considering only the most metal-poor stars (to around 70 per cent in most cases).

From a philosophical point of view, it becomes increasingly difficult to consider any single progenitor galaxy to be the *main* MW progenitor before a time of $\sim z = 6$. Many of the different progenitor components would be of a similar mass and star formation rate at these earliest times (see e.g. Appendix A), and so the distinction between *in* and *ex situ* becomes more and more meaningless. It can be argued that all stars formed within the deepest potential wells at $\sim z > 6$ should be considered *in situ*, regardless of whether any of those potential wells were the *main* progenitor or not (e.g. Rix et al. 2022). Alternatively, these early stellar populations could be distinguished based on whether they are chemically 'evolved' or 'unevolved' (e.g. Fernandes et al. 2022). Whatever their label, the stars from these earliest mergers offer a privileged view into processes that governed the early Universe, and so their detection is of paramount importance.

Investigating the centre of the MW incurs numerous challenges. There is high stellar crowding, obscuration from foreground populations and interstellar dust, ongoing active star formation, and low-metallicity stars are exceedingly rare. However, there are also distinct advantages. A large number of stars can be investigated with relatively few pointings and less volume coverage. The debris from ancient disrupted dwarfs will have lower orbital apocentres that are contained within the Galactic Centre, meaning that low-metallicity stars identified here are more likely to be genuine ancient debris, as opposed to at the solar radius and beyond where there are visitations from a wide range of eccentric orbits. Moreover, in the search for the very first stars, the Galactic Centre is far closer and more available than other likely environments – such as distant dwarf galaxies.

Observational surveys are beginning to resolve the inner MW in greater and greater detail (e.g. ARGOS Ness et al. 2013, EMBLA Howes et al. 2017, COMBS Lucey et al. 2019, PIGS Arentsen et al. 2020 and also APOGEE Rix et al. 2022), with increasing focus on metal poor stars below [Fe/H] $= -2$. These surveys reveal a high stellar density in the central few kpc of the MW, with a metallicity-dependent rotation that vanishes at around [Fe/H] $\simeq -2$ (i.e. Arentsen et al. 2020) – possibly revealing a transition to a pressure-supported classical bulge (e.g. Babusiaux et al. 2010). There is now evidence that some of these stars were not born *in situ*, and arrived via globular cluster or dwarf galaxy merger (Sestito et al.

2023). The Galactic Centre may therefore offer a resolved view of the stellar populations formed in pre-reionisation galaxies, which would be complementary to the insights provided by the *JWST* (Gardner et al. 2006). This, combined with cutting-edge spectroscopic surveys: SDSS-V (Kollmeier et al. 2017), 4MOST (de Jong et al. 2019), MOONS (Cirasuolo et al. 2020), and asteroseismic surveys focused on metal-poor and dense stellar fields (e.g. HAYDN Miglio et al. 2021b), promises an unprecedented understanding of the Galactic Centre and the physics that govern the first galaxies. Therefore, it is encouraging that the vast debris from the GSE and its satellites could potentially be screened with relatively simple parameter cuts.

5 CONCLUSIONS

We have investigated a selection of GSE-like merger events from nine MW-like galaxies in the AURIGA simulation suite. F19 originally identified radially anisotropic debris features within these simulations, with properties comparable to that of the velocity ellipsoid observed in the MW (Belokurov et al. 2018). They showed that the bulk of this debris can be traced back to progenitor galaxies that accreted onto their hosts over the range $2.9 > z > 0.75$, with peak stellar masses spanning roughly an order of magnitude $(3 \times 10^8 < M_\star(\text{peak})/\text{M}_\odot < 4 \times 10^9)$. We list our core results as follows:

(i) In six of the nine simulations, F19 show that the vast majority of the radially anisotropic stars (velocities of $|v_\phi| < 50 \, \text{km s}^{-1}$; $100 < |v_\text{r}|/\text{km s}^{-1} < 400$) are associated with a single merger event. We find that there are some instances where a second or even a third merger contributes significant mass fractions of their own. Some of these lesser contributions are also radially anisotropic, and it is not always possible to distinguish them by their chemodynamic properties. Similar results were found independently in Rey et al. (2023).

(ii) The GSE-type progenitor galaxies exhibit a wide range of properties, including differing metallicity gradients and both rotation and pressure supported systems. However, the final debris distributions do not strongly relate to the progenitor properties.

(iii) The GSE-type debris is initially prolate and tilted with respect to the disc, but dynamical processes drive an evolution towards a round or oblate shape and close alignment with the disc. This is in tension with the prolate and inclined shape of GSE debris reported in Iorio & Belokurov (2019), Han et al. (2022a), Lane et al. (2023), and may indicate that the early stellar mass enhancement in the AURIGA physics model leads to exaggerated torques between the disc and other galactic components.

(iv) Many of the GSE-type mergers are accompanied by their own luminous satellite populations, with between 0-8 of stellar mass $> 10^6 \, \text{M}_\odot$. The majority of these are fully disrupted by $z = 0$, and in most cases are not dragged to low orbital energies along with the main merger debris. Of those few satellites that survive as remnants, there is no clear indication in their orbital properties that they were once associated with the GSE-type merger.

(v) The progenitor mergers of GSE-like debris features contribute high mass fractions to the stellar halo around the solar neighbourhood, and in some cases out to larger radii. However, these fractions are in most cases reduced to below 20 per cent when considering the central few kpc and metallicities below [Fe/H] < -2.5. The fraction of stars from *ex-situ* sources is simultaneously increased when considering these low metallicities, increasing from less than 10 per cent in the inner few kpc to around 70 per cent. This highlights the centre of the MW as a promising environment to search for the

ancient stars that formed in pre-reionization dwarf galaxies, and a convenient alternative to distant objects at high redshifts.

DATA AVAILABILITY

Available upon request.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 597, 21
Amarante J. A. S., Debattista V. P., Beraldo e Silva L., Laporte C. F. P., Deg N., 2022, ApJ, 937, 12
Amorisco N. C., 2017, MNRAS, 464, 2882
Arentsen A. et al., 2020, MNRAS, 491, L11
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2022, ApJ, 935, 167
Babusiaux C. et al., 2010, A&A, 519, A77
Barbá R. H., Minniti D., Geisler D., Alonso-García J., Hempel M., Monachesi A., Arias J. I., Gómez F. A., 2019, ApJ, 870, L24
Bell E. F. et al., 2008, ApJ, 680, 295
Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689
Belokurov V. et al., 2006, ApJ, 642, L137
Belokurov V. et al., 2007a, ApJ, 657, L89
Belokurov V. et al., 2007b, ApJ, 658, 337
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020, MNRAS, 494, 3880
Berentzen I., Shlosman I., 2006, ApJ, 648, 807
Bignone L. A., Helmi A., Tissera P. B., 2019, ApJ, 883, L5
Binney J., 1980, MNRAS, 190, 873
Binney J., May A., 1986, MNRAS, 218, 743
Binney J., Tremaine S., 1987, Galactic Dynamics, Princeton University Press, Princeton, N.J.
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition, Princeton University Press, Princeton, N.J.
Blázquez-Calero G. et al., 2020, MNRAS, 491, 1800
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Bochanski J. J., Willman B., West A. A., Strader J., Chomiuk L., 2014, AJ, 147, 76
Bonaca A. et al., 2020, ApJ, 897, L18
Brauer K., Andales H. D., Ji A. P., Frebel A., Mardini M. K., Gómez F. A., O'Shea B. W., 2022, ApJ, 937, 14
Brook C. B., Kawata D., Gibson B. K., Flynn C., 2003, Ap&SS, 284, 845
Buist H. J. T., Helmi A., 2015, A&A, 584, A120
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931
Chandra V. et al., 2022, AJ, 951, 15
Chiba M., Beers T. C., 2000, AJ, 119, 2843
Cirasuolo M. et al., 2020, The Messenger, 180, 10
Cohen J. G., Sesar B., Bahnolzer S., He K., Kulkarni S. R., Prince T. A., Bellm E., Laher R. R., 2017, ApJ, 849, 150
Cooper A. P. et al., 2010, MNRAS, 406, 744
Cooper A. P., Parry O. H., Lowing B., Cole S., Frenk C., 2015, MNRAS, 454, 3185
de Jong R. S. et al., 2019, The Messenger, 175, 3
Das P., Hawkins K., Jofré P., 2020, MNRAS, 493, 5195
DeBuhr J., Ma C.-P., White S. D. M., 2012, MNRAS, 426, 983
Deason A. J. et al., 2012, MNRAS, 425, 2840
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, ApJ, 763, 113
Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, ApJ, 862, L1
Debattista V. P., van den Bosch F. C., Roškar R., Quinn T., Moore B., Cole D. R., 2015, MNRAS, 452, 4094
Dey A. et al., 2023, ApJ, 944, 1
Dillamore A. M., Belokurov V., Font A. S., McCarthy I. G., 2022, MNRAS, 513, 1867
Dodge B. C., Slone O., Lisanti M., Cohen T., 2023, MNRAS, 518, 2870
Donlon T., Newberg H. J., 2023, ApJ, 944, 169
Donlon Thomas I., Newberg H. J., Kim B., Lépine S., 2022, ApJ, 932, L16
Dubinski J., Carlberg R. G., 1991, ApJ, 378, 496
Dubinski J., Kuijken K., 1995, ApJ, 442, 492
Earp S. W. F., Debattista V. P., Macciò A. V., Wang L., Buck T., Khachaturyants T., 2019, MNRAS, 488, 5728
Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748
El-Badry K. et al., 2018, MNRAS, 480, 652
Erkal D., Sanders J. L., Belokurov V., 2016, MNRAS, 461, 1590
Evans T. A., Fattahi A., Deason A. J., Frenk C. S., 2022, MNRAS, 516, 3861
Fattahi A. et al., 2019, MNRAS, 484, 4471 (F19)
Fattahi A. et al., 2020, MNRAS, 497, 4459
Fernandes L. et al., 2022, MNRAS
Fernández-Alvar E. et al., 2018, ApJ, 852, 50
Feuillet D. K., Feltzing S., Sahlholdt C. L., Casagrande L., 2020, MNRAS, 497, 109
Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, MNRAS, 416, 2802
Forbes D. A., 2020, MNRAS, 493, 847
Fragkoudi F. et al., 2020, MNRAS, 494, 5936
Franx M., Illingworth G., de Zeeuw T., 1991, ApJ, 383, 112
Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487
Frenk C. S., Evrard A. E., White S. D. M., Summers F. J., 1996, ApJ, 472, 460
Gaia Collaboration et al., 2016, A&A, 595, A2
Gaia Collaboration et al., 2018, A&A, 616, A10
Gallart C., Zoccali M., Aparicio A., 2005, ARA&A, 43, 387
Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019a, Nat. Astron., 3, 932
Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019b, Nat. Astron., 3, 932
Gardner J. P. et al., 2006, Space Sci. Rev., 123, 485
Gargiulo I. D. et al., 2019, MNRAS, 489, 5742
Genel S. et al., 2014, MNRAS, 445, 175
Gómez F. A., Helmi A., Cooper A. P., Frenk C. S., Navarro J. F., White S. D. M., 2013, MNRAS, 436, 3602

Gómez F. A., White S. D. M., Grand R. J. J., Marinacci F., Springel V., Pakmor R., 2017a, MNRAS, 465, 3446

Gómez F. A. et al., 2017b, MNRAS, 472, 3722

Grand R. J. J. et al., 2017, MNRAS, 467, 179

Grand R. J. J. et al., 2018, MNRAS, 474, 3629

Grand R. J. J. et al., 2020, MNRAS, 497, 1603

Grand R. J. J. et al., 2021, MNRAS, 507, 4953

Han J. J. et al., 2022a, AJ, 164, 249

Han J. J. et al., 2022b, ApJ, 934, 14

Hayes C. R. et al., 2018, ApJ, 852, 49

Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, ApJ, 863, 113

Helmi A., de Zeeuw P. T., 2000, MNRAS, 319, 657

Helmi A., White S. D. M., de Zeeuw P. T., Zhao H., 1999, Nature, 402, 53

Helmi A., White S. D. M., Springel V., 2003, MNRAS, 339, 834

Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85

Horta D. et al., 2021, MNRAS, 500, 1385

Horta D. et al., 2023, MNRAS, 520, 5671

Howes L. M. et al., 2016, MNRAS, 460, 884

Huang S., Carlberg R. G., 1997, ApJ, 480, 503

Ibata R. A., Gilmore G., Irwin M. J., 1994, Nature, 370, 194

Ibata R. A., Lewis G. F., Irwin M. J., Quinn T., 2002, MNRAS, 332, 915

Iorio G., Belokurov V., 2019, MNRAS, 482, 3868

Jean-Baptiste I., Di Matteo P., Haywood M., Gómez A., Montuori M., Combes F., Semelin B., 2017, A&A, 604, A106

Jenkins A., 2013, MNRAS, 434, 2094

Jethwa P., Erkal D., Belokurov V., 2016, MNRAS, 461, 2212

Katz N., 1991, ApJ, 368, 325

Katz N., White S. D. M., 1993, ApJ, 412, 455

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Kazantzidis S., Abadi M. G., Navarro J. F., 2010, ApJ, 720, L62

Kollmeier J. A. et al., 2017, preprint (arXiv:1711.03234)

Koppelman H., Helmi A., Veljanoski J., 2018, ApJ, 860, L11

Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019, A&A, 631, L9

Koppelman H. H., Bos R. O. Y., Helmi A., 2020, A&A, 642, L18

Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019, MNRAS, 486, 3180

Kruijssen J. M. D. et al., 2020, MNRAS, 498, 2472

LSST Science Collaboration et al., 2009, preprint (arXiv:0912.0201)

Lancaster L., Koposov S. E., Belokurov V., Evans N. W., Deason A. J., 2019, MNRAS, 486, 378

Lane J. M. M., Bovy J., Mackereth J. T., 2022, MNRAS, 510, 5119

Lane J., Bovy J., Mackereth T., 2023, preprint (arXiv:2306.03084)

Limberg G. et al., 2021, ApJ, 907, 10

Limberg G., Souza S. O., Pérez-Villegas A., Rossi S., Perottoni H. D., Santucci R. M., 2022, ApJ, 935, 109

Lucey M. et al., 2019, MNRAS, 488, 2283

Lynden-Bell D., Lynden-Bell R. M., 1995, MNRAS, 275, 429

McLaughlin D. E., van der Marel R. P., 2005, ApJS, 161, 304

Mackereth J. T., Bovy J., 2020, MNRAS, 492, 3631

Mackereth J. T. et al., 2019, MNRAS, 482, 3426

Maffione N. P., Gómez F. A., Cincotta P. M., Giordano C. M., Cooper A. P., O'Shea B. W., 2015, MNRAS, 453, 2830

Maffione N. P. et al., 2018, MNRAS, 478, 4052

Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, ApJ, 599, 1082

Malhan K. et al., 2022, ApJ, 926, 107

Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750

Martínez-Delgado D., Peñarrubia J., Gabany R. J., Trujillo I., Majewski S. R., Pohlen M., 2008, ApJ, 689, 184

Martínez-Delgado D., D'Onghia E., Chonis T. S., Beaton R. L., Teuwen K., GaBany R. J., Grebel E. K., Morales G., 2015, AJ, 150, 116

Massana P. et al., 2022, MNRAS, 513, L40

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

Mayer L., Moore B., Quinn T., Governato F., Stadel J., 2002, MNRAS, 336, 119

Meza A., Navarro J. F., Abadi M. G., Steinmetz M., 2005, MNRAS, 359, 93

Miglio A. et al., 2017, Astron. Nachr., 338, 644

Miglio A. et al., 2021a, Exp. Astron., 51, 963

Miglio A. et al., 2021b, Exp. Astron., 51, 963

Mihos J. C., Hernquist L., 1994, ApJ, 425, L13

Monachesi A. et al., 2019, MNRAS, 485, 2589

Moore B., Kazantzidis S., Diemand J., Stadel J., 2004, MNRAS, 354, 522

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, ApJ, 856, L26

Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235

Myeong G. C., Belokurov V., Aguado D. S., Evans N. W., Caldwell N., Bradley J., 2022, ApJ, 938, 21

Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, ApJ, 901, 48

Naidu R. P. et al., 2021, ApJ, 923, 92

Naidu R. P. et al., 2022, ApJ, 926, L36

Ness M. et al., 2013, MNRAS, 432, 2092

Ngan W., Carlberg R. G., Bozek B., Wyse R. F. G., Szalay A. S., Madau P., 2016, ApJ, 818, 194

Nissen P. E., Schuster W. J., 2010, A&A, 511, L10

Orkney M. D. A. et al., 2022, MNRAS, 517, L138

Pace A. B., Erkal D., Li T. S., 2022, ApJ, 940, 136

Peñarrubia J., Benson A. J., Martínez-Delgado D., Rix H. W., 2006, ApJ, 645, 240

Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825

Perottoni H. D., Limberg G., Amarante J. A. S., Rossi S., Queiroz A. B. A., Santucci R. M., Pérez-Villegas A., Chiappini C., 2022, ApJ, 936, L2

Planck Collaboration XVI, 2014, A&A, 571, A16

Prada J., Forero-Romero J. E., Grand R. J. J., Pakmor R., Springel V., 2019, MNRAS, 490, 4877

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

Price-Whelan A. M., Johnston K. V., Valluri M., Pearson S., Küpper A. H. W., Hogg D. W., 2016, MNRAS, 455, 1079

Quinn P. J., Hernquist L., Fullagar D. P., 1993, ApJ, 403, 74

Re Fiorentin P., Spagna A., Lattanzi M. G., Cignoni M., 2021, ApJ, 907, L16

Read J. I., Lake G., Agertz O., Debattista V. P., 2008, MNRAS, 389, 1041

Renaud F., Agertz O., Read J. I., Ryde N., Andersson E. P., Bensby T., Rey M. P., Feuillet D. K., 2021a, MNRAS, 503, 5846

Renaud F., Agertz O., Andersson E. P., Read J. I., Ryde N., Bensby T., Rey M. P., Feuillet D. K., 2021b, MNRAS, 503, 5868

Rey M. P. et al., 2023, MNRAS

Rix H.-W. et al., 2022, ApJ, 941, 45

Robertson B., Bullock J. S., Font A. S., Johnston K. V., Hernquist L., 2005, ApJ, 632, 872

Rocha M., Peter A. H. G., Bullock J., 2012, MNRAS, 425, 231

Ruiz-Lara T. et al., 2020, A&A, 639, L3

Sales L. V., Wang W., White S. D. M., Navarro J. F., 2013, MNRAS, 428, 573

Sales L. V., Navarro J. F., Kallivayalil N., Frenk C. S., 2017, MNRAS, 465, 1879

Schaye J. et al., 2015, MNRAS, 446, 521

Schlaufman K. C., Casey A. R., 2014, ApJ, 797, 13

Schweizer F., 1986, Science, 231, 227

Searle L., Zinn R., 1978, ApJ, 225, 357

Sestito F. et al., 2023, MNRAS, 518, 4557

Simion I. T., Belokurov V., Koposov S. E., 2019, MNRAS, 482, 921

Spergel D. et al., 2015, preprint (arXiv:1503.03757)

Springel V., 2010, MNRAS, 401, 791

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V. et al., 2005, Nature, 435, 629

Starkenburg E., Oman K. A., Navarro J. F., Crain R. A., Fattahi A., Frenk C. S., Sawala T., Schaye J., 2017, MNRAS, 465, 2212

Tissera P. B., Dominguez-Tenreiro R., 1998, MNRAS, 297, 177

Tissera P. B., Domínguez-Tenreiro R., Scannapieco C., Sáiz A., 2002, MNRAS, 333, 327

Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371

Tomassetti M. et al., 2016, MNRAS, 458, 4477

Unavane M., Wyse R. F. G., Gilmore G., 1996, MNRAS, 278, 727

van den Bosch F. C., Jiang F., Campbell D., Behroozi P., 2016, MNRAS, 455, 158

Vasiliev E., 2019, MNRAS, 482, 1525

Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, AJ, 128, 1177

Vera-Casanova A. et al., 2022, MNRAS, 514, 4898

Villalobos Á., Helmi A., 2008, MNRAS, 391, 1806

Vincenzo F., Spitoni E., Calura F., Matteucci F., Silva Aguirre V., Miglio A., Cescutti G., 2019, MNRAS, 487, L47

Vivas A. K. et al., 2001, ApJ, 554, L33

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Warren M. S., Quinn P. J., Salmon J. K., Zurek W. H., 1992, ApJ, 399, 405

Wetzel A. R., 2011, MNRAS, 412, 49

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Yuan Z. et al., 2020a, ApJ, 891, 39

Yuan Z., Chang J., Beers T. C., Huang Y., 2020b, ApJ, 898, L37

Zhu Q., Hernquist L., Marinacci F., Springel V., Li Y., 2017, MNRAS, 466, 3876

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

suppl_data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: MERGER TREE

Here, we show a visualization of the LHALOTREE merger tree for Au-5 in Fig. A1. The formatting was made with the assistance of GRAPHVIZ and the PYDOT package for PYTHON. Each node represents a separate subhalo, with the connecting lines indicating the descendants (left) and progenitors (right). The node size represents the total mass of the subhalo as identified by SUBFIND. The node colour is the instantaneous star formation rate over all gas cells. We consider all mergers that are greater than 1:20, and exclude haloes with no progenitors greater than a halo mass of $10^9 \, M_\odot$.

The main progenitor lines of the final halo, and of the notable mergers from Fig. 1, are shown with distinct colours as described in the legend. These notable mergers are annotated with a merger mass ratio, which describes the ratio of the M_{200} virial mass between the main progenitor line and the merger progenitor line at a time before infall.

Merger tree visualizations for the other AURIGA simulations presented in this paper are included at this link, and see also the supplementary material.

APPENDIX B: HALO CONTRIBUTIONS WITHIN THE GSE E/L_z LOCUS

In Section 3.1, we argue that the contribution fraction of GSE-type merger debris is lowest for halo stars with metallicity [Fe/H] < −2.5, and this suggests that other merger debris would be more effectively distinguished from GSE-type debris in the low-metallicity regime. It is conceivable that this trend might fail for debris that shares the same locus in E/L_z as the GSE-type merger debris.

In Fig. B1, we reproduce the right-hand panels of Fig. 13, but for a selection in E/L_z that encompasses 90 per cent of the GSE-type stellar debris mass. We include a chemical cut of [Mg/Fe] > 0.2, intended to highlight merger debris associated with ancient dwarf galaxies. The left-hand panel shows a histogram of all *ex-situ* stars in E/L_z, where the thick black contour is the GSE-type debris selection. The right-hand panels show the relative contribution of the stars within that contour and with the same parameter cuts, using the same colour scheme as in Section 3.1.

Even though the contour is specifically limited to the E/L_z space that is dominated by GSE-type debris, there remains a large fraction of stars from other merger events. The GSE-type debris fraction at metallicities of [Fe/H] < −2.5 is 30 per cent, up from 24 per cent when no E/L_z selection is made. The results are similar across the rest of our AURIGA selection, with the E/L_z selection making little or no difference to the low fraction of GSE-type debris at low metallicity.
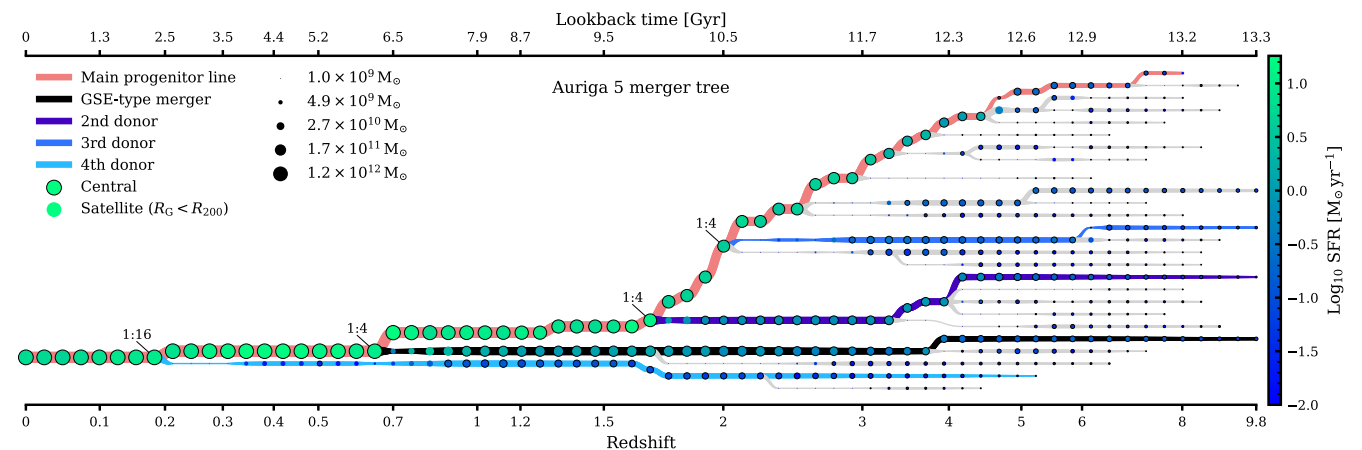

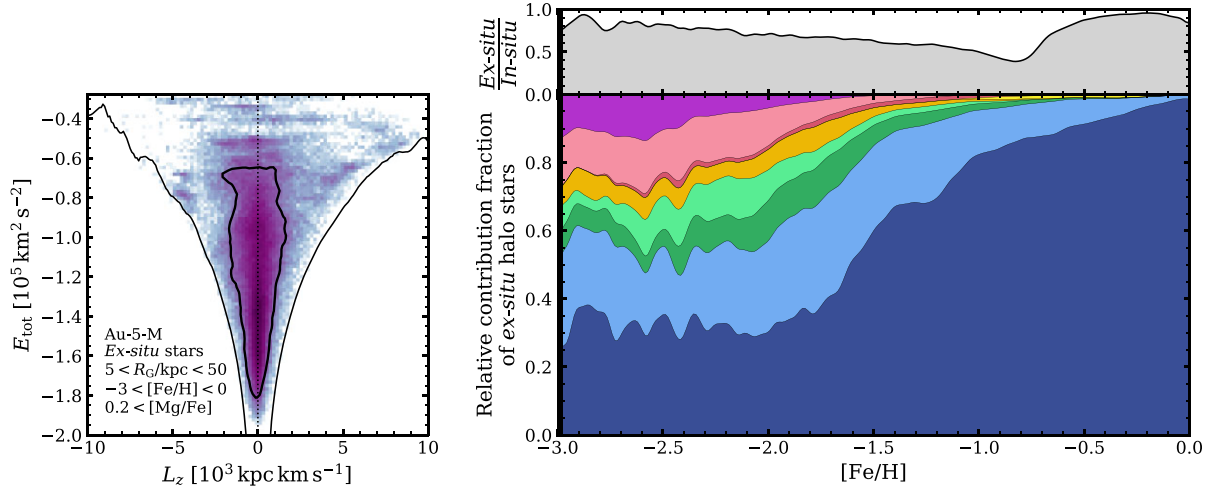**Figure A1.** A merger tree visualization for Au-5. See the main text for details.

Figure B1. Left-hand panel: a histogram of the *ex-situ* halo stars in Au-5, as shown in E/L_z parameter space. The parameter cuts are listed in the lower left-hand corner. A thick black contour encircles 90 per cent of the mass associated with the GSE-type merger event. Right-hand panels: the same as in Fig. 13, but for stars within the contour selection defined in the left-hand panel. This shows that the proportion of stars from the GSE-type merger (dark blue band) is reduced for metallicities [Fe/H] < −2.5, even within the E/L_z region dominated by GSE-type debris.

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