Gastro Library. I. The Simulated Chemodynamical Properties of Several Gaia–Sausage–Enceladus-like Stellar Halos

João A. S. Amarante[1,2,6] , Victor P. Debattista[2] , Leandro Beraldo E Silva[2,3] , Chervin F. P. Laporte[4] , and Nathan Deg[5]

[1] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E-08028 Barcelona, Spain; joaoant@gmail.com
[2] Jeremiah Horrocks Institute, University of Central Lancashire, Preston, PR1 2HE, UK
[3] Department of Astronomy, University of Michigan, 1085 S. University Ave., Ann Arbor, MI 48109, USA
[4] Institut de Ciencies del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E-08028 Barcelona, Spain
[5] Department of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, ON, K7L 3N6, Canada

Abstract

The Milky Way (MW) stellar halo contains relics of ancient mergers that tell the story of our galaxy's formation. Some of them are identified due to their similarity in energy, actions, and chemistry, referred to as the "chemodynamical space," and are often attributed to distinct merger events. It is also known that our galaxy went through a significant merger event that shaped the local stellar halo during its first billion years. Previous studies using N-body only and cosmological hydrodynamical simulations have shown that such a single massive merger can produce several "signatures" in the chemodynamical space, which can potentially be misinterpreted as distinct merger events. Motivated by these, in this work we use a subset of the GASTRO library, which consists of several smoothed particle hydrodynamics+N-body models of a single accretion event in a MW-like galaxy. Here, we study models with orbital properties similar to the main merger event of our galaxy and explore the implications to known stellar halo substructures. We find that (i) supernova feedback efficiency influences the satellite's structure and orbital evolution, resulting in distinct chemodynamical features for models with the same initial conditions; (ii) very retrograde high-energy stars are the most metal-poor of the accreted dwarf galaxy and could be misinterpreted as a distinct merger; (iii) the most bound stars are more metal-rich in our models, the opposite of what is observed in the MW, suggesting a secondary massive merger; and, finally, (iv) our models can reconcile other known apparently distinct substructures to a unique progenitor.

Unified Astronomy Thesaurus concepts: Milky Way dynamics (1051); Milky Way formation (1053); Milky Way stellar halo (1060); Hydrodynamical simulations (767)

1. Introduction

In the Lambda cold dark matter (ΛCDM) paradigm, the build-up process of the stellar halo of a Milky Way-mass galaxy includes the accretion of hundreds of luminous satellite galaxies into the main progenitor (e.g., Searle & Zinn 1978; Bullock & Johnston 2005; Rey & Starkenburg 2022). This view is supported by observations of high-redshift galaxy mergers (e.g., Patton et al. 2002; Lin et al. 2004) and also by numerical simulations of galaxy formation (e.g., Moore et al. 1999; Benson et al. 2000; Newton et al. 2018). The accreted stars will then build the host galaxy's stellar halo, and signatures in the chemodynamical space of these stars help trace back the accretion history of now-disrupted satellite galaxies.

Beyond 30 kpc in galactocentric radius, the Milky Way's (MW) outer stellar halo exhibits a large amount of past accretion events in the form of stellar streams (e.g., Belokurov et al. 2006; Shipp et al. 2018; Ibata et al. 2021) as well as a number of ongoing accretion events such as the Sagittarius dwarf (e.g., Ibata et al. 1994; Johnston et al. 1995; Majewski et al. 2003; Laporte et al. 2019) and the Large and Small Magellanic Clouds (e.g., Kallivayalil et al. 2006; Besla et al. 2007). While most of the stellar streams are detected as spatial overdensities in photometric surveys with relatively little spectroscopic follow-up, the inner stellar halo is currently being studied in greater detail thanks to the advent of large spectroscopic surveys and the Gaia mission.

With the increasing number of stars with full 6D phase-space information and measured chemical abundances, there are currently dozens of unambiguously[7] identified substructures in the MW's inner stellar halo (for an overview, see e.g., Helmi 2020, Naidu et al. 2020, Yuan et al. 2020b, Limberg et al. 2021, Malhan 2022, and Ruiz-Lara et al. 2022). In principle, different mergers should be distinguishable in the dynamical space, e.g., energy, angular momentum, and actions, in a slowly evolving potential (as shown in, e.g., Johnston et al. 1996 and Helmi & White 1999). This has been explored extensively in order to find signatures of merger event(s) with pure N-body (e.g., Jean-Baptiste et al. 2017; Koppelman et al. 2020; Naidu et al. 2021) and cosmological hydrodynamic simulations (e.g., Wu et al. 2022). These have shown that (i) different merger events can overlap in dynamical space even if they started with different orbit configurations and (ii) single merger events can leave their imprints on a large range of energy and actions. These results lead to the question whether all substructures have a separate origin or whether several are the relics of the same merger event.

Since the early 2000s there has been growing evidence of a significant merger event that shaped the MW's inner stellar

---

[6] Visiting Fellow at UCLan.

---

[7] Some substructures are reidentified and given different names.

halo ($r_{gc} \lesssim 30$ kpc during the galaxy's first billion years; Chiba & Beers 2000; Gilmore et al. 2002; Meza et al. 2005; Nissen & Schuster 2010; Deason et al. 2013). This event, recently dubbed the "Gaia–Sausage–Enceladus" (GSE), has been confirmed (Belokurov et al. 2018; Helmi et al. 2018) thanks to the exquisite data from the Gaia satellite (Gaia Collaboration et al. 2016, 2018). The GSE has a net rotational velocity close to zero and a large radial velocity dispersion, in contrast to the more isotropic velocity distribution of the metal-poor, [Fe/H] < −2, stellar halo (Belokurov et al. 2018). The merger is estimated to have happened during the first three Gyr of the MW's formation (e.g., Gallart et al. 2019; but see Donlon et al. 2020 for an alternative time of the merger). Its median [Fe/H] ranges from ∼−1.45 to −1.17 depending on the spectroscopic survey and sample selection (e.g., Fernández-Alvar et al. 2018; Amarante et al. 2020b; Das et al. 2020; Feuillet et al. 2020; Naidu et al. 2020; Bonifacio et al. 2021; Buder et al. 2022). Its stellar mass is estimated to be in the range $10^{8.5}$–$10^{9.7}\,M_{\odot}$ (e.g., Vincenzo et al. 2019; Feuillet et al. 2020; Naidu et al. 2020; Hasselquist et al. 2021; Limberg et al. 2022), which would correspond to a dark matter (DM) halo mass of ∼2–7 × $10^{11}\,M_{\odot}$ (e.g., Moster et al. 2013).

Besides explaining the observational features, simulations can also guide us to understand better the GSE merger. For instance, pure N-body models (Koppelman et al. 2020; Naidu et al. 2021) constrained the GSE dwarf to be on a retrograde orbit (with respect to the MW's disk rotation) with low-to-intermediate circularity, and with a low-inclination angle between its orbital plane and the MW's disk plane. GSE-like mergers have also been identified/studied in cosmological simulations of MW-like galaxies (e.g., Bignone et al. 2019; Fattahi et al. 2019; Mackereth et al. 2019; Elias et al. 2020; Grand et al. 2020; Dillamore et al. 2022). For example, Fattahi et al. (2019) selected MW-like galaxies in Auriga (Grand et al. 2017) and found that about one-third of them have a highly anisotropic, metal-rich, [Fe/H] ≈ −1, stellar halo, which becomes isotropic at lower metallicities, i.e., just as is observed in the MW. In the simulations, these stellar halos are mostly assembled during the first billion years of galaxy evolution. Moreover, the typical stellar mass of such mergers is of the order of 10^{9}–$10^{10}\,M_{\odot}$, similar to what is estimated in EAGLE simulations, $10^{8.5}$–$10^{9}\,M_{\odot}$ (Mackereth et al. 2019).

Given its relatively massive and extended nature, how is the GSE's debris dispersed in the MW? Could the GSE merger event result in more than one of the substructures observed in the chemodynamical space of the local stellar halo? The former question has been explored with pure N-body models by Koppelman et al. (2020) and Naidu et al. (2021), who showed that such a merger extends over a large range in the energy–angular momentum space. In this work, we will tackle the second question exploring a suite of N-body+smoothed particle hydrodynamics (SPH) models of a GSE-like dwarf merging into a MW-like galaxy. Due to the galaxy formation model that self-consistently follows the growth of the host galaxy and its merging satellite—which cannot be done with a pure N-body model—we are able to self-consistently study the chemical evolution of the merging dwarf and its chemodynamical properties at the end of the simulation, i.e., when the accreted stars are dynamically relaxed. In these simulations, a MW-like galaxy grows from the cooling of a hot gas halo and experiences a merger with a dwarf galaxy in a fine-tuned orbit.

Thus, these simulations are intended to reproduce the MW in a tailored manner.

In this paper, we will show the structural evolution of different satellites during their merger process into the host galaxy and how this depends on the strength of the supernova feedback of the models. Also, due to the disk-like nature of dwarfs, we will demonstrate that a GSE-like merger event produces a rich behavior in the accreted stars' chemodynamical space. This has important consequences for the known MW substructures as we are able to link some of them dynamically and chemically to the GSE merger.

The paper is organized as follows. Section 2 introduces the GASTRO library of simulations and the subset of the models analyzed in this work. In Section 3 we describe the evolution of the dwarfs during the merger process. We then explore the chemodynamical signatures of the GSE-like merger event in Section 4. We discuss the observational implications of our models in Section 5, and list our conclusions and final remarks in Section 6.

2. GASTRO Library

The Gaia–EncelAdus–Sausage Timing, chemistRy and Orbit (GASTRO)[8] library is a suite of N-body+SPH simulations exploring the parameter space for GSE-like accretion events into the MW and its impact on the formation history of the MW. In this work, we use a subset of the GASTRO library to explore the chemodynamical trends of the accreted stars. A complete description of the GASTRO library will be presented in a future paper. Throughout this paper we analyze four models (and four extra ones in Appendix B) where a GSE-like galaxy merges with a MW-like galaxy. In Section 2.1 we describe the input physics of the code used to run the models. In Sections 2.2 and 2.3, we describe the initial conditions of the host galaxy and the GSE-like dwarf galaxy models, respectively. Finally, in Section 2.4 we describe the setup of the dwarf's initial orbital parameters.

2.1. GASOLINE Input Physics

All the models are evolved with GASOLINE (Wadsley et al. 2004), which implements SPH into the N-body tree code PKDGRAV (Stadel 2001). The gas phase uses the metallicity-dependent cooling method of Shen et al. (2010). We set a pressure floor on gas particles $p_{floor} = 3G\epsilon^2\rho^2$, where G is the universal gravitational constant, ϵ is the softening length, and ρ is the gas particle's density (Agertz et al. 2009). Star formation begins wherever the temperature drops below 15,000 K and the density exceeds $1\,cm^{-3}$ in a convergent flow (Stinson et al. 2006). The feedback by supernova explosions follows the blast-wave implementation of Stinson et al. (2006). In this work, we present models which couple either 20% or 80% of the 10^{51} erg per supernova as thermal energy ejected into the interstellar medium. In isolated galaxy simulations, low-feedback models have been shown to produce geometrical and chemical thin and thick disks (Clarke et al. 2019; Beraldo e Silva et al. 2020), as well as a kinematically hot thick-disk population (Amarante et al. 2020a), very similar to those observed in the MW. Turbulent diffusion of the gas (Shen et al. 2010) is also included, allowing the gas to mix, which reduces the scatter in the age–metallicity relation (Pilkington et al. 2012).

---

8 http://amarante.netlify.app/GASTRO

The integration base time step is set to $\Delta t = 10$ Myr and is refined such that $\delta t = \Delta t/2^2 < \nu\sqrt{\epsilon/a_g}$, where a_g is the acceleration at a particle's position and the refinement parameter $\nu = 0.175$. The opening angle, which determines when to open a branch in the N-body tree code, is set to $\theta = 0.7$. The time step of gas particles satisfies the further condition $\delta t = \nu_{courant} h/[(1 + \alpha)c + \beta\mu_{max}]$, where the Courant number for SPH criteria $\eta_{courant} = 0.4$, h is the SPH smoothing length set over the nearest 32 particles, α and β are the linear and quadratic viscosity coefficients, and μ_{max} is the maximum viscous force measured between the gas particles (Wadsley et al. 2004; Springel 2010).

Each star particle represents a single stellar population following a Miller–Scalo initial mass function (Miller & Scalo 1979) with a initial mass one-third of the initial parent gas particle. Once a gas particle loses 80% of its mass, its mass is distributed among the nearest neighboring gas particles, which avoids having an inert population of low-mass gas particles. Type II and Type Ia supernova yields of oxygen and iron are taken from Raiteri et al. (1996). Type II and Type Ia supernova rates are determined using Padova stellar lifetimes (as in Raiteri et al. 1996), with the latter taking into account a binary evolution model.

2.2. Milky Way–like Galaxy Setup

The initial conditions set for the host galaxy are the same as in the isolated model used by our group, which describes several chemo-kinematic properties of the MW: the geometrical and chemical thin/thick disks (Clarke et al. 2019; Beraldo e Silva et al. 2020), the "Splash" population (Amarante et al. 2020a), the existence of a significant old thin-disk population in the solar neighborhood (Beraldo e Silva et al. 2021, with a higher mass resolution of the same model), and the bulge chemistry (V. P. Debattista et al. 2022, in preparation). In the following we summarize the initial conditions of the host galaxy.

The initial conditions are characterized by a spherical hot gas corona embedded in a DM halo with a Navarro–Frenk–White (NFW; Navarro et al. 1997) density profile with virial radius $r_{200} \approx 200$ kpc, mass $10^{12}\,M_\odot$, central density $2.4 \times 10^6\,M_\odot$ kpc^{-3}, and scale radius 30 kpc. DM particles come in two mass flavors of $10^6\,M_\odot$ and $3.5 \times 10^6\,M_\odot$ inside and outside 200 kpc, respectively; in total the model has 10^6 DM particles. The gas corona follows the same radial density profile but constitutes only 10% of the total mass of the system. There are 10^6 gas particles each with a mass $1.4 \times 10^5\,M_\odot$. The gas has an initial net rotation, with spin parameter $\lambda = 0.065$ (Bullock et al. 2001), cools via metal-line cooling (see Section 2.1), and settles into a disk. The softening length used for the DM and gas is 100 pc and 50 pc, respectively. The N-body+SPH nature of the models allows star particles to form self-consistently from the cold gas, inheriting the softening length and chemistry of the parent gas particle.

2.3. Dwarf Galaxy Setup

Currently the GASTRO library explores seven different dwarf galaxy models with varying initial gas and DM distributions. In this paper, we show the results for two dwarf galaxies (and a third one in Appendix B), and Table 1 summarizes some of their properties. The variations in their initial conditions have important consequences for the star formation, and thus to the chemical evolution, of each dwarf. Throughout the paper, we will use the words "dwarf" and "satellite" interchangeably, always referring to the less massive galaxy.

The dwarfs D1, D2, and D3 (shown only in Appendix B) are generated using the GALACTICS code (Kuijken & Dubinski 1995; Widrow & Dubinski 2005; Widrow et al. 2008). This particular iteration of the code allows for the generation of equilibrium, exponential gas disks. For a full description of the new features, we refer the reader to Deg et al. (2019). In our runs, all dwarfs are initialized with a DM and gaseous disk with 10^5 and 2×10^4 particles, respectively. These are parameterized by NFW halos with a scale velocity of 200 km s^{-1} and are truncated at 50 kpc. The gas disks of the three dwarfs are exponentials built with an initial kinematic temperature of 1000 K and have a scale radius of 4 kpc. The D1 and D3 DM halos have gas disk exponential scale radii of 1 kpc, while D2 has a scale radius of 5 kpc. The initial gas mass for each dwarf is $1.4 \times 10^9\,M_\odot$, $2.75 \times 10^9\,M_\odot$, and $5.5 \times 10^8\,M_\odot$ for D1, D2, and D3, respectively.

2.4. Merger Impact Parameters

In this paper, we focus on two orbital scenarios involving two merger timescales: one at ~ 3 Gyr and another at ~ 2 Gyr. Throughout the text we will refer to this timescale as the "merging time," which is when the dwarf is completely disrupted. We set the dwarf at a distance from the center of the host galaxy of 200 kpc and 150 kpc for the former and the latter. In both scenarios the initial circularity of the orbit, $\eta \equiv L_z/L_c(E)$, is set to 0.3, where L_z is the z-component of the angular momentum and $L_c(E)$ is the angular momentum of circular motion for an orbit of energy E in the Galactic plane. Finally, all dwarf models have their orbit inclination relative to the host galaxy's midplane set to $15°$.

We note that Naidu et al. (2021), using pure N-body models, found that the optimal initial orbital conditions were $r_{gc} = 150$ kpc, orbit circularity 0.5, and orbit inclination $15°$. While we show here scenarios with the same initial r_{gc} and orbit inclination, our models differ in orbit circularity. This particular choice is based on preliminary analysis of the effects on the host galaxy disk depending on the initial orbit circularity of the satellite. Nonetheless, in Figure B1 we show models in which satellites D2 and D3 are placed on orbits with $\eta = 0.5$, and it is clear that they also have qualitatively similar trends in the energy–L_z space as the models discussed in the main text.

In the next section, we study the evolution of the dwarf properties in each model. Given our present focus on the properties of the stellar halo, the study of the impact of the GSE-like mergers on the evolution of the Galactic disk is left to a separate paper.

3. Evolution of the Models

In this paper, we study four models, where each is set up with the initial conditions of the host galaxy plus a dwarf model. We use dwarfs D1 and D2 while also varying the feedback strength. The low- (high-)feedback models are FB20_D1 and FB20_D2 (FB80_D1 and FB80_D2), in which 20% (80%) of the supernova energy is ejected back into the interstellar medium (see Section 2.1). The models are evolved for 10 Gyr and we save snapshots every 0.1 Gyr. The host galaxy for all of our models, regardless of the dwarf merger or feedback prescription used, end up roughly with the same final
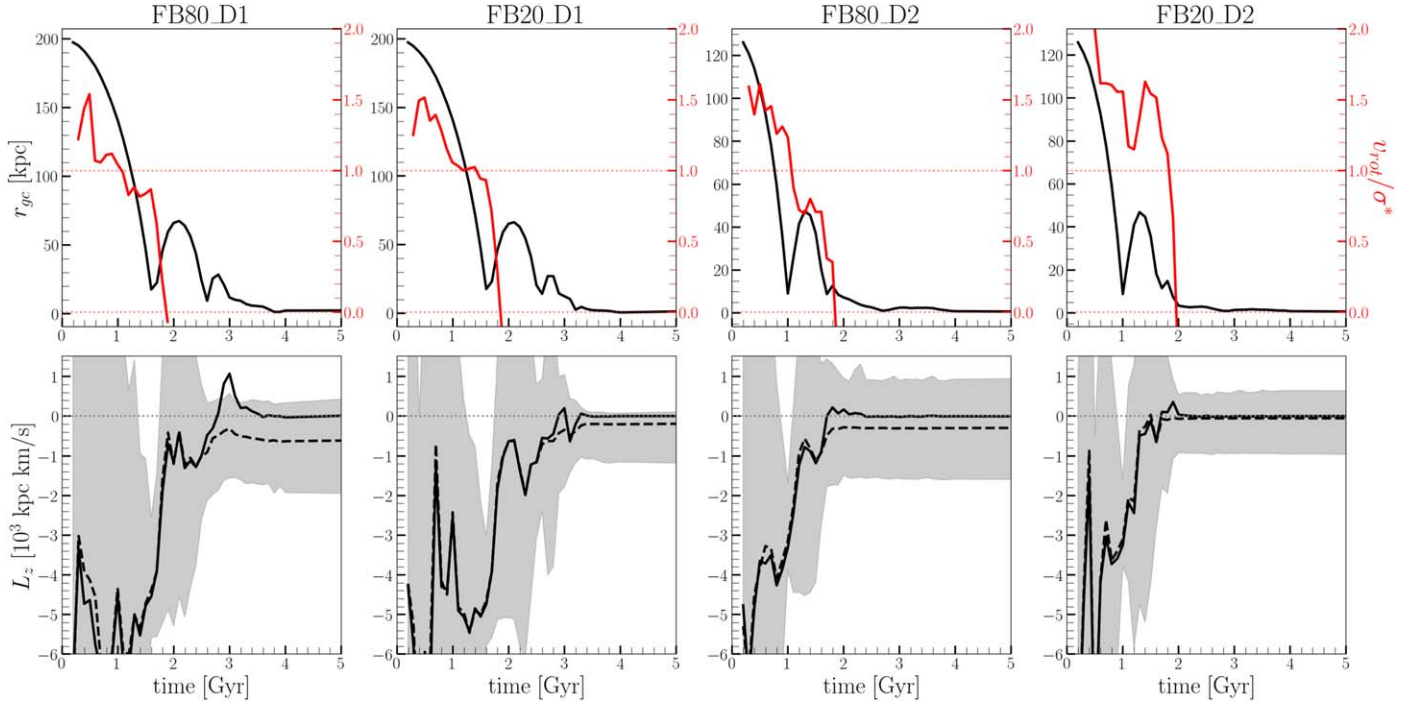
Figure 1. Top row: the evolution of the distance between the dwarf's and the host galaxy's centers represented by a black line with the scale shown on the left axis. We also show the dwarf's $v_{\rm rot}/\sigma^*$ evolution (red line and scale shown on the right axis). See text for details. Bottom row: the median L_z evolution of the dwarf's center of mass (COM) and all accreted stars, shown as solid and dashed lines, respectively. While the COM ends up with $L_{z,\rm orb} = 0$ kpc km s^{-1}, the overall distribution shows signs of bulk retrograde motion, except in model FB20_D2, shown in the last column. The shaded area corresponds to the 5% and 95% percentiles of the L_z distribution.

Table 1
Properties of the Dwarf in Each Model

Model	Mass ($M_\odot$)	$\langle$(Fe/H)$\rangle$	Δ(Fe/H)$/\Delta r$ (dex kpc^{-1})	$\langle$(O/Fe)$\rangle$	Δ(O/Fe)$/\Delta r$ (dex kpc^{-1})	VRHE Fraction (%)
FB80_D1	3.15×10^8	-1.14	-0.17	0.2	0.039	3.72
FB20_D1	8.97×10^8	-0.65	-0.29	0.31	0.034	1.5
FB80_D2	5.53×10^8	-1.08	-0.18	0.23	-0.015	2.06
FB20_D2	1.18×10^9	-0.77	-0.25	0.28	0.016	1.33

Notes. The mass, $\langle$(Fe/H)$\rangle$, and $\langle$(O/Fe)$\rangle$ are measured at the epoch of the merger. The (Fe/H) and (O/Fe) gradients are measured before the dwarfs reach their pericenter, i.e., when they are still rotationally supported. Finally, the very retrograde high-energy (VRHE) stellar mass fraction of the satellite is measured at $t = 10$ Gyr.

total stellar mass of $\approx 6 \times 10^{10} M_\odot$ ($\approx 2 \times 10^6$ star particles). The total rotation curve of the models is in good agreement with the MW's. For instance, its total circular velocity at the solar radius is $v_{c\odot} \approx 260$ km s^{-1} and has an approximately flat profile toward the outer disk. The evolution of the dwarf's structure prior and during the merger are more pertinent to the present work so we describe them in this section.

3.1. Dwarf Galaxy Orbital Decay

The dwarfs D1 and D2 start at a distance of $r_{gc} = 200$ kpc and $r_{gc} = 150$ kpc, respectively, from the center of the host galaxy. In all cases, the satellite's velocity is such that its initial orbit circularity is $\eta = 0.3$. As each model evolves and forms stars, we track the orbital evolution of the dwarf with respect to the host by following its center of mass (COM), distance, $r_{\rm com}$, and the vertical component of its orbit angular momentum, $L_{z,\rm orb}$, measured with respect to the host galaxy's plane. The top row of Figure 1 shows the evolution of $r_{\rm com}$ for all the models as solid black lines. The dwarfs in models FB20_D1 and FB80_D1 have merging timescales of $t \approx 3$ Gyr, and have two

apocenter passages. In contrast, models FB20_D2 and FB80_D2 have comparatively shorter merging timescales of $t \approx 2$ Gyr after only one apocenter.

The satellite's orbital decay can also be studied through the evolution of $L_{z,\rm orb}$, as shown by the solid line in the bottom row of Figure 1. The dwarfs D1 and D2 start with $|L_{z,\rm orb}| \approx 5.8$ and 4.5×10^3 kpc km s^{-1}, respectively. Their COM loses absolute angular momentum during the orbital decay until the orbit is radialized at $L_{z,\rm orb} \approx 0$ kpc km s^{-1}. The evolution of the median L_z of all the dwarf's stars (measured with respect to the host galaxy) is represented by the dashed black line. The shaded area corresponds to the interpercentile range 5%–95%. The accreted stars in the high-feedback models, FB80_D1 and FB80_D2, end up with a lower median L_z compared to their low-feedback counterparts. Moreover, the final L_z distribution is broader in the higher-feedback models. This is caused by feedback directly impacting the evolution of the internal structure of the dwarf galaxies, which in turn affect the final angular momentum distribution. Finally, the median L_z follows the $L_{z,\rm orb}$ until the last pericenter, at which point the dwarf is

completely disrupted. We will explore how the feedback impacts the internal structure of the dwarf galaxies in the following section.

3.2. Dwarf Internal Structural Evolution

In the previous section, we showed that the strength of the feedback does not affect the dwarfs' COM orbital evolution, but influences the final L_z distribution of the accreted stars. In order to understand this result, we now proceed to explore the effect of feedback on the internal structure of the satellites. We start by checking the evolution of $v_{\rm rot}/\sigma^*$, where $v_{\rm rot}$ and σ^* are the dwarf's mean rotational velocity and the total velocity dispersion, respectively, measured in the dwarf's frame. If $v_{\rm rot}/\sigma^* > 1$ the object is rotationally supported. When a satellite is orbiting a massive host galaxy, such as the ones explored here, Kazantzidis et al. (2011) demonstrated that rotationally supported dwarfs can be tidally stirred, by losing their rotational angular momentum during gravitational interactions with the host galaxy.

We study the evolution of the $v_{\rm rot}/\sigma^*$ of the dwarfs for the stars within 1 kpc of its center; the trend also holds at different radii. This is shown in the top row of Figure 1 as the solid red line. Models using D1, which is the less massive and radially extended model, show similar evolution of $v_{\rm rot}/\sigma^*$. Before the first pericenter, the centers of the dwarfs are mildly rotationally supported, becoming pressure supported after the first pericenter passage.

Models accreting dwarf D2 show a different $v_{\rm rot}/\sigma^*$ evolution depending on the feedback strength. Before the first pericenter, both models have similar $v_{\rm rot}/\sigma^* \approx 1.5$. However, soon after, the high-feedback model, FB80_D2, loses all of its rotational support due to the interaction with the host galaxy, already reaching $v_{\rm rot}/\sigma^* < 1$. On the other hand, the low-feedback model, FB20_D2, recovers part of its internal rotational support due to a stronger starburst in its central parts. As the dwarf approaches the final pericenter, it becomes pressure supported, i.e., $v_{\rm rot}/\sigma^* \approx 0$.

The effect of different feedback models can also be seen in the dwarf's rotation curve and density profile evolution. The first and second rows of Figure 2 show the evolution of satellite D1's circular rotation curve, v_c, which is calculated based on the total mass within its midplane, and spherical mass density profile, ρ, respectively. The DM, gas, and stellar mass content are shown from the left to right columns. Before the first pericenter passage, at $t = 1.4$ Gyr, both feedback models have similar behavior, as shown by the lighter colored lines. The dark lines show the profiles before total disruption at $t = 2.4$ Gyr. At this point, the mass profiles are drastically different depending on the feedback strength. The DM, gas, and stellar mass distributions in the low-feedback model are more centrally concentrated. This effect is enhanced by the tidal interaction of the dwarf with its host: the dwarf center is compressed and, in the lower-feedback model, together with its centrally concentrated gas density, allows a higher star formation rate (since high feedback suppresses star formation; e.g., Hopkins et al. 2012; Faucher-Giguère et al. 2013) in the central parts, leading to a more centrally concentrated stellar mass distribution compared to the high-feedback model. Finally, the higher stellar mass content leads to adiabatic contraction of the DM to the center (Blumenthal et al. 1986; Gnedin et al. 2004), making its distribution more centrally concentrated in the low-feedback regime.

The last two rows show the profile for dwarf D2, at $t = 0.9$ Gyr and $t = 1.4$ Gyr, i.e., before and after the first pericenter passage, with light and dark colors, respectively. This dwarf shows the same trends with feedback as discussed above, with the difference that its stellar content is more radially extended in comparison with dwarf D1. The fact that low-feedback models have a centrally concentrated mass distribution allows their stellar content to lose more L_z, becoming more radialized compared to the higher-feedback counterpart (as shown in Figure 1; see also a detailed discussion on the radialization of mergers in Vasiliev et al. 2022).

As previously mentioned, models with single merger events have already been extensively explored with N-body-only simulations. Now as we add more complexity to these models, by adding gas physics and star formation, we start to see extra dependence on the adopted input physics. For instance, so far we have shown that the feedback strength influences the satellite's internal structure and, consequently, affects the satellite stars' final L_z distribution. In the following section we will also discuss the effect of feedback on the chemical evolution of the dwarfs.

### 3.3. The Dwarf Galaxy Chemistry

As the models evolve and form stars self-consistently, there is a natural chemical evolution of the host and satellite galaxies. In particular for the dwarf, which is the main focus in the present paper, this will reflect on the final chemodynamical configuration of the accreted stars (see Section 4) and impact the interpretation of the known substructures in the MW (see Section 5). Therefore, we now explore the chemical evolution of the dwarfs.

We start by measuring the [Fe/H] and [O/Fe] radial abundance profiles for each model. While the [Fe/H] evolution is robust for different stellar yield models, the α-abundance, here tracked by [O/Fe], varies strongly with stellar yields and is in general used in a more qualitative[9] sense (as shown in Buck et al. 2021). We measure the gradients by fitting a linear function to the radial abundance profile before the dwarf's first pericenter, i.e., at $t = 1.4$ Gyr and $t = 0.9$ Gyr for dwarfs D1 and D2, respectively. At this time, the dwarf is still rotationally supported and has not been stirred by the tidal interactions (see Section 3.2). The measured gradients are presented in Table 1.

All the models have a significant radial metallicity gradient, where those with low feedback have steeper gradients compared to their high-feedback counterparts. This happens because stronger feedback suppresses star formation more efficiently, and thus delays the enrichment of the central parts of the dwarf, resulting in shallower metallicity gradients. We also measured the metallicity gradient just before the time of the merger and, as the satellite becomes more pressure supported (see Section 3.2), its gradient approaches ≈ -0.04 dex kpc^{-1} in all models, similar to what Naidu et al. (2021) estimated for GSE. Finally, we also verified that, when run in isolation up to $t = 10$ Gyr, the dwarf model's radial metallicity profile becomes shallow, ≈ -0.03 dex kpc^{-1}, in agreement with observations of local dwarfs (e.g., Kirby et al. 2013).

Similarly, we fit a linear function to the [O/Fe] radial abundance profile. The [O/Fe] radial gradient is shown in

---

[9] For instance, it is still possible to distinguish α-poor and α-rich stellar populations.
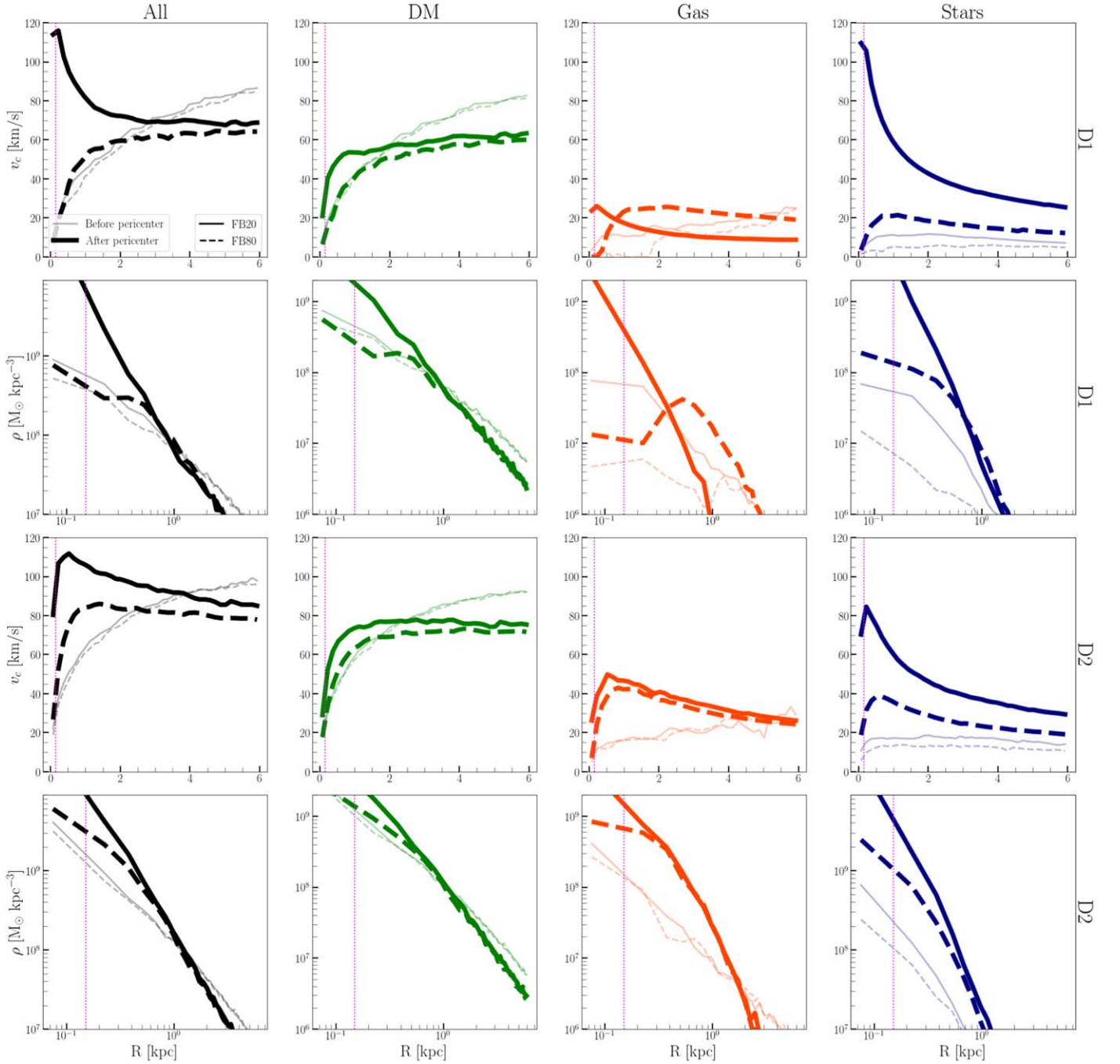
Figure 2. The circular velocity and mass density radial profiles for D1 (D2) are shown in the first (last) two rows. From left to right the total, dark matter, gas, and stellar mass profiles for the low- and high-feedback models are shown as solid and dashed lines, respectively. Before the first pericenter passage, the profiles for different feedback models, shown as the lighter colors, do not show significant difference. After the first pericenter and before the dwarf is completely disrupted, the low-feedback models have a more centrally concentrated mass profile, whereas the high-feedback models maintain a shallower distribution. The magenta dashed line delimits $r = 0.3$ kpc, which is three times the star-particle softening length.

Table 1. All the models exhibit a very mild gradient consistent with a flat [O/Fe] radial profile. We note, however, that the stars in the outskirts of the dwarfs have higher [O/Fe] in comparison to its central parts. This is due to the lower efficiency of star formation at larger radii, which inhibits star formation, leading to a slow chemical enrichment. We illustrate the chemical profiles of a dwarf in the first two rows of Figure 3. It shows a zoom in of the satellite in FB20_D2—all the dwarf models show similar trends—where each star particle is colored by its [Fe/H] and [O/Fe] (first and second rows, respectively).

We have described the structure and chemical evolution of the dwarfs up until just before they merge. In the following, we describe the accreted stars after the merger.
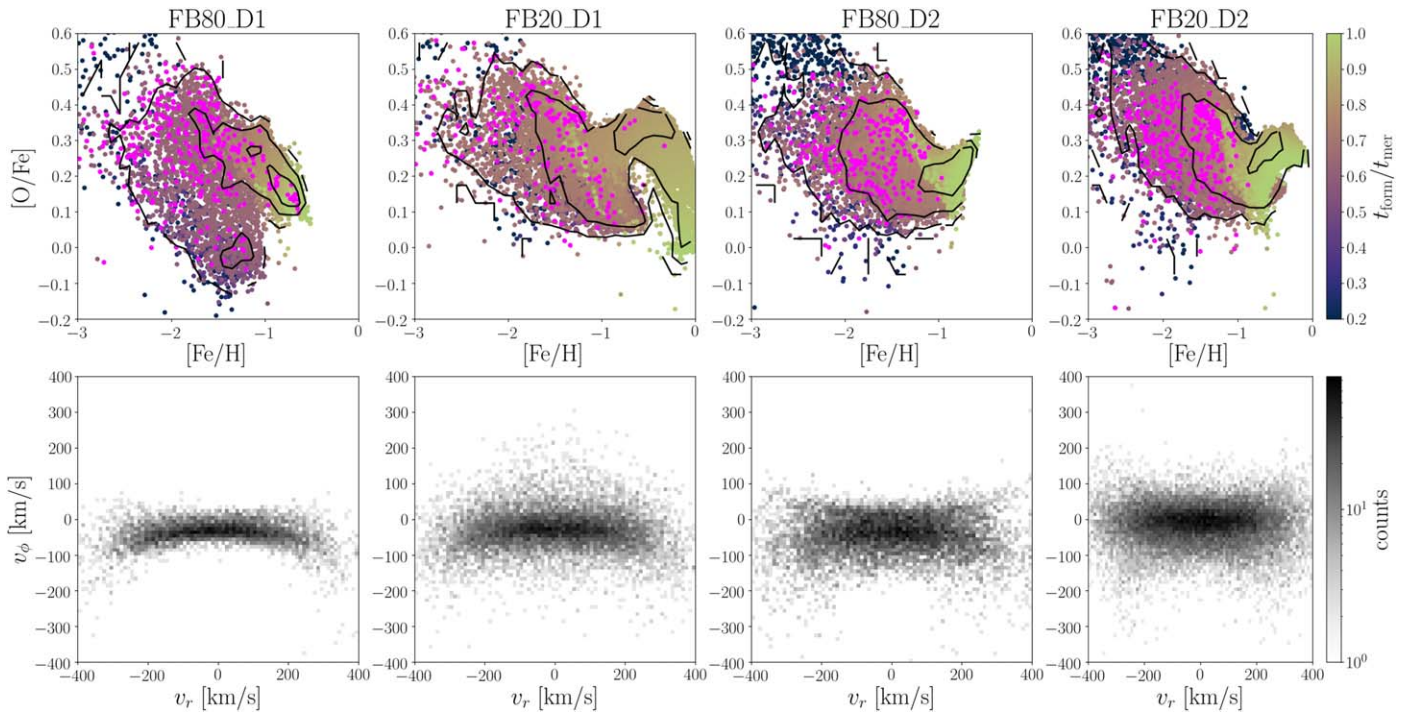
3.4. The Accreted Stars

Throughout the paper, we refer to "accreted stars" as those bound to the host galaxy potential at $t = 10$ Gyr originating
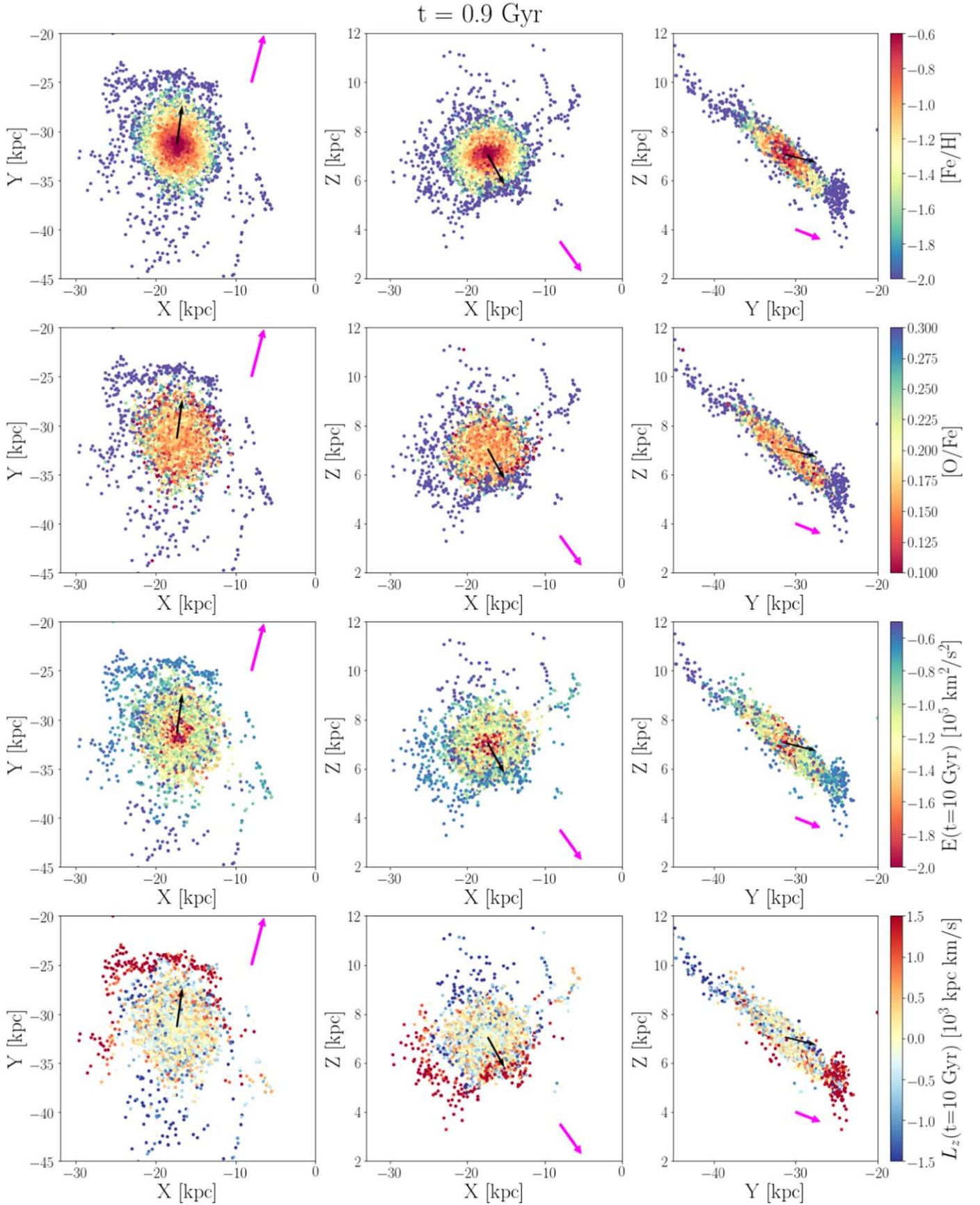
Figure 3. Zoom in on the satellite of FB20_D2 prior to its first pericenteric passage, at $t = 0.9$ Gyr. From top to bottom, rows are color-coded based on the stars' [Fe/H], [O/Fe], $E(t = 10$ Gyr), and $L_z(t = 10$ Gyr). The accreted stars that end up less bound to the host galaxy are those located on the outskirts of the dwarf. The dwarf center-of-mass velocity vector is shown as the black arrow. The magenta arrow points toward the host galaxy's center. An animation of the model's evolution can be seen here.

Figure 4. Top row: [Fe/H]–[O/Fe] plane for the accreted stars at $t = 10$ Gyr color-coded by the time of formation normalized by the dwarf's time of merger. The nontrivial chemical patterns of the dwarf evolution, e.g., the secondary younger sequence in the chemical abundance, are due to stars formed immediately before the second pericenter. The magenta points are the accreted stars with $L_z < -1.5 \times 10^3$ kpc km s^{-1} and spread over the same [Fe/H]–[O/Fe] space similar to what is observed in the MW stellar halo; see Section 5.2 for discussion. Bottom row: v_r–v_ϕ plane for the accreted stars at $t = 10$ Gyr. All the models have the generic "sausage" shape, characterized by a radially anisotropic velocity distribution, similar to the GSE merger in the MW. The larger radial velocity dispersion compared to the rotational velocity dispersion is the result of the merging dwarf's radial orbit. The lower-feedback models (second and fourth columns) present a larger spread in v_ϕ compared to higher-feedback models (first and third columns).

from the satellite. As the dwarfs in the models evolve, they not only have stars stripped by the host galaxy, but also form new stars after the first pericenter passage. Therefore, we flag as accreted stars those present in the dwarf immediately before its first pericenter passage, and also those born afterwards, up until the satellite disruption.

We start by analyzing their chemical abundances and in Table 1 we give the median values of [Fe/H] and [O/Fe] for all of our models. We also show, in the top row of Figure 4, the [O/Fe]–[Fe/H] plane for the accreted stars. Each star is colored by its formation time, $t_{\rm form}$, normalized by the time of the satellite's merger. The higher-density region, which is detached from the main chemical sequence, has the youngest stars in the dwarf, which are formed in the central starburst after the first pericenter passage (for details of this process, see Du et al. 2019). This effect is more intense in the low-feedback regime, where the stars can reach higher metallicities. Although these seem a rather unusual effect, we note the Large Magellanic Cloud also presents a "bump" in the [α/Fe]–[Fe/H] plane, which is associated with a burst in its star formation history (Nidever et al. 2020), and is also observed in MW in situ stars (prograde stars with ecc < 0.8, as seen in Conroy et al. 2022).

The models presented here are idealized to mimic the GSE merger event in the MW, in terms of the dwarf's mass, orbit, and time of merger. As expected, they reproduce the large radial anisotropy in the velocity distribution as observed in the MW. The bottom row of Figure 4 shows the distribution of the accreted stars in the v_r–v_ϕ plane. The radial orbit of the satellite prior to the merger and its relatively high stellar mass cause the final velocity distribution of the accreted stars to be very radial (as demonstrated by Amorisco 2017, with N-body-only

models). In fact, the "Sausage"-shape of the velocities distributions was a smoking gun to confirm that the inner stellar halo consists of stars mainly from a single massive progenitor (Belokurov et al. 2018), as several minor mergers are expected to fall with different radial directions assembling a more isotropic stellar halo.

The accreted component in FB20_D2 has a mean net zero rotational velocity, similarly to what is measured for the GSE (e.g., Belokurov et al. 2018; Helmi et al. 2018; Buder et al. 2022), but model FB80_D2, which has the same satellite orbit, shows a slight retrograde net motion. This is also the case for models with dwarf D1, with the stronger feedback leading to the most retrograde realization of the GSE. As shown in Section 3.2, the lower-feedback regime has a higher stellar density in its center, allowing it to be more efficiently radialized (also seen in the bottom row of Figure 1).

4. Present-day Chemodynamical Properties

We now explore the main chemodynamical trends of each model's host galaxy stellar halo, i.e., the accreted stars, after 10 Gyr of evolution. We calculate the gravitational potential and actions at 10 Gyr—assuming an axisymmetric system—and integrate orbits of star particles using AGAMA (Vasiliev 2019).

In our coordinate system, the vertical component of the angular momentum, L_z, represents prograde (retrograde) motion if $L_z > 0$ kpc km s^{-1} ($L_z < 0$ kpc km s^{-1}). Finally, throughout this section we also show a sample of chemically selected accreted stars from the Apache Point Observatory Galactic Evolution Experiment (APOGEE) Data Release 17
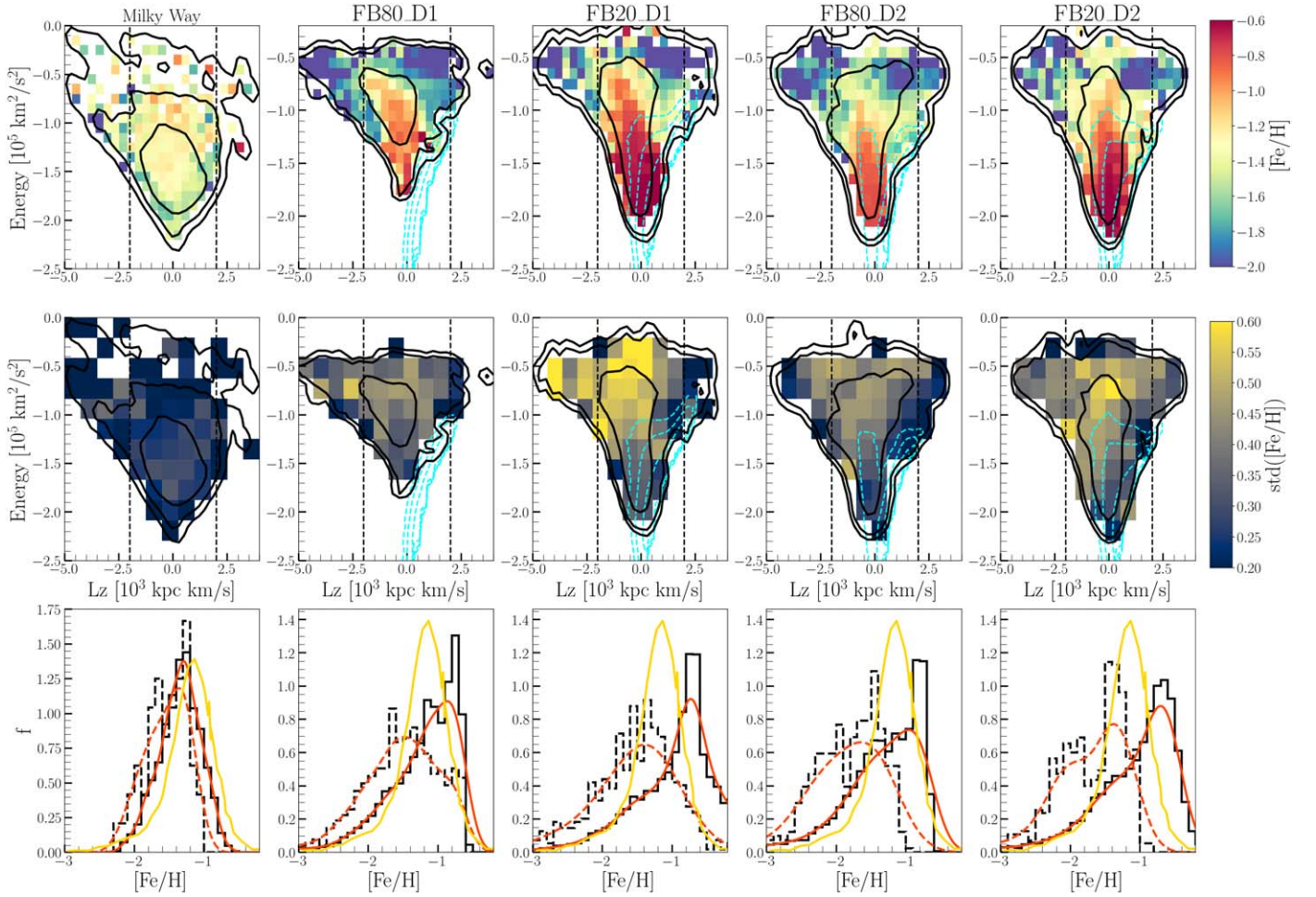
Figure 5. First row: properties of accreted stars in L_z-E space for each of the models at $t = 10$ Gyr. The log-density contours of the accreted stars and old stellar disk are shown in black and cyan lines, respectively. Each pixel on the plane is colored by the median [Fe/H] of the accreted stars. It shows how a single merger can spread in this plane with distinct chemical trends. Second row: same as the first row, but now colored by the standard deviation of [Fe/H] in each pixel. Third row: the [Fe/H] distribution of the accreted stars and its smoothed kernel-density estimator (KDE) version are shown as solid black and orange lines, respectively. For comparison, the MW observational data from the H3 survey (Naidu et al. 2020) is shown as the yellow line. The dashed black and orange lines show the distribution and its smoothed KDE version for accreted stars with very retrograde orbits, $L_z < -1.5 \times 10^3$ kpc km s^{-1}. See text for details.

(DR17; Abdurro'uf et al. 2022) to illustrate the general properties of the MW stellar halo in the chemodynamical spaces we explore with the models; we describe the data selection in Appendix A.

4.1. Energy–L_z Space

Energy, E, and L_z are integrals of motion in a time-independent axisymmetric potential. Thus, under the assumption that the host galaxy is not strongly perturbed after the merger, and as long as nonaxisymmetric features such as the bar and spiral arms can be neglected, E and L_z are approximately conserved. For this reason it is commonly said that the observed dynamic state of stars retains "memory of their origin" (e.g., Helmi 2020).

As each star will retain information about its origin, overdensities in the E–L_z plane may be associated to substructures with, in principle, distinct origins. In fact, this plane is commonly used to find substructures in the MW stellar halo (e.g., Helmi & de Zeeuw 2000; Koppelman et al. 2018; Horta et al. 2021). However, a single massive merger can spread over a large range of E and L_z (see e.g., Jean-Baptiste et al. 2017; Naidu et al. 2021). As we will see below, this is

clearly seen in our N-body+SPH GSE-like merger models, which also show similar chemical behavior.

The top panels of Figure 5 show the E–L_z plane of the accreted stars in the MW and in each of our models, at $t = 10$ Gyr, where each pixel is color-coded by the median metallicity. We opt to use the same [Fe/H] color bar scale for the MW and the models for a better direct comparison, although the models have a higher median metallicity compared to the MW, as we will discuss below. Black and cyan lines represent contours of the number-density distribution for all the accreted stars and old disk (stars born in the host galaxy within the first 3 Gyr), respectively. All models have a main overdensity of accreted stars concentrated at $L_z \approx 0$ kpc km s^{-1} and $E > -1.5 \times 10^5$ (-2×10^5) km^2 s^{-2} in the high-(low-)feedback models. Despite all the dwarfs having the same initial orbit circularity, $\eta = 0.3$, in Figure B1 we see that this overdensity is also present for a more circular initial orbit, $\eta = 0.5$. The most metal-rich stars are concentrated at $L_z \approx 0$ kpc km s^{-1}, and are more bound to the host galaxy's potential. On the other hand, the metal-poor stars are less bound to the potential and are often located on very retrograde orbits, $L_z < -2 \times 10^3$ kpc km s^{-1}, but some also have prograde motion, $L_z > 1 \times 10^3$ kpc km s^{-1}. In the MW, the accreted star

candidates in APOGEE also produce a main overdensity at $L_z \approx 0$ kpc km s^{-1}, but show a trend where the more bound stars are more metal-poor. These are stars associated with "Heracles," which we discuss in Section 5.3. The retrograde region, $L_z < -1.5 \times 10^3$ kpc km s^{-1}, associated with Sequoia (see Section 5.2) is overall at the more metal-poor regime.

A model's chemical trend is caused by the radial metallicity gradient present in the dwarf prior to the merger, as discussed in Section 3.3. We illustrate this behavior by returning to the zoom in of the dwarf in model FB20_D2 prior to its merger. The last two rows of Figure 3 show each star of the dwarf colored by its E and L_z at $t = 10$ Gyr. As discussed in Section 3, the most metal-rich stars are mainly located near the dwarf's center, and because they are more bound to the satellite's potential they are less easily stripped. These stars only become unbound in the very last stages of the merger, ending up being the most bound at $E(t=10) \approx -1.7 \times 10^5$ km^2 s^{-2}, with $\langle L_z(t=10) \rangle \approx 0$ kpc km s^{-1}.

On the other hand, the outskirts of the dwarf are populated by metal-poor stars. These get stripped during the first and second pericenter passages, which is why they end up being less bound to the galactic potential at $E > -1.2 \times 10^5$ km^2 s^{-2}, and with $|L_z| > 2 \times 10^3$ kpc km s^{-1}.

The features shown in the top row of Figure 5 demonstrate that a single massive merger does not have a unique "signature" in the chemodynamical space. For instance, all the models present a very retrograde region, $L_z < -1.5 \times 10^3$ kpc km s^{-1} and $E > -1 \times 10^5$ km^2 s^{-2}, with median [Fe/H] $\lesssim -1.8$ that could naively be misinterpreted as a distinct merger event from the main overdensity at $L_z \approx 0$ kpc km s^{-1} due to their distinct metallicities and orbital characteristics (see discussion in Section 5.2).

We can also see the effects of different feedback strengths. First, high feedback prevents the dwarf galaxy from sinking deeper into the potential compared with the low-feedback case. This happens because the high-feedback models have less dense centers, due to the reduced star formation (see Section 3.2), and thus are disrupted at larger radii. Second, and for similar reasons, low feedback allows the dwarf (and the progenitor) to have a faster metallicity enrichment. The faster increase of metallicity in the low-feedback regime also happens more intensively in the central parts of the dwarf (see Section 3.2), which produces a lower dispersion in [Fe/H] for the most bound accreted stars in these models, as shown in the second row of Figure 5. The difference in the efficiency of star formation is also seen when comparing the same model with different feedback strength. The lower-feedback-strength regime (FB20) has an overall higher [Fe/H] dispersion compared to the higher counterpart (FB80). This is also evident in the high-$|L_z|$ regions, although with a milder difference between the two feedback strengths. For instance, the [Fe/H] dispersion in the prograde/retrograde regions of FB20_D1 (FB20_D2) is higher by a factor of 0.03 (0.04) dex compared to FB80_D1 (FB80_D2). This is caused by the lower gas fraction in the satellite's outer regions (see Figure 2) where the majority of the high-$|L_z|$ stars originate from, as discussed previously in this section. Lastly, the old disk is hotter in the low-feedback-strength regime, i.e., it extends to lower L_z and higher E compared to the high-feedback-strength models, due to the presence of clumps in the progenitor during the first billion years of evolution.

The bottom panels of Figure 5 show the metallicity distribution function (MDF) of the accreted stars as solid histograms. The effect of feedback on the MDF of the dwarfs is again seen in the presence of an extended tail toward higher [Fe/H] in the MDF of the low feedback strength. These metal-rich stars are formed during the starburst at the final pericenter. Moreover, although the peaks of the MDFs in both feedback regimes are roughly similar, the low-feedback models have higher $\langle$[Fe/H]$\rangle$ compared to their high-feedback counterpart (see Table 1). We also show the [Fe/H] distribution of the H3 survey (Naidu et al. 2020) as the yellow line. Clearly, none of the models reproduces the exact MDF peak location for accreted stars in the MW; however, we note that the peak metallicity value is degenerate with the merger timescale. We stress that the goal of this work is not to find the exact match between the peak metallicity of a dwarf model and the GSE's, especially given the degeneracy between the time of the merger and the MDF.

Nonetheless, we can qualitatively explore trends in the models that are also in the MW, such as the MDF of the very retrograde stars, at $L_z < -1.5 \times 10^3$ kpc km s^{-1}. The H3 survey claims three distinct peaks for these stars which were associated with distinct substructures: Arjuna, Sequoia, I'itoi (Naidu et al. 2020). We show the MDF for the very retrograde stars in our models as the dashed histograms in the bottom panels of Figure 5. The very retrograde stars in all the models are dominated by the metal-poor accreted stars. In the bottom panel of Figure 3, we can see that these stars originated from the more metal-poor outer disk of the satellite, and were the first to be stripped, retaining some of their orbital angular momentum. We note that it is also tempting to visually identify "peaks" in the MDFs, especially for FB20_D2. However, if we either increase the bin size[10] or smooth the distribution with a kernel-density estimator (KDE, orange lines in the bottom panels of Figure 5), with the bandwidths chosen using Scott's rule (Scott 2015), the MDF shows a smoother behavior with a single peak and extended tail toward low [Fe/H] distribution. We also note that the MDF for the retrograde stars in APOGEE may also hint at visually identified peaks, but the smoothed KDE MDF simply demonstrates that these have more metal-poor stars compared to the whole distribution. We will discuss whether there is any statistical significance of these "peaks" in Section 5.2.

4.2. Eccentricity Distribution

We calculate the eccentricity as ecc $= (r_{\rm apo} - r_{\rm per})/(r_{\rm apo} + r_{\rm per})$, where $r_{\rm peri}$ and $r_{\rm apo}$ are the pericenter and apocenter in galactocentric spherical coordinates of the stellar orbit. In the MW GSE stars are found on very eccentric orbits, with ecc > 0.7, which is a consequence of the dwarf's initial radial orbit. For this reason, the eccentricity of stellar orbits is commonly used to select samples of GSE stars (e.g., Mackereth et al. 2019; Mackereth & Bovy 2020; Naidu et al. 2020). Besides that, Sequoia, which is in the very retrograde part of the accreted halo (e.g., Myeong et al. 2018a), also has less eccentric stellar orbits when compared with GSE. In this section we explore the chemical trend with the eccentricity of the accreted stars.

The first row of Figure 6 shows the L_z–ecc plane, where each pixel is colored by the median [Fe/H] of the accreted stars in

---

[10] The bin size is 0.1 dex, as in Naidu et al. (2020).
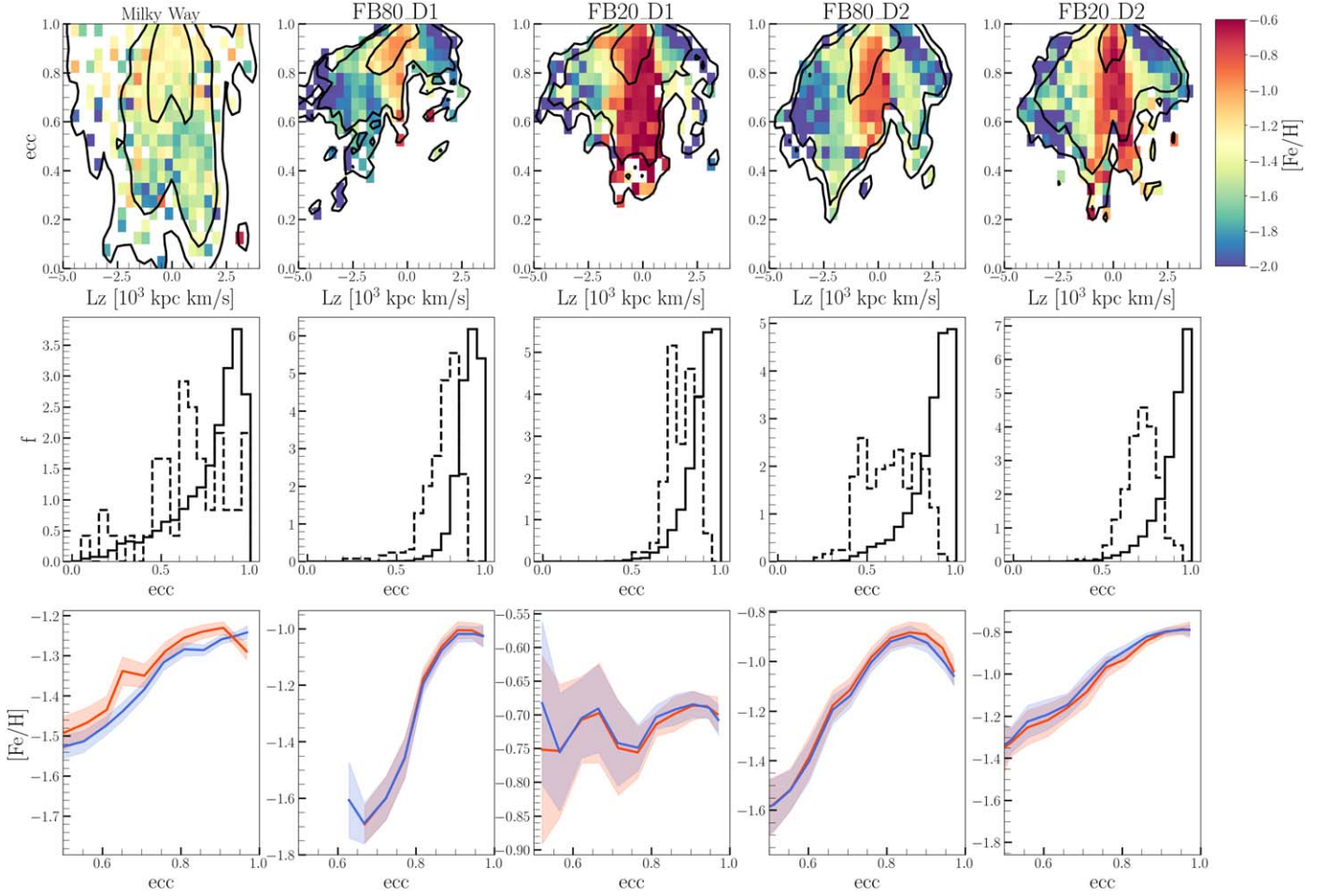
Figure 6. First row: properties of accreted stars in L_z–ecc space for each of our models at $t = 10$ Gyr. The log-density contours of the accreted stars are shown as black solid lines. Each pixel on the plane is colored by the median [Fe/H] of the accreted stars. Second row: the accreted stars' ecc density distribution. Solid and dashed lines show the distribution for all the stars and those with $L_z < -1.5 \times 10^3$ kpc km s^{-1}, respectively. While the bulk of stars have ecc > 0.7, the orbit for very retrograde stars are less eccentric, similar to what is observed in the MW. Third row: ecc–[Fe/H] relation for all the accreted stars and its retrograde ($L_z < 0$ kpc km s^{-1}) component within the solar neighborhood for each model, shown in blue and red, respectively. The shaded area corresponds to $\sigma/\sqrt{N}$, where σ and N are the [Fe/H] dispersion and total number of stars in each bin, respectively.

the MW and in the models (at $t = 10$ Gyr). The number-density distribution is shown as black solid lines. The chemically selected MW accreted star candidates have a main overdensity at ecc > 0.7 and a large tail toward lower eccentricities, possibly associated with other accretion events or with an ancient in situ population formed prior to the MW disk (Belokurov & Kravtsov 2022; Myeong et al. 2022). It is also evident that the lower-eccentricity stars, ecc < 0.7, are more metal-poor compared to the higher-eccentricity ones. In the models, the bulk of the distribution is centered at $L_z \approx 0$ kpc km s^{-1} and ecc > 0.8, which results from the radialization of the satellite's orbit during the merger process (see Section 3). Nonetheless, there is still a small but nonnegligible fraction of stars with ecc < 0.7.

The second row of Figure 6 shows the ecc distribution for all the accreted stars and those with $L_z < -1.5 \times 10^3$ kpc km s^{-1} as solid and dashed black lines, respectively. While the accreted stars have median ecc = 0.9 in all the models, those with $|L_z| > 1.5 \times 10^3$ kpc km s^{-1} tend to be on less eccentric orbits, with median ecc = 0.7, but with a significant tail toward ecc ≈ 0.5, despite being less bound to the galactic potential (notice, in the top row of Figure 5, that E tends to higher values for $|L_z| > 1.5 \times 10^3$ kpc km s^{-1}). This reflects the fact that the

stars in the outskirts of the dwarf are more rotationally supported, having $v_{\rm rot}/\sigma^* > 2$, and retain more of their angular momentum.

The third row of Figure 6 shows, for the accreted stars, the relation between the median [Fe/H] and their orbital eccentricity. In the MW, which is shown in the first column, [Fe/H] increases with stellar eccentricity (blue line). Kordopatis et al. (2020) have already shown this positive gradient for kinematically selected ($v_\phi < 0$ km s^{-1}) accreted star candidates. We further verify that the retrograde accreted stars in our chemically selected sample also have a positive gradient in the [Fe/H]–ecc relation, as shown by the red line.

For the models, we select the accreted stars within a mock solar neighborhood, i.e., $5 < R/{\rm kpc} < 12$ and $|z| < 3$ kpc, and show the relations for all the accreted stars and the retrograde ones, indicated by the blue and red lines, respectively. The shaded area represents the regions between the 16% and 84% percentiles of the [Fe/H] for a given eccentricity bin. Remarkably, for a given dwarf with the same initial orbital conditions, the feedback changes the [Fe/H]–ecc relation significantly. Both high-feedback models show the same general trend: an increase of [Fe/H] with ecc, until it appears to turn over at ecc ≈ 0.95. On the other hand, the low-feedback

models present different trends: FB20_D1, which ends the merger later, has a flat ecc–[Fe/H] relation, whereas FB20_D2 shows a positive trend over all the ecc range.

The difference in the ecc–[Fe/H] relation between different feedback models is again caused by the higher star formation rate in the low-feedback regime. As shown in Section 3.2, after the first pericenter the dwarf in the low-feedback models has a higher star formation rate compared to their high-feedback counterpart. This process forms more metal-rich stars in the dwarf's center and, as shown in Figure 3, these stars remain bound longer in the satellite and end up being the most radialized, thus having higher ecc. However, the very eccentric stars, ecc $\gtrsim 0.95$, are overall formed earlier (i.e., are older) compared to those with ecc < 0.95 and were more bound in the dwarf, causing the turnover in the ecc–[Fe/H] relation. This is enhanced in the high-feedback regime due to the more inefficient star formation compared to the low-feedback case, and the stars with ecc > 0.95 never reach higher [Fe/H] abundances.

4.3. Action Space

We now focus on the distribution of the merger debris in action space. In general, orbits in a nearly integrable gravitational potential can be described by three actions which are adiabatic invariants in a slowly varying potential (Binney & Tremaine 2008). The axisymmetry of the interpolated potential, as described in the beginning of this section, allows us to calculate the actions assuming the local Stäckel fudge approximation (Binney 2012). In such a gravitational potential, the action J_R measures the cylindrical radial excursion of an orbit, J_ϕ quantifies the circular motion and is equal to L_z, and J_z measures the vertical excursion of the orbit with respect to the galactic plane. These have been used to study the stars in the MW stellar halo (e.g., Posti et al. 2018; Feuillet et al. 2020, 2021; Lane et al. 2022). Thus in this section we explore the action space of the accreted stars in our models.

The top row of Figure 7 shows the L_z–J_R plane where each star particle is colored by its vertical action, J_z. Given the radial and low-inclination nature of the orbit of the merging dwarf, it is expected that most of the accreted stars will tend to have large values of J_R compared to J_z. Indeed, the accreted stars spread over a larger range of J_R, as seen in the second row of Figure 7, than what is typically used to define GSE in the MW ($J_R > 900$ kpc km s^{-1}; e.g., Feuillet et al. 2021). Indeed, our MW chemically selected accreted stars also spread across a large range of J_R.

Finally, the majority of accreted stars in our models have relatively low J_z, due to the low-inclination nature of the merger orbit, as seen in the third row of Figure 7. This is also observed in the MW accreted sample, where the majority of stars have $J_z \lesssim 800$ kpc km s^{-1}, in agreement with the models. The MW's tail toward higher J_z is due to the presence of stars originated from distinct merger event(s), such as the Helmi stream (Helmi et al. 1999; Yuan et al. 2020b). Finally, we note that, in the models, the accreted stars with $|L_z| > 2 \times 10^3$ kpc km s^{-1} also have typically low J_z, and are found with relatively high J_R, despite having on average lower eccentricity (as discussed in Section 4.2).

The fact that a merger like the GSE spreads over such a large range of J_R, and, as previously shown, in E and L_z, reinforces the need to be cautious when trying to find substructures in the MW halo. Nonetheless, the fact it is somewhat restricted in J_z

(which is also used as an extra selection criterion for GSE stars; e.g., Yuan et al. 2020b; Limberg et al. 2021), can help interpret the origins of some substructures in the MW, as will be discussed in the next section.

5. Implications for Milky Way Substructures

In the previous section we showed and studied the main chemodynamical properties of the stellar halos of several MW-like galaxies that went through a GSE-like merger during their first billion years. Now, we explore the implications of our results for the known substructures observed in the MW's stellar halo. Most importantly, we want to ask whether a single GSE-like merger can explain multiple substructures observed in the MW.

5.1. Gaia–Sausage–Enceladus

As shown in Section 3, the models presented are intended to reproduce general trends observed for GSE in the MW. While all models show a large radial velocity dispersion compared to the rotational velocity dispersion, we note that models with low feedback have a larger spread in v_ϕ compared to the high-feedback regime (see Figure 4). As shown throughout this study, stronger feedback inhibits star formation and thus there is a deficiency of stars at large radii in these models. The stars on the outskirts of the dwarf are the ones which keep more of their initial L_z and end up with $|L_z| \geqslant 1 \times 10^3$ kpc km s^{-1} (see Section 4.1), thus causing a larger spread in the v_ϕ distribution.

The MDF is another important constraint on the progenitor dwarf of GSE. However, observational measurements in the MW have not converged yet, as the GSE peak metallicity varies depending on the survey and the selection criteria for GSE star candidates (as pointed out in Section 1). In the models presented here, the final MDF is degenerate with both the time of the merger and the feedback model adopted in the model's subgrid physics. The idealized nature of the models allows us to adjust the orbit of the dwarf, which will lead to an early/late merger time, permitting us to fine-tune the MDF peak.

The [Fe/H]–ecc space also depends on the adopted feedback model, as shown in the third row of Figure 6 and discussed in Section 4.2. In the MW, Kordopatis et al. (2020) found a positive correlation for a sample of retrograde GSE candidate stars, similar to what we see in model FB20_D2. The other models either show a positive trend up to ecc ≈ 0.85, when the median [Fe/H] starts to decrease for higher ecc (FB80_D1 and FB80_D2), or a flat correlation (FB20_D1). These are related to feedback strength and the star formation history of the dwarfs (see Section 4.2). This suggests that the [Fe/H]–ecc correlation is a strong constraint on the initial conditions of the GSE merger and the feedback prescription.

5.2. Sequoia and Other Very Retrograde Substructures

Since the early 2000s, there has been growing evidence that very retrograde stars in the MW stellar halo are more α-element depleted compared to the bulk of the halo (Stephens & Boesgaard 2002; Venn et al. 2004). It was speculated that these stars originated from different accretion events during the MW stellar halo buildup. More recently, Myeong et al. (2018a, 2019) found evidence that the very retrograde high-energy (VRHE) stars are chemically distinct from the GSE and suggested they originated from a different progenitor, which they named Sequoia. The fact that Sequoia is more metal-poor
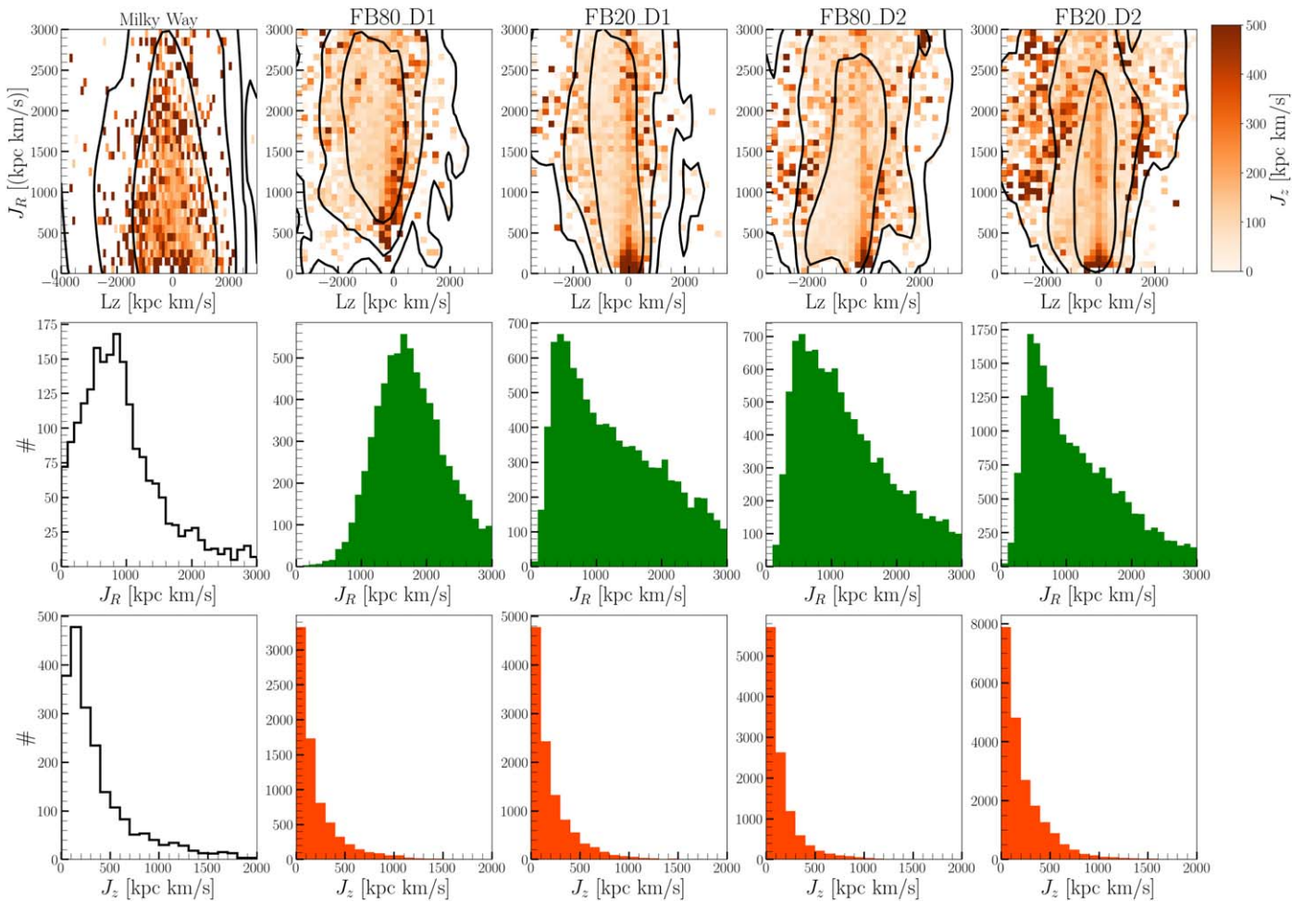
Figure 7. First row shows the L_z–J_R plane for the accreted stars at $t = 10$ Gyr. Each star is colored by its vertical action, J_z. Stars with net zero rotation spread over a large range of J_R and J_z, whereas the very retrograde ones, $L_z < 2 \times 10^3$ kpc km s^{-1}, tend to have a higher J_z component. We restrict ourselves to show stars with $J_R < 2000$ kpc km s^{-1} because of the scarcity of stars observed in the MW with high J_R values. The second and third rows show histograms of J_R and J_z, respectively. A GSE-like merger spreads over a large range in J_R while remaining relatively confined in J_z.

compared to the bulk of GSE stars has been confirmed by several other studies (Myeong et al. 2019; Ruiz-Lara et al. 2022). However, it is still debatable how Sequoia compares to GSE in terms of its α-element abundance. While Mackereth et al. (2019), Myeong et al. (2019), and Feuillet et al. (2021) all found the retrograde stars associated with Sequoia to be more α-rich compared to the bulk of the GSE, Matsuno et al. (2019), Monty et al. (2020), and Matsuno et al. (2021a) found evidence for Sequoia being more α-poor compared to GSE. Interestingly, Aguado et al. (2021), using high-resolution spectroscopy, found that Sequoia and GSE have rather similar abundances of neutron-capture elements, although there is a hint for a distinct [Ba/Fe] between them. Recently, Horta et al. (2022) and Limberg et al. (2022), using APOGEE DR17, showed that Sequoia and GSE are chemically indistinguishable from each other.

The [O/Fe]–[Fe/H] plane of the accreted stars in each model is shown in the top row of Figure 4. The VRHE stars in the models, defined as the accreted stars with $L_z < -1.5 \times 10^3$ kpc km s^{-1}, are shown as the magenta points. Similar to what is observed in the H3 survey (see Figure 14 in Naidu et al. 2020), they spread over a large range in this plane, and could be misinterpreted as originating from a different progenitor. As discussed in Section 3.3, the gradient present in the radial

abundance profile of the dwarfs prior to the merger is significant enough to create such intricate chemical patterns in the accreted halo.

One such pattern is the observed metallicity of the VRHE stars in the MW. These stars, which are, for instance, associated with the Sequoia, are more metal-poor compared to the GSE, and their MDF hints at the presence of three peaks, as observed with the H3 survey (see Figure 14 in Naidu et al. 2020). Due to this, Naidu et al. (2020) suggested the presence of two other substructures, Arjuna and I'itoi, besides the previously known Sequoia, each associated with the visually identified peaks. While all our models are consistent with the fact that the VRHE stars are more metal-poor compared to stars with $L_z \approx 0$ kpc km s^{-1} (Section 4.1), we further investigate whether any of them have a MDF with a statistically significant multiple-peak structure.

To do this, we use the test proposed by Silverman (1981),[11] in which there is no assumption on the form of the data probability distribution function, as opposed to the often-used Bayesian information criterion of a Gaussian mixture model (see a summary on the method in Beraldo e Silva et al. 2021). These tests reveal no significant evidence for the presence of

---

[11] The python code for the test is available at https://github.com/lberaldoesilva/silverman-test.

multiple peaks in the MDFs of VRHE stars in our models. This serves as a caution on the statistical significance of visually identified peaks. It remains to be seen if the peaks observed in the H3 survey are statistically significant and whether they correlate with the star formation history of the structures (in our case, we did not find any such correlation, as discussed in Section 4.1).

Another characteristic of Sequoia and other VRHE substructures that has been used to argue for distinct entities is the fact that they have lower eccentricity compared to the GSE (e.g., Myeong et al. 2019; Limberg et al. 2021). In Section 4.2 we showed that VRHE stars in our models are on less eccentric orbits compared to the bulk of the accreted stars (see Figure 6, second row). This is expected as these least-bound stars are the first to be stripped, retaining some of their orbital angular momentum, and thus end up on less eccentric orbits (as already shown in Naidu et al. 2021, with N-body-only models). The self-consistent chemistry of our models then demonstrates that a radial [Fe/H] gradient naturally appears in the merging dwarf, giving rise to the metal-poor nature of the VRHE stars.

The common origin of GSE, Sequoia, and perhaps other VRHE structures, then should result in a relatively small spread in the ages of the stars of each substructure. The mean age of GSE stars is found to be ≈10 Gyr (e.g., Kilic et al. 2019; Borre et al. 2022; Matsuno et al. 2021b; Montalbán et al. 2021). On the other hand, Sequoia stars are found to be slightly older than GSE stars (e.g., Feuillet et al. 2021) and with a hint of being accreted circa 0.3 Gyr earlier than GSE's (Kruijssen et al. 2020), but this small difference is within the error of the estimated values. Our models can naturally explain the slight difference in age and time of accretion measured in the literature. First, as shown in Section 4.1, the VRHE originated from the outskirts of the dwarf and these are the less-bound stars and thus are stripped first. Second, in Section 3.2 we show that, during the merger process, there is a burst of star formation mainly in the central parts of the dwarf. Consequently these relatively younger stars will end up in the region mainly associated with GSE, i.e., at $|L_z| < 0.5$ kpc km s^{-1}.

We can also constrain the relative mass fraction of VRHE substructures in comparison to GSE. Naidu et al. (2020) estimated the relative fraction of VRHE stars in comparison to GSE to be < 2.3%, i.e., much smaller than initially estimated (e.g., see Kruijssen et al. 2020). All our models show a relatively small fraction of VRHE compared to the rest of the accreted stars, with values ranging from 1.3% to 4.2% (see Table 1). We note that we are showing the overall contribution without taking into account possible effects of a survey-selection function.

We have demonstrated, with a set of GASTRO library models, that VRHE-accreted stars share the same qualitative properties as those observed in the MW: they are more [Fe/H]-poor and [α/Fe]-rich compared to the bulk of the accreted stars, but spread over a large range in the abundance plane (top row in Figure 4), have an apparent multipeak MDF, are on less eccentric orbits, and comprise similar mass fraction compared to GSE. We also note that the large spread in the chemical space is expected for relatively large dwarfs (see, e.g., Tolstoy et al. 2009, for a review). Thus, our results suggest that Sequoia, Arjuna, I'itoi, and other retrograde structures with these characteristics (e.g., Dodd et al. 2022; Oria et al. 2022) are naturally explained as stars originating from the outskirts of the GSE satellite. This was shown in Koppelman et al. (2020)

and Naidu et al. (2021) with N-body-only models; here we have demonstrated the same result with models where the satellite (and the disk) has a self-consistent metal enrichment.

5.3. Heracles

While most of the substructures are mainly observed at $R \geqslant 5$ kpc, the inner parts of our galaxy may also hide another significant merger event which happened in the first billion years of the MW. Evidence for such a merger was first seen as a low-energy group of globular clusters not associated with any known accreted substructure (Kruijssen et al. 2019; Massari et al. 2019; Forbes 2020).

Horta et al. (2021), using APOGEE DR16 and Gaia Data Release 2, found stronger evidence in the chemodynamical space of a merger remnant located in the inner part, $R < 5$ kpc, of the galaxy. They named this inner galaxy substructure Heracles and estimated its progenitor had a stellar mass of $\sim 5 \times 10^8 M_\odot$ (see also Kruijssen et al. 2020), i.e., as massive as the GSE merger event. More recently, Lane et al. (2022) showed that the APOGEE survey-selection function causes gaps in the E–L_z space. Given the galactocentric radii distribution of a stellar population, they will be constrained to a parabola in the E–L_z space, where stars at lower radii can reach lower E (see their Figure 6). This explains the presence of two groups, low and high energy, in the E–L_z plane of chemically defined accreted stars, as observed in Horta et al. (2021). Despite this dynamical argument of Heracles being the low-energy counterpart of GSE, one can still argue about the distinct chemistry of both structures. In the MW, the low-energy accreted stars, associated with Heracles, are on average more [Fe/H]-poor and [α/Fe]-rich compared to the high-energy accreted stars, associated with GSE (Horta et al. 2021; Naidu et al. 2022). Moreover, Heracles seems to have a flatter [α/Fe]–[Fe/H] relation compared to GSE stars.

In Section 4.1 we showed that the models' metal-rich accreted stars are located preferentially in the low-energy part of the E–L_z plane, i.e., the *opposite* of what is observed for Heracles. This result is independent of the dwarf and/or the merger impact parameter explored in the current paper. Therefore our results do not explain the [Fe/H]–energy relation observed for Heracles in the MW if it is associated with the GSE, as argued in Lane et al. (2022).

Given that the MW's accreted stars are clearly distinct in the [Mg/Mn]–[Al/Fe] plane from in situ stars (e.g., Das et al. 2020; Horta et al. 2021; Perottoni et al. 2021; Naidu et al. 2022)—this chemical selection of accreted stars also holds for the bulge region (Lucey et al. 2022)—we can conclude that a single merger event such as the GSE is not able to explain all the chemodynamical trends observed in the inner MW. Then, the origin of Heracles can be attributed to a distinct merger (as argued by, e.g., Kruijssen et al. 2019 and Horta et al. 2021), or as part of the in situ turbulent stellar system located in the inner part of the galaxy (Belokurov & Kravtsov 2022; Myeong et al. 2022). If Heracles was indeed accreted, such a merger event would imprint an increase in the star formation rate at $R < 5$ kpc (Orkney et al. 2022).

5.4. Other Substructures: a Cautionary Approach

There are several automated techniques used to search for substructures in the MW stellar halo (e.g., Koppelman et al. 2019; Yuan et al. 2020b; Limberg et al. 2021; Sofie Lövdal

et al. 2022; Malhan 2022), which are important to elucidate the accretion history of our galaxy. However, several substructures identified in the MW have a separation in energy, angular momentum, and [Fe/H] consistent with the single-merger-event models presented here. This calls into question the true nature of these substructures as distinct galactic entities.

We can use our fine-tuned GSE-like merger models to reconcile substructures into a single merger event. For instance, Malhan et al. (2022) identified Pontus as a new substructure in the dynamical space of a sample of globular clusters (Vasiliev & Baumgardt 2021). Limberg et al. (2021) also identified substructures, DTG-7, DTG-14, and DTG-18, in a sample of metal-poor halo stars with similar dynamical properties as Pontus. Despite being classified as distinct groups, Limberg et al. (2021) raised the possibility that these were part of GSE given their their low J_z and high orbital eccentricities.

Our results also suggest that Pontus, and consequently the other DTG groups, have the same origin as the GSE. For instance, both have $L_z \approx 0$ kpc km s^{-1} and $J_z < 500$ kpc km s^{-1}, and their difference in J_R is lower than ~ 1000 kpc km s^{-1}. These are within the expectations of a GSE-like merger, as shown in Figure 7, and reinforced by the fact that they also have similar chemical abundances (Malhan 2022; Malhan et al. 2022). Moreover, we verified in APOGEE DR17 that Pontus and GSE display a rather similar MDF, despite the latter having higher mean [Fe/H] (as shown in Malhan et al. 2022). It remains to be confirmed with more data whether the [Fe/H] difference is significant enough to rule out a common origin between Pontus and GSE, as suggested by our models.

Using the same line of reasoning as above, we can use our models to identify substructures which are *not* associated to the GSE merger event. For example, besides Heracles (see Section 5.3), we argue that LMS-1/Wukong (Naidu et al. 2020; Yuan et al. 2020b) and Cetus (Newberg et al. 2009; Yuan et al. 2019, 2022) also most likely have distinct progenitors other than the GSE. Malhan et al. (2022) showed that both substructures have $J_z \approx 2000$ kpc km s^{-1}, meaning they have large vertical excursions. Throughout this paper we have shown that a single GSE-like merger spreads over a large range in energy, L_z, and J_R, but in Figure 7 we show that the accreted stars will end up typically with $J_z < 1000$ kpc km s^{-1}, i.e., it cannot produce structures like LMS-1/Wukong and Cetus. Despite not being associated with GSE, we can question whether they do have a distinct progenitor, based on the similar arguments used for Pontus and GSE. Besides having similar J_z, the distance between LMS-1/Wukong and Cetus in L_z and J_R is less than ~ 600 kpc km s^{-1} and ~ 300 kpc km s^{-1}. This difference is well within the range of what a merger can produce (as shown in Figure 7). This argument is reinforced by the fact that they also have only a difference of 0.04 dex in [Fe/H] (see Figure 7 in Malhan 2022).

Finally, our models also predict the presence of a *prograde* metal-poor accreted population. This somewhat unexpected population would potentially be hard to distinguish from disk contaminants, such as the "Splash" population (Di Matteo et al. 2019; Belokurov et al. 2020), however the apocenters of its stars should be relatively high, $r_{\rm apo} \approx 25$ kpc, similar to those associated with its retrograde counterpart associated with Sequoia. In fact, we suggest Cand14/DTG-2 (Myeong et al. 2018b; Yuan et al. 2020b) as a candidate for being the prograde GSE remnant as it is on a prograde orbit, with $L_z \approx 1500$ kpc km s^{-1}, $J_R < 1000$ kpc km s^{-1}, $J_z < 300$ kpc km s^{-1}, and

ecc ≈ 0.6–0.8, and is more metal-poor, $\langle$[Fe/H]$\rangle \approx -1.45$, compared to the GSE peak metallicity.

6. Conclusion and Final Remarks

In this work, we have used a subset of the GASTRO library and explored the chemodynamical features of the stellar halo in models in which a GSE-like galaxy merges with a MW-like galaxy. We list below our main conclusions:

1. The mass density distribution in the dwarfs is more centrally concentrated in the low-feedback models after the first pericenter passage (Figure 2), allowing them to become more radialized, with $\langle L_z \rangle \approx 0$ kpc km s^{-1} (bottom row of Figure 1), in contrast to their high-feedback counterparts.

2. We confirm with a self-consistent metal-enrichment model the predictions from N-body-only models (under the assumption of an initial radial [Fe/H] gradient) that the accreted stars on low-energy orbits are more [Fe/H]-rich than those on high-energy orbits (top row of Figure 5).

3. VRHE stars are on average more [Fe/H]-poor and [O/Fe]-rich, but with a large spread in the chemical abundance plane, and on less eccentric orbits compared to the overall accreted stars distribution (bottom row of Figure 5 and second row of Figure 6), similar to what is observed in the MW. This confirms that Sequoia, Arjuna, and I'itoi are likely parts of the GSE merger event given its dynamical (as demonstrated in Koppelman et al. 2020 and Naidu et al. 2021) and chemical properties.

4. A GSE-like merger has a broad J_R distribution and a narrow one in J_z (Figure 7). Therefore, we suggest that Pontus (Malhan 2022) is most likely part of the GSE merger given its small spread in energy and action space, in addition to having the same [Fe/H] as GSE stars.

5. None of our GSE-like merger models can explain the chemistry of Heracles as part of the GSE merger event, as argued by Lane et al. (2022). This could give further credence to a scenario where an earlier massive merger event occurred in the lifetime of the galaxy (as in Horta et al. 2021 and Naidu et al. 2022), but it remains to be confirmed whether it has an accreted origin or is part of an old in situ population (Belokurov & Kravtsov 2022; Myeong et al. 2022).

6. With a similar line of reasoning, we also suggest that LMS-1/Wukong (Naidu et al. 2020; Yuan et al. 2020a) and Cetus (Malhan 2022) most likely share the same progenitor, but different from GSE, as both have a small spread in action space and have the same [Fe/H]. The high value of their J_z suggests their progenitor had an orbit with a high inclination angle with respect to the Galactic plane, as opposed to the low J_z values of the accreted stars in the low-inclination mergers shown here.

7. We also predict the presence of prograde substructures, with relatively low J_z, ecc $\approx 0.6 - 0.8$, and which are more metal-poor compared to the GSE, that originated from the outskirts of the merging satellite. A potential candidate for the prograde remnant of GSE is Cand14/DTG-2 (Myeong et al. 2018b; Yuan et al. 2020b).

The self-consistent star formation and chemical enrichment of the models presented here provide insights into the structural evolution of the satellite during the merger and is a step further

in complexity compared to pure *N*-body models. This clarifies how a single GSE-like merger is able to produce a rich chemodynamical space, thus helping to reconcile some known substructures to a single merger event. It is evident that a single merger event will not explain all the MW stellar halo substructures, but it serves as a cautionary tale when identifying potentially new substructures. Moreover, hydro-dynamical cosmological simulations show that MW-like galaxies have their inner ($r < 20$ kpc) stellar halo dominated by roughly three massive mergers, and low-mass accreted satellites only start to dominate the stellar halo fraction in the halo outskirts (Fattahi et al. 2020). These can set constraints on the probability of finding accreted structures from distinct mergers, and stress the need for a validation of substructure-finding algorithms using MW-like formation models.

Appendix A
APOGEE DR17 Data Selection

We use the cross-match of APOGEE DR17 (Abdurro'uf et al. 2022) with Gaia EDR3 (Gaia Collaboration et al. 2016, 2021) to select accreted star candidates based on their chemical abundances, derived from the APOGEE Stellar Parameters and Abundances Pipeline. We select only stars without flags, making use of ASPCAPFLAG (García Pérez et al. 2016).[12] We select stars with [Al/Fe] < -0.1 and [Mg/Mn] >0.2, as the [Al/Fe]–[Mg/Mn] diagram has been shown to efficiently segregate ex situ from in situ stars in the MW (as explored, for e.g., in Hawkins et al. 2015, Das et al. 2020, Feuillet et al. 2021, and Perottoni et al. 2021). We also require the uncertainty in Al, Mg, Mn, and Fe to be <0.2 dex. We opt to use a red giant sample, i.e., stars with surface gravity $\log g < 3$, and remove stars with invalid estimations of temperatures and $\log g$. Finally, we select stars with $\varpi/\varpi_{\rm err} > 5$; ϖ and $\varpi_{\rm err}$ are the parallax and its error, respectively, in order to have reliable estimates of their orbital parameters.

We calculate the position and velocities of the stars assuming the solar galactocentric distance $R_\odot = 8.27$ kpc and local circular velocity $v_c = 238$ km s^{-1} (Schönrich 2012). We assume the solar motion with respect to the local standard of rest to be $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25)$ km s^{-1} (Schönrich et al. 2010). Finally, the orbits and actions were calculated with AGAMA (details in Section 4) with the gravitational potential described in McMillan (2017).

Appendix B
Energy–L_z of Additional Models

Figure B1 presents four extra models and their [Fe/H] features in the energy–L_z plane. The second and third rows show models FB80_D2 and FB20_D2, which have the same initial conditions as the models presented in the main text but the satellite has an initial orbit circularity of $\eta = 0.5$. We also show two extra models with a dwarf D3, described in Section 2.3. The dwarfs in these models start at $r_{gc} = 120$ kpc and they are run in the low-feedback regime. The fourth and fifth columns of Figure B1 show the cases where the initial orbit circularity of the satellite is $\eta = 0.3$ and $\eta = 0.5$, respectively. This shows that the E–L_z space is qualitatively similar for different initial orbital circularity of the satellite.

---

[12] See examples in https://www.sdss.org/dr17/irspec/catalogs/.
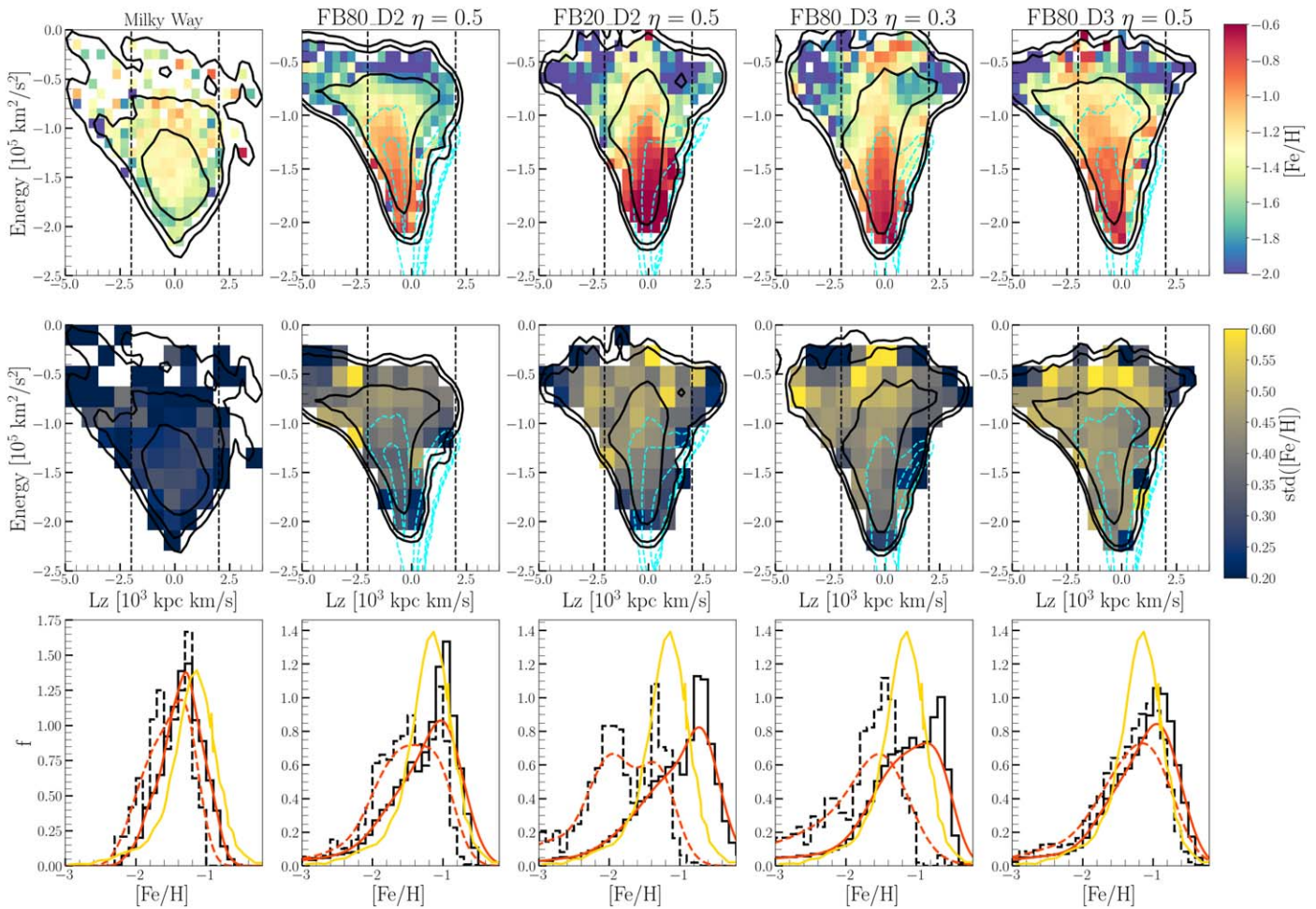
Figure B1. Same as Figure 5 but for models FB20_D2 and FB80_D2 with $\eta = 0.5$ and FB80_D3 with $\eta = 0.3$ and 0.5.

ORCID iDs

João A. S. Amarante ⓘ https://orcid.org/0000-0002-7662-5475
Victor P. Debattista ⓘ https://orcid.org/0000-0001-7902-0116
Leandro Beraldo E Silva ⓘ https://orcid.org/0000-0002-0740-1507
Chervin F. P. Laporte ⓘ https://orcid.org/0000-0003-3922-7336
Nathan Deg ⓘ https://orcid.org/0000-0003-3523-7633

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Agertz, O., Teyssier, R., & Moore, B. 2009, MNRAS, 397, L64
Aguado, D. S., Belokurov, V., Myeong, G. C., et al. 2021, ApJL, 908, L8
Amarante, J. A. S., Beraldo e Silva, L., Debattista, V. P., & Smith, M. C. 2020a, ApJL, 891, L30
Amarante, J. A. S., Smith, M. C., & Boeche, C. 2020b, MNRAS, 492, 3816
Amorisco, N. C. 2017, MNRAS, 464, 2882
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., & Kravtsov, A. 2022, MNRAS, 514, 689
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJL, 642, L137
Benson, A. J., Cole, S., Frenk, C. S., Baugh, C. M., & Lacey, C. G. 2000, MNRAS, 311, 793
Beraldo e Silva, L., Debattista, V. P., Khachaturyants, T., & Nidever, D. 2020, MNRAS, 492, 4716
Beraldo e Silva, L., Debattista, V. P., Nidever, D., Amarante, J. A. S., & Garver, B. 2021, MNRAS, 502, 260
Besla, G., Kallivayalil, N., Hernquist, L., et al. 2007, ApJ, 668, 949
Bignone, L. A., Helmi, A., & Tissera, P. B. 2019, ApJL, 883, L5
Binney, J. 2012, MNRAS, 426, 1324
Binney, J., & Tremaine, S. 2008, Galactic Dynamics (2nd edn.; Princeton, NJ: Princeton Univ. Press)
Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27
Bonifacio, P., Monaco, L., Salvadori, S., et al. 2021, A&A, 651, A79
Borre, C. C., Aguirre Børsen-Koch, V., Helmi, A., et al. 2022, MNRAS, 514, 2527
Buck, T., Rybizki, J., Buder, S., et al. 2021, MNRAS, 508, 3365
Buder, S., Lind, K., Ness, M. K., et al. 2022, MNRAS, 510, 2407
Bullock, J. S., Dekel, A., Kolatt, T. S., et al. 2001, ApJ, 555, 240
Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Clarke, A. J., Debattista, V. P., Nidever, D. L., et al. 2019, MNRAS, 484, 3476
Conroy, C., Weinberg, D. H., Naidu, R. P., et al. 2022, arXiv:2204.02989
Das, P., Hawkins, K., & Jofré, P. 2020, MNRAS, 493, 5195
Deason, A. J., Belokurov, V., Evans, N. W., & Johnston, K. V. 2013, ApJ, 763, 113
Deg, N., Widrow, L. M., Randriamampandry, T., & Carignan, C. 2019, MNRAS, 486, 5391
Di Matteo, P., Haywood, M., Lehnert, M. D., et al. 2019, A&A, 632, A4
Dillamore, A. M., Belokurov, V., Font, A. S., & McCarthy, I. G. 2022, MNRAS, 513, 1867
Dodd, E., Callingham, T. M., Helmi, A., et al. 2022, arXiv:2206.11248
Donlon, T. I., Newberg, H. J., Sanderson, R., & Widrow, L. M. 2020, ApJ, 902, 119
Du, M., Debattista, V. P., Ho, L. C., et al. 2019, ApJ, 875, 58
Elias, L. M., Sales, L. V., Helmi, A., & Hernquist, L. 2020, MNRAS, 495, 29

Fattahi, A., Belokurov, V., Deason, A. J., et al. 2019, MNRAS, 484, 4471
Fattahi, A., Deason, A. J., Frenk, C. S., et al. 2020, MNRAS, 497, 4459
Faucher-Giguère, C.-A., Quataert, E., & Hopkins, P. F. 2013, MNRAS, 433, 1970
Fernández-Alvar, E., Carigi, L., Schuster, W. J., et al. 2018, ApJ, 852, 50
Feuillet, D. K., Feltzing, S., Sahlholdt, C. L., & Casagrande, L. 2020, MNRAS, 497, 109
Feuillet, D. K., Sahlholdt, C. L., Feltzing, S., & Casagrande, L. 2021, MNRAS, 508, 1489
Forbes, D. A. 2020, MNRAS, 493, 847
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018, A&A, 616, A10
Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, NatAs, 3, 932
García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144
Gilmore, G., Wyse, R. F. G., & Norris, J. E. 2002, ApJL, 574, L39
Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Grand, R. J. J., Kawata, D., Belokurov, V., et al. 2020, MNRAS, 497, 1603
Hasselquist, S., Hayes, C. R., Lian, J., et al. 2021, ApJ, 923, 172
Hawkins, K., Jofré, P., Masseron, T., & Gilmore, G. 2015, MNRAS, 453, 758
Helmi, A. 2020, ARA&A, 58, 205
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Natur, 563, 85
Helmi, A., & de Zeeuw, P. T. 2000, MNRAS, 319, 657
Helmi, A., & White, S. D. M. 1999, MNRAS, 307, 495
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Natur, 402, 53
Hopkins, P. F., Kereš, D., Murray, N., Quataert, E., & Hernquist, L. 2012, MNRAS, 427, 968
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2021, MNRAS, 500, 1385
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2022, arXiv:2204.04233
Ibata, R., Malhan, K., Martin, N., et al. 2021, ApJ, 914, 123
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Natur, 370, 194
Jean-Baptiste, I., Di Matteo, P., Haywood, M., et al. 2017, A&A, 604, A106
Johnston, K. V., Hernquist, L., & Bolte, M. 1996, ApJ, 465, 278
Johnston, K. V., Spergel, D. N., & Hernquist, L. 1995, ApJ, 451, 598
Kallivayalil, N., van der Marel, R. P., Alcock, C., et al. 2006, ApJ, 638, 772
Kazantzidis, S., Łokas, E. L., Callegari, S., Mayer, L., & Moustakas, L. A. 2011, ApJ, 726, 98
Kilic, M., Bergeron, P., Dame, K., et al. 2019, MNRAS, 482, 965
Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102
Koppelman, H., Helmi, A., & Veljanoski, J. 2018, ApJL, 860, L11
Koppelman, H. H., Bos, R. O. Y., & Helmi, A. 2020, A&A, 642, L18
Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019, A&A, 631, L9
Kordopatis, G., Recio-Blanco, A., Schultheis, M., & Hill, V. 2020, A&A, 643, A69
Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019, MNRAS, 486, 3180
Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, MNRAS, 498, 2472
Kuijken, K., & Dubinski, J. 1995, MNRAS, 277, 1341
Lane, J. M. M., Bovy, J., & Mackereth, J. T. 2022, MNRAS, 510, 5119
Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019, MNRAS, 485, 3134
Limberg, G., Souza, S. O., Pérez-Villegas, A., et al. 2022, ApJ, 935, 109
Limberg, G., Rossi, S., Beers, T. C., et al. 2021, ApJ, 907, 10
Lin, L., Koo, D. C., Willmer, C. N. A., et al. 2004, ApJL, 617, L9
Lucey, M., Hawkins, K., Ness, M., et al. 2022, MNRAS, 509, 122
Mackereth, J. T., & Bovy, J. 2020, MNRAS, 492, 3631
Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, MNRAS, 482, 3426
Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, ApJ, 599, 1082
Malhan, K. 2022, ApJL, 930, L9
Malhan, K., Ibata, R. A., Sharma, S., et al. 2022, ApJ, 926, 107
Massari, D., Koppelman, H. H., & Helmi, A. 2019, A&A, 630, L4

Matsuno, T., Aoki, W., & Suda, T. 2019, ApJL, 874, L35
Matsuno, T., Hirai, Y., Tarumi, Y., et al. 2021a, A&A, 650, A110
Matsuno, T., Aoki, W., Casagrande, L., et al. 2021b, ApJ, 912, 72
McMillan, P. J. 2017, MNRAS, 465, 76
Meza, A., Navarro, J. F., Abadi, M. G., & Steinmetz, M. 2005, MNRAS, 359, 93
Miller, G. E., & Scalo, J. M. 1979, ApJS, 41, 513
Montalbán, J., Mackereth, J. T., Miglio, A., et al. 2021, NatAs, 5, 640
Monty, S., Venn, K. A., Lane, J. M. M., Lokhorst, D., & Yong, D. 2020, MNRAS, 497, 1236
Moore, B., Ghigna, S., Governato, F., et al. 1999, ApJL, 524, L19
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Myeong, G. C., Belokurov, V., Aguado, D. S., et al. 2022, arXiv:2206.07744
Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018a, ApJL, 856, L26
Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018b, MNRAS, 478, 5449
Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, MNRAS, 488, 1235
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2021, ApJ, 923, 92
Naidu, R. P., Ji, A. P., Conroy, C., et al. 2022, ApJL, 926, L36
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Newberg, H. J., Yanny, B., & Willett, B. A. 2009, ApJL, 700, L61
Newton, O., Cautun, M., Jenkins, A., Frenk, C. S., & Helly, J. C. 2018, MNRAS, 479, 2853
Nidever, D. L., Hasselquist, S., Hayes, C. R., et al. 2020, ApJ, 895, 88
Nissen, P. E., & Schuster, W. J. 2010, A&A, 511, L10
Oria, P.-A., Tenachi, W., Ibata, R., et al. 2022, ApJL, 936, L3
Orkney, M. D. A., Laporte, C. F. P., Grand, R. J. J., et al. 2022, arXiv:2206.09246
Patton, D. R., Pritchet, C. J., Carlberg, R. G., et al. 2002, ApJ, 565, 208
Perottoni, H. D., Amarante, J. A. S., Limberg, G., et al. 2021, ApJL, 913, L3
Pilkington, K., Gibson, B. K., Brook, C. B., et al. 2012, MNRAS, 425, 969
Posti, L., Helmi, A., Veljanoski, J., & Breddels, M. A. 2018, A&A, 615, A70
Raiteri, C. M., Villata, M., & Navarro, J. F. 1996, A&A, 315, 105
Rey, M. P., & Starkenburg, T. K. 2022, MNRAS, 510, 4208
Ruiz-Lara, T., Matsuno, T., Sofie Lövdal, S., et al. 2022, arXiv:2201.02405
Schönrich, R. 2012, MNRAS, 427, 274
Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829
Scott, D. W. 2015, Multivariate Density Estimation: Theory, Practice, and Visualization (Hoboken, NJ: John Wiley & Sons)
Searle, L., & Zinn, R. 1978, ApJ, 225, 357
Shen, S., Wadsley, J., & Stinson, G. 2010, MNRAS, 407, 1581
Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, ApJ, 862, 114
Silverman, B. W. 1981, J. R. Stat. Soc, 43, 97
Sofie Lövdal, S., Ruiz-Lara, T., Koppelman, H. H., et al. 2022, arXiv:2201.02404
Springel, V. 2010, ARA&A, 48, 391
Stadel, J. G. 2001, PhD thesis, Univ. Washington
Stephens, A., & Boesgaard, A. M. 2002, AJ, 123, 1647
Stinson, G., Seth, A., Katz, N., et al. 2006, MNRAS, 373, 1074
Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371
Vasiliev, E. 2019, MNRAS, 482, 1525
Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
Vasiliev, E., Belokurov, V., & Evans, N. W. 2022, ApJ, 926, 203
Venn, K. A., Irwin, M., Shetrone, M. D., et al. 2004, AJ, 128, 1177
Vincenzo, F., Spitoni, E., Calura, F., et al. 2019, MNRAS, 487, L47
Wadsley, J., Stadel, J., & Quinn, T. 2004, NewA, 9, 137
Widrow, L. M., & Dubinski, J. 2005, ApJ, 631, 838
Widrow, L. M., Pym, B., & Dubinski, J. 2008, ApJ, 679, 1239
Wu, Y., Valluri, M., Panithanpaisal, N., et al. 2022, MNRAS, 509, 5882
Yuan, Z., Chang, J., Beers, T. C., & Huang, Y. 2020a, ApJL, 898, L37
Yuan, Z., Smith, M. C., Xue, X.-X., et al. 2019, ApJ, 881, 164
Yuan, Z., Myeong, G. C., Beers, T. C., et al. 2020b, ApJ, 891, 39
Yuan, Z., Malhan, K., Sestito, F., et al. 2022, ApJ, 930, 103