

UC Davis

UC Davis Previously Published Works

Title

The origin of metal-poor stars on prograde disc orbits in FIRE simulations of Milky Way-mass galaxies

Permalink

<https://escholarship.org/uc/item/19j0t8z7>

Journal

Monthly Notices of the Royal Astronomical Society, 505(1)

ISSN

0035-8711

Authors

Santistevan, IB
Wetzel, A
Sanderson, RE
et al.

Publication Date

2021-07-01

DOI

10.1093/mnras/stab1345

Peer reviewed

The origin of metal-poor stars on prograde disk orbits in FIRE simulations of Milky Way-mass galaxies

Isaiah B. Santistevan¹★, Andrew Wetzel¹, Robyn E. Sanderson^{2,3}, Kareem El-Badry⁴, Jenna Samuel¹, Claude-André Faucher-Giguère⁵

¹*Department of Physics & Astronomy, University of California, Davis, CA 95616, USA*

²*Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*

³*Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*

⁴*Department of Astronomy, Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA*

⁵*Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

In hierarchical structure formation, metal-poor stars in and around the Milky Way (MW) originate primarily from mergers of lower-mass galaxies. A common expectation is therefore that metal-poor stars should have isotropic, dispersion-dominated orbits that do not correlate strongly with the MW disk. However, recent observations of stars in the MW show that metal-poor ($[\text{Fe}/\text{H}] \lesssim -2$) stars are preferentially on prograde orbits with respect to the disk. Using the FIRE-2 suite of cosmological zoom-in simulations of MW/M31-mass galaxies, we investigate the prevalence and origin of prograde metal-poor stars. Almost all (11 of 12) of our simulations have metal-poor stars preferentially on prograde orbits today and throughout most of their history: we thus predict that this is a generic feature of MW/M31-mass galaxies. The typical prograde-to-retrograde ratio is $\sim 2 : 1$, which depends weakly on stellar metallicity at $[\text{Fe}/\text{H}] \lesssim -1$. These trends predicted by our simulations agree well with MW observations. Prograde metal-poor stars originate largely from a single LMC/SMC-mass gas-rich galaxy merger, typically 7 – 12.5 Gyr ago, which deposited both existing metal-poor stars and significant gas on an orbital vector that sparked the formation of and/or shaped the orientation of a long-lived stellar disk, giving rise to a prograde bias for all low-metallicity stars. We also find sub-dominant contributions from in-situ stars formed in the host galaxy before this merger, and in some cases, additional massive mergers. We find few clear correlations between any properties of our MW/M31-mass galaxies at $z = 0$ and the degree of this prograde bias as a result of diverse merger scenarios.

Key words: stars – metal-poor, galaxies – formation

1 INTRODUCTION

The populations of stars in different regions of a galaxy, such as the disk, the bulge, and the halo, can reveal its formation history. In the standard picture of galaxy formation, unenriched gas collapses in dark matter (DM) halos to form stars and galaxies (e.g. Ostriker & Tremaine 1975; Rees & Ostriker 1977; White & Rees 1978; Fall & Efstathiou 1980). Many works have studied the hierarchical nature of galaxy formation, such as determining the ages of globular clusters throughout the MW’s stellar halo, and the formation of disk galaxies within their DM halos (e.g. Searle & Zinn 1978; White & Frenk 1991; Mo et al. 1998). Eggen et al. (1962) first suggested that old, metal-poor stars orbit the MW in elliptical orbits compared to

metal-rich stars, which move in more circular orbits, a picture that broadly holds today (e.g. Chiba & Beers 2000). Many authors have explored how metal-poor stars were deposited into the stellar halo of the MW by mergers of low-mass dwarf galaxies (e.g. Bullock et al. 2001; Bullock & Johnston 2005; Ibata et al. 1994; Newberg et al. 2003; Nissen & Schuster 2010). We now understand that most of the mass in the stellar halo came from a few massive mergers, which is a natural consequence of the steepness of the stellar mass halo mass relation. Deason et al. (2016) showed that only 1 – 2 satellite galaxies with $M_{\text{star}} \sim 10^8 - 10^{10} M_{\odot}$ deposited a majority of the stars in the halo, and dwarf galaxies with $M_{\text{star}} \sim 10^5 - 10^8 M_{\odot}$ contributed the bulk of metal-poor ($[\text{Fe}/\text{H}] \lesssim -2$) stars. A recent observational study by Naidu et al. (2020) suggests that at distances greater than ~ 15 kpc from the galactic center, ~ 80 per cent of halo stars came from 2 mergers: Gaia-Enceladus (Helmi et al. 2018;

★ E-mail: ibsantistevan@ucdavis.edu

Belokurov et al. 2018) and Sagittarius (Ibata et al. 1994; Newberg et al. 2003; Majewski et al. 2003).

Many studies focused on the properties and origins of these old and/or metal-poor stars, such as their spatial and elemental distributions within the MW, the masses of the accreted dwarf galaxies they came from, and the formation of the oldest Population III stars (e.g. Scannapieco et al. 2006; Brook et al. 2007; Deason et al. 2016; Griffen et al. 2018). Spectroscopic surveys that have observed stars with $[\text{Fe}/\text{H}] \lesssim -2$, such as RAVE, GALAH, LAMOST, and APOGEE, provide the elemental abundances and ages of stars in different regions of the MW and now allow for an investigation into the early formation period (Steinmetz et al. 2006; De Silva et al. 2015; Li et al. 2015; Majewski et al. 2017). The Pristine survey also has pushed metallicity measurements down to $[\text{Fe}/\text{H}] < -3$, with the potential of reaching the ‘ultra metal-poor’ ($[\text{Fe}/\text{H}] < -4$) population (Starkenburg et al. 2017b). Finally, the recent Hectochelle in the Halo at High Resolution (H3) Survey is targeting stars outside of the disk and bulge regions of the MW, and is reaching metallicities down to $[\text{Fe}/\text{H}] \sim -3$ (Conroy et al. 2019).

Recent simulation work on the spatial distributions of old, metal-poor stars within a MW-mass galaxy suggests that they are less centrally concentrated than younger stars. For example, Scannapieco et al. (2006) used DM-only simulations of MW-mass halos and concluded that both unenriched stars, and their descendants, occupy a wide range of radii throughout the MW. In a similar study, Brook et al. (2007) used cosmological baryonic simulations of 4 MW-mass galaxies and found that metal-free stars are distributed throughout the MW’s halo, though the authors suggest that the oldest stars are more centrally concentrated. Using a sample of old ($z_{\text{form}} > 5$), metal-poor ($[\text{Fe}/\text{H}] < -2$) stars in 3 of the FIRE-2 simulations that we examine here, El-Badry et al. (2018b) proposed that this spatially extended distribution of metal-poor stars in a MW-mass galaxy results from a combination of kinematic heating of stars to larger orbits in the host galaxy from a time-varying galactic potential and from early mergers that deposited these stars on radial orbits. However, they did not examine any preference for metal-poor stars to be on prograde orbits aligned with the disk. Similarly, using the APOSTLE simulations, Starkenburg et al. (2017a) found that the fraction of stars with $[\text{Fe}/\text{H}] < -2.5$, older than $\gtrsim 13$ Gyr, increases with increasing distance from the center of the galaxy. Although these studies agree that the number density of old stars is largest near the central bulge region, the overabundance of younger stars can make finding older stars difficult.

The origin of the MW’s stellar halo and its disk are likely connected, particularly for the inner stellar halo and the thick disk. Gallart et al. (2019) showed that stellar populations in the inner halo and thick disk partially overlap in kinematic space. The thick disk extends $\gtrsim 1$ kpc from the midplane and is comprised of primarily older stars across a wide range of iron abundances ($-2 \lesssim [\text{Fe}/\text{H}] \lesssim -0.5$; e.g. Gilmore & Reid 1983; Freeman & Bland-Hawthorn 2002). Several authors have proposed different formation mechanisms for the thick disk such as the kinematic heating of thin disk stars to larger vertical or radial distances, and the disruption and accretion of satellite galaxies (e.g. Walker et al. 1996; Villalobos & Helmi 2008; Loebman et al. 2011; Haywood et al. 2018) However, many works now show that the disk formed thick and settled over time. Bird et al. (2013) used the cosmological MW-mass simulation ‘Eris’ (Guedes et al. 2011) and concluded that the MW’s disk formed ‘upside-down’, where older stars formed farther from the midplane with larger vertical velocity dispersions compared to younger stars, which formed in the thin disk with smaller velocity dispersions. More recent cosmological simulations strongly support this idea

(e.g. Grand et al. 2016; Bird et al. 2020), including Ma et al. (2017), who analyzed one of the simulated galaxies in our sample (m12i). The authors concluded that older stars in the thick disk formed during the early, bursty phase of the MW when the galaxy was gas-rich, while younger stars formed during the more calm period of MW growth and that the disk becomes thinner over time.

A great deal of work is dedicated to understanding the existence of net rotation in the stellar halo, and its origins. In an early study, Norris & Ryan (1989) found that, in a population of ~ 500 halo stars, the majority showed net retrograde motion. We now know that the halo is comprised of two components, an inner and outer component, each with different elemental abundances and kinematics (Zinn 1993; Carney et al. 1996; Bonaca et al. 2017). Deason et al. (2011) found that metal-rich stars in the halo ($[\text{Fe}/\text{H}] > -2$) favored prograde motion and were likely accreted from a satellite, while metal-poor stars ($[\text{Fe}/\text{H}] < -2$) were more retrograde on average and could have been part of the early-forming halo. They also noted that many globular clusters in the MW halo orbit in the prograde direction. Another study by Bonaca et al. (2017) concluded that metal-rich stars ($[\text{Fe}/\text{H}] > -1$) near the MW have prograde orbits, while metal-poor stars have no average orbit direction. This suggests the metal-rich halo stars formed in-situ in the MW galaxy, while the metal-poor stars likely were accreted. The authors found similar features in one of the simulated galaxies in our sample, m12i. Finally, recent work focused on kinematically, or elementally, coherent structures throughout the MW suggest that many previous mergers currently comprise most of the retrograde population in the MW halo (Myeong et al. 2019; Naidu et al. 2020).

Recently, Sestito et al. (2019) analyzed all known MW stars with $[\text{Fe}/\text{H}] < -4$ and concluded that 11 of the 42 stars (26 per cent of their sample) have orbits that are confined to within ± 3 kpc of the disk mid-plane. Of these stars on disk-like orbits, 10 are prograde ($J_{\phi} > 0$), and only one shows retrograde motion ($J_{\phi} < 0$). They determined that the remaining 31 metal-poor stars in their sample have halo-like orbits, bringing them far away from the disk, which is typical for accretion-based origin. Subsequently, Sestito et al. (2020) selected stars with $[\text{Fe}/\text{H}] \lesssim -2$ from both the Pristine survey (Starkenburg et al. 2017b) and LAMOST DR3 (Li et al. 2018), increasing their sample size from 42 to 1027 (a factor of ~ 24). They again found that stars in the prograde region of action space outnumber those in the retrograde region, and ~ 31 per cent of the sample have prograde orbits that are confined to ± 3 kpc of the disk mid-plane. New observational studies continue to support the results from Sestito et al. (2019, 2020), particularly the discovery of metal-poor stars near the plane of the MW (e.g. Venn et al. 2020; Di Matteo et al. 2020). Sestito et al. (2019) posit 3 possible explanations as to the origins of these metal-poor, disk-like stars: (i) they formed in-situ within the MW disk itself and remained on prograde, near-circular orbits; (ii) another galaxy merged into the MW on an orbit that was aligned with the (pre-existing) MW disk, depositing its stars on prograde orbits; or (iii) these stars formed in one or more progenitor galaxies that merged into the MW as it was forming.

Our goal in this paper is to test the origin and frequency of metal-poor stars on prograde disk orbits, using 12 MW/M31-mass galaxies from the FIRE-2 suite of cosmological zoom-in simulations. The main questions that we address are:

- i) How commonly do MW/M31-mass galaxies show a preference for prograde motions in their metal-poor stars, as observed in the MW, and what is the range in strength of this prograde bias?

ii) How does this prograde bias depend on the metallicity and/or age of the stars?

iii) What process(es) cause this prograde bias among metal-poor stars?

iv) Do any properties of the MW/M31-mass galaxy or their galaxy mergers correlate with this prograde bias?

2 METHODS

2.1 FIRE-2 Simulations

We use cosmological zoom-in baryonic simulations of MW/M31-mass galaxies from the Feedback In Realistic Environments (FIRE) project¹ (Hopkins et al. 2018). We ran all of the simulations using the same N -body gravitational plus hydrodynamics code GIZMO (Hopkins 2015), with the mesh-free finite-mass (MFM) hydrodynamics method and the FIRE-2 physics model (Hopkins et al. 2018). Across a temperature range of $10 - 10^{10}$ K, the FIRE-2 model includes several radiative cooling and heating processes for gas such as free-free emission, photoionization/recombination, Compton scattering, photoelectric, metal-line, molecular, fine-structure, dust-collisional, and cosmic-ray heating, including the spatially uniform, redshift-dependent cosmic UV background from Faucher-Giguère et al. (2009), where HI reionization occurs at $z_{\text{reion}} \sim 10$. The simulations self-consistently generate and track 11 elemental abundances (H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe), including the sub-grid diffusion/mixing of these elements in gas via turbulence (Hopkins 2016; Su et al. 2017; Escala et al. 2018).

Because these FIRE simulations do not model the initial metal enrichment from Pop III stars, we initialize gas elements in the simulations with a metallicity floor of $[\text{Fe}/\text{H}] = -4$. We choose this metallicity floor as a rough model for the level of enrichment we expect from Population III stars, which also corresponds to the lowest metallicity in one of the observational studies we compare our results to, Sestito et al. (2019). As we will show, the strength of the prograde bias does not depend on metallicity at low metallicities, so the choice of the metallicity floor is unimportant. Stars form in gas that is self-gravitating, Jeans unstable, molecular (following Krumholz & Gnedin 2011), and dense ($n_{\text{H}} > 1000 \text{ cm}^{-3}$). A star particle inherits the mass and elemental abundances from its progenitor gas particle, and represents a single stellar population, assuming a Kroupa (2001) initial mass function. It then evolves along stellar population models from STARBURST99 v7.0 (Leitherer et al. 1999). Furthermore, the FIRE-2 simulations include the following feedback processes: core-collapse and Ia supernovae, stellar winds, radiation pressure, photoionization, and photo-electric heating.

We generated the cosmological zoom-in initial conditions for each simulation embedded within periodic cosmological boxes of lengths 70.4–172 Mpc using the code MUSIC (Hahn & Abel 2011), at $z \approx 99$. Each simulation saves of 600 snapshots down to $z = 0$, spaced every ≈ 25 Myr, and all assume flat Λ CDM cosmology with the following range in cosmological parameters, consistent with Planck Collaboration et al. (2018): $\Omega_{\text{m}} = 0.266 - 0.31$, $\Omega_{\text{b}} = 0.0449 - 0.048$, $\sigma_8 = 0.801 - 0.82$, $n_{\text{s}} = 0.961 - 0.97$, and $h = 0.68 - 0.71$.

We analyze the same 12 MW/M31-mass galaxies in Santistevan et al. (2020), and Table 1 lists their stellar masses at $z = 0$. Half of our sample comes from the Latte suite of isolated MW/M31-mass galaxies first introduced in Wetzel et al. (2016). The Latte suite has

gas and initial star particle masses of $7100 M_{\odot}$, but because of stellar mass loss, the average star particle mass is $\approx 5000 M_{\odot}$. The mass of DM particles in the zoom-in region is $3.5 \times 10^4 M_{\odot}$. The gravitational softening lengths for star and DM particles are fixed at 4 and 40 pc (Plummer equivalent), comoving at $z > 9$ and physical thereafter, while gas elements use adaptive force softening, equal to their hydrodynamic smoothing, which adapts down to 1 pc. We chose the halos in the Latte suite to have $M_{200\text{m}} = 1 - 2 \times 10^{12} M_{\odot}$ and have no other similar-mass halos within $5 \times R_{200\text{m}}$, but we imposed no additional selection criteria beyond this.

The other half of our sample comes from the ‘ELVIS on FIRE’ suite of LG-like MW+M31 pairs, introduced in Garrison-Kimmel et al. (2019a,b), which have approximately $2\times$ better mass resolution, with gas and initial star particle masses of $3500 - 4000 M_{\odot}$. Each pair of halos was chosen based on their masses (each with $M_{200\text{m}} = 1 - 3 \times 10^{12} M_{\odot}$ and total LG mass between $2 - 5 \times 10^{12} M_{\odot}$), current separation (600 – 1000 kpc) and radial velocities at $z = 0$ ($v_{\text{rad}} < 0$), and that they have no other massive halos within 2.8 Mpc of either host center.

We selected all halos without prior knowledge of their formation/merger histories or satellite populations, except m12w, which we selected to have an LMC-mass satellite analog at $z \sim 0$ (Samuel et al. 2020b). Therefore, our galaxy sample should reflect random/typical formation histories. The host galaxies in our sample reflect a class of morphologies (Garrison-Kimmel et al. 2018; El-Badry et al. 2018a), which host stellar-to-halo mass relations (Hopkins et al. 2018), disk morphologies and metallicity gradients (Ma et al. 2017; Sanderson et al. 2020), and stellar halos (Bonaca et al. 2017; Sanderson et al. 2018) akin to that of the MW and M31. Furthermore, each host galaxy possesses realistic satellite galaxy populations, with stellar masses and velocity dispersions (Wetzel et al. 2016; Garrison-Kimmel et al. 2019a), radial and 3-D distributions (Samuel et al. 2020b,a), and star-formation histories (Garrison-Kimmel et al. 2019b) that broadly agree with observations of the LG and the local Universe.

2.2 ROCKSTAR Halo Catalogs & Merger Trees

We generate (sub)halo catalogs using only DM particles, at all 600 snapshots, using the ROCKSTAR 6-D halo finder (Behroozi et al. 2013a). To generate merger trees, we use CONSISTENT-TREES (Behroozi et al. 2013b). All (sub)halos that we examine have zero contamination by low-resolution DM particles, because of the large zoom-in region that we generate for each host.

We assign star particles to (sub)halos in post-processing, which we outline below, but refer the reader to Samuel et al. (2020b) for more details. We identify star particles with positions within $0.8 \times R_{\text{halo}}$ (out to a maximum distance of 30 kpc) and velocities within $2 \times V_{\text{circ,max}}$ of a (sub)halo’s center-of-mass velocity. Following these initial criteria, we keep star particles if they are within $1.5 \times R_{\text{star,90}}$ (the radius enclosing 90 per cent of the stellar mass) of the current member star particle’s center-of-mass and halo center position. This guarantees that the centers-of-mass of both the galaxy and (sub)halo are consistent with one another. We then keep star particles within $2 \times \sigma_{\text{vel,star}}$ (the velocity dispersion of current member star particles) of the center-of-mass velocity of member star particles, and we iterate on both criteria until the (sub)halo’s stellar mass converges to within 1 per cent.

We use two publicly available analysis packages: HALOANALY-

¹ FIRE project web site: <http://fire.northwestern.edu>

Table 1. Properties of the 12 MW/M31-mass galaxies in the FIRE-2 simulation suite that we analyze, ordered by decreasing prograde-to-retrograde ratio for metal-poor stars. Simulations with ‘m12’ names are isolated hosts from the Latte suite, while the others are from the ELVIS on FIRE suite of LG-like galaxies. We measure all masses and mass fractions at $z = 0$, using only metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars (except for the total stellar mass of the galaxy, which uses all stars). Columns: name; $M_{\text{star},90}$ is the host’s stellar mass within $R_{\text{star},90}$, the disk radius enclosing 90 per cent of the stellar mass within 20 kpc; $M_{\text{star,pro}}/M_{\text{star,ret}}$ is the prograde bias, defined as the mass ratio of prograde to retrograde stars; $t_{\text{lb,merger},1}$ is the lookback time of the merger that contributed the most prograde stars; $f_{\text{merger},1}$ is the fraction of all prograde stars that came from the primary (most contributing) merger; $f_{\text{merger,top 3}}$ is the fraction of prograde stars from the 3 most contributing mergers; $f_{\text{in-situ}}$ is the fraction of prograde stars that formed in-situ (within 15 kpc); $f_{\text{merger},1} + f_{\text{in-situ}}$ is the sum; $f_{\text{merger,top 3}} + f_{\text{in-situ}}$ is the sum of fractions from the 3 most contributing mergers and the fraction that formed in-situ.

Name	$M_{\text{star},90}$ [$10^{10} M_{\odot}$]	$M_{\text{star,pro}}/M_{\text{star,ret}}$	$t_{\text{lb,merger},1}$ [Gyr]	$f_{\text{merger},1}$	$f_{\text{merger,top 3}}$	$f_{\text{in-situ}}$	$f_{\text{in-situ}} + f_{\text{merger},1}$	$f_{\text{merger,top 3}} + f_{\text{in-situ}}$
m12m ^a	10.0	9.14	9.56	0.44	0.74	0.21	0.65	0.95
m12w ^b	4.8	3.02	8.17	0.20	0.50	0.13	0.32	0.63
m12c ^c	5.1	2.67	8.99	0.48	0.75	0.21	0.70	0.96
Romeo ^c	5.9	2.35	12.52	0.24	0.38	0.19	0.43	0.57
m12b ^c	7.3	2.29	2.84	0.27	0.59	0.06	0.34	0.65
Juliet ^c	3.3	2.01	12.25	0.14	0.28	0.15	0.30	0.43
Thelma ^c	6.3	1.96	8.61	0.28	0.57	0.10	0.38	0.67
Romulus ^d	8.0	1.56	7.60	0.53	0.77	0.05	0.58	0.82
m12f ^e	6.9	1.48	11.91	0.20	0.35	0.16	0.36	0.51
Remus ^d	4.0	1.42	12.02	0.23	0.45	0.10	0.34	0.55
Louise ^c	2.3	1.28	10.75	0.15	0.34	0.31	0.46	0.65
m12i ^f	5.5	0.98	10.50	0.24	0.57	0.24	0.48	0.81

Note: Simulation introduced in: a: [Hopkins et al. \(2018\)](#), b: [Samuel et al. \(2020b\)](#), c: [Garrison-Kimmel et al. \(2019b\)](#), d: [Garrison-Kimmel et al. \(2019a\)](#), e: [Garrison-Kimmel et al. \(2017\)](#), f: [Wetzel et al. \(2016\)](#).

sis² ([Wetzel & Garrison-Kimmel 2020a](#)) for assigning star particles to halos and for reading and analyzing halo catalogs/trees, and GIZMOANALYSIS³ ([Wetzel & Garrison-Kimmel 2020b](#)) for reading and analyzing particles from Gizmo snapshots.

2.3 Selecting & Tracking Metal-poor Stars

For each MW/M31-mass host galaxy, we define the ‘prograde bias’ as the ratio of the total stellar mass of prograde stars to that of retrograde stars, $M_{\text{star,pro}}/M_{\text{star,ret}}$. To measure the prograde bias, we first select star particles in the simulations at $z = 0$ based on the following two criteria, motivated by [Sestito et al. \(2020\)](#): (1) within ± 3 kpc vertically of the disk midplane and within 4 – 12 kpc radially from the center of the galaxy, which ensures that the sample is not contaminated significantly from bulge stars; (2) with iron abundance $[\text{Fe}/\text{H}] < -2.5$. These conditions define our fiducial sample, however, we explore how our results vary with $[\text{Fe}/\text{H}]$, and with how we spatially and kinematically select star particles in Section 3.1. We then define prograde or retrograde motion based on the star particle’s action variables, J_{ϕ} and J_z . J_{ϕ} is equal to the z -component of angular momentum, L_z . We use the approximation $J_z \approx |z v_z|$, where z is a star particle’s vertical position and v_z is its vertical velocity, relative to the host’s disk. Our choice is motivated by the definition for specific action (per unit mass),

$$J_z \equiv \oint v_z dz \quad (1)$$

where the integral is over one orbit in z . In the epicyclic approximation ($J_z \ll |J_{\phi}|$),

$$J_z \approx \frac{|v_z| z_{\text{max}}}{2}. \quad (2)$$

Because z follows simple harmonic motion, $\langle |z| \rangle = z_{\text{max}}/2$, where $\langle |z| \rangle$ is the time-averaged height. Therefore, for a population of orbits described by the epicyclic approximation, trading time-averaging for spatial-averaging, we get $J_z \approx |v_z z|$ for the population. Even though this approximation is good only for stars near the plane of the disk on near-circular orbits, our geometric selection of stars avoids stars that orbit more than 3 kpc above or below the disk at $R = 8$ kpc. As we show in Section 3.1 the strength of the prograde bias does not change much for the different kinematic selections that we investigate.

[Sestito et al. \(2020\)](#) defined prograde stars with $0.5 < J_{\phi}/J_{\phi\odot} < 1.2$ and $J_z/J_{z\odot} < 1250$, and retrograde stars with $-1.2 < J_{\phi}/J_{\phi\odot} < -0.5$ and $J_z/J_{z\odot} < 1250$, where the solar values are $J_{\phi\odot} = 2009 \text{ km s}^{-1} \text{ kpc}$ and $J_{z\odot} = 0.35 \text{ km s}^{-1} \text{ kpc}$. Our simulated MW/M31-mass galaxies have a range of rotational velocities ($v_{\text{rot}} \sim 130 - 230 \text{ km s}^{-1}$), so we normalize J_{ϕ} by $8 \times v_{\text{rot}} \text{ km s}^{-1} \text{ kpc}$, where v_{rot} is the rotational velocity of the host galaxy, for stars at a similar galactocentric distance as the Sun. Given that we approximate the J_z action variable, we use the same MW values to normalize J_z as in [Sestito et al. \(2020\)](#), and keep stars with $|J_z| < 438 \text{ km s}^{-1} \text{ kpc}$. As we will show below, our qualitative results do not depend on the details of our selection window, if for example, we select all prograde ($J_{\phi} > 0$) and all retrograde ($J_{\phi} < 0$) stars.

Figure 1 shows the action space coordinates of all metal-poor ($[\text{Fe}/\text{H}] < -2.5$) star particles within our selection region, for one of our simulated galaxies, m12w. The black rectangles show our fiducial selection windows to define prograde and retrograde stars. An overabundance of metal-poor stars lie in the prograde region (right box) compared to the retrograde region (left box), similar to the MW results in [Sestito et al. \(2020\)](#).

We track these stars, which we select at $z = 0$, back in time to track their origin, specifically, to determine whether they formed (a) in-situ, within 15 kpc physical of the most massive progenitor of the host or (b) in another galaxy that merged into the host. At

² https://bitbucket.org/awetzel/halo_analysis

³ https://bitbucket.org/awetzel/gizmo_analysis

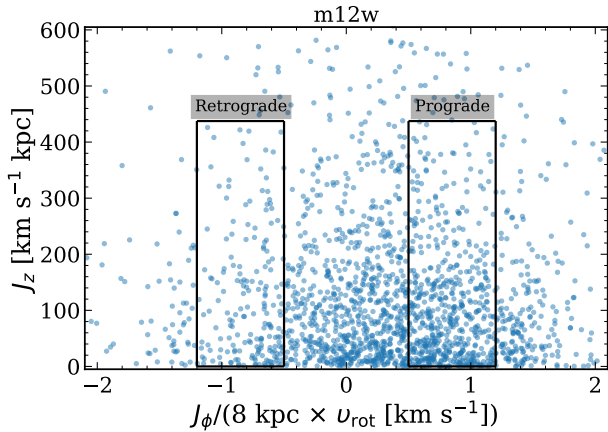


Figure 1. Metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars in the disk ($4 \leq R \leq 12$ kpc, $|Z| \leq 3$ kpc) of the m12w simulation, in action space. Blue dots show individual star particles, and the right and left boxes show regions in action space that we select for prograde and retrograde orbits, respectively. We adopt a similar selection region as in Sestito et al. (2020) for the Pristine survey, which focused on metal-poor stars ($[\text{Fe}/\text{H}] < -2.5$) in the MW. We normalize J_ϕ by what ‘solar’ values would be in the simulations, where v_{rot} is the rotational velocity of the galaxy at 8 kpc. Metal-poor stars in m12w have a strong preference for prograde orbits, with $M_{\text{star,pro}}/M_{\text{star,ret}} \approx 3$.

early times, this distance may encompass other nearby low-mass progenitor galaxies, so we tested how the in-situ fractions changed using stars within 10 and 20 kpc, and we found that they varied by approximately ± 4 per cent on average. To assign membership to another progenitor galaxy, we require that a star particle must have been a member of that galaxy for at least 3 consecutive snapshots. We also tested requiring 2, 4, or 5 consecutive snapshots, and we found no significant differences in our results.

3 RESULTS

3.1 How does the prograde bias depend on stellar metallicity & age?

3.1.1 Comparison with the Pristine survey

We first examine how the present prograde bias varies with iron abundance, $[\text{Fe}/\text{H}]$. Figure 2 shows the cumulative prograde bias for all stars below a given $[\text{Fe}/\text{H}]$, and the prograde bias for stars within 0.5 dex bins of $[\text{Fe}/\text{H}]$, for all of our hosts, as a function of $[\text{Fe}/\text{H}]$ at $z = 0$. The thick black lines show the median across all 12 hosts, and we overlay Pristine observations of the MW from Sestito et al. (2019) in black crosses labeled ‘MW (S19)’. Solid lines show our 6 isolated hosts, while dashed lines show our 6 hosts in LG-like pairs. We show stars only up to $[\text{Fe}/\text{H}] = -0.5$ to avoid the metal-rich stars in the thin disk that are on overwhelmingly prograde orbits, and we remind the reader of the metallicity floor of $[\text{Fe}/\text{H}] = -4$ in the simulations.

The median prograde bias is fairly flat at $[\text{Fe}/\text{H}] < -1$ for the cumulative plot (top panel) and at $[\text{Fe}/\text{H}] < -1.75$ for the binned (middle panel) population. A few hosts (such as m12b) show a slight rise at $[\text{Fe}/\text{H}] \lesssim -3$. The prograde bias rapidly rises at higher iron abundance, as the stellar population quickly transitions to the more metal-rich, younger, thin disk (and thus highly prograde) component of each host. The observed values from Sestito et al. (2019) are close to our median lines, so the MW is both consistent and typical compared with our simulation suite.

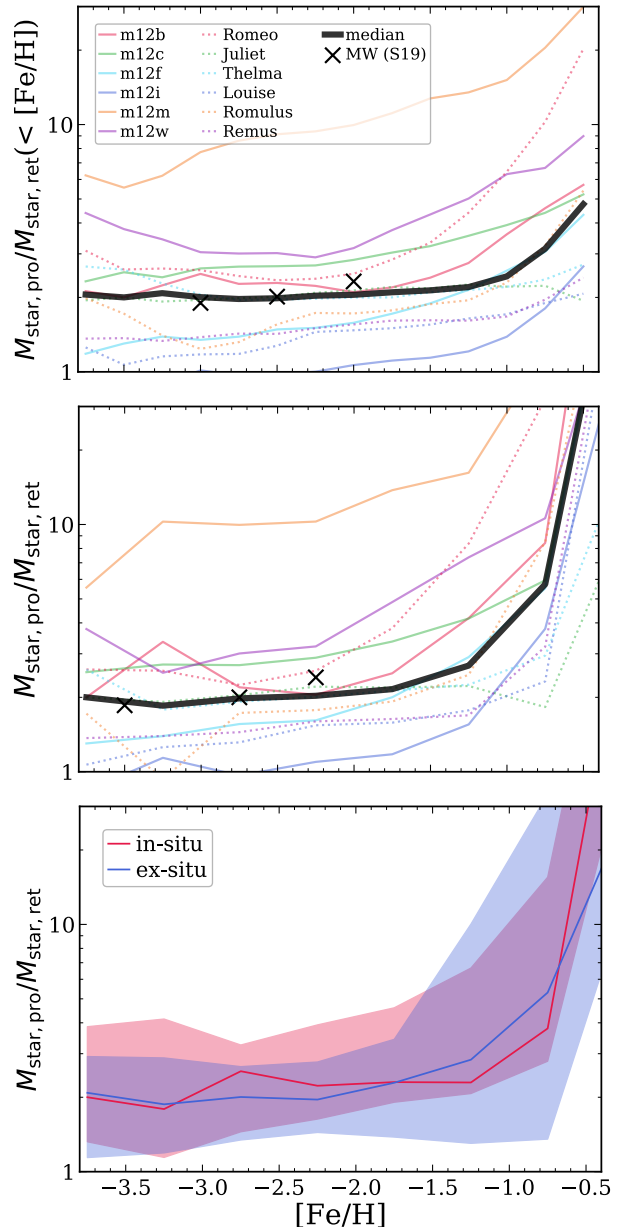


Figure 2. The prograde bias, defined as the mass ratio of prograde to retrograde stars, $M_{\text{star,pro}}/M_{\text{star,ret}}$, versus stellar iron abundance for our 12 MW/M31-mass galaxy simulations at $z = 0$. **Top and middle panels:** the prograde bias as a function of cumulative $[\text{Fe}/\text{H}]$ and binned $[\text{Fe}/\text{H}]$, respectively. Solid lines show the 6 isolated hosts, while dotted lines show the 6 LG-like hosts. The thick black lines show the medians across all 12 hosts, which are consistent with MW observations from the Pristine survey (Sestito et al. 2019, S19), which we show with black crosses. 11 of the 12 galaxies in our sample show a significant prograde bias (> 1) in metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars. For most hosts, the prograde bias is nearly constant with iron abundance at $[\text{Fe}/\text{H}] \lesssim -1.5$, and it rises rapidly at higher $[\text{Fe}/\text{H}]$. **Bottom:** The binned prograde bias, split into in-situ (red) and ex-situ (blue) stars, where we select ‘in-situ’ stars as formed within 15 kpc of the host. The lines show the medians, and the shaded regions show the 68 per cent scatter across 12 hosts. Ex-situ and in-situ stars have similar prograde bias at $[\text{Fe}/\text{H}] \lesssim -1.7$, while ex-situ stars show a slightly higher prograde bias at $-1.7 \lesssim [\text{Fe}/\text{H}] \lesssim -0.7$. As we will show, a single SMC/LMC-mass merger typically drives this trend. At higher $[\text{Fe}/\text{H}]$, the in-situ component quickly rises in prograde bias, as stars transition to the thin-disk component.

Figure 2 shows that nearly all (11 of 12) of our hosts have significant prograde bias at all $[\text{Fe}/\text{H}]$. The key exception is m12i, which does not have any significant prograde bias at $[\text{Fe}/\text{H}] \lesssim -2.5$. However, m12i does have a prograde bias at higher $[\text{Fe}/\text{H}]$, so this feature is ubiquitous (within our suite) at all $[\text{Fe}/\text{H}] \gtrsim -2$.

Isolated and LG-like hosts differ slightly in Figure 2. Isolated hosts typically have higher prograde bias curves than LG-like hosts, with median values of 2.48 and 1.76 at $[\text{Fe}/\text{H}] = -2.5$, respectively. There are differences in the in-situ formation times between LG-like and isolated galaxies, with LG-like galaxies forming around $z \sim 4.9$ ($t_{\text{lb}} \sim 12.5$ Gyr ago) and isolated galaxies forming later, around $z \sim 2.7$ ($t_{\text{lb}} \sim 11.3$ Gyr ago), so the lack of strong differences in prograde bias may be surprising (Santistevan et al. 2020). However, as we discuss below, the prograde bias is primarily driven by mergers and sub-dominantly by in-situ formation; we do not explore these differences further throughout the rest of the paper.

Figure 2 (bottom panel) shows the same binned prograde bias, but we now split the stars into two origins: in-situ (formed within 15 kpc of the host’s main progenitor), and ex-situ stars that formed in another galaxy that merged into the host. The lines show the median values across all 12 hosts, and the shaded regions show the 68 per cent scatter. One might naively expect that the in-situ population would have a larger prograde bias, to the degree that one expects that in-situ stars always formed on disk-like prograde orbits, and because a merger’s orbit need not align itself with the host’s existing disk. However, we find only weak differences between in-situ and ex-situ stars, which are well within the host-to-host scatter, at all $[\text{Fe}/\text{H}] < -0.5$. At $[\text{Fe}/\text{H}] \lesssim -1.75$, in-situ stars tend to be a little more prograde, but at higher $[\text{Fe}/\text{H}]$ up to ~ -0.75 , ex-situ stars in fact show a slightly stronger preference for prograde orbits. At $[\text{Fe}/\text{H}] > -0.75$, in-situ stars quickly become the more prograde population, because we now transition to stars that primarily formed in the host’s thin disk.

Figure 2 shows that the prograde bias depends only weakly on iron abundance at $[\text{Fe}/\text{H}] \lesssim -2$. Thus, to maintain consistency with the selection in Sestito et al. (2020), we present results in Section 3.2 and after (and Table 1) using $[\text{Fe}/\text{H}] < -2.5$ as our fiducial selection for metal-poor stars. Our results do not change significantly if we use a slightly different threshold in $[\text{Fe}/\text{H}]$.

3.1.2 Comparison with the H3 survey

Recently, the H3 Spectroscopic Survey (Conroy et al. 2019) also examined the orbits of metal-poor ($[\text{Fe}/\text{H}] < -2$) stars in the MW’s halo (out of the disk; Carter et al. 2020). They also found a strong prograde bias, with nearly 70 per cent of their stars on prograde orbits (defined as $J_\phi > 0$), in general agreement with (Sestito et al. 2020), though Carter et al. (2020) found some evidence that the prograde bias increases at lower $[\text{Fe}/\text{H}]$ as well. Figure 3 (crosses with uncertainties) shows their results.

Figure 3 also shows the results from our simulation suite when applying a selection window and prograde bias metric similar to Carter et al. (2020). Specifically, we select star particles above the disk, at $1 < |Z| < 3$ kpc, and at high Galactic latitude, $|b| > 30^\circ$. Following Carter et al. (2020), we calculate the prograde bias as simply the ratio of the mass of star particles with $J_\phi > 0$ to those with $J_\phi < 0$. As in Figure 2, Figure 3 shows each of our hosts separately, and the thick black line shows the median.

The results in Figure 3 are similar to those in Figure 2 (middle). The median prograde bias is flat across a wider range of iron abundance, $-3.75 < [\text{Fe}/\text{H}] < -0.75$, before rising at higher $[\text{Fe}/\text{H}]$. The median prograde bias is consistently lower, with a value of ~ 1.6

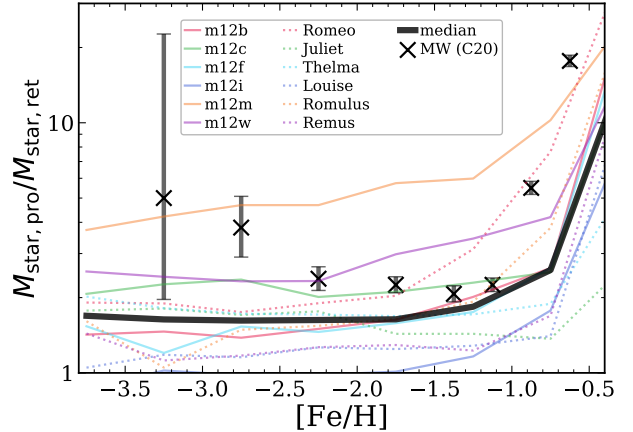


Figure 3. The prograde bias versus stellar iron abundance, similar to Figure 2 (middle panel), but using using a ‘stellar halo’ selection function to mimic the H3 survey ($1 < |Z| < 3$ kpc and $|b| > 30^\circ$) and computing the prograde bias as the mass ratio of stars with $J_\phi > 0$ to those with $J_\phi < 0$. Crosses show MW observations from the H3 survey (Carter et al. 2020, C20). Despite the differing selection function and prograde metric, the trends in our simulations remain similar to those in Figure 2. Our simulations qualitatively agree with the MW observations from C20, though now the prograde bias at $[\text{Fe}/\text{H}] < -2.5$ is lower in most simulations than the MW, and most simulations do not show as strong an increase as the MW at $[\text{Fe}/\text{H}] \gtrsim -1$, although typically at least 2 simulations have prograde biases at least as strong as these MW observations at any iron abundance.

at $[\text{Fe}/\text{H}] = -2.5$ compared to ~ 2 in Figure 2. By selecting stars at higher vertical distance from the disk and without imposing a selection on J_z , we now select stars with orbits farther from the disk, which leads to a somewhat smaller prograde bias.

Again, our simulation suite broadly agrees with MW observations from Carter et al. (2020), although less well than with the results of Sestito et al. (2020). At $-2.5 < [\text{Fe}/\text{H}] < -1$, the observations are consistent with being constant, but they show an upward trend at lower $[\text{Fe}/\text{H}]$. The values below $[\text{Fe}/\text{H}] < -2.5$ are higher than most of the simulated sample, but the uncertainties encompass much of the region spanned by our simulations. The rise of the MW at $[\text{Fe}/\text{H}] > -1$ is sharper than most of the simulations, but the simulations show similar qualitative trends.

To investigate further how this metallicity dependence changes with how we kinematically select star particles, we implemented the H3 kinematic selection ($J_\phi > 0$ versus $J_\phi < 0$) using the Pristine spatial selection window in Section 3.1.1 and saw nearly identical results to Figure 3. The medians are essentially the same up to $[\text{Fe}/\text{H}] \sim -0.5$, where the H3 selection has a slightly smaller value of ~ 10 compared to ~ 15 when using the Pristine window. This suggests that implementing a more stringent spatial selection and ignoring the disk midplane, as in the H3 survey, does not change the prograde bias results when compared to our *entire* disk selection. As a final check, we examined how the prograde bias varies using the stellar populations solely above versus below the disk: both populations have nearly identical behavior.

3.1.3 Dependence on stellar age

In general, a star’s metallicity correlates with its age, though with significant scatter. Thus we investigate how the prograde bias at $z = 0$ depends on stellar age, with no selection based on metallicity. We once again select stars as in Section 3.1.1, following the approach of Sestito et al. (2020) from the Pristine survey, though we find

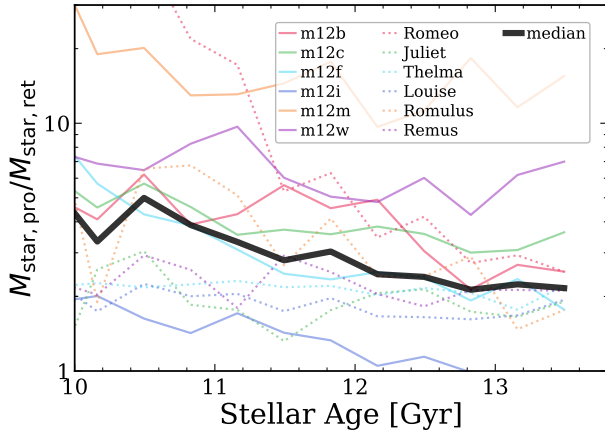


Figure 4. The prograde bias versus stellar age (regardless of iron abundance) at $z = 0$. Solid lines show the 6 isolated hosts, dotted lines show the 6 LG-like hosts, and the thick black line shows the median across all 12 hosts. The prograde bias decreases with stellar age, such that younger stars tend to be on more prograde orbits, consistent with the expectation that younger stars typically form on diskier orbits. However, even stars in our oldest age bin ($\gtrsim 13.5$ Gyr) have a median prograde bias of ~ 2.2 . All but one host (m12i) shows significant prograde bias to arbitrarily old stellar ages, and all hosts show prograde bias > 1 at stellar ages $\lesssim 12.5$ Gyr. Thus, the strong preference for prograde orbits in our simulations extends to arbitrarily old and/or arbitrarily metal-poor stars.

broadly similar results following the methodology of Carter et al. (2020) from the H3 Survey.

Figure 4 shows the the prograde bias at $z = 0$ as a function of stellar age, using age bins 0.5 Gyr wide. The prograde bias increases for younger stars, with a median prograde bias of ~ 4.5 at 10 Gyr of age, as stars transition increasingly to thin-disk formation in each host. This trend is more dramatic for even younger stars (not shown) where the prograde bias reaches ~ 100 for stars of age ~ 7 Gyr. All host galaxies except m12i have a median prograde bias of ~ 2.2 even at the oldest ages ($\gtrsim 13.5$ Gyr), and m12i does show prograde bias starting at ~ 12.5 Gyr age. We thus conclude that the prograde bias at $z = 0$ is nearly ubiquitous for star at all ages, including arbitrarily old stars. While this result might be surprising, as we will show below, it can be understood as arising from the fact that the prograde bias arises largely from a single major merger, which brings in old stars with it.

3.2 Which galaxies contributed to the prograde bias?

We next investigate the origin of the prograde bias in metal-poor stars, where we define ‘metal-poor’ as $[\text{Fe}/\text{H}] < -2.5$ hereon. Using the tracking method in Section 2.3, we first attribute all star particles as either formed in-situ (formation within 15 kpc from the host) or ex-situ. For ex-situ star particles, we track which progenitor galaxy they were a member of prior to merging into the host, and we track the total stellar mass that each progenitor contributed to the prograde metal-poor population of each host. We examine the three progenitors that contributed the most mass to the prograde metal-poor stars (or, put differently, the three progenitors that contributed the largest number of prograde metal-poor stars), and we refer to the galaxy that contributed the most, second most, and third most as the ‘primary’, ‘secondary’, and ‘tertiary’ merger, respectively. These mergers, however, are not always the top three most massive mergers to have ever contributed prograde metal-poor stars to the host. Table 1 lists values for the in-situ fractions, $f_{\text{in-situ}}$, primary

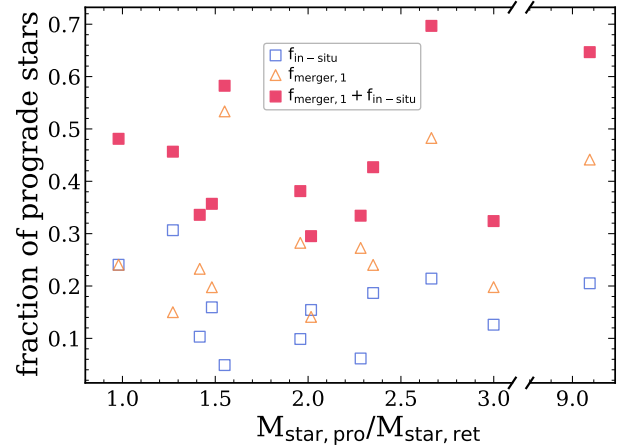


Figure 5. For each host, the fraction of all of its prograde metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars that formed in different components as a function of the host’s prograde bias at $z = 0$. One host, m12m, has a particularly high prograde bias of ≈ 9 , as the break in the x-axis indicates. Blue open squares show the fraction from in-situ ($d_{\text{form}} < 15$ kpc) stars that formed in the (most massive) progenitor of the host, while orange open triangles show the fraction from the ‘primary’ merger that contributed the most prograde stars. In nearly all (10 of 12) hosts, the primary merger contributed more prograde stars than in-situ formation. Red squares show the sum of these two, which accounts for 30 – 70 per cent of all prograde metal-poor stars.

merger fractions, $f_{\text{merger},1}$, and fraction of stars contributed from the primary, secondary, and tertiary mergers combined, $f_{\text{merger},\text{top}3}$.

Figure 5 shows the fractions of all prograde metal-poor stars that came from the primary merger (empty orange triangles) and that formed in-situ (empty blue squares) versus the prograde bias at $z = 0$ for each host. The red filled squares show the sum of these two fractions. Perhaps surprisingly, none of these fractions show any clear correlation with the strength of the prograde bias.

For all hosts except Louise and Juliet, the primary merger fractions are higher than the in-situ fractions, with median values of $f_{\text{merger},1} \approx 0.24$ and $f_{\text{in-situ}} \approx 0.16$. Thus, the primary merger generally was responsible for more prograde metal-poor stars than the (most massive progenitor of the) host galaxy. The primary merger fractions range from ≈ 15 –55 per cent: the merger in Romulus contributed the most, and the mergers in Juliet and Louise contributed the least. The in-situ fractions range from ≈ 5 –30 per cent. Furthermore, on average, the primary merger contributed more prograde stars (~ 24 per cent) than retrograde stars (~ 16 per cent) to the host galaxy.

Figure 6 shows the ratio of the stellar masses contributed by the primary (red), secondary (green), and tertiary (blue) mergers, relative to in-situ formation, of prograde metal-poor stars at $z = 0$. The square points show the median values across our 12 hosts, and the dark and light vertical bands show the 68 per cent scatter and full range, respectively. The full scatter extends up to ≈ 10 in the primary merger to in-situ ratio, though we truncate the plot for clarity. The primary merger contributed twice as many stars as in-situ formation, so the primary merger is the single most important contributor to prograde metal-poor stars. Furthermore, the secondary merger contributed comparable stellar mass as in-situ formation. Finally, the tertiary merger contributed typically only $\sim 1/2$ that of in-situ formation. We emphasize that most metal-poor stars formed early, $\gtrsim 12$ Gyr ago; this includes both stars that form in-situ and ex-situ. Almost all metal-poor stars formed prior to these three mergers

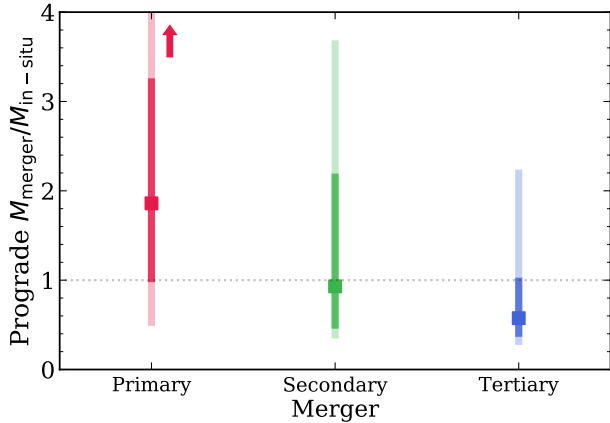


Figure 6. The ratio of the mass of prograde metal-poor stars contributed from a merger to that from in-situ formation, for the primary (red), secondary (green), and tertiary (blue) galaxy mergers, defined as the 3 progenitor galaxies that deposited the most mass to the prograde metal-poor population. Squares show the median values across all 12 hosts, and the dark and light vertical bars indicate the 68 per cent scatter and full range, respectively. The red arrow indicates that the full range of values for the primary-to-in-situ ratios extend beyond the axis (to ~ 11). Relative to in-situ stars, the primary merger typically contributed $\approx 2\times$ more stars, the secondary merger was comparable to in-situ, and the tertiary merger contributed only $\sim 1/2$ as much as in-situ. The sum of these 4 components nearly always accounts for the majority of prograde metal-poor stars (see also Table 1), and any single merger beyond the tertiary merger is typically negligible. Thus, the primary source of prograde metal-poor stars at $z = 0$ is from a single galaxy merger.

($\gtrsim 93$ per cent), thus, these mergers did not induce in-situ formation that contributed to the prograde bias.

Figure 5 and Table 1 show that in-situ formation together with the primary merger account for 30 – 70 per cent of all prograde metal-poor stars. Considering the in-situ fraction along with the top 3 mergers, this accounts for 43 – 96 per cent. Juliet has the smallest summed fraction, so an unusually large number of additional mergers contributed significantly to its prograde metal-poor stars. In m12m, m12c, and Romulus, the combined fractions reach ≈ 82 per cent or higher, and the top 3 mergers account for roughly ≈ 75 per cent. We also find no clear differences in these fractions for isolated versus paired hosts; m12c (isolated) and Romulus (LG-like) have the 2 highest primary merger fractions, and m12b (isolated) and the Romulus have the 2 lowest in-situ fractions. Furthermore, m12m, which has the highest prograde bias, has a relatively high primary merger fraction and average in-situ fraction, while m12i, which has the smallest (no) prograde bias, has average values for both. These results highlight the diversity of merger/growth histories that lead to a prograde bias for metal poor stars at $z = 0$, and that the fractions of such stars from in-situ formation or the primary merger do not predict the strength of this kinematic feature.

We thus conclude that mergers dominate the origin of prograde metal-poor stars, and the primary merger dominates over any other merger. We thus focus most of our subsequent analysis on the properties of the primary merger.

3.3 How did the prograde bias evolve over time?

We next investigate how the prograde bias evolved over time. We calculate the prograde bias at each snapshot spaced ≈ 500 Myr, selecting stars based on the criteria in Section 2.3. We measured the orientation of the host’s disk separately at each snapshot, defined

according to the moment of inertia tensor of all star particles in the host. We then calculated the prograde bias at each snapshot based on the host’s rotational velocity, v_{rot} , for normalizing J_{ϕ} . Finally, we do not limit this analysis across time to star particles that are confined to the plane of the disk throughout their lifetimes, only those that are within the action selection windows discussed above.

Figure 7 (top sub-panels) shows the evolution of the prograde bias for 6 of our hosts as a function of lookback time. These hosts show a wide range of prograde biases and histories, and they are representative of the other 6 hosts that we do not show. The vertical arrows show when each primary, secondary, and tertiary merger took place (darkest to lightest, respectively).

Figure 7 (top left) shows m12m, the host with the strongest prograde bias, ≈ 9 . As with many hosts, the prograde bias at early times was highly variable, caused by rapid merger/accretion activity and no well-defined long-lived disk. The prograde bias increased dramatically after the secondary and primary mergers, with continued gradual growth after the primary merger ~ 6 Gyr ago.

Figure 7 (top right) shows our one host, m12i, that has no significant prograde bias over the last 10.5 Gyr. The reason may be that all 3 mergers occurred at nearly the same time, possibly canceling out any coherent effects. Both m12m and m12i highlight the extremes of the prograde biases found in our simulated sample.

Both m12w (middle left) and Romeo (middle right) are interesting because their prograde bias steadily increased across most of cosmic time, and their the most recent merger was not their primary merger. The primary merger in m12w occurred ~ 8.2 Gyr ago and significantly increased the prograde bias, while the primary merger in Romeo occurred ~ 12.5 Gyr ago and had a more mild effect on the prograde bias, which subsequently decreased. Interestingly, the tertiary merger in m12w continued to drive the prograde bias upward, and more importantly, the secondary merger corresponded to the main increase for Romeo. Thus, secondary and tertiary mergers, which contributed fewer prograde metal-poor stars than the primary merger, still can effect the prograde bias. Juliet, Louise, Remus, and m12f (not shown) are similar to Romeo, because they have gradually and monotonically increasing asymmetry biases over time.

m12b (bottom left) has the latest primary merger, ~ 3 Gyr ago, which increased the prograde bias to ≈ 2 . Here, the (early) secondary and tertiary mergers did not have much of an effect. m12b shows that a galaxy can have no significant long-term prograde bias throughout most of its history, until a late merger induces one. Both Thelma and m12c (not shown) show similar behavior, with a negligible prograde bias until $\sim 6 - 7$ Gyr ago, which then continued to increase. In Thelma, there was no prograde bias until ~ 5 Gyr ago, however, the primary merger occurred ~ 8.5 Gyr ago and had no lasting affect on its prograde bias. The primary merger in m12c occurred 9 Gyr and did cause the prograde bias to spike to ~ 3.5 , however, the secondary merger caused the prograde bias to completely vanish around 0.5 Gyr later. In both m12c and Thelma, it is likely that we do not track the mergers that cause their prograde biases to increase and become long-lived.

Finally, figure 7 (bottom right) shows Romulus, which had two distinct periods of significant prograde bias. The secondary and tertiary mergers induced the first period ~ 11.5 Gyr ago, but it quickly went away. The primary merger then occurred ~ 7.5 Gyr ago, which drove the prograde bias to ~ 1.5 , where it remained to $z = 0$. This primary merger in Romulus contributed the highest fraction of prograde metal-poor stars compared to all other hosts (53 per cent). Both m12b and Romulus show how the timing of these mergers affects the prograde bias.

We tested several variations in how to measure the prograde

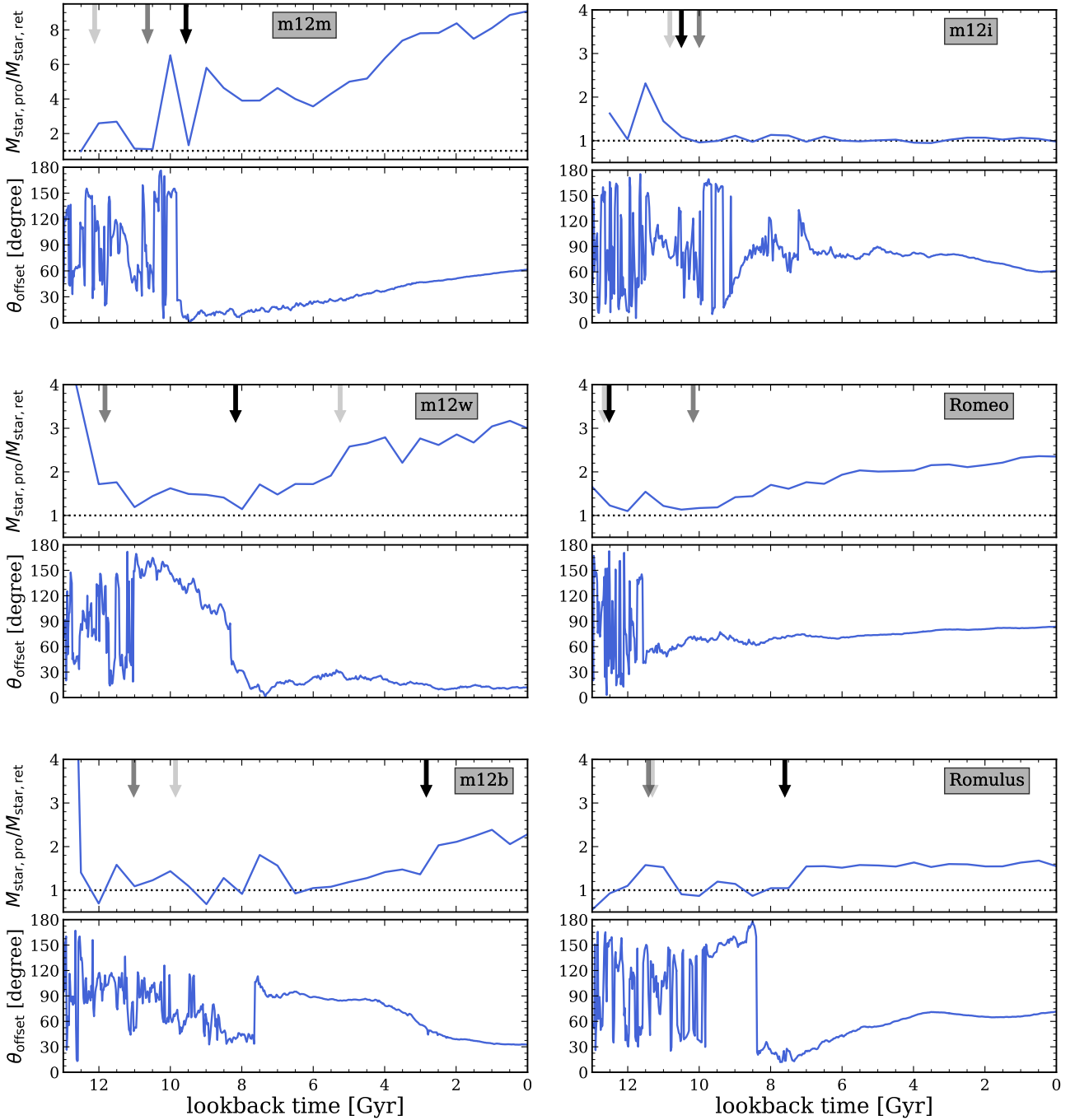


Figure 7. The formation histories for 6 galaxies in our sample. We show half of our simulated galaxies, which represent the diverse origins of the prograde bias across our suite; the other galaxies have similar behavior. Arrows indicate the times of the top three galaxy mergers, ranked by the mass that they contribute to prograde metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars at $z = 0$, with arrow darkness indicating the rank, such that the primary merger is the darkest. For each galaxy, the upper sub-panel shows the evolution of the prograde bias: the dotted horizontal line at 1 indicates no bias (equal mass in prograde and retrograde stars). At early times, the prograde bias was typically near 1, though this reflects the fact that the disk had not formed yet, so prograde and retrograde are not well defined; see discussion below. For all galaxies except m12i (top right), the prograde bias increased sharply following either the primary merger (as in m12m, m12w, m12b, Romulus) or the secondary merger (as in Romeo). m12w also shows an additional increase after the tertiary merger. For each galaxy, the lower sub-panel shows the evolution of the offset angle between the angular momentum vector of the galaxy’s disk and the angular momentum vector of the primary merger’s orbit, where we freeze the latter at its final orientation vector after the merger coalesces with the host (see Section 3.4). At early times, the offset angle of the host’s disk was un-aligned with the (future) merger, and it rapidly shifted given the high merger/accretion rates, and that the disk itself was only marginally defined. However, the host’s disk typically aligned itself with the orbital plane of the primary merger (as in m12m, m12w, m12b, Romulus) during/after the merger event. These mergers were typically gas-rich and drove (via torquing and gas deposition) the formation and/or prograde orientation of the host’s disk, seeding the origin of prograde metal-poor stars.

bias for each host across time, and we found similar overall trends using all of them. First, we tested using the same J_ϕ selection region for each host at all snapshots, instead of scaling to each host's v_{rot} at each snapshot. Second, while Figure 7 shows the prograde bias using all metal-poor stars within the host's selection region at each snapshot, we tested following back only the star particles that are in the selection region at $z = 0$. Finally, we examined how the prograde bias evolves selecting all metal-poor stars within a radius of 15 kpc from the host, as opposed to our fiducial 'disk' spatial selection. For all of these variations, we found the same general evolutionary trends, which reinforces that the prograde bias is a global feature of these MW/M31-mass galaxies.

3.4 How did the primary merger affect the orientation of the host's disk?

We next investigate the relation between the primary merger's orbit and the orientation of the stellar disk of the host galaxy. In particular, we examine whether the merger was on an orbit that was aligned with the host's pre-existing disk, whether the merger torqued the direction of a pre-existing disk, and/or whether the merger helped to seed the formation and orientation of the disk.

The bottom sub-panels in Figure 7 show the evolution of the offset angle, θ_{offset} , for each host, defined as the angle between the angular momentum of the primary merger's orbit and the angular momentum of the host galaxy's disk. We store the final angular momentum vector of the primary merger just it merged into the host, and we compare it with the evolving angular momentum of the disk after the merger.

At early times, θ_{offset} changed rapidly, caused by rapid accretion and mergers with many galaxies that built the host galaxy (e.g. Santistevan et al. 2020). As a result, the angular momentum of the host's (often poorly defined) disk quickly changed. All hosts in Figure 7 except m12b show a rapid transition from a rapidly varying θ_{offset} to a settled long-lived disk orientation. This transition often coincides with one of the major mergers.

In both m12m and m12w (top left and middle left panels, respectively), just prior to the primary merger θ_{offset} dipped significantly, and after the merger it went to near 0° . This means the merging galaxy was aligned with the *resultant* host's disk *after the merger*, that is, the galaxy merged into the host, deposited the metal-poor stars that remain until $z = 0$, and defined the prograde direction of the host's disk. These events occurred early and helped seed the formation of a stable long-lived stellar disk in the host. After these merger events, the disk continued to gradually change its orientation, presumably from interaction with other galaxies and/or accretion events. In m12m, the disk ends up tilted by $\approx 60^\circ$ with respect to the primary merger orbit. However, in m12w, the disk did not rotate much, remaining at $\approx 10^\circ$ with respect to the primary merger orbit.

Romulus (bottom right panel) shows similar behavior as m12m and m12w; starting ~ 500 Myr before the merger, the host's disk flipped from $\approx 180^\circ$ down to $\approx 20^\circ$ and stayed there for ~ 1 Gyr during/after the merger, after which it gradually increased to $\approx 70^\circ$, similar to both Romeo and m12i (middle right and top right panels, respectively). Both Juliet and m12c are similar to m12m, m12w, and Romulus, because there were temporary merging events that drove θ_{offset} down, while subsequent merger events caused θ_{offset} to rise/rotate again.

Romeo and m12i are different from the rest of the hosts shown here, because θ_{offset} remained rapidly variable both before and after the primary merger, stabilizing only $\sim 1-2$ Gyr after. The remaining

hosts not shown here (m12f, Thelma, Louise, and Remus) reflect the same behavior as Romeo and m12i: rapid change in θ_{offset} at early times, including before, during, and after the primary merger, but eventually the disk settled and θ_{offset} remained fairly constant at $30 - 140^\circ$.

Finally, m12b (bottom left panel) is unique among our suite, because the primary merger occurred late, ~ 2.9 Gyr ago. By this time, the host's disk already formed and established itself, and because primary merger was only < 20 per cent of the stellar mass of the host, the merger only moderately torqued the orientation of the host's disk. m12b provides our single example that a *late-time* merger can drive significant prograde bias as well.

As we discuss in further detail in Section 3.5, the strength of the prograde bias does not correlate with the gas or stellar masses of the primary mergers, nor the gas or stellar mass ratios of the primary mergers to their host galaxies. At first glance, it may seem surprising that there is no correlation of these merger/host properties with the prograde bias because these mergers largely source the prograde metal-poor population, however, as we mentioned above, they also deposit gas that contributes to the formation of the host's disk and set the prograde direction. These lacks of correlations hold true across different times (when the primary merger occurred, when the primary merger was at its peak stellar mass, 300 Myr before the primary merger) and across different ways we spatially and kinematically select gas particles in the galaxies.

3.5 Correlations with galaxy properties

Finally, we investigate correlations between properties of the MW/M31-mass host galaxies, or their galaxy mergers, and their prograde biases. Figure 8 shows distributions of key properties of the primary mergers and lists the Spearman correlation coefficients and p-values with respect to the prograde bias. Table 2 lists all of the correlations that we tested.

Given that both in-situ stars and the primary merger contribute significantly to the prograde stars, we explore whether hosts whose primary merger's orbit was more aligned with its disk orientation *before the merger* show a stronger prograde bias. Specifically, we examine the median θ_{offset} over the 500 Myr before the primary merger, and Figure 8 (top left) shows its relation to the prograde bias at $z = 0$ for all hosts. $\theta_{\text{offset, median}}$ ranges from $\sim 20 - 145^\circ$, with Romulus having the smallest and m12m having the largest offset. In particular, m12m has both the largest pre-merger offset angle and the largest prograde bias. m12w and m12i have values in the middle of the sample, even though they have the second-most and least prograde biases, respectively. *We thus conclude that the orientation of the primary merger's orbit with the host's pre-existing disk has no significant role in driving prograde bias for metal-poor stars.*

Figure 8 (top right) shows the prograde bias versus the look-back time of the primary merger. Almost all primary mergers occurred > 7 Gyr ago, with the sole exception of m12b, whose merger occurred ≈ 3 Gyr ago. We find a weak correlation that hosts with an earlier primary merger tend to have a smaller prograde bias, although this correlation is weak. This correlation may arise because hosts with earlier mergers have had more time to subsequently phase-mix their star via subsequent mergers, for example.

Figure 8 (bottom left) shows the prograde bias versus the stellar masses of the primary merger. Because stars can start to be stripped from the merging galaxy well before coalescence, we measure the peak stellar mass of the merging galaxy throughout its history. Among our sample, the peak stellar mass of the primary merger

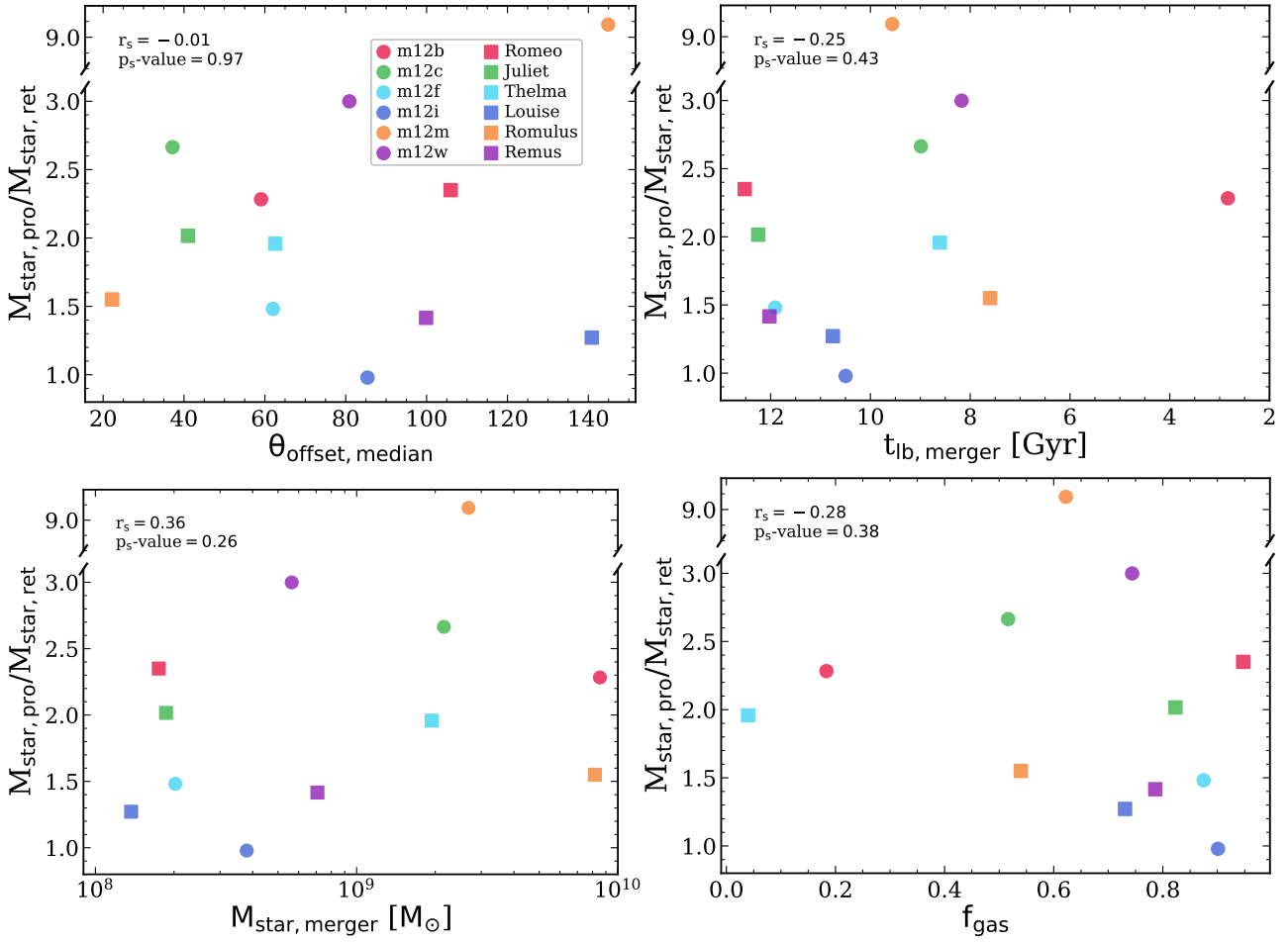


Figure 8. The prograde bias of metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars in each MW/M31-mass galaxy at $z = 0$ versus various properties of its primary merger galaxy. While we do not find strong correlations for these properties, these panels indicate the range of primary mergers to highlight their diversity. **Top left:** the median offset angle between the angular momentum vector of the host’s disk and the angular momentum vector of the primary merger’s orbit, $\theta_{\text{offset, median}}$, over 500 Myr prior to the merger. $\theta_{\text{offset, median}}$ indicates how prograde the merger’s orbit was with respect to the host’s existing disk, with 0° indicating complete alignment. The prograde bias does not correlate with $\theta_{\text{offset, median}}$; rather, the primary merger drove the orientation of the host’s disk after the merger. **Top right:** the lookback time when the primary merger occurred, $t_{\text{lb, merger}}$. Most mergers occurred 7.5 – 12.5 Gyr ago, although m12b is an example of a late merger, only ≈ 3 Gyr ago. We find a weak negative correlation with $t_{\text{lb, merger}}$, such that earlier primary mergers cause a lower prograde biases, likely because the host galaxies have had more time to merge with other galaxies, which can phase-mix the stars. **Bottom left:** the stellar mass of the primary merger, $M_{\text{star, merger}}$, which ranges from $10^8 - 10^{10} M_{\odot}$, and which ranges from 0.1 – $1\times$ the stellar mass of the host at the time of the merger. The prograde bias increases weakly with $M_{\text{star, merger}}$, as expected. **Bottom right:** the gas fraction of the primary merger, $f_{\text{gas, merger}} = M_{\text{gas, merger}}/(M_{\text{gas, merger}} + M_{\text{star, merger}})$. This ranges from $\approx 0.05 - 0.95$, though most (10 of 12) primary mergers have $f_{\text{gas, merger}} > 0.5$. The prograde bias decreases weakly with $f_{\text{gas, merger}}$. Each panel lists the Spearman correlation coefficient and p-value: we find no strong correlation between the prograde bias and these or almost any other property that we tested (see Table 2). Thus, while metal-poor stars nearly ubiquitously prefer prograde orbits in our simulations, the strength of this prograde bias has a complex dependence on formation/merger history.

spans $10^8 - 10^{10} M_{\odot}$, roughly evenly in log mass. Given the stellar masses of the host galaxies at these merger times (Santistevan et al. 2020), the primary merger galaxy was $\sim 7-95$ per cent as massive as the host during the merger. As expected, the prograde bias increases with the stellar mass of the primary merger, but the strength of this correlation is weak, as is the correlation with the ratio of stellar masses between the primary merger and the host (see Table 2).

We also examine whether the gas content of the primary merger correlates with the prograde bias. The prograde metal-poor stars formed $\gtrsim 12$ Gyr ago, before the primary merger occurred, so the gas content during the merger event did not contribute to this population. Rather, the (more metal-rich) gas deposited by the merger could contribute to the formation/stabilization of the host’s disk, because the merger deposited significant gas mass on a single angular momentum vector. Thus, one might expect that higher gas content in the

primary merger would drive stronger prograde-ness of the deposited metal-poor stars with respect to the resultant disk. Because, gas can be stripped as the merger orbits around and coalesces with the host, we measure the gas mass of the merging galaxy 300 Myr before merging. We include all gas that is within $R_{\text{star, 90}}$, the radius that encloses 90 per cent of the stellar mass of the merging galaxy, and that has relative velocity $< 2\times\sigma_{\text{star}}$, the stellar velocity dispersion of the merging galaxy. Most of these primary mergers were gas-rich, bringing in between $\approx 10^8 - 10^{10} M_{\odot}$ of gas. We then calculate their gas fractions as $f_{\text{gas, merger}} = M_{\text{gas, merger}}/(M_{\text{gas, merger}} + M_{\text{star, merger}})$.

Figure 8 (bottom right) shows the prograde bias versus the primary merger’s $f_{\text{gas, merger}}$. Most (10 of 12) primary mergers were gas-rich, with $f_{\text{gas, merger}} > 0.5$, and more than half had values > 0.75 . The two exceptions are m12b, which merged most recently, and Thelma, whose primary merger occurred ~ 9 Gyr ago but still

Table 2. Properties that we tested for a correlation with the prograde bias of metal-poor stars at $z = 0$. Column list: property name; Spearman correlation coefficient, r_s ; Spearman p-value. We calculate the mass properties/ratios at the following different times: present-day (t_0); at the time of the primary merger ($t_{\text{merger},1}$); at the time the primary galaxy merger had its peak stellar mass ($t_{\text{merger, peak}}$); and 300 Myr before the primary merger (t_{300}).

Property	r_s	p_s value
$f_{\text{in-situ}}$	-0.11	0.73
$f_{\text{merger},1}$	0.31	0.33
$M_{\text{star, host}}(t_0)$	0.31	0.33
$M_{\text{halo, host}}(t_0)$	0.10	0.76
$M_{\text{gas, host}}(t_{300})$	0.23	0.47
$M_{\text{star, host}}/M_{\text{halo, host}}(t_0)$	0.39	0.22
$M_{\text{star, merger}}(t_{\text{merger, peak}})$	0.36	0.26
$M_{\text{gas, merger}}(t_{300})$	-0.09	0.78
$M_{\text{star, merger}}/M_{\text{star, host}}(t_{\text{merger}})$	0.20	0.53
$M_{\text{star, merger}}/M_{\text{star, host}}(t_{\text{merger, peak}})$	0.06	0.85
$M_{\text{gas, merger}}/M_{\text{gas, host}}(t_{300})$	-0.57	0.05
$M_{\text{gas, total}}(t_{300})$	0.18	0.57
$f_{\text{gas, merger}}$	-0.28	0.38
$t_{\text{lb, merger}}$	-0.25	0.43
$t_{\text{lb, 50}}$	-0.43	0.16
$t_{\text{lb, 90}}$	-0.28	0.38
$t_{\text{lb, form}}$	-0.34	0.27
N_{sat}	0.52	0.09
Primary merger ordering	0.31	0.32
v_{rot}/σ	0.29	0.35
$\theta_{\text{offset, median}}$	-0.01	0.97

was gas poor. Perhaps surprisingly, we find that more gas-rich primary mergers correspond to slightly weaker prograde bias, though again this correlation is weak. For example, m12i has the smallest prograde bias but also the second-highest $f_{\text{gas, merger}} (\approx 0.9)$, while m12m has the largest prograde bias but also an intermediate $f_{\text{gas, merger}}$.

In total, we examined correlations of the prograde bias at $z = 0$ with 21 properties of the host (both today and near the time of the primary merger) or of the primary galaxy merger, which we show in Table 2. Because we measure each property in Table 2 at different times, we calculate the Spearman correlation coefficient (r_s) and p-value at present-day (t_0), at the time of the primary merger ($t_{\text{merger},1}$), at the time the primary galaxy merger had its peak stellar mass ($t_{\text{merger, peak}}$), and 300 Myr before the primary merger (t_{300}). We briefly address each below:

- $f_{\text{in-situ}}$ and $f_{\text{merger},1}$: Figure 5 also shows these correlations. While $f_{\text{merger},1}$ does positively correlate with the prograde bias, the strength of its correlation is weak, indicating additional dependence, such as time and gas-richness of the merger. Importantly, $f_{\text{in-situ}}$ shows no meaningful correlation with the prograde bias.

- The ratio of stellar mass of the primary merger and the host does not correlate significantly with prograde bias. We tested measuring this ratio both at the time when the primary merger reached its peak stellar mass ($M_{\text{star, merger}}/M_{\text{star, host}}(t_{\text{merger, peak}})$), and at the time immediately before merging into the host galaxy ($M_{\text{star, merger}}/M_{\text{star, host}}(t_{\text{merger}})$).

- $M_{\text{gas, merger}}/M_{\text{gas, host}}(t_{300})$: the ratio of the gas mass in the primary merger to that in the host galaxy. We explored four ways

of measuring gas masses, including 300 Myr or 600 Myr before the primary merger, using only cold gas ($T < 2 \times 10^4 \text{K}$), or measuring gas out to $2 \times R_{90}$. Interestingly, this ratio 300 Myr before the merger is the most significant correlation that we find (p-value = 0.05). However, we caution the reader to over-interpret this result, because the strength of this correlation varies greatly with how we measure gas masses (p-values = 0.23-0.78). In particular, the gas mass of the primary merger is relatively robust across our methods, but the gas mass of the host galaxy varies more significantly, which may be because of ambiguity in mass association during or just before a merger.

- Host mass: $M_{\text{star, host}}(t_0)$, $M_{\text{halo, host}}(t_0)$, $M_{\text{gas, host}}(t_{300})$, and $M_{\text{star, host}}/M_{\text{halo, host}}(t_0)$, correlate at most weakly with prograde bias. The strongest correlation is the current ratio of stellar mass to dark-matter halo mass.

- $M_{\text{gas, total}}(t_{300})$: the total gas mass, summing that of the primary merger and the host galaxy, does not correlate significantly with prograde bias.

- $t_{\text{lb, 50}}$, $t_{\text{lb, 90}}$, and $t_{\text{lb, form}}$: we investigated three metrics of host galaxy formation: the time when the galaxy formed 50 per cent of its final mass, the time when it reached 90 per cent of its final mass, and the time when it transitioned from having primarily ex-situ to primarily in-situ stellar mass. All of these correlate negatively with prograde bias, however, the significance of these correlations remains weak. $t_{\text{lb, 50}}$ correlates most strongly, perhaps because it correlates best with typical times of the most important mergers. Despite differences in the formation histories between galaxies in LG-like pairs, and galaxies in isolated environments mentioned in Santistevan et al. (2020), and the small positive correlation between the primary merger times ($t_{\text{lb, merger}}$) and host galaxy formation times ($t_{\text{lb, form}}$; p-value=0.05), we found no significant differences in the metal-poor, prograde biases between LG-like and isolated host galaxies, which suggests prograde bias does not correlate with host galaxy formation time.

- N_{sat} : the number of satellite dwarf galaxies ($M_{\text{star}} > 10^5 M_{\odot}$) within the host halo at $z = 0$ is curiously one of the strongest correlation that we find (p-value = 0.09). It remains unclear whether this is physically meaningful or a statistical fluke among our modest sample size. Taken at face value, it would imply that M31 has an even stronger prograde bias than the MW.

- Primary merger ordering: we compare the relative timing of the primary merger to the secondary or tertiary mergers, that is, whether the primary merger occurred first, second, or third, which we define as the primary merger ordering. We find only a weak correlation, such that hosts whose primary merger happened the most recently out of the 3 have a somewhat stronger prograde bias.

- v_{rot}/σ : the ratio of the rotational velocity to the velocity dispersion of stars in the host at $z = 0$, v_{rot}/σ is a kinematic metric of ‘diskiness’. We calculated both v_{rot} and σ using all stars in the disk, however, we also calculated the ratio using only stars that are metal-poor (threshold from $-3.5 < [\text{Fe}/\text{H}] < -0.5$). We find a marginal increase in prograde bias with v_{rot}/σ , but it is weak.

4 SUMMARY & DISCUSSION

4.1 Summary

We investigated the prevalence and origin of the preference of metal-poor stars to be on prograde disk orbits using 12 MW/M31-mass galaxies from the FIRE-2 simulation suite. We reiterate the questions that we articulated in the introduction, and we summarize

our answers to them:

i) *How commonly do MW/M31-mass galaxies show a preference for prograde motions in their metal-poor stars, as observed in the MW, and what is the range in strength of this prograde bias?*

- 11 of 12 of our MW/M31-mass galaxies show a preference for prograde orbits among metal-poor stars (Figure 1).
- Our simulations predict that this prograde bias is a general feature of MW/M31-mass galaxies throughout most of their history, and its presence does not depend significantly on the way that we spatially or kinematically select metal-poor stars (Figures 2, 3, 4).
- The present-day prograde biases for our sample range from 0.98 – 9.14, with a median value of 2.0 (Table 1).
- We find no significant differences in prograde bias between LG-like and isolated host galaxies (Figure 2).

ii) *How does this prograde bias depend on the metallicity and/or age of the stars?*

- We find little-to-no dependence of the prograde bias on the iron abundance of stars at $[\text{Fe}/\text{H}] \lesssim -1$ (Figure 2).
- Our results broadly agree with recent observations of the MW from the Pristine and H3 surveys (Figures 2 & 3).
- 11 of 12 of our galaxies have a prograde bias for arbitrarily old stars ($\gtrsim 13.5$ Gyr) regardless of metallicity, and all hosts have a significant prograde bias for stars $\lesssim 12.5$ Gyr old.

iii) *What process(es) cause this prograde bias among metal-poor stars?*

- We find no large difference in the prograde bias for stars that formed in-situ versus ex-situ (Figure 2). The sample of prograde metal-poor stars (both in-situ and ex-situ) almost entirely formed $\gtrsim 12$ Gyr ago, prior to the primary mergers.
- A single galaxy merger is typically the most important cause of the prograde bias, with additional but sub-dominant contributions from in-situ star formation and lesser mergers (Figures 5 & 6). This primary merger typically contributed ~ 24 per cent of all prograde metal-poor stars, while the fraction that formed in-situ is ~ 16 per cent. Combining in-situ stars with those from the top 3 mergers accounts for ~ 70 per cent, on average (Table 1).
- We do not find any significant correlation in orientation between the primary merger's orbit and any pre-existing disk in the host (Figure 7).
- The primary merger occurred typically 7 – 12.5 Gyr ago and typically was gas rich ($f_{\text{gas,merger}} \gtrsim 0.5$). It thus deposited a significant population of old metal-poor stars and significant gas into the host on the same orbital vector, which typically seeded/shaped the formation of a long-lived disk in the host, giving rise to metal-poor stars preferentially on a prograde disk orbit (Figures 7 & 8).
- In 3 of our simulated hosts, the host stellar disk orientation became settled/long-lived within ~ 500 Myr of 1 of the 3 mergers we tracked, and in two cases, the orientation became settled, but it subsequently changed via other mergers (Figure 7). This suggests that mergers can be a dominant factor in the formation of stellar disks.

iv) *Do any properties of the MW/M31-mass galaxy or their galaxy mergers correlate with this prograde bias?*

- We explored correlations with 21 properties of the host galaxy or primary merger, but we find few clear correlations with

the strength of the prograde bias (Table 2). This indicates that the strength of the prograde bias has a complicated dependence on galaxy formation/merger history, which limits our immediate interpretation of how to translate recent observations of the prograde bias for the MW into a robust constraint on its formation history.

4.2 Discussion

A few galaxy mergers predominantly drive the growth of the metal-poor component of MW/M31-mass galaxies/halos. For example, using a suite of dark-matter-only simulations of MW-mass halos and applying abundance matching to assign stellar mass to (sub)halos, Deason et al. (2016) found that most of the accreted stellar mass in the stellar halo comes from only 1-2 dwarf galaxies with $M_{\text{star}} = 10^8 - 10^{10} M_{\odot}$. Similarly, in Santistevan et al. (2020) we used the same 12 simulated galaxies in this paper and found that stars that form ex-situ (in galaxies other than the most massive progenitor) and accrete into the host galaxy largely come from the most massive progenitor galaxies, and that this is true for progenitors at all redshifts. Because just a few galaxies are responsible for most of the accreted (metal-poor) population, they leave a kinematic imprint, as we have shown. Accreted substructures populate our own galaxy, such as Gaia-Enceladus (Helmi et al. 2018; Belokurov et al. 2018), the Sagittarius stream (Ibata et al. 1994; Newberg et al. 2003; Majewski et al. 2003), Sequoia (Barbá et al. 2019; Myeong et al. 2019; Matsuno et al. 2019), and a few others (e.g., see Naidu et al. 2020). Therefore, we consider it likely that one (or more) of these mergers corresponds the origin of the MW's prograde bias, though we defer a more detailed connection to future work.

In their analysis of the prograde bias in the H3 survey, Carter et al. (2020) suggest that these metal-poor stars likely came from either (1) a combination of in-situ and ex-situ star formation during the early assembly and formation of the MW, which is broadly similar to what we find, or (2) late accretion of dwarf galaxies on prograde orbits with respect to the disk. However, because the stars in the MW have not been kinematically heated enough to change their prograde orbits, the authors claim that the angular momentum of the MW has been in place for at least 12 Gyr. From the bottom sub-panels of Figure 7, we see find changes in θ_{offset} at early times, with a sharp transition, after which the angle slowly precesses over time. This transition marks when the stellar disk orientation stabilized in the host, and although the disk can precess afterwards, θ_{offset} remains relatively stable. We find a range of times when this transition occurs, across $\approx 5 - 11.5$ Gyr ago. Romeo is the host galaxy that has the oldest transition akin to the MW, at ~ 11.5 Gyr ago. However, in most of our hosts, this transition occurs more recently, which suggests that the MW may be uncommon in this regard if the angular momentum of the disk has been stable for the last 12 Gyr. If instead we compare the \hat{z} orientation of the disk at $z = 0$ to the orientation at early time, we see the same qualitative behavior in Figure 7: rapid changes likely driven by early assembly, then a sharp transition to a more gradual change from disk precession.

Although mergers are violent processes, they can result in the formation of a disk (e.g. Springel & Hernquist 2005; Robertson et al. 2006; Sparre & Springel 2017; Tapia et al. 2017) or change the properties of existing disks (e.g. Toomre & Toomre 1972; Bournaud et al. 2004, 2005; Hopkins et al. 2009). For example, following a series of papers aimed at understanding disk formation through major mergers (Athanasoula et al. 2016; Rodionov et al. 2017; Peschken et al. 2017), Peschken et al. (2020) analyzed galaxies in

the Illustris simulation (Vogelsberger et al. 2014) and showed that even though 38 of them experienced a $\gtrsim 4:1$ merger since $z \sim 1.5$, they eventually formed disks by $z = 0$. The most important factor in determining the final outcome of such a merger is the amount of gas, given that stars can form from this gas after the merger to build a disk, while the older pre-existing stars can be heated during the merger to form a thick-disk/halo/ellipsoidal component in the galaxy. In our simulated sample, 300 Myr prior to the primary mergers, the gas mass ratios of the primary merging galaxy to the host span values from ~ 0.002 in Thelma, to as high as ~ 1.6 in Louise. Similarly, the total gas masses (primary merging galaxy + host) 300 Myr prior to the primary mergers span values from $\sim 5 \times 10^8 - 4 \times 10^{10} M_{\odot}$. In 3 of our hosts (m12m, m12w, and Romulus), the orientation of the disk settles from the rapid variations caused by early accretion/mergers within ≈ 500 Myr of the three mergers we investigated here, suggesting that these mergers drove the formation and prograde orientation of the disk. In 2 other hosts (m12c and Juliet), 1 of the 3 mergers caused the orientation of the disk to settle, but other subsequent mergers caused the disk to then change orientation again. Our results qualitatively agree with Peschken et al. (2020) and suggest that mergers were a dominant mode of promoting the transition to thin-disk morphology, with much of the gas mass mentioned above likely contributing to subsequent star formation in the disk region.

Recently, Sestito et al. (2021) performed a similar study as ours using 5 of the NIHAO-UHD cosmological zoom-in simulations of MW-mass galaxies (Buck et al. 2020). Building on their previous work, they found that the simulated galaxies also show a prevalence for a prograde bias. The authors also showed that a rotating spheroid of metal-poor stars, akin to a halo-like or pressure supported distribution of stars from the early accretion phase, cannot explain the origin of the prograde bias in a galaxy. Rather, the early accretion and mergers of progenitor galaxies is the dominant source of metal-poor planar stars, with a smaller additional contribution from later accretion of satellite galaxies on prograde orbits that get dragged and disrupted into the proto-disk. Furthermore, they suggest that $\sim 88 - 93$ per cent of metal-poor stars ($[\text{Fe}/\text{H}] < -2$) in those simulations are older than ~ 12 Gyr, where the retrograde planar population traces the early accretion of the main galaxy and the prograde planar population better traces the full history. Our results qualitatively agree with this picture, and we discuss the similarities and differences below.

Both our analysis and Sestito et al. (2021) find that this population of metal-poor stars within the disks of MW-mass galaxies has a preference of prograde motion as opposed to retrograde. Sestito et al. (2021) find a prograde bias in all 5 NIHAO-UHD simulated galaxies, and we find a prograde bias in 11 of our 12 galaxies. This reinforces that the prograde bias for metal-poor stars is a nearly (but not completely) ubiquitous feature of MW/M31-mass galaxies. Our sample of 12 host galaxies is larger than that in Sestito et al. (2021), which is likely why we find one host without a prograde bias. The strengths of the prograde biases differ significantly between our analysis and theirs. Sestito et al. (2021) find prograde biases of $\sim 3.5 - 9.1$, higher than the MW, while our simulations have $0.98 - 9.1$, with a median value nearly identical to the MW. In particular, 11 of our hosts have prograde biases $\lesssim 3$, below the smallest prograde bias in Sestito et al. (2021), which may indicate the importance on these results from differences in physical and/or numerical implementations between our FIRE-2 and the NIHAO-UHD simulations.

We qualitatively agree about the origins of prograde metal-poor stars. These star particles primarily came from a number of

mergers, both from the early accretion and late growth phases. Even though a non-negligible fraction of these stars form in-situ (5 – 30 per cent) in our simulations, they are too old to have formed in a disk-like component within the host. In Figure 5 of Sestito et al. (2021), massive satellite galaxies ($M_{\text{star}} \gtrsim 10^8 M_{\odot}$) merged into the host on orbits aligned with the angular momentum of the stellar disk around the time the host galaxy formed 25 per cent of its final mass (t_{25}). At these times, they report that $\sim 37 - 90$ per cent of the prograde planar stars are already in the host galaxy from the early assembly phase. However, from the bottom sub-panels in Figure 7, we show that the orientation of the proto-disk is not well defined until the primary merger (or another lesser merger) occurs, implying that the primary merger helped to seed the formation and set the orientation of the host's disk by depositing a significant supply of gas on a coherent angular momentum vector. Thus, this means that the primary merger determined the prograde direction in the host, and that these progenitor galaxies do not merge into the host on orbits aligned with the disk because the disk was not well-defined before this. Our answer to the origin of prograde metal-poor stars in the disks of MW-mass galaxies is then most similar to explanation (iii) from Sestito et al. (2019): these stars formed inside one or more progenitor galaxies that merged into the MW-mass host as it was forming.

As a final comparison, we note important differences between our analysis and Sestito et al. (2021). The authors do not explore the stability of the stellar disk orientations in their Figure 5 at t_{25} . Therefore, it is unclear what effect, if any, the satellite progenitor galaxies had on the orientation of the host. Although they quantify the growth of the accreted stellar mass of the host galaxy, they do not examine the gas mass of the merging galaxies, which plays a large role in shaping the disk. In this paper, we only focus on the evolution of the ratio of the stellar mass of prograde to retrograde stars, but Sestito et al. (2021) suggests that 74 – 90 per cent of the retrograde planar population are primarily accreted before t_{25} , while the prograde population is accreted throughout most of cosmic time. This paper also investigated several properties of the host galaxy to look for possible correlations with the strength of its prograde bias. Finally, we do not analyze other regions in kinematic space, such as the high eccentricity feature in Sestito et al. (2021).

Evidence for the prograde, metal-poor population in the MW disk is further supported by a recent study involving stars located in the stellar halo and thick disk from the *SkyMapper Survey for extremely metal-poor stars* (Da Costa et al. 2019). In a sample of 475 stars with iron abundances in the range of $-6.5 \lesssim [\text{Fe}/\text{H}] \lesssim -2$, Cordoni et al. (2020) reported that ~ 21 per cent of their sample have disk-like orbits confined to 3 kpc of the disk's midplane. The authors suggest that this sub-sample of metal-poor stars with disk-like orbits consists of a high-eccentricity and low-eccentricity population, where the low-eccentricity population are on primarily prograde orbits (defined as $J_{\phi} > 0$) with a prograde bias of ~ 3.1 . Cordoni et al. (2020) postulate that the prograde and retrograde low-eccentricity stars confined to $|Z| < 3$ kpc likely have different origins: the prograde stars possibly formed in the thick disk of the MW, while the retrograde were accreted from a disrupted satellite galaxy. These results broadly agree with what we presented here, however, it is likely that some of the prograde, low-eccentricity stars formed both in-situ and were accreted early in the MW's formation history.

Similar studies focused on old, metal-poor stars in the MW also halo have investigated their origin and kinematics. By selecting stars within 3 kpc of the Sun with halo-like kinematics, Bonaca et al. (2017) found that there are an excess of metal-rich halo stars

($[\text{Fe}/\text{H}] > -1$) with prograde orbits compared to the metal-poor halo stars ($[\text{Fe}/\text{H}] < -1$) which have no orbital preference. The authors compared this with one of the simulations in our sample, m12i, and found similar results. They found that the two halo populations also have different origins, with the metal-rich stars primarily forming in-situ, and the metal-poor population accreting from mergers (ex-situ). A similar study using the APOSTLE simulations by Starkenburg et al. (2017a) suggests that the fractions of metal-poor ($[\text{Fe}/\text{H}] < -2.5$) and old ($t_{\text{form}} < 0.8$ Gyr) stars increase for larger galactocentric radii, with ~ 60 per cent of the stars orbiting outside of the solar circle. We did not examine the prograde bias of the halo region in our simulations, but rather, we compared how the prograde bias evolution changed when selecting *all* metal-poor ($[\text{Fe}/\text{H}] < -2.5$) stars within 15 kpc, radially. We saw the same qualitative evolution of the prograde biases as in the top sub-panels of Figure 7. This is not a strict ‘halo selection’ of stars, but given that we showed that the prograde bias is a general feature, independent of the kinematic and spatial selection windows we examined, we do not expect our results to change significantly by making a more strict selection of halo stars. Similarly, El-Badry et al. (2018b) investigated the spatial distribution and dynamics of the oldest stars ($z_{\text{form}} \gtrsim 5$) in MW-mass galaxies and found that the majority formed ex-situ. The authors found that metal-rich stars ($[\text{Fe}/\text{H}] > -1$) typically have disk-like orbits, while the more metal-poor stars reflect an isotropic distribution. Furthermore, when only selecting the oldest stars ($z_{\text{form}} \gtrsim 5$), the in-situ and ex-situ populations in m12f have nearly identical velocity distributions, however, both populations do show a slight prograde bias, which are centered at $v_{\phi} > 0 \text{ km s}^{-1}$.

Instead of selecting stars based on their distributions in velocity space, as in a Toomre diagram, one recent study tested a Gaussian mixture model to deconstruct the different components of the MW using the star’s velocities and iron abundances (Nikakhtar et al., in prep). The authors used data from APOGEE (Majewski et al. 2017) and Gaia DR2 (Gaia Collaboration et al. 2018), and created mock catalogs using 3 of the simulations in our study (m12f, m12i, m12m), and showed that the MW is best described by five different components: the stellar halo, the thin disk, the metal-rich thick disk, and a 2-component metal-poor thick disk. Nikakhtar et al. (in prep) suggest that the 3-component thick disk originates from a combination of kinematic heating (via gravitational interactions) and radial migration of early forming stars: we show that a significant fraction of the most metal-poor stars in the disk come from multiple mergers during the galaxy’s early assembly. In a different approach, using the same 12 simulated galaxies as in our sample, Yu et al. (in prep) differentiated thin disk and thick disk stars based on their circularity (angular momenta relative to the angular momentum of a circular orbit) and suggest that the transition from thick disk to thin disk formation correlates with the transition from early, bursty star formation to steady, near constant star formation. This transition occurred some time between $\sim 0.6 - 6.5$ Gyr ago, and galaxies that had earlier transitions often have older thick disk stars and more prominent thin disks.

One interesting avenue for future work is to investigate the correlation of the prograde bias with larger-scale structure. To first order, we found no distinct differences in the metal-poor prograde bias between isolated and LG-like environments, which suggests that the prograde bias is not inherent in LG-like environments only, as in the MW.

ACKNOWLEDGEMENTS

We greatly appreciate interesting and fruitful discussions with both Federico Sestito and Nicolas Martin, as well as Courtney Carter and Charlie Conroy. We also express our gratitude for their sharing of observational data. IBS, AW, and JS received support from NASA through ATP grants 80NSSC18K1097 and 80NSSC20K0513; HST grants GO-14734, AR-15057, AR-15809, and GO-15902 from STScI; a Scialog Award from the Heising-Simons Foundation; and a Hellman Fellowship. AW performed this work in part at KITP, supported by NSF grant PHY-1748958. CAFG was supported by NSF through grants AST-1715216 and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant HST-AR-16124.001-A; and by a Cottrell Scholar Award and a Scialog Award from the Research Corporation for Science Advancement. RES acknowledges support from NASA grant 19-ATP19-0068, NSF grant AST-2009828, and HST-AR-15809 from the Space Telescope Science Institute (STScI), which is operated by AURA, Inc., under NASA contract NAS5-26555. We ran simulations using: XSEDE, supported by NSF grant ACI-1548562; Blue Waters, supported by the NSF; Pleiades, via the NASA HEC program through the NAS Division at Ames Research Center.

This paper used various python packages including NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), and MATPLOTLIB (Hunter 2007), as well as NASA’s Astrophysics Data System.

DATA AVAILABILITY

[Fill out later]

REFERENCES

- Athanassoula E., Rodionov S. A., Peshken N., Lambert J. C., 2016, *ApJ*, **821**, 90
- Barbá R. H., Minniti D., Geisler D., Alonso-García J., Hempel M., Monachesi A., Arias J. I., Gómez F. A., 2019, *ApJ*, **870**, L24
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, *ApJ*, **762**, 109
- Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013b, *ApJ*, **763**, 18
- Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, *Monthly Notices of the Royal Astronomical Society*, **478**, 611
- Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, *ApJ*, **773**, 43
- Bird J. C., Loebman S. R., Weinberg D. H., Brooks A., Quinn T. R., Christensen C. R., 2020, arXiv e-prints, p. arXiv:2005.12948
- Bonaca A., Conroy C., Wetzel A., Hopkins P. F., Kereš D., 2017, *ApJ*, **845**, 101
- Bournaud F., Combes F., Jog C. J., 2004, *A&A*, **418**, L27
- Bournaud F., Jog C. J., Combes F., 2005, *A&A*, **437**, 69
- Brook C. B., Kawata D., Scannapieco E., Martel H., Gibson B. K., 2007, *ApJ*, **661**, 10
- Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, *MNRAS*, **491**, 3461
- Bullock J. S., Johnston K. V., 2005, *ApJ*, **635**, 931
- Bullock J. S., Kravtsov A. V., Weinberg D. H., 2001, *ApJ*, **548**, 33
- Carney B. W., Laird J. B., Latham D. W., Aguilar L. A., 1996, *AJ*, **112**, 668
- Carter C., et al., 2020, arXiv e-prints, p. arXiv:2012.00036
- Chiba M., Beers T. C., 2000, *AJ*, **119**, 2843
- Conroy C., et al., 2019, *ApJ*, **883**, 107
- Cordoni G., et al., 2020, *MNRAS*,
- Da Costa G. S., et al., 2019, *MNRAS*, **489**, 5900
- De Silva G. M., et al., 2015, *MNRAS*, **449**, 2604
- Deason A. J., Belokurov V., Evans N. W., 2011, *MNRAS*, **411**, 1480
- Deason A. J., Mao Y.-Y., Wechsler R. H., 2016, *ApJ*, **821**, 5

- Di Matteo P., Spite M., Haywood M., Bonifacio P., Gómez A., Spite F., Caffau E., 2020, *A&A*, **636**, A115
- Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, *ApJ*, **136**, 748
- El-Badry K., et al., 2018a, *MNRAS*, **473**, 1930
- El-Badry K., et al., 2018b, *Monthly Notices of the Royal Astronomical Society*, **480**, 652
- Escala I., et al., 2018, *MNRAS*, **474**, 2194
- Fall S. M., Efstathiou G., 1980, *MNRAS*, **193**, 189
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, **703**, 1416
- Freeman K., Bland-Hawthorn J., 2002, *ARA&A*, **40**, 487
- Gaia Collaboration et al., 2018, *A&A*, **616**, A1
- Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, *Nature Astronomy*, **3**, 932
- Garrison-Kimmel S., et al., 2017, *MNRAS*, **471**, 1709
- Garrison-Kimmel S., et al., 2018, *MNRAS*, **481**, 4133
- Garrison-Kimmel S., et al., 2019a, *MNRAS*, **487**, 1380
- Garrison-Kimmel S., et al., 2019b, *MNRAS*, **489**, 4574
- Gilmore G., Reid N., 1983, *MNRAS*, **202**, 1025
- Grand R. J. J., Springel V., Gómez F. A., Marinacci F., Pakmor R., Campbell D. J. R., Jenkins A., 2016, *MNRAS*, **459**, 199
- Griffen B. F., Dooley G. A., Ji A. P., O’Shea B. W., Gómez F. A., Frebel A., 2018, *MNRAS*, **474**, 443
- Guedes J., Callegari S., Madau P., Mayer L., 2011, *ApJ*, **742**, 76
- Hahn O., Abel T., 2011, *MNRAS*, **415**, 2101
- Harris C. R., et al., 2020, *Nature*, **585**, 357
- Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, *ApJ*, **863**, 113
- Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, *Nature*, **563**, 85
- Hopkins P. F., 2015, *MNRAS*, **450**, 53
- Hopkins P. F., 2016, *MNRAS*, **455**, 89
- Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, *ApJ*, **691**, 1168
- Hopkins P. F., et al., 2018, *MNRAS*, **480**, 800
- Hunter J. D., 2007, *Computing in Science Engineering*, **9**, 90
- Ibata R. A., Gilmore G., Irwin M. J., 1994, *Nature*, **370**, 194
- Kroupa P., 2001, *MNRAS*, **322**, 231
- Krumholz M. R., Gnedin N. Y., 2011, *ApJ*, **729**, 36
- Leitherer C., et al., 1999, *ApJS*, **123**, 3
- Li H.-N., Zhao G., Christlieb N., Wang L., Wang W., Zhang Y., Hou Y., Yuan H., 2015, *ApJ*, **798**, 110
- Li H., Tan K., Zhao G., 2018, *ApJS*, **238**, 16
- Loebman S. R., Roškar R., Debattista V. P., Ivezić Ž., Quinn T. R., Wadsley J., 2011, *ApJ*, **737**, 8
- Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Quataert E., 2017, *MNRAS*, **467**, 2430
- Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, *ApJ*, **599**, 1082
- Majewski S. R., et al., 2017, *AJ*, **154**, 94
- Matsuno T., Aoki W., Suda T., 2019, *ApJ*, **874**, L35
- Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, **295**, 319
- Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, *MNRAS*, **488**, 1235
- Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, arXiv e-prints, p. [arXiv:2006.08625](https://arxiv.org/abs/2006.08625)
- Newberg H. J., et al., 2003, *ApJ*, **596**, L191
- Nissen P. E., Schuster W. J., 2010, *A&A*, **511**, L10
- Norris J. E., Ryan S. G., 1989, *ApJ*, **336**, L17
- Ostriker J. P., Tremaine S. D., 1975, *ApJ*, **202**, L113
- Peschken N., Athanassoula E., Rodionov S. A., 2017, *MNRAS*, **468**, 994
- Peschken N., Lokas E. L., Athanassoula E., 2020, *MNRAS*, **493**, 1375
- Planck Collaboration et al., 2018, arXiv e-prints, p. [arXiv:1807.06209](https://arxiv.org/abs/1807.06209)
- Rees M. J., Ostriker J. P., 1977, *MNRAS*, **179**, 541
- Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, *ApJ*, **645**, 986
- Rodionov S. A., Athanassoula E., Peschken N., 2017, *A&A*, **600**, A25
- Samuel J., Wetzel A., Chapman S., Tollerud E., Hopkins P. F., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., 2020a, arXiv e-prints, p. [arXiv:2010.08571](https://arxiv.org/abs/2010.08571)
- Samuel J., et al., 2020b, *MNRAS*, **491**, 1471
- Sanderson R. E., et al., 2018, *ApJ*, **869**, 12
- Sanderson R. E., et al., 2020, *ApJS*, **246**, 6
- Santistevan I. B., Wetzel A., El-Badry K., Bland-Hawthorn J., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., Benincasa S., 2020, *MNRAS*, **497**, 747
- Scannapieco E., Kawata D., Brook C. B., Schneider R., Ferrara A., Gibson B. K., 2006, *ApJ*, **653**, 285
- Searle L., Zinn R., 1978, *ApJ*, **225**, 357
- Sestito F., et al., 2019, *MNRAS*, **484**, 2166
- Sestito F., et al., 2020, *MNRAS*, **497**, L7
- Sestito F., et al., 2021, *MNRAS*, **500**, 3750
- Sparre M., Springel V., 2017, *MNRAS*, **470**, 3946
- Springel V., Hernquist L., 2005, *ApJ*, **622**, L9
- Starkenbourg E., Oman K. A., Navarro J. F., Crain R. A., Fattahi A., Frenk C. S., Sawala T., Schaye J., 2017a, *MNRAS*, **465**, 2212
- Starkenbourg E., et al., 2017b, *MNRAS*, **471**, 2587
- Steinmetz M., et al., 2006, *AJ*, **132**, 1645
- Su K.-Y., Hopkins P. F., Hayward C. C., Faucher-Giguère C.-A., Kereš D., Ma X., Robles V. H., 2017, *MNRAS*, **471**, 144
- Tapia T., Eliche-Moral M. C., Aceves H., Rodríguez-Pérez C., Borlaff A., Querejeta M., 2017, *A&A*, **604**, A105
- Toomre A., Toomre J., 1972, *ApJ*, **178**, 623
- Venn K. A., et al., 2020, *MNRAS*, **492**, 3241
- Villalobos Á., Helmi A., 2008, *MNRAS*, **391**, 1806
- Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
- Vogelsberger M., et al., 2014, *MNRAS*, **444**, 1518
- Walker I. R., Mihos J. C., Hernquist L., 1996, *ApJ*, **460**, 121
- Wetzel A., Garrison-Kimmel S., 2020a, HaloAnalysis: Read and analyze halo catalogs and merger trees (ascl:2002.014)
- Wetzel A., Garrison-Kimmel S., 2020b, GizmoAnalysis: Read and analyze Gizmo simulations (ascl:2002.015)
- Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, *ApJ*, **827**, L23
- White S. D. M., Frenk C. S., 1991, *ApJ*, **379**, 52
- White S. D. M., Rees M. J., 1978, *MNRAS*, **183**, 341
- Zinn R., 1993, in Smith G. H., Brodie J. P., eds, *Astronomical Society of the Pacific Conference Series Vol. 48, The Globular Cluster-Galaxy Connection*. p. 38

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.