CrossMark

Reconstructing the Last Major Merger of the Milky Way with the H3 Survey

Rohan P. Naidu[1] , Charlie Conroy[1] , Ana Bonaca[1] , Dennis Zaritsky[2] , Rainer Weinberger[1] , Yuan-Sen Ting
(丁源森)[3,4,5,6,10] , Nelson Caldwell[1] , Sandro Tacchella[1] , Jiwon Jesse Han[1], Joshua S. Speagle[7,8,9] , and
Phillip A. Cargile[1]

[1] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA; rohan.naidu@cfa.harvard.edu
[2] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA
[3] Institute for Advanced Study, Princeton, NJ 08540, USA
[4] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[5] Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[6] Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, Cotter Road, Weston Creek, ACT 2611, Canberra, Australia
[7] University of Toronto, Department of Statistical Sciences, Toronto, ON M5S 3G3, Canada
[8] University of Toronto, David A. Dunlap Department of Astronomy & Astrophysics, Toronto, ON M5S 3H4, Canada
[9] Dunlap Institute for Astronomy & Astrophysics, Toronto, ON M5S 3H4, Canada

Abstract

Several lines of evidence suggest that the Milky Way underwent a major merger at $z \sim 2$ with the Gaia-Sausage-Enceladus (GSE) galaxy. Here we use H3 Survey data to argue that GSE entered the Galaxy on a retrograde orbit based on a population of highly retrograde stars with chemistry similar to the largely radial GSE debris. We present the first tailored N-body simulations of the merger. From a grid of ≈ 500 simulations we find that a GSE with $M_\star = 5 \times 10^8\, M_\odot$, $M_{\mathrm{DM}} = 2 \times 10^{11}\, M_\odot$ best matches the H3 data. This simulation shows that the retrograde stars are stripped from GSE's outer disk early in the merger. Despite being selected purely on angular momenta and radial distributions, this simulation reproduces and explains the following phenomena: (i) the triaxial shape of the inner halo, whose major axis is at $\approx 35°$ to the plane and connects GSE's apocenters; (ii) the Hercules-Aquila Cloud and the Virgo Overdensity, which arise due to apocenter pileup; and (iii) the 2 Gyr lag between the quenching of GSE and the truncation of the age distribution of the in situ halo, which tracks the lag between the first and final GSE pericenters. We make the following predictions: (i) the inner halo has a "double-break" density profile with breaks at both ≈ 15–18 kpc and 30 kpc, coincident with the GSE apocenters; and (ii) the outer halo has retrograde streams awaiting discovery at >30 kpc that contain $\approx 10\%$ of GSE's stars. The retrograde (radial) GSE debris originates from its outer (inner) disk—exploiting this trend, we reconstruct the stellar metallicity gradient of GSE (-0.04 ± 0.01 dex r_{50}^{-1}). These simulations imply that GSE delivered $\approx 20\%$ of the Milky Way's present-day dark matter and $\approx 50\%$ of its stellar halo.

Unified Astronomy Thesaurus concepts: Galaxy stellar halos (598); Halo stars (699); Galaxy dark matter halos (1880); Milky Way stellar halo (1060); Milky Way dark matter halo (1049); Galaxy mergers (608); Metallicity (1031); Galaxy evolution (594); Milky Way evolution (1052); Milky Way Galaxy physics (1056); Milky Way formation (1053); Milky Way Galaxy (1054)

1. Introduction

A hallmark feature of ΛCDM cosmology is hierarchical assembly, in which galaxies continually assimilate smaller systems (e.g., White & Frenk 1991). Nowhere in the universe do we have a clearer view of this hierarchical buildup than in the stellar halo of the Milky Way (MW). At this very moment the Sagittarius dwarf galaxy is being tidally disrupted (e.g., Ibata et al. 1994), the Magellanic Clouds are on first infall (e.g., Besla et al. 2007), and dozens of globular cluster streams encircle the Galaxy (e.g., Bonaca et al. 2021).

While these ongoing mergers are apparent on the sky, the record of even more mixed, ancient mergers can be extracted from the stellar halo. Due to the long relaxation time in the halo, stars that were accreted as part of the same galaxy can be connected through their shared integrals of motion (e.g., angular momenta, energies) even several Gyr after their arrival (e.g., Helmi & de Zeeuw 2000; Font et al. 2011; Simpson et al. 2019). We are also aided by the shared chemical abundance patterns expected of stars born in the same system (e.g., Freeman & Bland-Hawthorn 2002; Venn et al. 2004; Lee et al.

2015). Integrals of motion and chemical information have recently been obtained for millions of stars in the solar neighborhood thanks to the Gaia mission (Gaia Collaboration et al. 2018) and stellar spectroscopic surveys such as APOGEE (Majewski et al. 2017), RAVE (Steinmetz et al. 2006), SEGUE (Yanny et al. 2009), LAMOST (Cui et al. 2012), GALAH (De Silva et al. 2015), and H3 (Conroy et al. 2019a). These data have allowed us to piece together the history of the Galaxy in unprecedented detail.

A single dwarf galaxy that merged with the MW at $z \approx 2$—Gaia-Sausage-Enceladus (GSE)—constitutes the bulk of the inner halo (e.g., Belokurov et al. 2018; Helmi et al. 2018; Naidu et al. 2020). The lines of evidence for this accretion event are numerous and compelling. Kinematics of halo stars show a preponderance of eccentric, radial orbits (e.g., Eggen et al. 1962; Chiba & Beers 2000; Koppelman et al. 2018; Mackereth et al. 2019; Carollo & Chiba 2021; Yuan et al. 2020; Limberg et al. 2021), exactly as expected for debris from a major merger that is radialized owing to dynamical friction (e.g., Amorisco 2017). The ages and abundances of these eccentric stars point to the same, ancient ($\gtrsim 8$–10 Gyr old) progenitor (e.g., Haywood et al. 2018; Gallart et al. 2019;

---

[10] Hubble Fellow.

Conroy et al. 2019b; Das et al. 2020; Feuillet et al. 2020; Bonaca et al. 2020; Gudin et al. 2021). A large number of MW globular clusters (≈20–30) are eccentric and clustered in the age–metallicity plane, suggesting that they accompanied GSE to the MW (e.g., Myeong et al. 2018; Kruijssen et al. 2019; Massari et al. 2019; Forbes 2020). A break in the halo density and anisotropy profiles at ≈25–30 kpc has been associated with an apocenter in the GSE orbit (e.g., Deason et al. 2018; Bird et al. 2019; Lancaster et al. 2019; Iorio & Belokurov 2021). These observations are supported by cosmological simulations that show that the inner halos of MW-like galaxies are often built out of a handful of massive progenitors and that a large fraction of debris from these mergers often ends up on eccentric orbits (e.g., Deason et al. 2015; Fattahi et al. 2019; Grand et al. 2020; Santistevan et al. 2020).

A fundamental open question is the configuration of the merger—did GSE collide with our Galaxy head-on, or was it on an initially circular orbit that decayed? Recovering the configuration hinges on whether GSE debris today is purely radial or it extends to orbits with significant angular momentum (e.g., Evans 2020; Helmi 2020; Koppelman et al. 2020). The configuration informs a variety of issues, e.g., the expected velocity and spatial distributions of the dark matter (DM) that arrived with GSE (≈20% of the MW's DM; see Section 5.1). The velocity distribution modulates the expected signal in DM detection experiments, and a significant nonradial component could influence efforts seeking directional signatures (e.g., O'Hare et al. 2018; Evans et al. 2019; Necib et al. 2019; O'Hare et al. 2020; Vahsen et al. 2020). Similarly, a merger configuration resulting in a nonplanar mass distribution that breaks axisymmetry would have important implications for the MW potential, particularly in the inner halo ($<30\,$kpc), which is essentially entirely composed of GSE (Naidu et al. 2020).

The debate around the radial or retrograde nature of GSE has largely relied on local halo samples that are limited to a few kiloparsecs from the Sun. However, the first stars stripped from GSE, which contain the most information about its initial orbit, are likely at larger distances and higher energies than stars that pass through the solar neighborhood. Capturing this early debris, which retains the most pristine memory of the merger configuration, requires forging beyond the local halo.

The H3 Stellar Spectroscopic Survey (Conroy et al. 2019a) is designed to study the distant halo. Combined with Gaia, H3 is measuring full 6D phase-space coordinates and chemical abundances for ≈200,000 stars at $r_{gal} \approx 3$–$100\,$kpc. Using these data, Naidu et al. (2020) presented a comprehensive inventory of structure in the halo out to 50 kpc, including the largest sample of GSE ($N = 2684$) stars with integrals of motion and abundances from high-resolution spectroscopy. This sample is unique in encompassing the farthest reaches of the merger and in being largely metallicity unbiased (unlike, e.g., RR Lyrae or blue horizontal branch stars (BHBs) or color-selected samples). In this work, we build on Naidu et al. (2020) to explore the retrograde halo with a view to chart the full extent of GSE.

Tailored simulations of the other significant MW mergers—Sagittarius (e.g., Law & Majewski 2010; Dierickx & Loeb 2017; Laporte et al. 2018) and the Magellanic Clouds (e.g., Besla et al. 2010; Garavito-Camargo et al. 2019; Vasiliev et al. 2021)—have proven crucial in interpreting phenomena across the Galaxy such as the phase-space spiral (e.g., Antoja et al. 2018; Bland-Hawthorn et al. 2019) and the reflex motion

of the outer halo (e.g., Erkal et al. 2021). However, the retrograde or radial nature of GSE is still unclear, and most existing constraints on the merger are derived from the local halo, inhibiting the production of a high-fidelity model. Consequently, GSE has been studied largely qualitatively via analogs in cosmological simulations (Bignone et al. 2019; Elias et al. 2020), MW zooms (Fattahi et al. 2019; Grand et al. 2020), and existing merger simulations (Helmi et al. 2018; Koppelman et al. 2020). These studies have been immensely successful in demonstrating how a major merger can produce eccentric debris and reshape the early MW disk. Equipped with constraints from the H3 Survey, we are well positioned to build on these results and produce a tailored model for the merger as has been done for Sagittarius and the LMC.

The plan for the paper is as follows. In Section 2 we argue that a subset of the retrograde halo stars are associated with the GSE merger. In Section 3 we summarize existing observational constraints on the merger. Section 4 describes the numerical simulations. Section 5 is based on the fiducial simulation—here we interpret the origin of GSE's highly retrograde debris (Section 5.2), the shape of the inner halo (Section 5.3), the all-sky distribution of GSE debris (Section 5.4), the inner halo density profile (Section 5.5), the time line of the GSE merger (Section 5.6), the net rotation of GSE (Section 5.7), and the relationship between GSE and other retrograde accreted galaxies (Section 5.8). In Section 6 we use our fiducial simulation to reconstruct the stellar metallicity gradient measurement in a $z \approx 2$ star-forming galaxy (GSE). A summary follows in Section 7.

We adopt a Planck Collaboration et al. (2020) cosmology. To describe central values of distributions, we generally report the median, along with 16th and 84th percentiles. We use r_{gal} to denote 3D Galactocentric distance, X_{gal}, Y_{gal}, Z_{gal} to denote Galactocentric Cartesian distances, and d_{helio} to refer to 3D heliocentric distance. We use V_r, V_ϕ, V_θ for velocities in a right-handed spherical coordinate system with origin at the Galactic center. Prograde stars have negative V_ϕ and L_z. Unless mentioned otherwise, total orbital energy (E_{tot}) is always reported in units of $10^5\,\mathrm{km^2\,s^{-2}}$ and angular momenta (L_x, L_y, L_z) in units of $10^3\,\mathrm{kpc\,km\,s^{-1}}$. These quantities are always computed in a Galactocentric frame tied to the center of the MW in both the data and the simulations.

2. Revealing the Full Extent of GSE

2.1. Data: The H3 Survey

The H3 Survey (Conroy et al. 2019a) is a high-latitude ($|b| > 30°$), high-resolution ($R = 32,000$) spectroscopic survey of the distant ($d_{helio} \approx 2$–$100\,$kpc) Galaxy. Targets are selected purely on their Gaia parallax ($\pi < 0.4$–0.5 mas, evolving with Gaia data releases), brightness ($15 < r < 18$), and observability (decl. $> -20°$) from the 6.5 m MMT in Arizona, USA. H3 is measuring radial velocities precise to $\lesssim 1\,\mathrm{km\,s^{-1}}$, [Fe/H] and [$\alpha$/Fe] abundances precise to $\lesssim 0.1$ dex, and spectrophotometric distances precise to $\lesssim 10\%$ (see Cargile et al. 2020 for details of the stellar parameter pipeline). Combined with Gaia proper motions (PMs), H3 thus provides the full 6D phase space and 2D chemical space for all stars in the sample.

Naidu et al. (2020) used the sample of H3 giants ($N = 5684$, $|b| > 40°$, $d_{helio} = 3$–$50\,$kpc) to assign almost the entire distant Galaxy to various structures (summarized in their Table 1). These authors systematically identified debris from known

accreted galaxies (e.g., Sgr, GSE, Sequoia), as well as new structures (e.g., I'itoi, Arjuna, Wukong). Pertinent to the matter at hand, they tagged ≈3000 stars as belonging to GSE or the high-energy retrograde halo that we focus on here. All quantities sourced from Naidu et al. (2020) (e.g., L_z and r_{gal} distributions) describe the $|b| > 40°$ Galaxy and have been corrected for the H3 selection function—in particular, for the survey magnitude limit and targeting algorithm (see their Section 2.3). We make no corrections for the window function and compare models and data only within the survey footprint. We use the Naidu et al. (2020) parameters derived from Gaia DR2—note that the distance errors dominate the error budget for quantities of interest, so the improved PMs from Gaia EDR3 make little difference to quantities like L_z studied here (see Appendix A of Naidu et al. 2020).

2.2. Arjuna as the Retrograde Debris of GSE

After excluding Sgr, the high-α disk, and the in situ halo, Naidu et al. (2020) attributed stars on highly eccentric orbits ($e > 0.7$) to GSE ($N = 2684$). This is essentially the head-on "Gaia-Sausage" from Belokurov et al. (2018). The resulting, well-sampled metallicity distribution function (MDF) is unimodal (median $[Fe/H] = -1.15$), is consistent with a simple chemical evolution model ("Best Accretion Model"; Lynden-Bell 1975; Kirby et al. 2011), and resembles the narrow MDFs of local dwarfs like Fornax and Leo I (Kirby et al. 2013).

The high-energy retrograde halo is defined in Naidu et al. (2020) as excluding GSE and by the following condition: $(\eta > 0.15) \wedge (L_z > 0.7) \wedge (E_{tot} > -1.25)$, where η is the orbital circularity computed as $L_z/|L_{z,max}(E_{tot})|$, where $L_{z,max}(E_{tot})$ is the maximum L_z achievable for an orbit of energy E_{tot}. This definition generously selects stars on retrograde orbits and excludes the Thamnos structure (Koppelman et al. 2019a) at lower energy. In Figure 1 we further limit the high-energy retrograde halo to $L_z > 1.5$ to make it clear that the radial locus of stars typically associated with GSE (distributed around $L_z = 0$) is not responsible for the features discussed below. Three chemical populations compose the high-energy, highly retrograde halo: Arjuna ($[Fe/H] \approx -1.2$), Sequoia ($[Fe/H] \approx -1.6$), and I'itoi ($[Fe/H] < -2$). We emphasize that these three chemical populations do not just occur along the margins of GSE in $E - L_z$, but extend to highly retrograde orbits. The Sequoia MDF peaks exactly where other studies have found it to peak (e.g., Matsuno et al. 2019; Monty et al. 2019; Myeong et al. 2019), and I'itoi is a distinct metal-poor population. As foreshadowed in Naidu et al. (2020), we argue here that the metal-rich Arjuna is the retrograde debris of GSE.

The Arjuna MDF closely tracks the GSE MDF (bottom panel of Figure 1), with a similar mode and similar mean metallicity but fewer metal-rich stars. In addition to this, the α abundances of GSE and Arjuna are virtually identical—median $[\alpha/Fe]$ of 0.21 and 0.24, respectively. Due to these similarities, we associate Arjuna with GSE. The Arjuna debris extends to very retrograde orbits ($L_z \approx 4$), is more distant (median $r_{gal} \approx 23$ kpc vs. 18 kpc for GSE), and is less eccentric ($e = 0.55$). In the sections that follow, we demonstrate through numerical simulations that Arjuna's properties are consistent with it being material from the outer regions of GSE. This material may have been stripped before the satellite was radialized and so retains the high retrograde angular momentum of the early orbit of GSE. Further, since this material is shed during the early phase of the merger, it has a larger mean
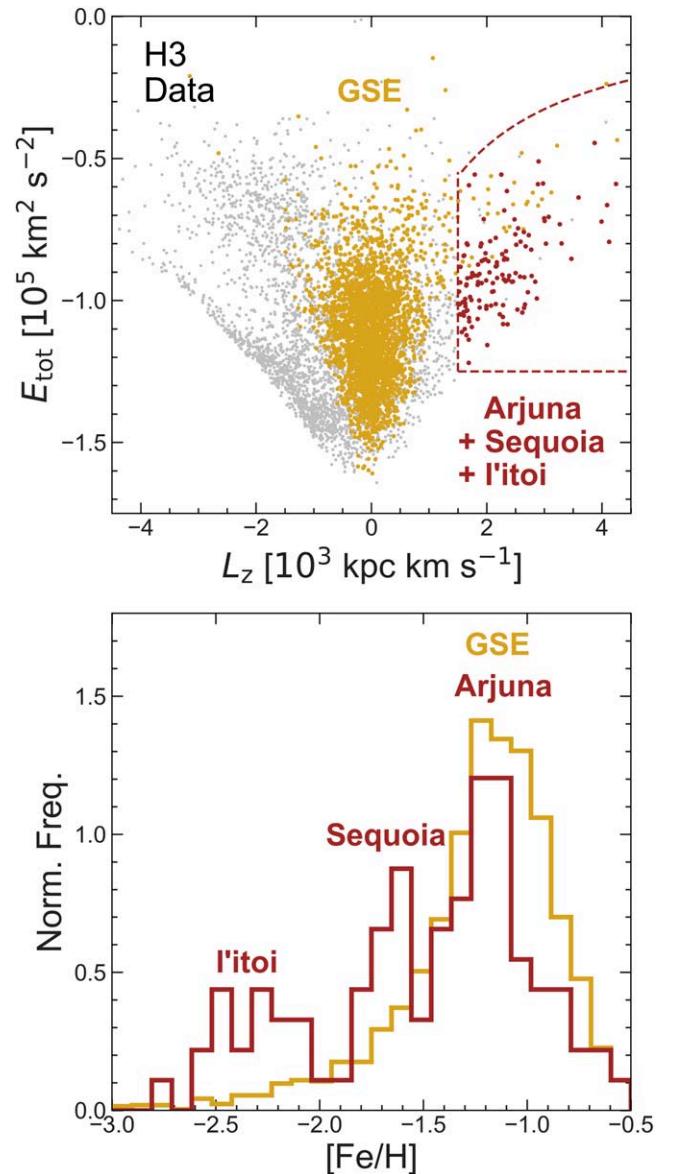

Figure 1. Arjuna as the highly retrograde debris of GSE. Top: $E - L_z$ diagram plotting the total energies of the H3 giants (gray) against the z-component of their angular momenta. GSE, defined to lie at eccentricities >0.7, is shown in gold, whereas the high-energy retrograde halo defined by the dashed lines is shown in brown. Bottom: MDF of the GSE stars compared with retrograde stars. The retrograde MDF shows three populations—I'itoi at $[Fe/H] < -2$, Sequoia at $[Fe/H] \approx -1.6$, and Arjuna at $[Fe/H] \approx -1.2$. The Arjuna MDF closely tracks the GSE MDF.

distance from the Galactic center compared to debris stripped at later times.

Before moving on, we briefly consider an alternative scenario: Arjuna as an $[Fe/H] = -1.2$ dwarf galaxy that, despite having virtually identical abundances, has nothing to do with GSE. It would be a significant coincidence for two distinct accreted dwarf galaxies to have mean $[Fe/H]$ and mean $[\alpha/Fe]$ within 0.05 dex. Further, the relative star counts from H3 imply that Arjuna is only ≈5% of the GSE stellar mass (i.e., ≈10^7 $M_\odot$). This stellar mass and the measured metallicity (-1.2) together constrain the accretion epoch of the hypothetical Arjuna dwarf to be $z \approx 0$ according to the redshift evolution of the mass–metallicity relation (Kirby et al. 2013; Ma et al. 2016). However, we would then expect the very recently

accreted ($z \approx 0$) Arjuna to be rather coherent on the sky (a la Sgr), but this is not the case. For these reasons we disfavor the interpretation of Arjuna as an unrelated dwarf galaxy.

A natural question is why the highly retrograde Arjuna, the most dominant component of the high-energy retrograde halo ($2\times$ as many stars as Sequoia), was not prominent in the local halo data sets (typically limited to $d_{helio} \lesssim 5$ kpc) used to study GSE (e.g., Myeong et al. 2019; Koppelman et al. 2019a; Helmi 2020). From orbit integration we find that the Arjuna stars observed by H3 spend $\approx 20\times$ less time in the solar neighborhood ($d_{helio} < 5$ kpc) than the local halo GSE samples used in these studies. The H3 Arjuna stars on average have larger apocenters and higher energies and are at higher Galactic latitudes. The discovery of Arjuna underscores the value of surveying the distant halo.

3. Summary of Constraints on the GSE Merger

Here we list measurements pertaining to the merger that we will use to guide our numerical experiments in Section 4. While various data sets have been mined to shed light on GSE, we will constrain our simulations purely to measurements from the H3 Survey for consistency. Other measurements are used as independent cross-checks. The H3 constraints listed below apply to GSE as it appears within the survey footprint, and they have been corrected for the selection function (Section 2.1).

1. *Existence of Arjuna*: While the bulk of GSE debris is on highly eccentric, radial orbits that appear as the "sausage" overdensity centered at $V_r \sim 0$ in the $V_r - V_\phi$ plane, in this work we argue that the highly retrograde Arjuna also belongs to GSE. In particular, $\approx 75\%$ of GSE debris is radial, with $|L_z| < 0.5$, while $\approx 5\%$ extends to highly retrograde, high-energy orbits with $L_z > 1.5$. See Section 2.2 for details.

2. *Spatial distribution of GSE debris*: At <50 kpc, $\approx 90\%$ of GSE debris is contained within $r_{gal} \approx 30$ kpc, $\approx 60\%$ within $r_{gal} \approx 20$ kpc, and $\approx 10\%$ within $r_{gal} \approx 10$ kpc. Profiles of the halo using other data sets also show a break at 25–30 kpc that has been associated with an apocenter of GSE (e.g., Deason et al. 2018; Lancaster et al. 2019). Further, the shape of the inner halo, which is dominated by GSE, has been measured by several authors (e.g., Jurić et al. 2008; Xue et al. 2015; Das & Binney 2016). We will use the recent all-sky Gaia RR Lyrae constraints from Iorio et al. (2018) and Iorio & Belokurov (2019), who found that the inner halo defines a triaxial ellipsoid (axis ratios 10: 7.9: 4.5) as a cross-check on the debris geometry.

3. *Hercules-Aquila Cloud (HAC) and Virgo Overdensity (VOD)*: The HAC and VOD are large, diffuse stellar overdensities occurring on either side of the plane that have been known for more than a decade (Vivas et al. 2001; Newberg et al. 2002; Belokurov et al. 2007; Jurić et al. 2008; Bonaca et al. 2012). Recently, thanks to Gaia, both these structures have been linked to GSE based on the integrals of motion and eccentric orbits of their constituent stars (Simion et al. 2018, 2019; but see Donlon et al. 2019, 2020). We will use the emergence of HAC and VOD-like structures at the appropriate locations as an independent cross-check on the models that best reproduce the H3 data.

4. *Stellar mass of GSE*: Estimates of the stellar mass of GSE are in the range of $\sim(2-7) \times 10^8$ $M_\odot$ and have been derived using the mass–metallicity relation assuming $z_{acc.} \approx 2$ ($\approx(4-7) \times 10^8 M_\odot$; Naidu et al. 2020), the age–metallicity relation and dynamical clustering of accreted GSE GCs ($\approx(2-4) \times 10^8 M_\odot$; Kruijssen et al. 2020), counts of metal-poor ([Fe/H] < -1) eccentric ($e > 0.7$) stars ($\approx(2-5) \times 10^8 M_\odot$; Mackereth & Bovy 2020), and chemical evolution models ($\approx(5-6) \times 10^8 M_\odot$; Fernández-Alvar et al. 2018; Helmi et al. 2018).

5. *Spatial extent of the in situ halo*: A substantial fraction of the local kinematic halo is composed of stars that have chemistry identical to the high-α disk but that are on orbits with eccentricities higher than typical disk stars (e.g., Nissen & Schuster 2010; Bonaca et al. 2017; Haywood et al. 2018; Belokurov et al. 2020; An & Beers 2021). This "in situ halo"/"splash" is composed of stars kicked out of the primordial disk during the GSE merger, and its properties are sensitive to the mass of the MW and GSE at the time of the merger (e.g., Fattahi et al. 2019; Grand et al. 2020). Naidu et al. (2020) chart the spatial extent of the in situ halo (defined to have $e > 0.5$ and high-α disk-like chemistry) and find that $>90\%$ of it is confined to $r_{gal} < 20$ kpc and $|Z_{gal}| < 15$ kpc.

6. *Timing and duration of the merger*: Using the H3 main-sequence turnoff sample with precise ages ($\approx 10\%$), Bonaca et al. (2020) report the star formation history (SFH) of the "accreted halo," which is essentially composed of GSE. The GSE SFH abruptly declines at ≈ 10 Gyr ($z \sim 2$). Interestingly, the youngest stars kicked into the in situ halo are ≈ 8 Gyr old ($z \sim 1$). One possible interpretation of these findings is that GSE began interacting with the MW at $z \approx 2$ and that the merger concluded by $z \approx 1$.

4. Numerical Simulations

We aim to reconstruct the GSE merger through controlled, collisionless, *N*-body simulations. Our strategy is to systematically explore a large grid of simulations spanning reasonable orbital and structural parameters to identify configurations that satisfy the constraints in Section 3. We generate galaxy models for GSE and the MW with the GalICv1.1 (Yurin & Springel 2014) initial condition generator and then run merger simulations with the smoothed particle hydrodynamics codes Gadget-2 (Springel 2005) for the initial low-resolution simulations and Gadget-4 (Springel et al. 2021) for the final high-resolution simulations. In all our numerical choices we closely follow recent, similar high-resolution merger simulations (Amorisco 2017; Laporte et al. 2018; Garavito-Camargo et al. 2019). In what follows we motivate the grid we explore and our simulation setup.

4.1. Structural Parameters

4.1.1. GSE

Our starting point is the GSE stellar mass ($M_\star = (2-7) \times 10^8$ $M_\odot$) and accretion redshift ($z \approx 2$) discussed in Section 3. We consider three different models that bracket the literature mass range—"M0," "M1," and "M2" with stellar masses of 2×10^8 $M_\odot$, $5 \times 10^8 M_\odot$, and $7 \times 10^8 M_\odot$, respectively. Extrapolating the size–mass relation (SMR) at $z = 2$ from Mowla et al. (2019)

Table 1
$z \sim 2$ GSE Structural Parameters

Component	Parameter	M0	M1	M2
DM halo	Mass, M_{200} (10^{11} $M_{\odot}$)	1.3	2.0	2.5
(Hernquist)	Concentration (c_{200})	4.2	4.0	4.0
	Scale length (kpc)	24.8	29.7	31.6
Disk	Mass (10^8 $M_{\odot}$)	2.0	5.0	7.0
(Exponential)	Spin fraction (10^{-3})	1.5	2.5	2.8
	r_{50} (kpc)	1.5	1.7	1.7
		2.3	2.5	2.6
		3.0	3.3	3.5

Note. For a model of a given $M_\star$ we consider three scale lengths such that the half-mass radius, r_{50}, is $1 \times$ SMR, $1.5 \times$ SMR, and $2 \times$ SMR, where SMR is set by the $z = 2$ size–mass relation (SMR) in Mowla et al. (2019).

to lower masses, we obtain half-light radii (r_{50}). At $z \sim 2$ half-light radii and half-mass radii are approximately equal (e.g., Mosleh et al. 2017; Suess et al. 2019). To account for the significant scatter in size at fixed mass observed at $z \sim 2$ (e.g., van der Wel et al. 2014), as well as the fact that the SMR has not been measured at masses below $M_\star \approx 5 \times 10^9$, we consider three sizes: $1\times$, $1.5\times$, and $2\times$ the r_{50} from the extrapolated SMR. In total we have nine models for GSE (three stellar masses times three sizes; see Table 1).

We model GSE stars as an exponential disk embedded in a spherical, Hernquist (1990) DM halo (GalIC's "Model D1"). We set the disk scale height to 60% of the disk scale length motivated by simulations that find that $z \approx 2$ disks are born thick from turbulent gas and stay thick (e.g., Bournaud et al. 2009; Forbes et al. 2012; Bird et al. 2013; Ma et al. 2017b; Park et al. 2021). We model GSE as a pure disk without a bulge, motivated by studies that measure declining bulge-to-total (B/T) ratios toward lower masses, such that B/T is $\approx$20% for $M_\star \approx 10^{10}$ $M_{\odot}$ star-forming galaxies at $z \approx 2$, and likely lower for GSE-like $M_\star \approx (2-7) \times 10^8$ $M_{\odot}$ galaxies (e.g., Lang et al. 2014). In the Appendix we nonetheless explore GSE models with bulges and find that these structural choices play a secondary role to the total mass and orbital parameters. We set the disk spin fraction equal to the disk mass fraction—this determines the overall kinematics of the disk and is the fraction of the total angular momentum of the galaxy concentrated in the disk.

For the mass of the DM halo we appeal to the $z = 2$ stellar mass–halo mass relation from the UniverseMachine empirical model (Behroozi et al. 2019). We set the size of the DM halo based on the $z = 2$ concentration–mass relation from Diemer & Joyce (2019) that is based on N-body DM simulations. The resulting parameters are listed in Table 1. Note that internally GalIC maps a specified Navarro–Frenk–White halo mass and concentration to a Hernquist halo of the same mass with a scale length such that the shape of the density profile in the inner regions is identical (Yurin & Springel 2014, their Equation (48)).

4.1.2. Milky Way

We require a faithful representation of the MW during the epoch of the merger ($z \approx 1$–2). Ages gleaned from a variety of methods suggest that almost the entirety of the present-day high-α/thick disk, as well as the present-day bulge, assembled at $z > 1$, whereas the low-α/thin disk largely grew at $z < 1$

Table 2
$z \sim 2$ Milky Way Structural Parameters

Component	Parameter	Value
DM halo	Mass, M_{200} (10^{11} $M_{\odot}$)	5.0
(Hernquist)	Concentration (c_{200})	3.8
	Scale length (kpc)	40.4
Disk	Mass (10^9 $M_{\odot}$)	6.0
(Exponential)	Scale length (kpc)	2.0
	Scale height (kpc)	1.0
	Spin fraction (10^{-2})	1.2
Bulge	Mass (10^{10} $M_{\odot}$)	1.4
(Hernquist)	Scale length (kpc)	1.5

(e.g., Gallart et al. 2019; Surot et al. 2019; Lian et al. 2020; Ruiz-Lara et al. 2020; Bonaca et al. 2020). We therefore model the $z \sim 1$–2 MW as a combination of the present-day thick disk and bulge with a total stellar mass of 2×10^{10} $M_{\odot}$ and scale lengths following Bland-Hawthorn & Gerhard (2016). Our adopted disk (i.e., the present-day thick disk) has a $\approx$30% smaller scale length than the present-day thin disk, accounting for the smaller size of the MW at $z \sim 1$–2. The total stellar mass is consistent with the $z \sim 1$–2 expectation from look-back studies of MW progenitors (e.g., van Dokkum et al. 2013).

The disk and bulge are embedded in a spherical, Hernquist (1990) DM halo (GalIC's Model M1). The mass of the DM halo is set to half the $z = 0$ mass (5×10^{11} $M_{\odot}$; Cautun et al. 2020; Vasiliev et al. 2021; Zaritsky et al. 2020; Deason et al. 2021) motivated by the average growth history of MW-like DM halos seen in simulations (e.g., Wechsler et al. 2001). The concentration is determined by the $z = 2$ concentration–mass relation from Diemer & Joyce (2019).

4.2. Orbital Parameters

All our simulations begin with GSE at the MW's virial radius. The initial velocity is set such that the total energy is the energy of a circular orbit with radius equal to the MW's virial radius, consistent with satellites in cosmological DM simulations (Jiang et al. 2015; Amorisco 2017). The radial component of the velocity is varied so that the circularity, η, ranges between 0.1-0.9 in uniform steps of 0.1. Pure radial orbits have $\eta = 0$ while perfectly circular orbits have $\eta = 1$. The orbital inclination with respect to the MW disk plane, θ, is set to one of $0°$, $30°$, $60°$, $90°$. The plane of the satellite disk always coincides with the satellite orbital plane. We consider prograde and retrograde orbits and allow the spin of GSE's disk to be co-rotating or counterrotating with respect to the MW.

4.3. Merger Simulations

We follow a two-step procedure. We first run $\approx$500 simulations exploring the grid of orbital and structural parameters summarized in Tables 1, 2, and 3 at a particle resolution of $m_{DM} = m_{baryon} = 10^5$ $M_{\odot}$ ("low-res") with the Gadget-2 code. We then identify the most promising configurations and simulate another grid around them at a resolution of $m_{DM} = m_{baryon} = 10^4$ $M_{\odot}$ ("high-res") with Gadget-4 that was released during the course of this project. In the high-res simulations the stellar component of GSE is represented by $\approx$50,000 particles.

Property	Parameters
Initial Grid ($m_{particle} = 10^5\,M_\odot$)	
Circularity (η)	0.1, 0.3, 0.5, 0.7, 0.9
Inclination (θ)	0°, 30°, 60°, 90°
Sense of orbit	prograde, retrograde
Disk spin	prograde, retrograde
Refined Grid ($m_{particle} = 10^4\,M_\odot$)	
Circularity (η)	0.40, 0.45, 0.50, 0.55, 0.60
Inclination (θ)	0°, 15°, 30°, 45°
Sense of orbit	retrograde
Disk spin	retrograde

All simulations are run for 10 Gyr, i.e., from $z = 2$ to $z = 0$. Time steps (Δt) are assigned in an adaptive scheme to individual particles via $\Delta t = \sqrt{2\zeta\epsilon/a}$, where $\zeta = 0.025$ is an accuracy parameter, ϵ is the softening length, and a is the gravitational acceleration of the particle under consideration. The softening lengths adopted for all particles are $\epsilon = 250$ (80) pc for the low-res (high-res) simulations following the Power et al. (2003) criteria for the optimal softening length (their Equation (15)). The maximum time step is limited to 20 Myr. We choose an opening angle of $\theta = 0°.5$ for the tree algorithm.

4.4. Comparing Models with Data

Since the MW in the simulations is less massive than the present-day MW, we need to account for the deeper $z = 0$ potential before comparing with $z = 0$ data. Following Villalobos & Helmi (1806) and Koppelman et al. (2020), we measure the mean rotational velocity of the MW disk in our simulations at $2.4 \times$ the scale length and compare this with the observed rotation velocity of the MW thick disk at the corresponding distance ($\approx 170\,\mathrm{km\,s^{-1}}$). Based on this comparison, we scale our $z = 0$ velocities by $1.37 \times$ (similar to Koppelman et al. 2020, who scale by $\approx 1.3 \times$, and equivalent to scaling the mass by $\approx 1.9 \times$). This scaling implies that there are other sources of mass growth in the inner Galaxy that are not accounted for in our simulation (e.g., the gas from GSE, subsequent accretion events like Sgr, the emergence of the low-α disk). The satisfactory reconstruction of the shape and extent of the $V_r - V_\phi$ "sausage" (Belokurov et al. 2018) in our fiducial simulation is a consistency check of the applied scaling (see Figure 6 below).

We add observational errors to match the properties of the H3 sample. The PM and RV errors of the sample under consideration are a negligible contribution to the error budget. The distance error is the primary source of uncertainty (see Appendix A of Naidu et al. 2020). We assume a 10% distance error for all stars, well matched to the data at hand ($8\% \pm 4\%$). In every simulation snapshot, Galactocentric positions and velocities of stars are computed with respect to the center of mass of the MW bulge stars. Dynamical quantities like angular momenta, energies, and Galactocentric velocities are computed exactly as is done for the data in Naidu et al. (2020). For all model–data comparisons, unless otherwise mentioned, we select only the simulation particles with $d_{helio} = 3$–50 kpc that fall within the survey's fields at $|b| > 40°$, decl. $> -20°$. Note that the data compared to below are already corrected for the

survey selection function (i.e., the photometric magnitude limit and targeting strategy). This means that the distribution of, say, r_{gal} or L_z from the simulations can now be directly compared with the data.

4.5. General Trends

Here we describe how the various structural and orbital parameters explored in our simulations produce varied debris distributions. We were able to rule out a few regions of parameter space quickly from our initial set of experiments with the M1, SMR model. As one might intuitively expect, none of the prograde simulations produce anything like the strongly retrograde Arjuna debris. We also found that retrograde mergers with counterrotating disk spin are highly efficient at producing retrograde debris (as seen in, e.g., Bignone et al. 2019). For the rest of this work we focus on such retrograde mergers. Trends with size, mass, and orbital parameters are shown in Figures 2, 3, 4 to give readers a sense of how these parameters translate into $E - L_z$ and r_{gal} distributions.

Size: At fixed mass, a larger size leads to more retrograde debris (Figure 2). Since a higher fraction of stars inhabit the outer, less-bound regions of the satellite's disk, they are stripped easily early in the merger. This debris from the outer regions retains memory of the initial (retrograde) orbit of the satellite.

Mass: At higher mass, dynamical friction operates more efficiently, satellites sink faster, and deposit a larger fraction of their stars deep in the potential (Figures 2 and 4). In particular, $t_{DF} \propto M_{sat}^{-1}$, where t_{DF} is the dynamical friction timescale and M_{sat} is the satellite mass (Mo et al. 2010). While quickly radialized, the massive satellites are nonetheless also able to produce a significant fraction of retrograde debris owing to their extended size. Low-mass satellites, on the other hand, experience prolonged mergers and deposit large fractions of their stars at distant radii and higher average energy (top row of Figure 2).

Circularity: Along with mass, η is the key moderator of the timing of the merger (Figure 4). This is a well-known result: $t_{DF} \propto \eta^s$, $s \approx 0.3$–0.5 (Mo et al. 2010). That is, circular orbits decay the slowest and leave a larger fraction of debris at higher angular momenta.

Inclination: Once we limit ourselves to retrograde mergers at fixed mass and circularity, the orbital inclination has minimal impact on physical aspects of the merger (e.g., it has little effect on the orbital decay profile). However, θ strongly moderates the final spatial distribution of the merger debris, and thus how the debris is observed by various surveys (Figure 3). For instance, a highly radial, entirely in-plane ($\theta = 0$) merger would be barely observable at $|b| > 40°$ in a survey like H3 but for mergers on inclined orbits the observable fraction is boosted (e.g., by $>3 \times$ between $\theta = 0°$ and $\theta = 90°$ for a radial $\eta = 0.1$ merger).

4.6. Selecting the Fiducial Model

To select models that best reproduce the $z = 0$ GSE+Arjuna debris we focus on the observed L_z and r_{gal} distributions described in Section 3, and use the other constraints (e.g., the extent of the in situ halo) as validation checks. L_z and r_{gal} require minimal assumptions (e.g., their computation is independent of the potential), and are sensitive discriminators
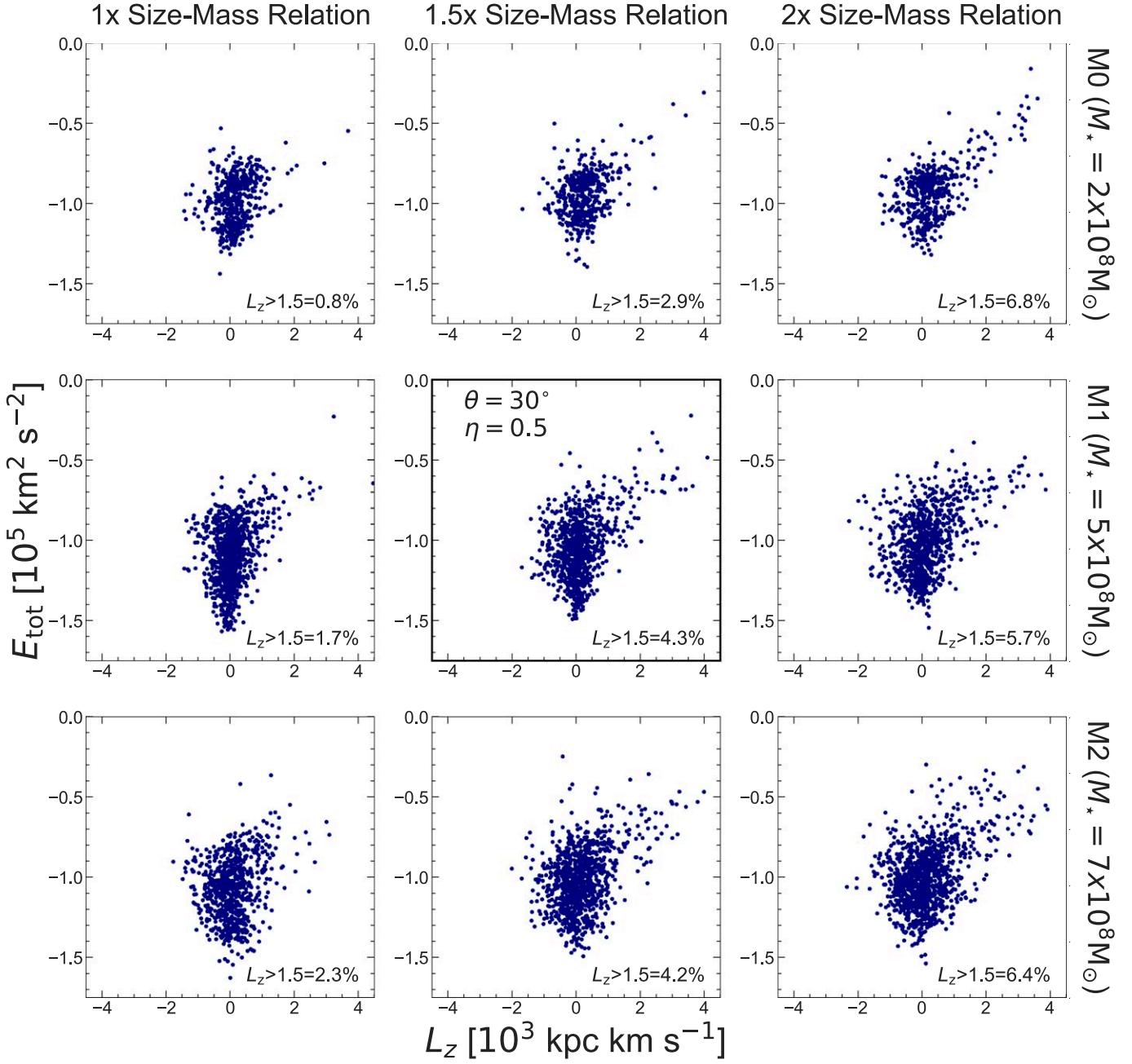
Figure 2. Variations in $E - L_z$ of accreted satellites with stellar mass (rows) and size (columns) while keeping orbital inclination ($\theta = 30°$) and circularity ($\eta = 0.5$) fixed. H3 uncertainties and the survey footprint are applied to the simulation, enabling direct comparison between the two. The central panel shows the most promising model from our initial grid, which reproduces the H3 r_{gal} and L_z distributions. The fraction of debris at $L_z > 1.5$, corresponding to the location of Arjuna, and measured to be $\approx 5\%$ in the data, is indicated at the bottom right of each panel. The M0 galaxies (top row) are spread out over a relatively smaller area and do not deposit debris as deep in the potential as the M1 and M2 galaxies (bottom rows), which due to their higher mass rapidly lose energy to dynamical friction. At fixed mass, more extended galaxies produce higher fractions of retrograde debris (compare first and third columns) that arises from their outer, loosely bound regions in the early stages of the merger.

of merger configurations (Figures 2, 3, and 4). In detail, we require the $L_z > 1.5$ fraction and $|L_z| < 0.5$ fractions to fall between the 16th and 84th percentiles of the observed distribution, i.e., the majority of the debris should be radial, but $\approx 5\%$ must extend to highly retrograde orbits. We also require the $L_z < -1.5$ fraction to be $< 1\%$ since we observe no stars with GSE chemistry on highly prograde orbits. We make a similar demand of the debris fraction at $r_{gal} = 10$–20 kpc, $r_{gal} = 20$–30 kpc, and $r_{gal} > 30$ kpc.

2The log-likelihood of the entire grid of counterrotating, retrograde configurations computed against our r_{gal} and L_z

requirements listed previously (assuming Gaussian errors) is shown in Figure 5. In detail, each of the fractions for L_z (radial, prograde, retrograde) and r_{gal} (10–20 kpc, 20–30 kpc, >30 kpc) contributes to the likelihood equally as six random normal variables—this weighting makes for a likelihood that is much more sensitive to the Arjuna component and the break at 25–30 kpc compared to a classical likelihood computed against the full r_{gal} and L_z distributions that is more sensitive to the peak of the distributions. Interestingly, most of the grid is easily ruled out with our sparse set of r_{gal} and L_z constraints. The low-mass models (M0) are heavily disfavored no matter their orbital configuration.
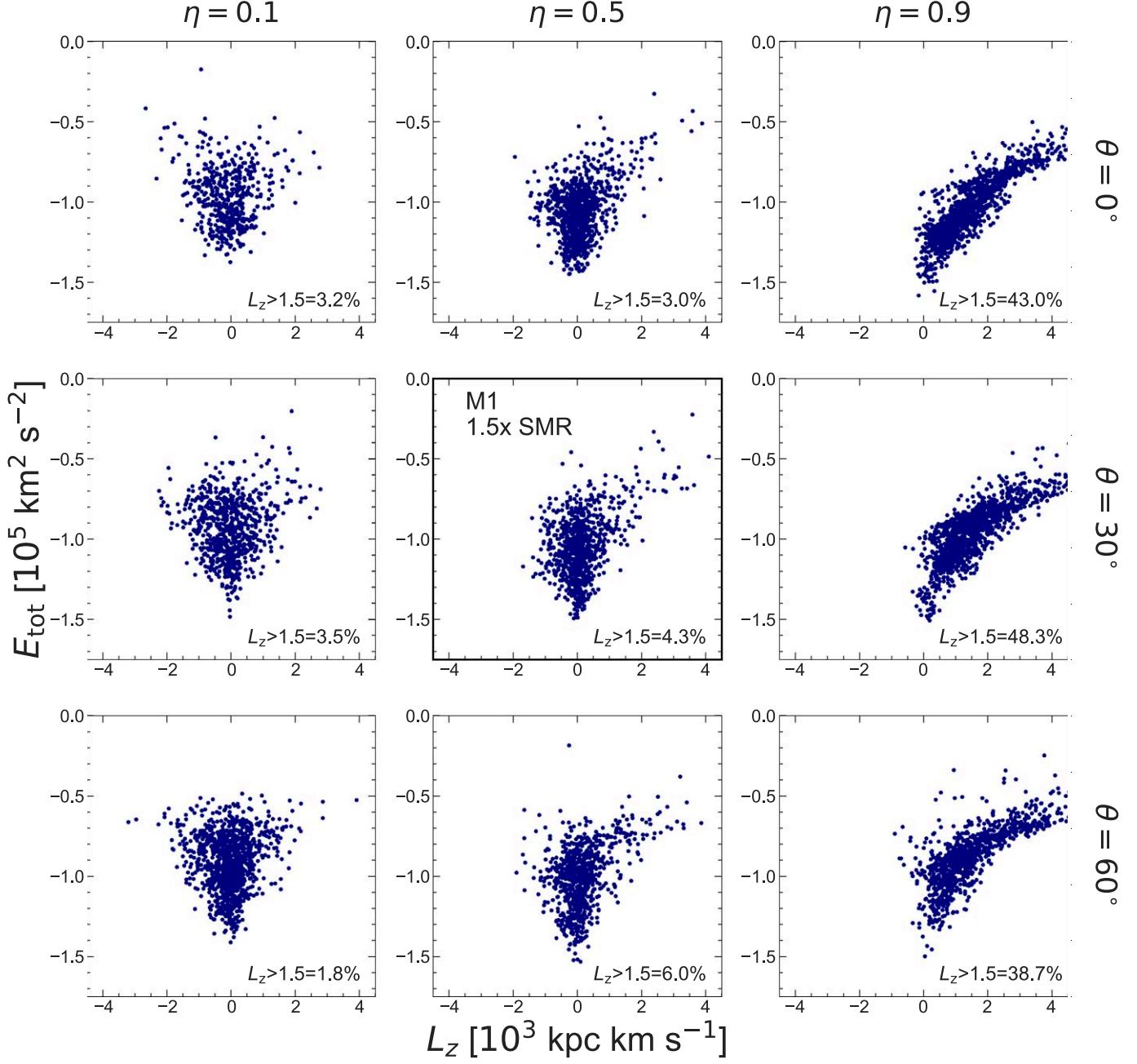
Figure 3. Same as Figure 2, but varying orbital parameters—the circularity (η) and inclination (θ)—while keeping mass and size fixed (M1, $1.5 \times$ SMR). The circularity, η, strongly influences the retrograde fraction of the observed debris. Circular orbits have longer dynamical friction timescales and so a smaller fraction of their debris is on radial orbits. θ sets the amount and energy of the debris that would rise into the field of view of a high-galactic field survey like H3. For instance, all the models in the left column produce mergers with indistinguishable orbital decay profiles but θ sets the number of stars that makes it into the H3 Survey fields.

Only one configuration out of the many hundred simulated satisfies these constraints: $M = 5 \times 10^{8}\ M_{\odot}$, $1.5 \times$ SMR, $\theta = 30°$, $\eta = 0.5$. This configuration has "Goldilocks" parameters: it is neither too radial nor too circular, has moderate orbital inclination, and has an intermediate mass/size. We explore a finer grid around this set of parameters ($\eta = [0.4, 0.45, 0.5, 0.55, 0.6]$, $\theta = [0°, 15°, 30°, 45°]$) at $10 \times$ resolution (i.e., particle mass of $10^{4}\ M_{\odot}$). We find that the $\eta = 0.5$, $\theta = 15°$ model best matches the data, and we focus on this model (the "fiducial model") for the rest of this work.

We note that there are large swathes of the merger configuration parameter space left unexplored in this work—

e.g., we have kept the MW structural parameters and the initial GSE orbital energy fixed, and we have assumed no scatter in our adopted stellar mass–halo mass relation or mass–concentration relation for GSE. We have also not accounted for the significant amounts of gas GSE likely brought into the Galaxy (potentially $>100\%$ of its stellar mass; Tacconi et al. 2020) that may have fueled the growth of both the high-α and low-α disks and altered the MW potential (e.g., Grand et al. 2020; Bonaca et al. 2020). However, note that the overall dynamics of the merger (e.g., the orbital decay profile) are essentially set by the much more massive DM halos of the MW and GSE. As discussed in subsequent sections, our adopted fiducial model is an excellent
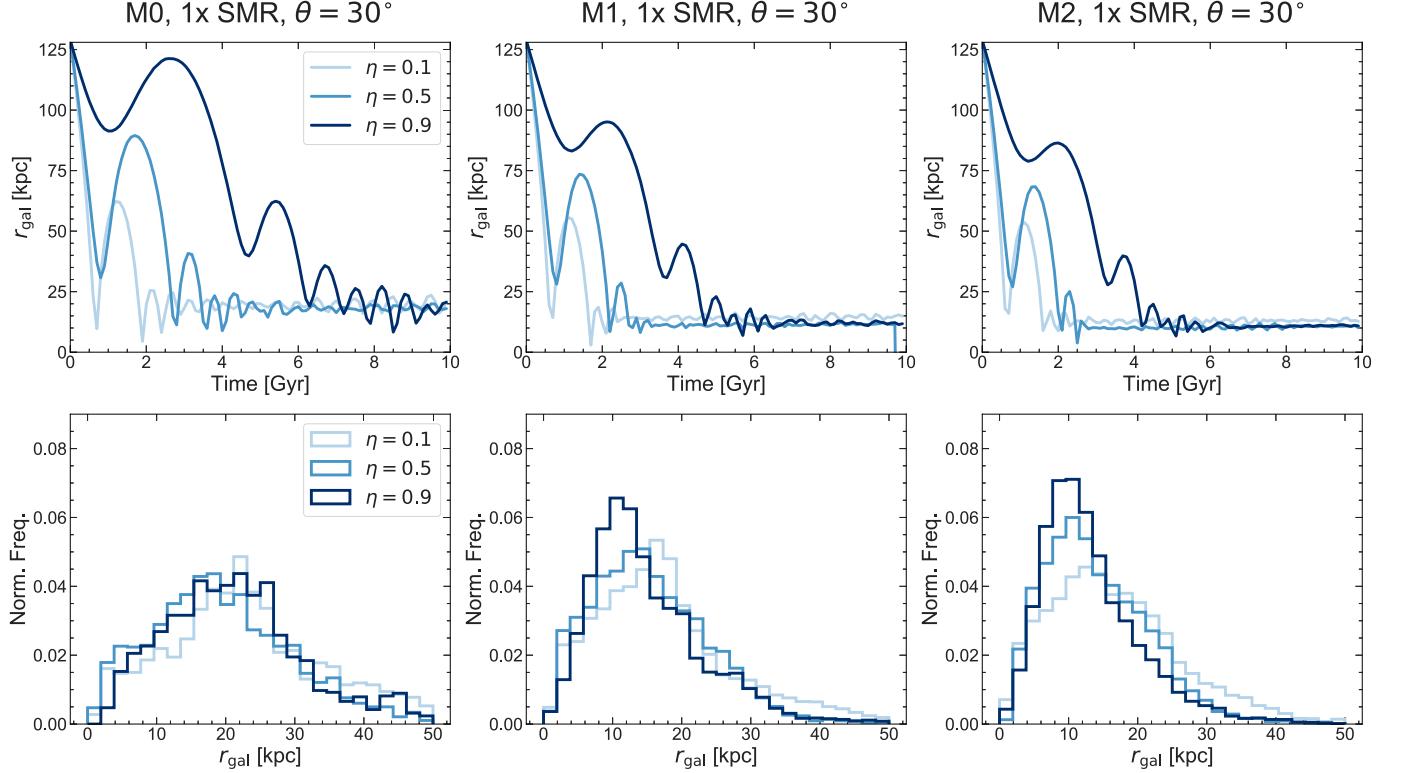
Figure 4. Top: Orbital decay as a function of time for mergers of varying mass and circularity. The 5% most bound particles at $t = 0$ Gyr are tagged and their mean r_{gal} is used to track orbital decay. Lower mass satellites and higher circularity orbits make for prolonged mergers as expected from dynamical friction considerations. The mean r_{gal} of the particles tagged and tracked from $t = 0$ converges to an approximately flat value after the merger is complete—this value is a measure of the final few apocenters where the most bound stars are stripped, and is lower for the higher mass mergers (i.e., the satellites sink deeper in the potential). Bottom: Spatial distribution of merger debris (all-sky, not limited to H3 footprint). High-circularity, low-mass satellites deposit their stars at larger distances whereas more massive satellites deposit their stars in the inner regions of the host as seen in the distributions growing peakier and shifting left across the panels.

match to the H3 data and satisfies a wide variety of constraints, but given these caveats, we are in no position to claim that its parameters are the only ones that match these constraints.

5. Results

5.1. Preferred Configuration: A 2.5:1 Merger on an Inclined, Retrograde Orbit

The fiducial merger configuration ($M_\star = 5 \times 10^8\ M_\odot$, $M_{DM} = 2 \times 10^{11}\ M_\odot$, $1.5 \times$ SMR, $\theta = 15°$, $\eta = 0.5$) selected from the high-resolution grid is summarized in Figure 6. The L_z and r_{gal} distributions are an excellent match to the H3 data by construction. The "sausage" in $V_r - V_\phi$ where GSE was first discovered with Gaia is satisfactorily reproduced (Belokurov et al. 2018). The orbit has an apocenter at 28 kpc, which is where several studies have found a break in the density profile in the halo. The slope of the GSE density profile within 25 kpc is a good match to that found for the inner halo (e.g., Deason et al. 2014, discussed further in Section 5.5). This is exactly as expected given that GSE dominates the halo within 25 kpc (e.g., Naidu et al. 2020). At larger distances, other components become more prominent, and the GSE density profile falls faster than that of the overall halo. The merger is fairly rapid, with the gap between first and final pericenter being a mere 2 Gyr, in agreement with the timing constraints (discussed further in Section 5.6).

We also confirm that >90% of the stars kicked out of the MW disk (the in situ halo) are contained within $|Z| < 15$ kpc. We do not analyze the perturbed MW disk and in situ halo beyond this check because we are unsure whether their detailed

properties are physically meaningful given the subsequent assembly history (e.g., growth of the massive thin disk around the thick disk) that we have not modeled here. Nonetheless, we direct interested readers to J. Han et al. (in preparation), where we discuss aspects of the in situ halo from the simulations.

In Figure 7 we show that the fiducial model also produces overdensities analogous to the HAC and the VOD at their exact observed locations. Furthermore, the direction along which the debris is spread out is an excellent match to that of the major axis of the triaxial ellipsoid fit by Iorio & Belokurov (2019) to describe the inner halo. The agreement between our fiducial model and these spatial constraints is particularly remarkable since we select the model only based on the H3 L_z and r_{gal} distributions (implicitly, the H3 window function has a spatial aspect). Now that we have a model that is in excellent agreement with the constraints in Section 3, we can use it to extract further physical insights about the merger.

The total GSE mass of our fiducial model ($2 \times 10^{11}\ M_\odot$) is ≈50% higher than that of the Large Magellanic Cloud ($\approx 1.3 \times 10^{11}\ M_\odot$; Erkal et al. 2019; Vasiliev et al. 2021) and represents as much as 20% of the MW's present-day virial mass ($\approx 10^{12}\ M_\odot$; Cautun et al. 2020; Zaritsky et al. 2020; Deason et al. 2021). The stellar mass constitutes ≈50% of the MW's stellar halo ($\approx 10^9\ M_\odot$; Deason et al. 2019; Mackereth & Bovy 2020). Within the ambit of our grid, this finding is particularly robust, since the lower-mass GSE models are strongly ruled out by the r_{gal} and L_z constraints (top row, Figure 5)—these models produce slowly decaying mergers that
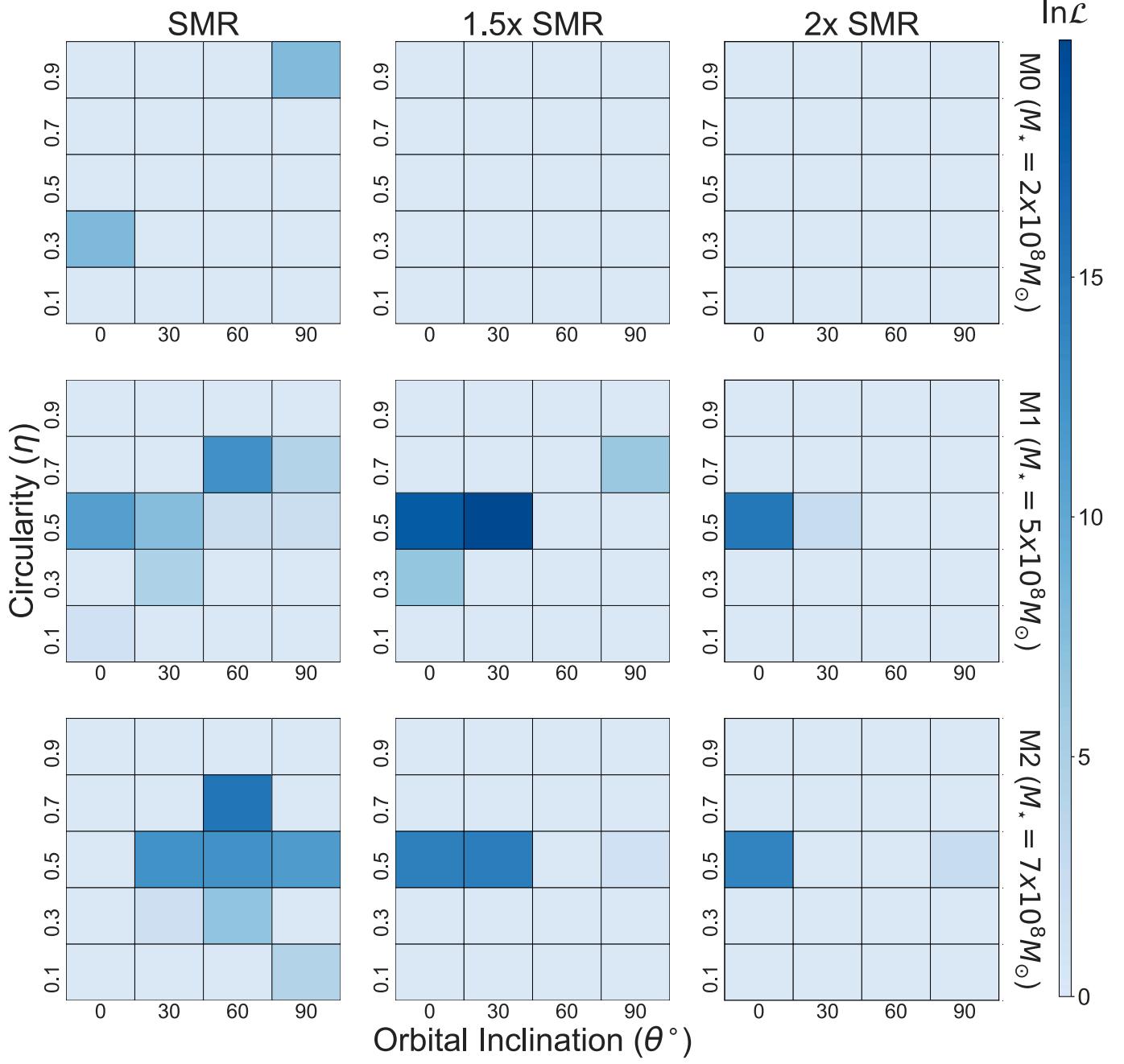
Figure 5. Log-likelihood of the initial, low-resolution ($10^5\,M_\odot$) grid of retrograde simulations as a function of GSE mass (rows) and GSE size (columns). Each 4×5 grid charts circularity (η) against the orbital inclination (θ). The log-likelihood is computed against the observed H3 r_{gal} and L_z distributions. Remarkably, only these two constraints eliminate the vast majority of the grid. The lowest-mass models (M0, top row), regardless of orbital parameters, are heavily disfavored (note that the color bar shows log-likelihood). The highest likelihood region occurs in the M1, $1.5 \times$ SMR simulation at intermediate circularity ($\eta = 0.5$) and moderate inclination ($\theta = 30°$). We resimulate a finer grid around these parameters at higher resolution ($10^4\,M_\odot$) and select our fiducial model ($\theta = 15°$, $\eta = 0.5$) from these simulations.

deposit a high fraction of their debris at larger distances than seen in the data (Figure 4).

In Figure 8 we plot the orbit of GSE from the fiducial model. This orbit is computed based on the center of mass of the 5% most bound GSE stars prior to the merger and shown for the first 3 Gyr of the simulation (i.e., covering the duration of the merger). The orbit is not radial right away—for $\approx$2 Gyr GSE journeys through the Galaxy with significant angular momentum before ending up on a radial track at <30 kpc

between its final two apocenters (shown as stars). We will refer to this orbit at various points in subsequent sections while interpreting, e.g., the spatial distribution of the GSE debris.

5.2. The Origin of Arjuna

In Figure 9 we trace the origins of the highly retrograde Arjuna stars. We tag stars that are at $L_z > 1.5$ at $z = 0$ and
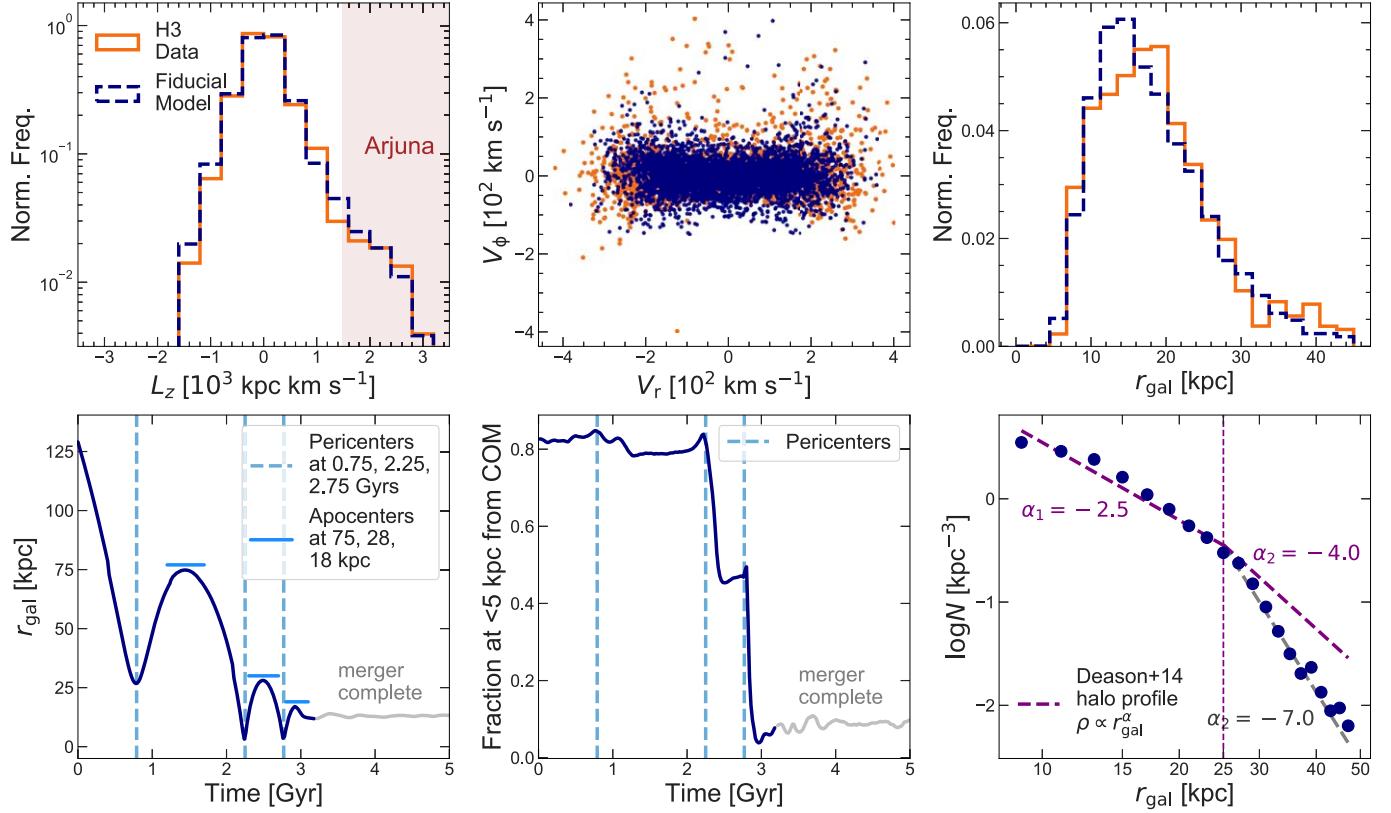
Figure 6. Summary of fiducial model. By construction the model is an excellent match to the observed L_z (top left) and r_{gal} distributions (top right). The characteristic $V_r - V_\phi$ "sausage" is satisfactorily reproduced (top middle). The orbital decay profile (bottom left) shows a rapid merger. Only 2 Gyr separate the first and final pericenter. Pericenters (apocenters) are marked by vertical (horizontal) blue lines. In the bottom middle panel we show the fraction of stars within 5 kpc of the COM of the 5% most bound GSE stars with time. The second pericenter is when half the stars are stripped, and the remaining are lost at third pericenter. The all-sky density profile of GSE (bottom right) in our model (blue points) and the observed stellar halo profile (Deason et al. 2014; purple) agree well within 25 kpc and diverge at larger distances where other substructures become important. The break in the profile occurs around the 28 kpc apocenter in our model.
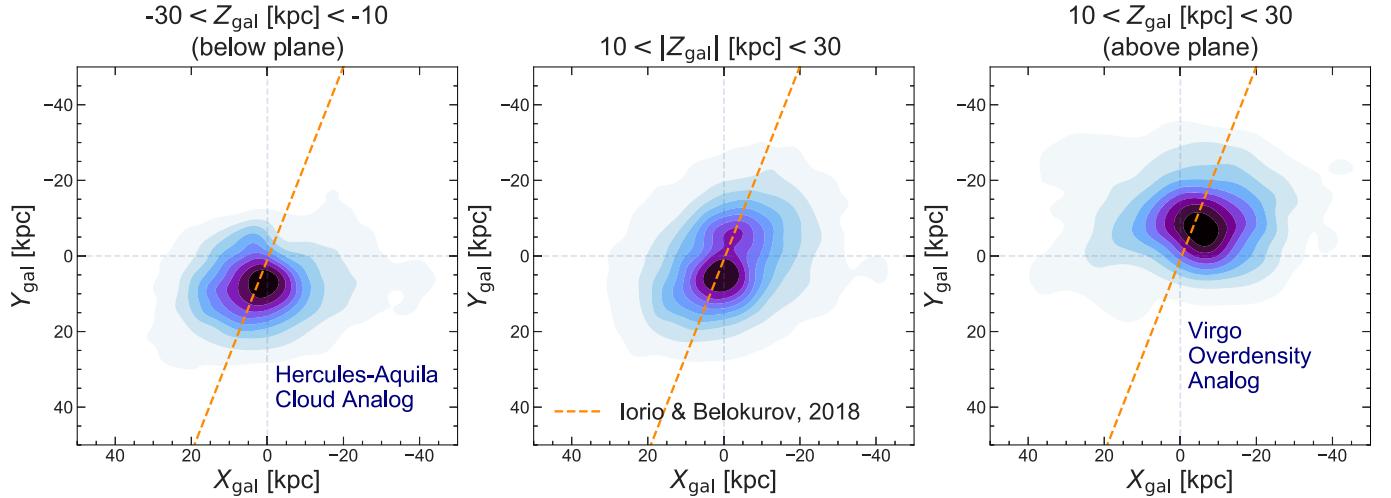
Figure 7. Comparison of the fiducial model (blue) with the triaxial inner halo found in Iorio et al. (2018) and Iorio & Belokurov (2019) using Gaia RR Lyrae. The density contours of our model line up almost exactly with the major axis of their triaxial profile (orange line). Further, the model produces analogs of the HAC (left) and the VOD (right) at the right locations—on opposite sides of the plane at either end of the major axis. We emphasize that the fiducial model was selected purely on the H3 L_z and r_{gal}, and so this close alignment should be viewed as independent corroboration of the model.

follow them through the simulation. In the top panel of Figure 9 we see that before the merger these stars occupy the outer regions of the disk of GSE and lie at a median radius of $\approx 2.5 \times r_{50}$. These relatively loosely bound stars from the outer

disk are stripped earlier in the merger, and so they retain the larger angular momenta and higher energy that the satellite initially arrived with. On the other hand, the majority of stars from the inner disk are stripped after the bulk motion of the

11
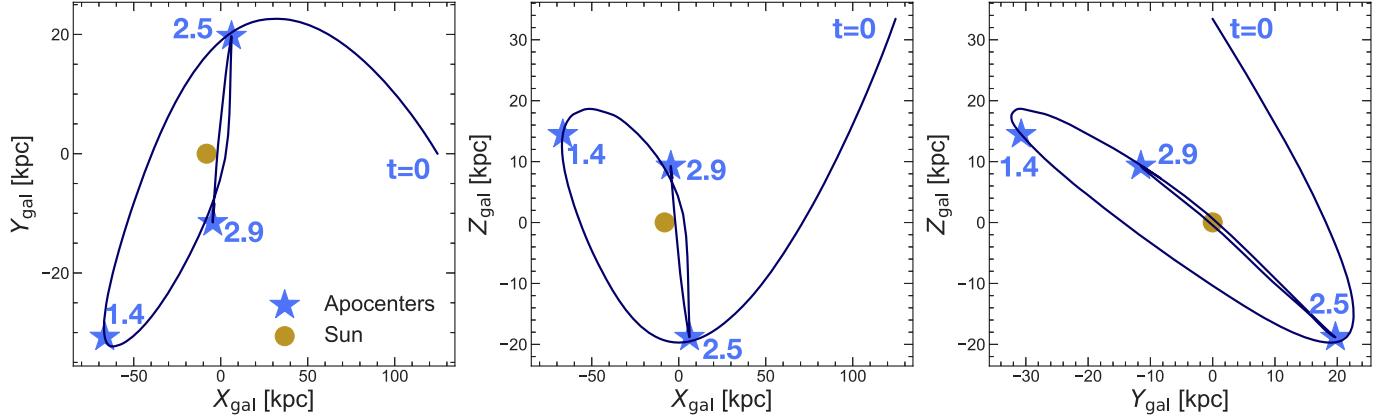
Figure 8. Orbit of fiducial model. Apocenters are marked as stars along with their time of occurrence in Gyr.
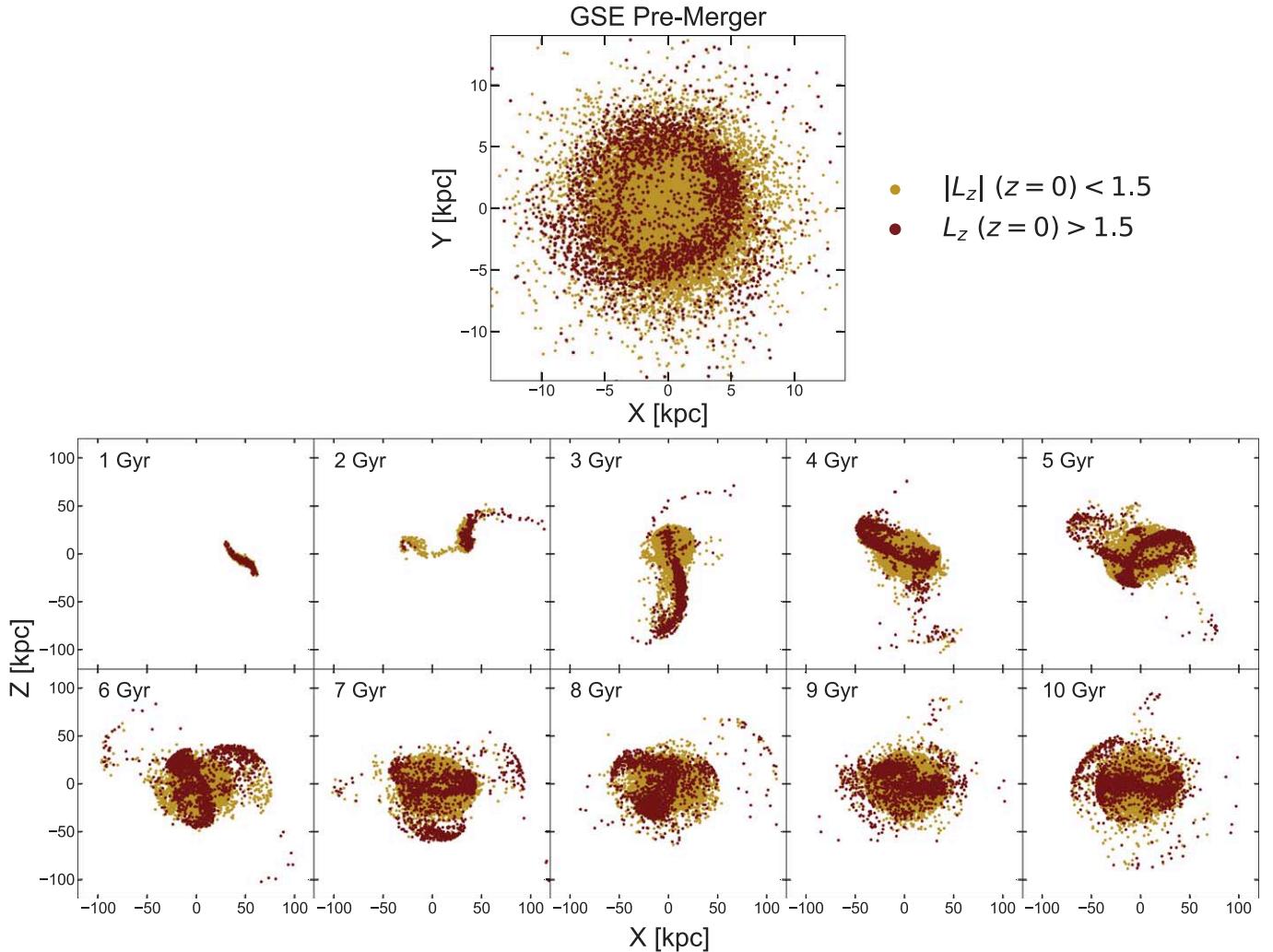

Figure 9. The origin of Arjuna according to the fiducial model. Top: the GSE disk before the merger. Stars are colored brown if they end up on highly retrograde, Arjuna-like orbits ($L_z > 1.5$), and gold otherwise. The Arjuna-like stars preferentially inhabit the outer disk ($\langle r_{gal} \rangle \approx 2.5 \times r_{50}$). Bottom: progression of the merger in the X-Z plane. Arjuna stars are stripped early, at larger distances, when the satellite still has its initial retrograde angular momentum. The other stars are stripped after the satellite is radialized, within a ≈ 25 kpc golden ball that corresponds to the second apocenter.

satellite has been radialized, and hence they are found on $|L_z| < 0.5$, eccentric orbits and appear as the $V_r - V_\phi$ "sausage." An implication of this exercise is that information about the

detailed spatial structure of a galaxy that was disrupted ≈ 10 Gyr ago is still retained in the present-day angular momenta distribution of its debris in the halo.
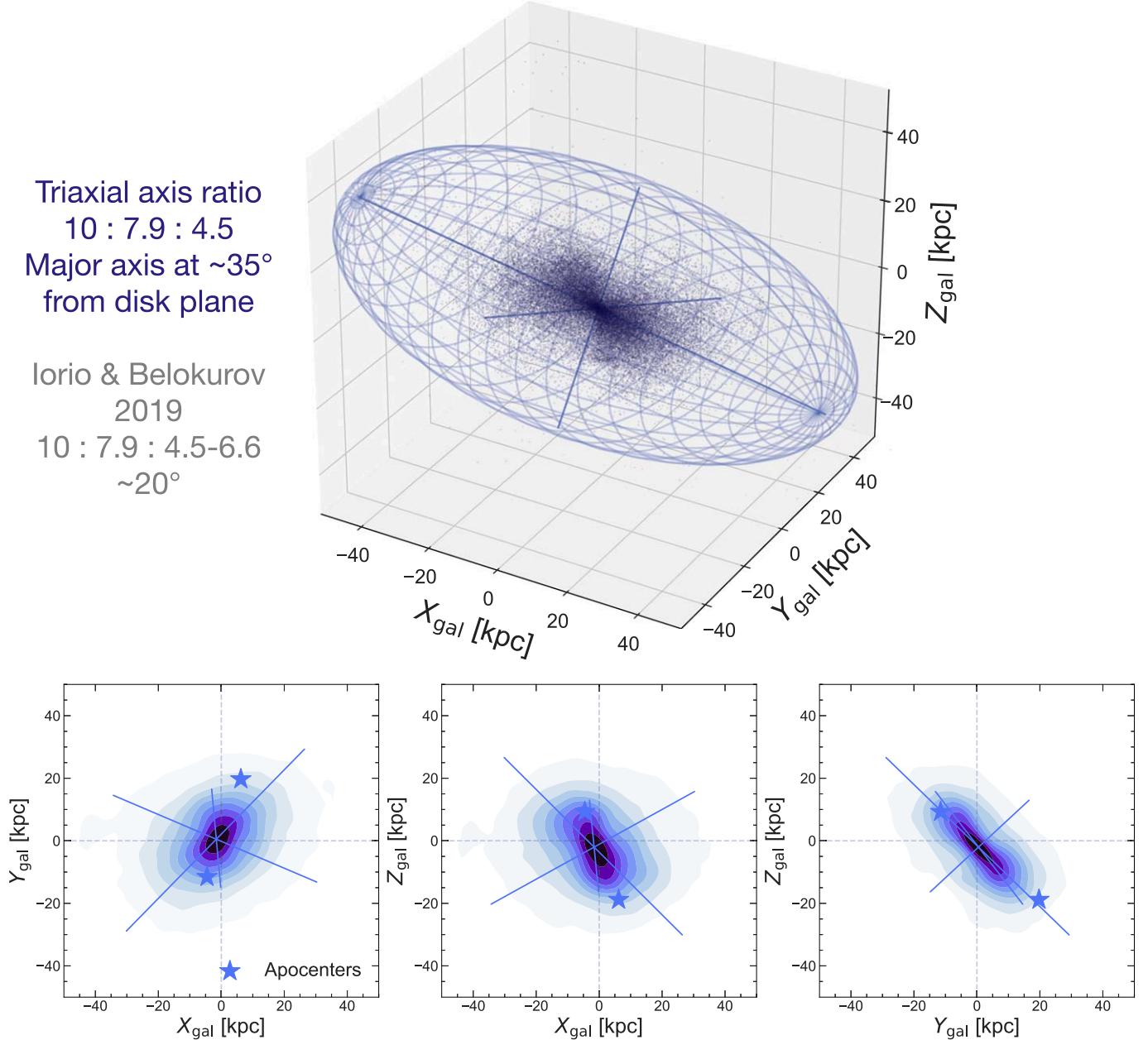
Figure 10. Shape of the MW inner halo in the fiducial model. Top: within 35 kpc, GSE is by far the most dominant component of the halo, and so the geometry of its debris sets the shape of the inner halo. We fit a trixial ellipsoid (light-blue grid) to describe the GSE debris (dark-blue points). The major axis sticks out of the Galactic plane at ≈35°. Our derived triaxial halo parameters agree very well with those found using Gaia RR Lyrae (Iorio et al. 2018; Iorio & Belokurov 2019) even though no shape information is used to constrain the model. Bottom: 2D projections of the stellar density with the ellipsoid axes overplotted. In each panel one of the three axes closely tracks the debris density. The tilt of the ellipsoid out of the plane and its elongated morphology are clearly seen in the bottom right panel. The major axis in this panel tracks almost exactly the line joining the penultimate (2.5 Gyr) and final (2.9 Gyr) apocenters depicted as stars.

5.3. The Shape of the Inner Halo

A large body of literature has pursued the morphology of the stellar halo since it is expected to be a reasonable tracer of the much more massive DM halo (e.g., Newberg & Yanny 2006; Miceli et al. 2008; Watkins et al. 2009; Sesar et al. 2013; Posti & Helmi 2019). With Gaia it has become apparent that the inner halo (<30 kpc) is essentially built out of GSE, and so these studies were in fact measuring the distribution of GSE debris in great detail. Here we connect our fiducial model to the halo morphology literature.

In Figure 10 we show a triaxial ellipsoid fit to the simulated debris distribution. We first fix the orientation of the orthogonal axes via principal component analysis and then use the `nestle`[11] ellipsoid bounding routine to measure the relative axis ratios. To ensure the robustness of the fit, we remove the most distant 1%, 5%, and 15% of the debris and find that the axis ratios are stable to <10%. The ellipsoid is centered on the Galactic center. The orientation of the axes is described by the rotation matrix $R(\gamma, \beta, \alpha) = R_Z(\gamma)R_Y(\beta)R_X(\alpha)$, where $\gamma \approx -35°$, $\beta \approx -5°$, and $\alpha \approx -135°$ are counterclockwise yaw, pitch, and roll angles, respectively. The axis ratios, in terms of the pre-rotation axes, are $X:Y:Z = 7.9:10:4.5$. The

---

[11] https://github.com/kbarbary/nestle
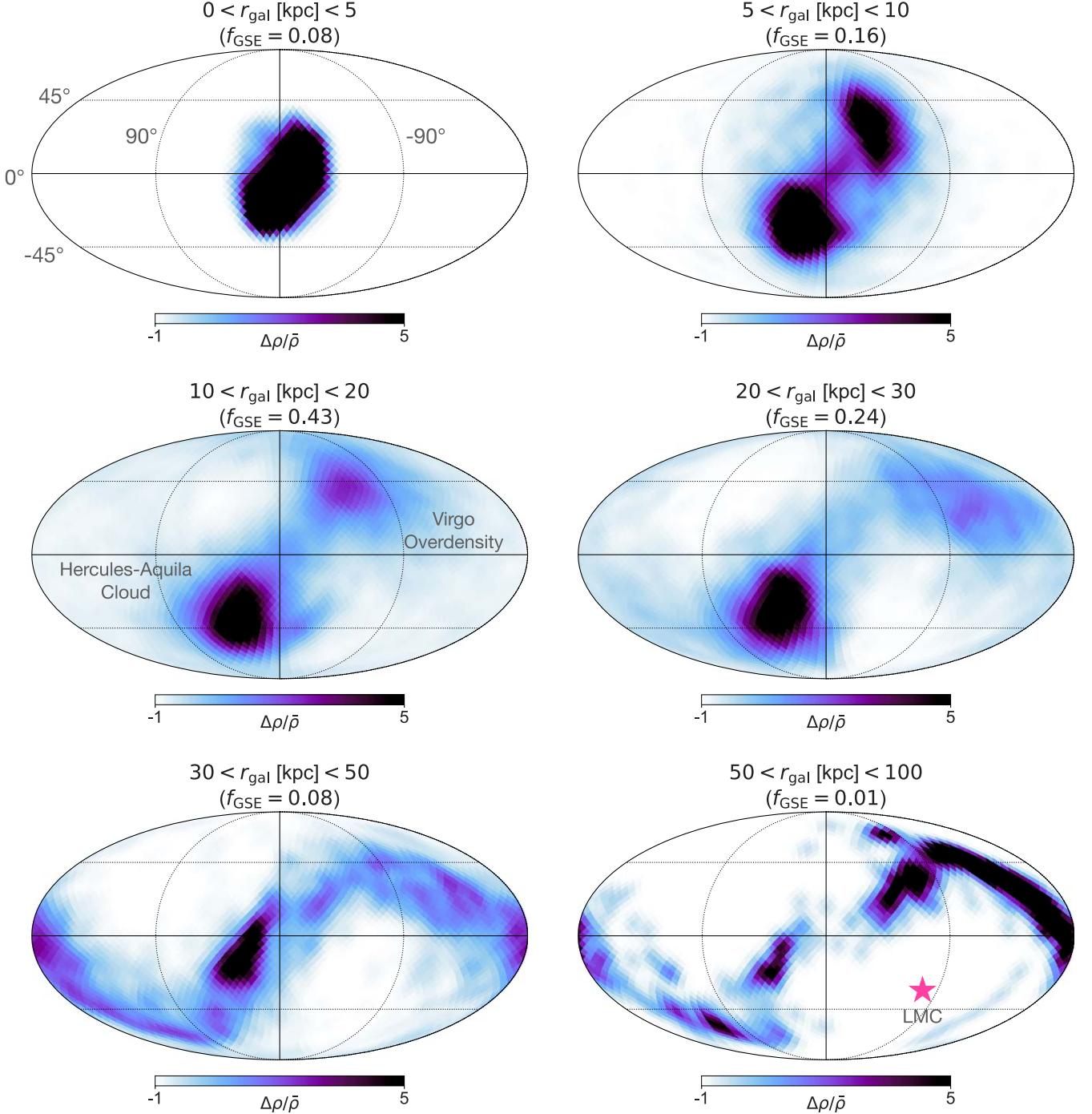
Figure 11. All-sky debris density maps from the fiducial model in Molleweide projection and Galactic coordinates smoothed with an FWHM $= 10°$ kernel. We show bins in Galactocentric distance (r_{gal}) between 0 and 100 kpc, with the fraction of debris in each bin (f_{GSE}) indicated in the title. These maps are richly structured. At $r_{gal} < 20$ kpc two prominent lobes are apparent, one above the plane and one below—these correspond to the locations of the HAC and the VOD. The northwest and southeast quadrants contain the bulk of GSE stars at all distances, underscoring the strong spatial anisotropy of the debris. The position of the LMC is indicated with a pink star in the bottom right panel.

major axis of the ellipsoid is at $\approx 35°$ to the plane. Perched on either end of the major axis are overdensities analogous to the HAC and the VOD. To visualize this debris geometry, imagine the $Y = X$ line and then lift it out of the plane by $\approx 35°$ such that it points from $(+Y, -Z)$ to $(-Y, +Z)$.

The elongated morphology of GSE debris is particularly evident in the Y-Z plane, where GSE stars almost entirely lie in two quadrants (Figure 10, bottom right panel). This geometry is

set by the locations of the final two apocenters that the bulk of stars are stripped between. The apocenter locations are highlighted as stars in the bottom panel of Figure 10—one lies above the plane, the other below. The major axis of the triaxial ellipsoid closely tracks the line joining these two apocenters.

These derived structural parameters are in good agreement with the triaxial ellipsoid model of Iorio et al. (2018) and Iorio & Belokurov (2019), who inferred the shape of the inner stellar
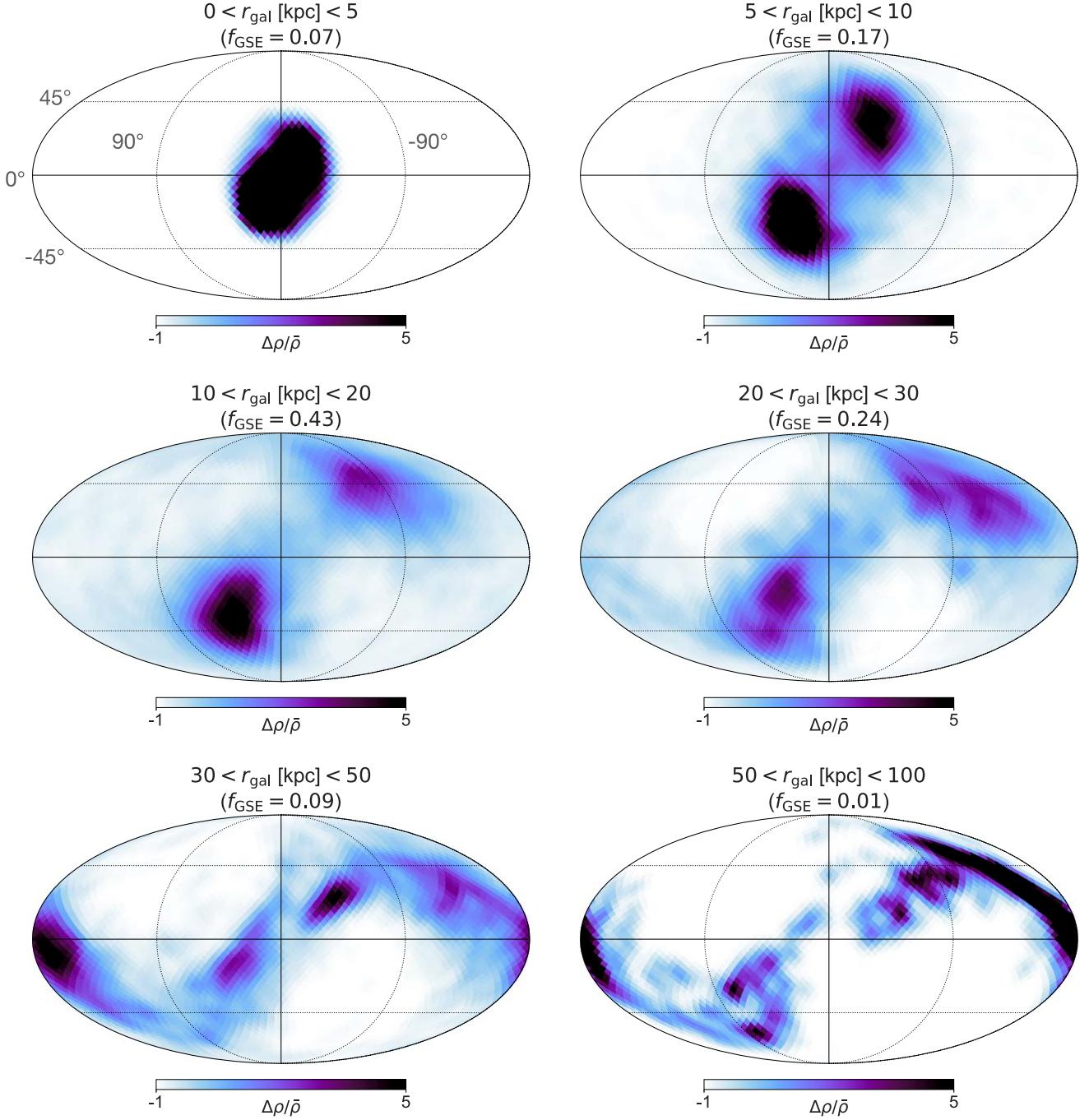
Figure 12. Same as Figure 11, but 500 Myr earlier ($t = 9.5$ Gyr) in the simulation. While the regions of the sky inhabited by GSE debris are largely unchanged and the integrals of motion (such as L_z) are stable, the relative density in these regions fluctuates as stars orbit between them. These density fluctuations are particularly dramatic in the $20 < r_{gal}$ [kpc] < 30 (middle right) and $30 < r_{gal}$ [kpc] < 50 (bottom left) bins (compare with Figure 11).

halo (<30 kpc) using a homogeneously selected all-sky data set (Gaia RR Lyrae) for the first time. These authors report remarkably similar axis ratios to those measured in our fiducial simulation (7.9:10:4.5–6.6; they allow the minor-axis ratio to vary with distance). Further, they find that the halo is at a 20° angle to the disk plane.

As hinted in Iorio & Belokurov (2019), we note that the major axis of GSE debris points in the same direction that the Magellanic Clouds entered the Galaxy from (based on the LMC orbit in Garavito-Camargo et al. 2019). A tantalizing possibility is that both GSE and the LMC traveled along the

same cosmic web filament that feeds the MW. This hypothesis would be strengthened if GSE merged on a purely radial orbit along the major axis. However, our fiducial model disfavors this scenario: the retrograde GSE enters the MW from a different direction and eventually ends up along the major axis only after being radialized (see Figure 8).

5.4. GSE throughout the Halo

Figures 11 and 12 present all-sky density maps of GSE debris from our fiducial model, split in radial bins. The defining
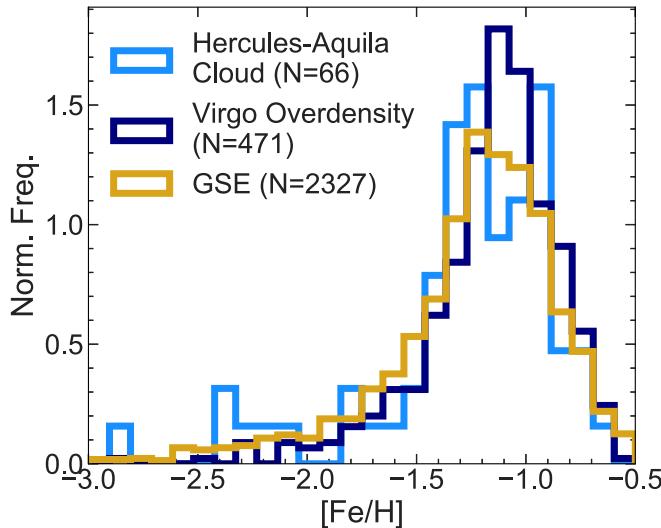
Figure 13. MDFs for the VOD (navy), HAC (light blue), and GSE (gold). HAC and VOD stars are excluded from the GSE sample to make for a clear comparison. The MDFs are well sampled and indistinguishable from each other, strongly supporting the scenario emerging from our simulations wherein the VOD and HAC are products of apocentric pileup from the GSE merger.

feature of these maps is that the GSE debris is structured and far from isotropic/axisymmetric. Figure 11 depicts the debris at $z = 0$ in our model, while Figure 12 shows the debris 500 Myr ago. Comparing these figures shows that while the regions occupied by GSE debris are stable, their relative densities fluctuate as stars orbit between these regions. Detailed density comparisons with all-sky data must take this time variability into account.

In the inner halo ($r_{\rm gal} < 30$ kpc) the GSE debris is spread across an inclined axis that runs through $l = 0°$, $b = 0°$. This is the major axis of the ellipsoid fit in Section 5.3. Occurring on either end of it are overdensities that correspond to the HAC and VOD. Stars in the present-day HAC composed the VOD 0.5 Gyr ago, and vice versa. The HAC/VOD are where the stars slow down and come to a halt before they turn around to descend/ascend the plane, and so these are the regions where stars pile up into on-sky overdensities.

As per this picture, the HAC and VOD should be chemically indistinguishable from each other and must resemble GSE. In Figure 13 we demonstrate that this is indeed the case. HAC and VOD stars are known to be eccentric (Simion et al. 2019), and their locations in Galactocentric coordinates are also known (Iorio & Belokurov 2020). Based on this, we select HAC and VOD stars from the H3 Survey giants sample as follows: HAC, $e > 0.75$, $-30 < Z_{\rm gal} < -10$, $0 < X_{\rm gal} < 20$, $0 < Y_{\rm gal} < 20$; VOD, $e > 0.75$, $10 < Z_{\rm gal} < 30$, $-20 < X_{\rm gal} < 0$, $-20 < Y_{\rm gal} < 0$. Debris from the Sagittarius dwarf galaxy and stars from the in situ halo are excluded as per Naidu et al. (2020). In Figure 13 the HAC and VOD stars have been excluded from the GSE MDF. The median metallicity and median $[\alpha/{\rm Fe}]$ of the HAC ($[{\rm Fe/H}] = -1.20^{+0.04}_{-0.02}$, $[\alpha/{\rm Fe}] = 0.21^{+0.01}_{-0.01}$) and VOD ($[{\rm Fe/H}] = -1.11^{+0.01}_{-0.02}$, $[\alpha/{\rm Fe}] = 0.20^{+0.01}_{-0.01}$) are virtually identical to GSE ($[{\rm Fe/H}] = -1.15$, $[\alpha/{\rm Fe}] = 0.21$; Naidu et al. 2020), strongly supporting a common origin for these structures.

Beyond 30 kpc, GSE stars trace stream-like patterns across the sky (bottom panels of Figures 11, 12) and are retrograde ($\langle L_z \rangle (r_{\rm gal} > 30$ kpc$) = 1.1$, $\langle L_z \rangle (r_{\rm gal} > 50$ kpc$) = 2.3$). This stream-like debris at >30 kpc arises from $(2-3) \times r_{50}$ in the GSE disk. We predict that all-sky maps of metal-rich

([Fe/H] ≈ -1.2), retrograde stars at these distances will show the diffuse "leading arm" and "trailing arm" of GSE seen in the bottom two panels. Without velocity information, this detection might be made challenging by the on-sky overlap with Sgr —$\approx 50\%$ of GSE debris beyond 30 kpc is at $|B_{\rm Sgr}| < 20°$, where $B_{\rm Sgr}$ is latitude in the Sgr plane defined in Belokurov et al. (2014). The $|B_{\rm Sgr}| < 20°$ fraction rises to $\approx 75\%$ when considering the $|l| > 90°$ regions. However, with PMs and velocities, distinguishing between the highly retrograde/radial GSE debris and the prograde Sgr debris that has high L_y (Johnson et al. 2020) will be trivial.

Also indicated in the bottom right panel is the location of the LMC. The LMC, due to its significant mass, and because it is on first infall, is predicted to induce large-scale features across the sky. In particular, Garavito-Camargo et al. (2019) forecast a "collective response" overdensity in the northern hemisphere and a dynamical friction wake overdensity in the southeast quadrant at $r_{\rm gal} \geqslant 45$ kpc. These quadrants are predicted to also harbor GSE debris ($\approx 10\%$ of the total mass) at $r_{\rm gal} = 30–100$ kpc. Similarly, efforts to constrain the barycentric motion of the MW due to the LMC by comparing radial velocities in the northern and southern hemispheres could be impacted by GSE stars at these distances (e.g., Erkal et al. 2021). At $r_{\rm gal} = 40–100$ kpc, the northern GSE stars have $V_{\rm GSR} \approx 85$ km s^{-1}, and the southern stars have $V_{\rm GSR} \approx -65$ km s^{-1}, i.e., the radial velocities of GSE stars mimic the expected LMC-induced redshift and blueshift signals. In detail these signals should be separable both because the predicted GSE debris is confined to relatively cold streams on-sky and because the predicted PM signals will differ.

5.5. A Second Apocenter at ≈15 kpc and a Double-break Inner Halo Profile

In Figure 14 we examine the density profile of GSE as a function of sky position (left panel) and integrated over the sky (right panel). We propose a "double-break" profile for the inner halo, with a prominent break at the penultimate apocenter ($\approx 28–30$ kpc) of the GSE orbit, and another break close to its final apocenter ($\approx 15–18$ kpc). Since GSE is by far the most dominant component of the inner halo ($r_{\rm gal} < 30$ kpc), we expect the overall halo density profile at these distances to largely trace the GSE profile.

Interestingly, several studies have found a "single-break" profile for the inner halo (e.g., Watkins et al. 2009; Deason et al. 2014; Xue et al. 2015). These single-break profiles do provide a reasonable fit to our model—an example (Deason et al. 2014) is shown in the bottom right panel of Figure 6. We also observe that the "single-break" inner halo profiles in the literature are divided about the location of the break, with some favoring $\approx 25–30$ kpc (e.g., Watkins et al. 2009; Deason et al. 2011; Sesar et al. 2011; Faccioli et al. 2014) and others finding $\approx 15–20$ kpc (e.g., Sesar et al. 2013; Pila-Díez et al. 2015; Xue et al. 2015). This unsettled state of affairs may be due to the fact that the halo density profile is not being well represented by a "single-break" function and also that it shows large-scale variation across the sky (left panel of Figure 14). The variation is expected from the tilted ellipsoid geometry of the debris—not only does the normalization of the profile vary by $\approx 5–10\times$, but also the shape of the profile shifts significantly from region to region. Our simulation motivates remeasuring the halo density profile, allowing for an extra break. We provide power-
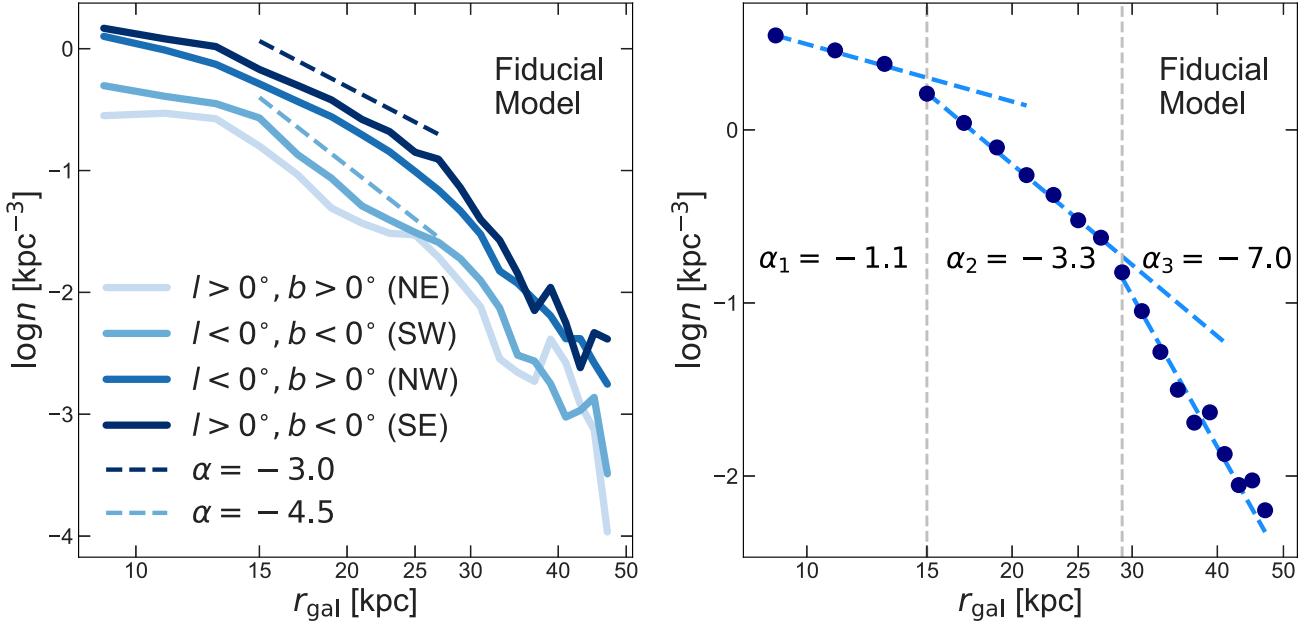
Figure 14. Left: variation of GSE density profile with on-sky location according to the fiducial model. There is a 5–10 × lower density in the NE and SW quadrants as a consequence of the triaxial GSE debris distribution (see Figure 11). Further, the density profile and location of the breaks in the profile shift across the sky—we highlight the different slopes in the 15–25 kpc range for the southeast ($\alpha = -3.0$) and southwest ($\alpha = -4.5$) quadrants of the sky. These variations have important implications for halo density profile measurements that typically probe only a fraction of the sky. Right: we propose a "double-break" all-sky density profile for GSE, with one break at ≈ 15–18 kpc and another at ≈ 30 kpc. This profile is motivated by the location of the apocenters in our simulation (Figures 6, 8).

law coefficients for our proposed $\rho \propto r_{\mathrm{gal}}^{\alpha}$ profile as a promising, physically motivated launching point for future measurements: $\alpha\,(<15\,\mathrm{kpc}) = -1.1$, $\alpha\,(15\text{–}30\,\mathrm{kpc}) = -3.3$.

5.6. Interpreting the Time Line of the GSE Merger

By measuring the SFHs of GSE and the in situ halo with precise (10% median uncertainty) ages of MSTO stars, Bonaca et al. (2020) uncovered a 2 Gyr offset between the quenching of GSE at ≈ 10 Gyr and the age of the youngest stars in the in situ halo (≈ 8 Gyr). In Figure 15 we compare the GSE orbital decay profile from our fiducial model with the observed SFHs—a unified picture that accounts for the 2 Gyr offset emerges. In particular, the offset is the gap between the first and final pericentric passages.

While GSE did not lose too many stars at its first pericentric passage (≈ 25 kpc; bottom middle panel of Figure 6), it likely lost a good fraction of its gas. This first pericentric passage is when the SFH of GSE abruptly declines (≈ 10 Gyr ago, and at 0.75 Gyr in the simulation). The final pericentric passage occurs exactly 2 Gyr later (≈ 8 Gyr ago, 2.75 Gyr in the simulation), and this is when the youngest stars in the in situ halo are kicked out of the disk. This time line also accounts for why the in situ halo contains a negligible fraction of low-α stars, since at >8 Gyr the high-α sequence was the dominant component of the disk (e.g., Lian et al. 2020). Note that the second and third pericenters in the fiducial model occur in rapid succession, separated by only ≈ 0.5 Gyr—it is thus also possible that the second pericenter produced the bulk of the in situ halo but the ages are not yet precise enough to be conclusive. The larger point is that there is a ≈ 1.5–2 Gyr lag between first pericenter and the deeper plunging orbit through the disk at the second and final pericenters, and this time lag agrees well with the age difference between GSE and the in situ halo.

Another tantalizing aspect of the Bonaca et al. (2020) in situ halo SFH is that it is not entirely smooth and shows three to four sharp, bursty spikes just after the GSE SFH begins declining. These might be genuine starbursts sparked in the early disk by GSE, as expected from simulations (e.g., Bignone et al. 2019) and seen in the case of Sagittarius's predicted orbit crossing the disk (e.g., Lian et al. 2020; Ruiz-Lara et al. 2020). Larger samples of stars with precise ages in the in situ halo will help confirm these bursts, pinpoint their exact timing, and thus provide a completely independent test of our proposed GSE orbit.

5.7. Net Rotation of GSE

The existence of Arjuna provides compelling evidence that GSE entered the MW on a highly retrograde orbit, even though the bulk of its present-day debris is radial. We demonstrate this in Figure 16, where in the top panel we plot the evolution of $\langle L_z \rangle$ for GSE stars with time. GSE has an initial $\langle L_z \rangle \approx 6000\,\mathrm{kpc\,km\,s^{-1}}$, but in a few Gyr it is radialized to $\langle L_z \rangle \approx 0$.

In the bottom panel of Figure 16 we plot $\langle L_z \rangle$ for the GSE debris as a function of r_{gal}. While there is very little mean rotation within 25 kpc, at larger distances GSE debris grows increasingly retrograde, reaching $\approx 750\,\mathrm{kpc\,km\,s^{-1}}$ by $r_{\mathrm{gal}} \approx 30$–50 kpc. This increase corresponds to a larger fraction of stars that were stripped early in the merger, when the bulk motion of GSE was still retrograde. Interestingly, the all-sky net rotation is higher than in the H3 sample by a factor of $\approx 2\times$. This can be understood via the bottom panels of Figure 11 that depict all-sky maps of the GSE debris at $r_{\mathrm{gal}} > 30$ kpc. The highly retrograde "arms" of debris at $|l| > 90°$ lie at $|b| < 40°$, resulting in a less retrograde $\langle L_z \rangle$ within the H3 footprint.

The transition from radial to retrograde rotation in the bottom panel of Figure 16 may remind readers of the
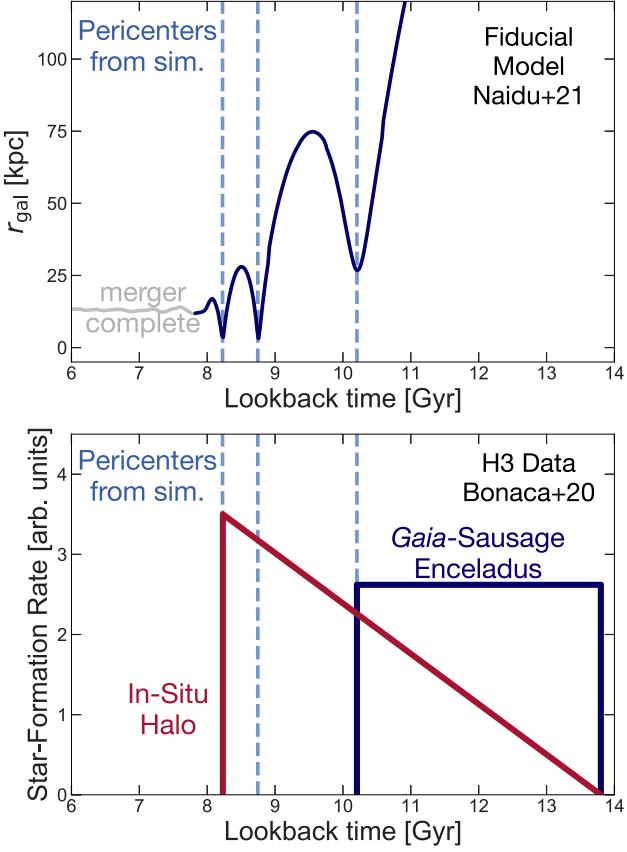
Figure 15. Orbital decay profile of GSE from our fiducial simulation (top) compared with the Bonaca et al. (2020) SFHs of GSE and the in situ halo (bottom). The first pericenter from our simulation is assigned a look-back time coincident with the quenching of GSE in the data (10.2 Gyr). The final pericenter, occurring ≈2 Gyr later, lines up remarkably well with the truncation of the SFH of the in situ halo, suggesting a causal relationship. After this pericenter, what is left of GSE is no longer massive/dense enough to kick stars out of the disk into the in situ halo (bottom middle panel, Figure 6).

"dual-halo" scenario (Carollo et al. 2007, 2010; Beers et al. 2012). These authors integrated orbits of local halo samples ($d_{helio} < 4$ kpc) to infer that the halo was composed of an "inner halo" ($r_{gal} \lesssim 15$ kpc, [Fe/H] = −1.6, small net prograde rotation) and an "outer halo" ($r_{gal} \sim 20$–50 kpc, [Fe/H] = −2.2, mean retrograde rotation). Updating this analysis with Gaia, Carollo & Chiba (2021) observe that GSE stars and other disk populations may constitute the inner halo, while the outer halo is composed of a variety of retrograde structures. The radial trend in GSE $\langle L_z \rangle$ seen in our data and simulation is in qualitative agreement with the dual-halo scenario. However, the metallicity of GSE ($\langle [Fe/H] \rangle = -1.15$) and its flat gradient (see Section 6) do not fit neatly with either the Carollo et al. (2010) inner or outer halo. Further work is needed to understand the relationship between local halo samples and the global halo. For now it is clear that directly surveying the distant halo is critical to recovering the highly retrograde debris of GSE, since it is much more prominent beyond the solar circle (the $L_z > 1.5$ fraction at $d_{helio} < 5$ kpc is ≈10× lower than at $d_{helio} > 30$ kpc).

5.8. Sequoia and I'itoi as Disrupted Satellites of GSE

Apart from Arjuna, Sequoia and I'itoi are the other prominent structures in the high-energy retrograde halo.
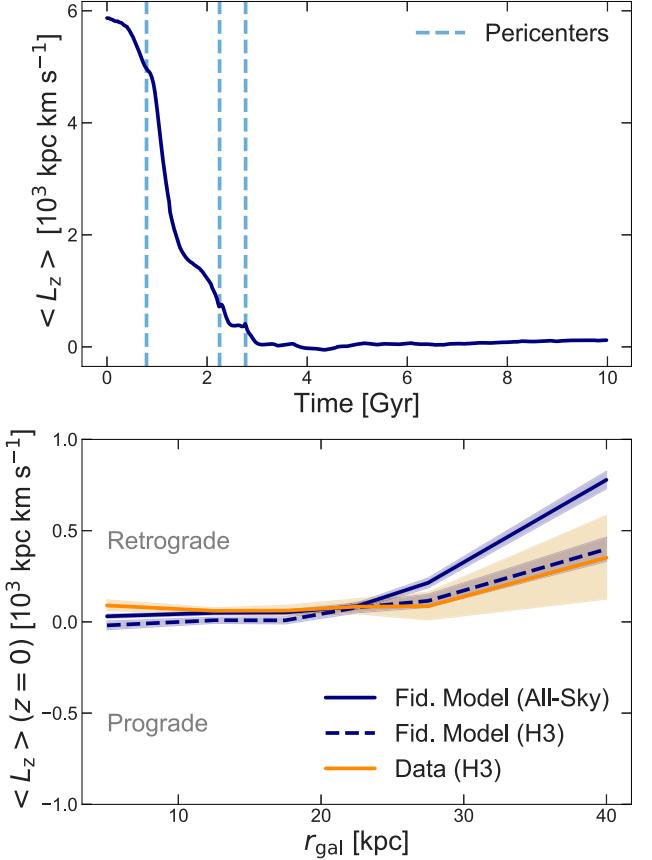

Figure 16. Top: evolution of the GSE mean angular momentum, $\langle L_z \rangle$, in the fiducial simulation. The mean momentum is highly retrograde at infall, but following radialization by dynamical friction, the debris today has $\langle L_z \rangle$ of ≈170 kpc km s^{-1}. Bottom: present-day $\langle L_z \rangle$ as a function of distance. $\langle L_z \rangle$ is very weakly retrograde within 25 kpc (<100 kpc km s^{-1}). At larger distances, as the fraction of debris stripped early in the merger increases, the net rotation grows increasingly retrograde. Interestingly, $\langle L_z \rangle$ averaged over the sky (solid navy) is more retrograde than within the H3 footprint (dashed navy)—this can be understood via Figure 11, where we see that the stream-like, highly retrograde debris at $|l| > 90°$ occurs outside the survey footprint.

Chemical analyses show Sequoia to have an [Fe/H] ≈ −1.6, exactly as seen in Figure 1, and that it has a "knee" characteristic of dwarf galaxies, distinct from the GSE knee, in the [Fe/H] versus [α/Fe] plane (e.g., Matsuno et al. 2019; Monty et al. 2020; Aguado et al. 2021). Nonetheless, it has been argued that Sequoia might not be a dwarf galaxy at all, and that it may in fact be debris from the outer regions of GSE, which in this scenario has a steep metallicity gradient (Koppelman et al. 2019b; Helmi 2020; Koppelman et al. 2020).

We offer an alternative scenario, wherein GSE has a rather flat metallicity gradient. This is supported by Arjuna's [Fe/H] that arises from the outer disk being similar to the radial debris' [Fe/H] that arises from the inner disk (Figure 9). If GSE has a stellar mass of $5 \times 10^8 M_\odot$ as our numerical experiments and literature constraints suggest, then the stellar mass of Sequoia must be ≈$10^7 M_\odot$ as per the relative star counts of these two structures (<1/42) in Naidu et al. (2020). Note that this is 5–10× less massive than estimated in Myeong et al. (2019) and Kruijssen et al. (2020). The low mass and highly retrograde phase-space position of Sequoia are consistent with it being a satellite of GSE (≈1:10 by total mass according to the Behroozi et al. 2019 stellar mass–halo mass relation). As we have shown
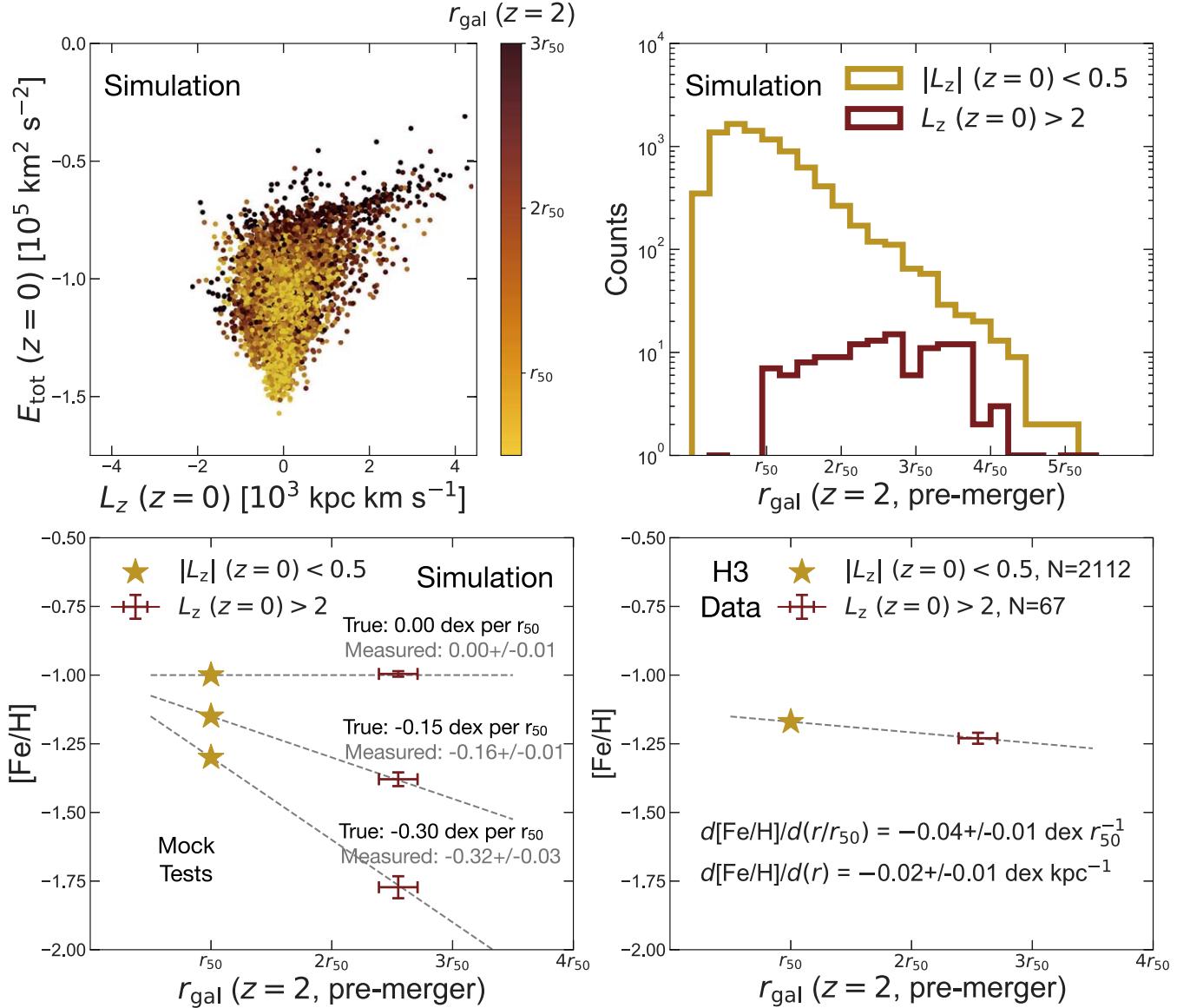
Figure 17. Inferring the [Fe/H] gradient of GSE. Top left: $E - L_z$ diagram of GSE debris from the fiducial simulation, with stars colored by their location in the pre-merger GSE disk. Highly retrograde, high-energy stars preferentially arise from the outer disk. Top right: radial distribution of GSE debris in the pre-merger disk ($r_{gal}(z=2)$) in two bins of present-day angular momenta ($L_z(z=0)$). These distributions quantify the trend seen in the left panel and motivate a mapping between $r_{gal}(z=2)$ and $L_z(z=0)$. Bottom left: we exploit the trend between $L_z(z=0)$ and $r_{gal}(z=2)$ to measure [Fe/H] gradients across the GSE disk. Radial debris ($|L_z| < 0.5$) that traces r_{50} is shown as a golden star, whereas highly retrograde debris that traces $2.5 \times r_{50}$ ($|L_z| < 0.5$) is shown in brown. The radial debris is sampled by >2000 stars, so the error on the mean is smaller than a point on this plot, while the retrograde debris is sampled with 67 stars. The mock tests show excellent recovery of the true gradient (see Section 6.1). Bottom right: the [Fe/H] gradient in GSE is measured between the highly retrograde debris that preferentially arises from the outer disk and radial debris that largely arises from the inner disk. The inferred gradient is fairly shallow and is interpreted in Section 6.2.

in this work, stars from the outer regions of GSE end up preferentially on high-energy, retrograde orbits (e.g., Figure 17) —one would expect this trend to also hold for a satellite stripped from the outer regions of GSE. This argument also applies to the more metal-poor I'itoi in Figure 1, which is chemically distinct and consistent with being a dwarf, while showing integrals of motion indistinguishable from those of Arjuna and Sequoia. Detailed satellite-of-satellite simulations and ages for Sequoia and I'itoi stars are important to test this scenario.

6. The Metallicity Gradient of GSE

In the previous section we discussed the properties of a particular simulation that matches a variety of observational

constraints. A key feature of this simulation is that the retrograde debris was stripped first and was on average at greater distances within the progenitor system than the radial debris. In this section we exploit this property to infer the metallicity *gradient* within the progenitor system.

6.1. Method and Measurement

In Figure 17 we quantify the radial distribution of stars in the pre-merger disk as a function of present-day angular momenta in our fiducial model. We find that the $|L_z| < 0.5$ stars within the H3 footprint arise from a mean radius of $1.00^{+0.02}_{-0.02} \times r_{50}$ in the GSE disk, whereas the $L_z > 2$ stars arise from a mean radius of $2.55^{+0.16}_{-0.15} \times r_{50}$. The choice to compare $L_z > 2$ stars with

$L_z < 0.5$ stars is to maximize the contrast in the average pre-merger disk location.

We perform mock tests where we inject flat (0 dex per r_{50}), moderate (-0.1 dex per r_{50}), and steep (-0.3 dex per r_{50}) metallicity gradients into our model and then attempt to recover them with an H3-like survey. We assume 0.05 dex uncertainties on individual measurements of [Fe/H] (the median uncertainty of our sample), assume a 10% distance error while computing L_z, and draw the exact number of stars as in the data (2112 at $|L_z| < 0.5$ and 67 at $L_z > 2$). In all cases we were able to recover the true metallicity gradient across the disk by comparing the observed $|L_z| < 0.5$ and $L_z > 2$ stars within 10% (bottom left panel, Figure 17). In the regime of very steep, unphysical gradients (e.g., -1 dex r_{50}^{-1}) the method overestimates the steepness of the metallicity gradient by $\approx 10\%$ because the averages no longer capture the rapid changes across the disk. A hint of this is seen in the -0.3 dex per r_{50} curve in Figure 17. Nonetheless, for the physical regimes of interest, these tests show that angular momenta act as a superb proxy for the average location of stars in the pre-merger disk.

From Figure 1 it is already apparent that the metallicity gradient between the inner and outer disk populations, i.e., $|L_z| < 0.5$ ($\approx 1 \times r_{50}$) and $|L_z| > 2$ ($\approx 2.5 \times r_{50}$) stars, must be weak. The bootstrapped mean metallicity of the $|L_z| < 0.5$ GSE debris is [Fe/H] $= -1.17_{-0.01}^{+0.01}$, and that of the $|L_z| > 2$ GSE +Arjuna debris is [Fe/H] $= -1.23_{-0.02}^{+0.02}$. This translates to a weak metallicity gradient of -0.04 ± 0.01 dex r_{50}^{-1}, comparable to Fornax (-0.02 ± 0.02 dex r_{50}^{-1}) and Ursa Minor (-0.06 ± 0.05 dex r_{50}^{-1}), which have the weakest [Fe/H] gradients of the Kirby et al. (2011) dwarfs. The corresponding [α/Fe] gradient, also very weak, is $+0.02 \pm 0.01$ dex r_{50}^{-1}.

Our definition for Arjuna truncates its MDF at [Fe/H] < -1.5 to avoid contamination from Sequoia, which introduces a bias in the mean [Fe/H]. Simultaneously, we also need to account for Sequoia stars at [Fe/H] > -1.5 that bias us to lower [Fe/H]. By fitting two Gaussians to the Sequoia and Arjuna MDFs in Figure 1, we find that these effects cancel out and the mean metallicity of the GSE+Arjuna $L_z > 2$ debris shifts imperceptibly within our stated errors ([Fe/H] $= -1.22_{-0.02}^{+0.02}$). Also note that the translation between L_z and r_{50} is entirely dependent on our merger model—our error bars do not reflect the systematic uncertainty that our model may not be the only model that fits the constraints.

6.2. Broader Context: Reconstructing the Stellar Metallicity Gradient of a z ≈ 2 Galaxy

Radial metallicity gradients probe the interplay between star formation, feedback, and inflows and thus provide a sensitive test for how galaxies assemble their baryons across cosmic time (for recent reviews see Kewley et al. 2019; Maiolino & Mannucci 2019; Sánchez 2020). While relatively well studied locally, gradients at higher redshifts are challenging owing to the difficulty of resolving faint objects at kiloparsec scales. Further, existing high-z gradients are derived via nebular emission and are thus susceptible to the many systematics inherent to measuring gas-phase metallicities (e.g., uncertainties in the ionization parameter, biases from sampling only active star-forming regions, probing only the instantaneous metallicity).

Through our rough reconstruction of the GSE disk structure we are able to effectively access the *stellar* metallicity gradient of a $z \approx 2$ star-forming galaxy via its $z = 0$ debris. The star

formation of GSE was abruptly truncated at $z \approx 2$ (Bonaca et al. 2020), around its first pericenter, shortly before it was shredded (Figure 15). GSE stars inhabiting the MW halo today retain a snapshot of their $z \approx 2$ chemical state. At similar redshifts, a stellar metallicity gradient has been measured in only one other galaxy—a highly lensed ($>10\times$), very bright ($H = 17.1$, $M_\star \approx 6 \times 10^{11}\ M_\odot$), $z = 1.98$ quiescent system (Jafariyazani et al. 2020). We also note that the quenched ultrafaint dwarfs retain a similar chemical record of the very early universe, though they likely probe higher redshifts (e.g., $z > 6$, corresponding to the epoch of reionization; Brown et al. 2014) and several dex lower halo masses (e.g., Simon 2019).

The weak, negative metallicity gradient of GSE validates the emerging observational picture that (gas-phase) gradients are generally flat at high z and grow steeper with time (e.g., Leethochawalit et al. 2016; Förster Schreiber et al. 2018; Curti et al. 2020). It is also in line with dwarf simulations that produce steep gradients only toward lower redshift owing to the accumulated effect of feedback-driven puffing of centrally concentrated, ancient metal-poor populations (e.g., El-Badry et al. 2016; Ma et al. 2017a; El-Badry et al. 2018; Mercado et al. 2021). This picture also fits well with the steep gradient inferred for the Sgr dwarf galaxy (e.g., Hayes et al. 2020), which has a comparable stellar and halo mass to GSE (e.g., Johnson et al. 2020) but a lower accretion redshift ($z \approx 0.5$) and $\gtrsim 5$ additional Gyr of star formation and evolution (e.g., Alfaro-Cuello et al. 2019; Lian et al. 2020; Ruiz-Lara et al. 2020).

Stellar metallicity gradients at the redshift and mass range studied in this work ($z \approx 2$, $M_\star \approx 5 \times 10^8\ M_\odot$) will be generally inaccessible even to JWST and upcoming ELTs (Extremely Large Telescopes), underscoring the immense promise of "near-field galaxy evolution" as a complementary route to studying the high-z universe through halo debris (e.g., Weisz et al. 2014; Boylan-Kolchin et al. 2015, 2016). GSE stars are beginning to be used in "$z \approx 2$" studies to understand the chemistry of the early universe (Molaro et al. 2020; Simpson et al. 2021; Matsuno et al. 2021)—our simulations will add rich context to these works (e.g., by mapping the phase space of GSE stars to their pre-merger disk location and the time they were stripped). Similar reconstructions will soon be possible for other less massive dwarfs accreted at a variety of redshifts as our census of halo substructure grows more complete.

7. Summary

We have used the H3 Survey to study the $z \approx 2$ GSE merger. Our unique sample of ≈ 2800 GSE stars has full 6D phase-space data and abundances ([Fe/H], [α/Fe]), is unbiased in metallicity, and encompasses the farthest reaches of GSE debris. We systematically explore a large grid (≈ 500) of high-resolution ($10^5\ M_\odot$) N-body simulations to reproduce the H3 constraints (summarized in Section 3) and resimulate the most promising configurations ($N \approx 20$) at an even higher resolution ($10^4\ M_\odot$). Our grid spans a plausible range of physical (size, mass) and orbital (circularity, inclination, disk spin, sense of orbit) parameters (Tables 1, 2, 3). We find that the merger and its resultant debris have the following characteristics:

1. GSE arrived in the Galaxy on a highly retrograde orbit. The Arjuna structure identified in Naidu et al. (2020) is the retrograde debris of GSE. Despite harboring some of the most retrograde stars in the halo, the [Fe/H] and [α/Fe] of this structure are within ≈ 0.05 dex of the

radial, eccentric ($e > 0.7$) GSE stars, strongly suggesting that they are associated. (Figure 1, Section 2)

2. A GSE of mass $M_\star = 5 \times 10^8\ M_\odot$, $M_{DM} = 2 \times 10^{11}\ M_\odot$ with an $r_{50} = 2.5$ kpc that merges on a retrograde orbit with a circularity of 0.5, inclination of 15°, and retrograde disk spin best reproduces the H3 constraints. The GSE merger was a 2.5:1 merger—its debris composes $\approx 20\%$ of the $z = 0$ MW DM halo and $\approx 50\%$ of its stellar halo. (Figures 5, 6, Section 5.1)

3. The retrograde Arjuna stars preferentially arise from the outer disk of GSE ($\approx 2.5 \times r_{50}$) in our fiducial simulation. These loosely bound stars are stripped early in the merger, before GSE has been radialized by dynamical friction, and so they reflect the initial retrograde angular momentum. (Figure 9, Section 5.2)

4. The net rotation of GSE at $z \approx 2$ was $\langle L_z \rangle \approx$ 6000 kpc km s^{-1}, but by $z = 0$ it is largely radialized ($\langle L_z \rangle \approx$ 170 kpc km s^{-1}). As a function of distance, rotation of GSE debris is weak within 25 kpc ($\langle L_z \rangle < 100$ kpc km s^{-1}) but grows to $\langle L_z \rangle \approx 750$ kpc km s^{-1} at $r_{gal} \approx 30$–50 kpc as the fraction of debris stripped early in the merger grows with distance. (Figure 16, Section 5.7)

Even though our fiducial simulation is selected purely based on the H3 Survey r_{gal} and L_z distributions, it self-consistently reproduces and explains the following phenomena:

1. The shape of the inner halo (<30 kpc) is set by GSE, which is its most dominant constituent. GSE debris defines an elongated triaxial ellipsoid with axis ratios 10:7.9:4.5, in remarkable agreement with Gaia RR Lyrae constraints (Iorio & Belokurov 2019). The major axis of the debris is at $\approx 35°$ to the disk plane and at $\approx 45°$ to the Galactic X/Y direction. The orientation tracks the second and final apocenters that occur on either side of the plane. GSE loses most of its stars between these apocenters. The tilted triaxial halo is a significant departure from planar, prolate models typically used to model the Galactic stellar and DM halos. (Figures 10, 11, Section 5.3)

2. The HAC and VOD occur on either end of the triaxial ellipsoid's major axis and emerge owing to apocenter pileup of GSE debris. The HAC and VOD are spatially proximal to the penultimate and final apocenter of the GSE orbit, respectively. We provide strong evidence for this association by showing that the MDFs of the HAC, the VOD, and GSE are indistinguishable from each other. (Figures 7, 13)

3. The ≈ 2 Gyr gap between the quenching of GSE and the cessation of star formation in the in situ halo (Bonaca et al. 2020) is precisely the gap between the first and final pericentric passages. At first pericenter the star formation within GSE is truncated, and by final pericenter it is no longer massive enough to kick stars out of the primordial disk into the in situ halo. (Figure 15, Section 5.6)

We make the following predictions based on our fiducial simulation:

1. The inner halo has a "double-break" density profile. The two breaks (at ≈ 15 kpc, ≈ 30 kpc) correspond to the second and final apocenters and are described by the following power-law ($\rho \propto r_{gal}^\alpha$) coefficients: $\alpha\,(<15\ \mathrm{kpc}) = -1.1$, $\alpha\,(15$–$30\ \mathrm{kpc}) = -3.3$. The profile shows strong variations across the sky in both the normalization (≈ 5–$10 \times$) and the

shape ($\Delta\alpha \approx 1.5$). Interestingly, several studies have fit single-break profiles for the inner halo but disagree about the break location, with some finding a break at ≈ 15–20 kpc and others at ≈ 25–30 kpc. Our proposed profile may resolve this tension. (Figure 14, Section 5.5)

2. The outer halo ($r_{gal} > 30$ kpc) contains $\approx 10\%$ of the GSE stellar mass. This debris manifests as highly retrograde, stream-like structures that await discovery. Approximately 50% of this debris lies within 20° of the Sgr orbital plane. (Figures 11, 12, Section 5.4)

3. The Sequoia and I'itoi dwarfs, which have integrals of motion virtually indistinguishable from Arjuna, may also have been stripped from the outer regions of GSE. These systems may have once constituted a group like the Magellanic Clouds. (Section 5.8)

Finally, we use our fiducial simulation to reconstruct the stellar metallicity gradient in a $z \approx 2$ star-forming galaxy (GSE):

1. Radial GSE debris ($|L_z| < 0.5$) originates from the inner disk ($\approx r_{50}$), while retrograde debris ($L_z > 2$) arises from the outer disk ($\approx 2.5 \times r_{50}$). Capitalizing on this trend, we measure a stellar metallicity gradient of -0.04 ± 0.01 dex r_{50}^{-1} and [α/Fe] gradient of $+0.02 \pm 0.01$ dex r_{50}^{-1}. Stellar abundance gradients for star-forming galaxies at $z \approx 2$ will be inaccessible even to JWST—our measurement underscores the immense promise of "near-field galaxy evolution" with halo debris as a complementary route to the high-z universe. (Figure 17, Section 6.1)

We once again emphasize that our fiducial simulation is *a* possible configuration for the GSE merger and not necessarily *the* configuration. However, it is quite successful at not only replicating the H3 data but also reproducing and explaining disparate phenomena across the Galaxy. Further, it makes specific, verifiable predictions that can be tested with existing and upcoming data sets. We foresee this model being used to drive progress on multiple fronts. For instance, the detailed phase-space distribution of the enormous amount of GSE DM left unexplored in this work could prove critical to designing DM detection experiments and informing realistic models of the MW potential. The in situ halo/splashed disk can be developed into a sensitive probe of the physical and chemical structure of the primordial MW disk. Taken together, GSE ($2 \times 10^{11}\ M_\odot$; this work), Sgr ($\approx 1 \times 10^{11}\ M_\odot$; e.g., Johnson et al. 2020), and the LMC ($\approx 1.3 \times 10^{11}\ M_\odot$; e.g., Erkal et al. 2019) account for almost the entirety of the growth of the MW since $z \approx 2$—the census of the MW's significant mergers in the last 10 Gyr is now complete.

Figure 18. Effect of varying the bulge fraction of GSE—100% of stars are in a bulge (top row), 70% of stars are in a bulge (bottom row). See Figure 6 for description of panels. We conclude that the detailed structure of the satellite plays a secondary role to the total mass and orbital parameters in determining the two key variables on which we judge the quality of models (L_z and r_{gal}).

Appendix A
Satellite Models with Bulges

In Figure 18 we present results from two $10^5 \, M_\odot$ resolution simulations exploring variations in the structure of GSE. To bracket the realm of possibilities, we consider one model wherein all the stars reside in a bulge (top row), another in which 30% of the stars reside in a disk and 70% in a bulge (bottom row, similar to our MW model), whereas in the fiducial model all GSE stars are in a disk. The scale radius of the bulge is adjusted such that its r_{50} matches that of our fiducial model. All other structural and orbital parameters are kept fixed to those of the fiducial simulation. These two simulations produce broadly similar r_{gal} and L_z distributions as our fiducial model. The upshot of this exercise is that the details of the satellite structure play a secondary role to the total mass and orbital parameters we zeroed in on, whose robustness we have validated through the diverse set of constraints we are able to match.

ORCID iDs

Rohan P. Naidu ⓘ https://orcid.org/0000-0003-3997-5705
Charlie Conroy ⓘ https://orcid.org/0000-0002-1590-8551
Ana Bonaca ⓘ https://orcid.org/0000-0002-7846-9787
Dennis Zaritsky ⓘ https://orcid.org/0000-0002-5177-727X
Rainer Weinberger ⓘ https://orcid.org/0000-0001-6260-9709
Yuan-Sen Ting (丁源森) ⓘ https://orcid.org/0000-0001-5082-9536
Nelson Caldwell ⓘ https://orcid.org/0000-0003-2352-3202

Sandro Tacchella ⓘ https://orcid.org/0000-0002-8224-4505
Joshua S. Speagle ⓘ https://orcid.org/0000-0003-2573-9832
Phillip A. Cargile ⓘ https://orcid.org/0000-0002-1617-8917

References

Aguado, D. S., Belokurov, V., Myeong, G. C., et al. 2021, ApJL, 908, L8
Alfaro-Cuello, M., Kacharov, N., Neumayer, N., et al. 2019, ApJ, 886, 57
Amorisco, N. C. 2017, MNRAS, 464, 2882
An, D., & Beers, T. C. 2021, ApJ, 907, 101
Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, Natur, 561, 360
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Beaumont, C., Goodman, A., & Greenfield, P. 2015, in ASP Conf. Ser. 495, Astronomical Data Analysis Software an Systems XXIV (ADASS XXIV), ed. A. R. Taylor & E. Rosolowsky (San Francisco, CA: ASP), 101
Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, ApJ, 746, 34
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., Evans, N. W., Irwin, M. J., et al. 2007, ApJ, 658, 337
Belokurov, V., Koposov, S. E., Evans, N. W., et al. 2014, MNRAS, 437, 116
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Besla, G., Kallivayalil, N., Hernquist, L., et al. 2007, ApJ, 668, 949
Besla, G., Kallivayalil, N., Hernquist, L., et al. 2010, ApJL, 721, L97
Bignone, L. A., Helmi, A., & Tissera, P. B. 2019, ApJL, 883, L5
Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, ApJ, 773, 43
Bird, S. A., Xue, X.-X., Liu, C., et al. 2019, AJ, 157, 104
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bland-Hawthorn, J., Sharma, S., Tepper-Garcia, T., et al. 2019, MNRAS, 486, 1167
Bonaca, A., Conroy, C., Cargile, P. A., et al. 2020, ApJL, 897, L18
Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Kereš, D. 2017, ApJ, 845, 101
Bonaca, A., Jurić, M., Ivezić, Ž., et al. 2012, AJ, 143, 105
Bonaca, A., Naidu, R. P., Conroy, C., et al. 2021, ApJL, 909, L26
Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, ApJL, 707, L1
Boylan-Kolchin, M., Weisz, D. R., Bullock, J. S., & Cooper, M. C. 2016, MNRAS, 462, L51
Boylan-Kolchin, M., Weisz, D. R., Johnson, B. D., et al. 2015, MNRAS, 453, 1503
Brown, T. M., Tumlinson, J., Geha, M., et al. 2014, ApJ, 796, 91
Cargile, P. A., Conroy, C., Johnson, B. D., et al. 2020, ApJ, 900, 28
Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Natur, 450, 1020
Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692
Carollo, D., & Chiba, M. 2021, ApJ, 908, 29
Cautun, M., Benítez-Llambay, A., Deason, A. J., et al. 2020, MNRAS, 494, 4291
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Conroy, C., Bonaca, A., Cargile, P., et al. 2019a, ApJ, 883, 107
Conroy, C., Naidu, R. P., Zaritsky, D., et al. 2019b, ApJ, 887, 237
Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, RAA, 12, 1197
Curti, M., Maiolino, R., Cirasuolo, M., et al. 2020, MNRAS, 492, 821
Das, P., & Binney, J. 2016, MNRAS, 460, 1725
Das, P., Hawkins, K., & Jofré, P. 2020, MNRAS, 493, 5195
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
Deason, A. J., Belokurov, V., & Evans, N. W. 2011, MNRAS, 416, 2903
Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018, ApJL, 862, L1
Deason, A. J., Belokurov, V., Koposov, S. E., & Rockosi, C. M. 2014, ApJ, 787, 30
Deason, A. J., Belokurov, V., & Sanders, J. L. 2019, MNRAS, 490, 3426
Deason, A. J., Belokurov, V., & Weisz, D. R. 2015, MNRAS, 448, L77
Deason, A. J., Erkal, D., Belokurov, V., et al. 2021, MNRAS, 501, 5964
Diemer, B., & Joyce, M. 2019, ApJ, 871, 168
Dierickx, M. I. P., & Loeb, A. 2017, ApJ, 836, 92
Donlon, T. I., Newberg, H. J., Sanderson, R., & Widrow, L. M. 2020, ApJ, 902, 119
Donlon, T. I., Newberg, H. J., Weiss, J., Amy, P., & Thompson, J. 2019, ApJ, 886, 76
Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

El-Badry, K., Bland-Hawthorn, J., Wetzel, A., et al. 2018, MNRAS, 480, 652
El-Badry, K., Wetzel, A., Geha, M., et al. 2016, ApJ, 820, 131
Elias, L. M., Sales, L. V., Helmi, A., & Hernquist, L. 2020, MNRAS, 495, 29
Erkal, D., Belokurov, V., Laporte, C. F. P., et al. 2019, MNRAS, 487, 2685
Erkal, D., Deason, A. J., Belokurov, V., et al. 2021, MNRAS, 506, 2677
Evans, N. W. 2020, in Galactic Dynamics in the Era of Large Surveys, Proc. IAU 353 (Cambridge: Cambridge Univ. Press)
Evans, N. W., O'Hare, C. A. J., & McCabe, C. 2019, PhRvD, 99, 023012
Faccioli, L., Smith, M. C., Yuan, H. B., et al. 2014, ApJ, 788, 105
Fattahi, A., Belokurov, V., Deason, A. J., et al. 2019, MNRAS, 484, 4471
Fernández-Alvar, E., Carigi, L., Schuster, W. J., et al. 2018, ApJ, 852, 50
Feuillet, D. K., Feltzing, S., Sahlholdt, C., & Casagrande, L. 2020, MNRAS, 497, 109
Font, A. S., McCarthy, I. G., Crain, R. A., et al. 2011, MNRAS, 416, 2802
Forbes, D. A. 2020, MNRAS, 493, 847
Forbes, J., Krumholz, M., & Burkert, A. 2012, ApJ, 754, 48
Förster Schreiber, N. M., Renzini, A., Mancini, C., et al. 2018, ApJS, 238, 21
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, NatAs, 3, 932
Garavito-Camargo, N., Besla, G., Laporte, C. F. P., et al. 2019, ApJ, 884, 51
Grand, R. J. J., Kawata, D., Belokurov, V., et al. 2020, MNRAS, 497, 1603
Gudin, D., Shank, D., Beers, T. C., et al. 2021, ApJ, 908, 79
Hayes, C. R., Majewski, S. R., Hasselquist, S., et al. 2020, ApJ, 889, 63
Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, ApJ, 863, 113
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
Hernquist, L. 1990, ApJ, 356, 359
Hunter, J. D. 2007, CSE, 9, 90
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Natur, 370, 194
Iorio, G., & Belokurov, V. 2019, MNRAS, 482, 3868
Iorio, G., & Belokurov, V. 2021, MNRAS, 502, 5686
Iorio, G., Belokurov, V., Erkal, D., et al. 2018, MNRAS, 474, 2142
Jafariyazani, M., Newman, A. B., Mobasher, B., et al. 2020, ApJL, 897, L42
Jiang, L., Cole, S., Sawala, T., & Frenk, C. S. 2015, MNRAS, 448, 1674
Johnson, B. D., Conroy, C., Naidu, R. P., et al. 2020, ApJ, 900, 103
Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511
Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102
Kirby, E. N., Lanfranchi, G. A., Simon, J. D., Cohen, J. G., & Guhathakurta, P. 2011, ApJ, 727, 78
Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, in Positioning and Power in Academic Publishing: Players, Agents and Agendas, ed. F. Loizides & B. Schmidt (Amsterdam: IOS Press), 87
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. 2019a, A&A, 631, L9
Koppelman, H. H., Helmi, A., Massari, D., Roelenga, S., & Bastian, U. 2019b, A&A, 625, A5
Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, MNRAS, 498, 2472
Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019, MNRAS, 486, 3180
Lancaster, L., Koposov, S. E., Belokurov, V., Evans, N. W., & Deason, A. J. 2019, MNRAS, 486, 378
Lang, P., Wuyts, S., Somerville, R. S., et al. 2014, ApJ, 788, 11
Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018, MNRAS, 481, 286
Law, D. R., & Majewski, S. R. 2010, ApJ, 714, 229
Lee, D. M., Johnston, K. V., Sen, B., & Jessop, W. 2015, ApJ, 802, 48
Leethochawalit, N., Jones, T. A., Ellis, R. S., et al. 2016, ApJ, 820, 84
Lian, J., Thomas, D., Maraston, C., et al. 2020, MNRAS, 494, 2561
Limberg, G., Rossi, S., Beers, T. C., et al. 2021, ApJ, 907, 10
Lynden-Bell, D. 1975, VA, 19, 299
Ma, X., Hopkins, P. F., Feldmann, R., et al. 2017a, MNRAS, 466, 4780
Ma, X., Hopkins, P. F., Kasen, D., et al. 2016, MNRAS, 459, 3614
Ma, X., Hopkins, P. F., Wetzel, A. R., et al. 2017b, MNRAS, 467, 2430
Mackereth, J. T., & Bovy, J. 2020, MNRAS, 492, 3631
Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, MNRAS, 482, 3426
Maiolino, R., & Mannucci, F. 2019, A&ARv, 27, 3
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
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. 2021, A&A, 650, A110

Mercado, F. J., Bullock, J. S., Boylan-Kolchin, M., et al. 2021, MNRAS, 501, 5121

Miceli, A., Rest, A., Stubbs, C. W., et al. 2008, ApJ, 678, 865

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

Molaro, P., Cescutti, G., & Fu, X. 2020, MNRAS, 496, 2902

Monty, S., Venn, K. A., Lane, J. M. M., Lokhorst, D., & Yong, D. 2020, MNRAS, 497, 1236

Mosleh, M., Tacchella, S., Renzini, A., et al. 2017, ApJ, 837, 2

Mowla, L. A., van Dokkum, P., Brammer, G. B., et al. 2019, ApJ, 880, 57

Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018, ApJL, 863, L28

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

Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48

Necib, L., Lisanti, M., & Belokurov, V. 2019, ApJ, 874, 3

Newberg, H. J., & Yanny, B. 2006, JPhCS, 47, 195

Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, ApJ, 569, 245

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

O'Hare, C. A. J., Evans, N. W., McCabe, C., Myeong, G., & Belokurov, V. 2020, PhRvD, 101, 023006

O'Hare, C. A. J., McCabe, C., Evans, N. W., Myeong, G., & Belokurov, V. 2018, PhRvD, 98, 103006

Oliphant, T. 2006, NumPy: A guide to NumPy (USA: Trelgol Publishing) , http://www.numpy.org/

Park, M. J., Yi, S. K., Peirani, S., et al. 2021, ApJS, 254, 2

Pérez, F., & Granger, B. E. 2007, CSE, 9, 21

Pila-Díez, B., de Jong, J. T. A., Kuijken, K., van der Burg, R. F. J., & Hoekstra, H. 2015, A&A, 579, A38

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Posti, L., & Helmi, A. 2019, A&A, 621, A56

Power, C., Navarro, J. F., Jenkins, A., et al. 2003, MNRAS, 338, 14

Price-Whelan, A., Sipocz, B., Major, S., & Oh, S. 2017, adrn/gala: v0.2.1, Zenodo doi:10.5281/zenodo.833339

Price-Whelan, A. M. 2017, JOSS, 2, 388

Robitaille, T., Beaumont, C., Qian, P., Borkin, M., & Goodman, A. 2017, glueviz v0.13.1: multidimensional data exploration, 0.13.1, Zenodo doi:10.5281/zenodo.1237692

Ruiz-Lara, T., Gallart, C., Bernard, E. J., & Cassisi, S. 2020, NatAs, 4, 965

Sánchez, S. F. 2020, ARA&A, 58, 99

Santistevan, I. B., Wetzel, A., El-Badry, K., et al. 2020, MNRAS, 497, 747

Sesar, B., Ivezić, Ž., Stuart, J. S., et al. 2013, AJ, 146, 21

Sesar, B., Jurić, M., & Ivezić, Ž. 2011, ApJ, 731, 4

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

Simion, I. T., Belokurov, V., Koposov, S. E., Sheffield, A., & Johnston, K. V. 2018, MNRAS, 476, 3913

Simon, J. D. 2019, ARA&A, 57, 375

Simpson, C. M., Gargiulo, I., Gómez, F. A., et al. 2019, MNRAS, 490, L32

Simpson, J. D., Martell, S. L., Buder, S., et al. 2021, MNRAS, 507, 43

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Pakmor, R., Zier, O., & Reinecke, M. 2021, MNRAS, 506, 2871

Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685

Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645

Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019, ApJ, 877, 103

Surot, F., Valenti, E., Hidalgo, S. L., et al. 2019, A&A, 623, A168

Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157

Vahsen, S. E., O'Hare, C. A. J., Lynch, W. A., et al. 2020, arXiv:2008.12587

van der Velden, E. 2020, JOSS, 5, 2004

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28

van Dokkum, P. G., Leja, J., Nelson, E. J., et al. 2013, ApJL, 771, L35

Vasiliev, E., Belokurov, V., & Erkal, D. 2021, MNRAS, 501, 2279

Venn, K. A., Irwin, M., Shetrone, M. D., et al. 2004, AJ, 128, 1177

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

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261

Vivas, A. K., Zinn, R., Andrews, P., et al. 2001, ApJL, 554, L33

Watkins, L. L., Evans, N. W., Belokurov, V., et al. 2009, MNRAS, 398, 1757

Wechsler, R. H., Somerville, R. S., Bullock, J. S., et al. 2001, ApJ, 554, 85

Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014, ApJ, 789, 148

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

Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, ApJ, 809, 144

Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377

Yuan, Z., Myeong, G. C., Beers, T. C., et al. 2020, ApJ, 891, 39

Yurin, D., & Springel, V. 2014, MNRAS, 444, 62

Zaritsky, D., Conroy, C., Zhang, H., et al. 2020, ApJ, 888, 114