Title

The Galaxy Progenitors of Stellar Streams around Milky Way-mass Galaxies in the FIRE Cosmological Simulations

Permalink

https://escholarship.org/uc/item/8t98k6cx

Authors

Panithanpaisal, Nondh
Sanderson, Robyn E
Wetzel, Andrew
et al.

Publication Date

2021-04-19

Peer reviewed

The Galaxy Progenitors of Stellar Streams around Milky Way-mass Galaxies in the FIRE Cosmological Simulations

Nondh Panithanpaisal [ID],[1] Robyn E. Sanderson [ID],[1, 2] Andrew Wetzel [ID],[3] Emily C. Cunningham [ID],[2]
Jeremy Bailin,[4] and Claude-André Faucher-Giguère[5]

[1] Department of Physics & Astronomy, University of Pennsylvania, 209 S 33rd St., Philadelphia, PA 19104, USA
[2] Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA
[3] Department of Physics & Astronomy, University of California, Davis, CA 95616, USA
[4] Department of Physics and Astronomy, University of Alabama, Box 870324, Tuscaloosa, AL, 35487, USA
[5] Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

ABSTRACT

Stellar streams record the accretion history of their host galaxy. We present a set of simulated streams from disrupted dwarf galaxies in 13 cosmological simulations of Milky Way (MW)-mass galaxies from the FIRE-2 suite at $z = 0$, including 7 isolated Milky Way-mass systems and 6 hosts resembling the MW-M31 pair (full dataset at: https://flathub.flatironinstitute.org/sapfire). In total, we identify 106 simulated stellar streams, with no significant differences in the number of streams and masses of their progenitors between the isolated and paired environments. We resolve simulated streams with stellar masses ranging from $\sim 5 \times 10^5$ up to $\sim 10^9 M_\odot$, similar to the mass range between the Orphan and Sagittarius streams in the MW. We confirm that present-day simulated satellite galaxies are good proxies for stellar stream progenitors, with similar properties including their stellar mass function, velocity dispersion, [Fe/H] and [α/H] evolution tracks, and orbital distribution with respect to the galactic disk plane. Each progenitor's lifetime is marked by several important timescales: its infall, star-formation quenching, and stream-formation times. We show that the ordering of these timescales is different between progenitors with stellar masses higher and lower than $\sim 2 \times 10^6 M_\odot$. Finally, we show that the main factor controlling the rate of phase-mixing, and therefore fading, of tidal streams from satellite galaxies in MW-mass hosts is non-adiabatic evolution of the host potential. Other factors commonly used to predict phase-mixing timescales, such as progenitor mass and orbital circularity, show virtually no correlation with the number of dynamical times required for a stream to become phase-mixed.

Keywords: stellar stream, tidal stream, cosmological simulations

1. INTRODUCTION

Cosmological simulations predict that galaxies on the scale similar to the Milky Way (MW) contain significant number of gravitationally bound halos within their vicinity. However, the mass profiles and distribution of these halos are dependent on the detailed properties of the underlying dark matter model, especially its temperature. The scales where the primodial matter power spectrum is suppressed set the lower limit on their mass.

In the cold dark matter (CDM) model, structures are formed in a hierarchical fashion. Small halos that are

Corresponding author: Nondh Panithanpaisal
nondh@sas.upenn.edu

bound under gravitational influence of more massive halos, so called subhalos, merge to form smoother and more massive halos through tidal disruptions. CDM models, such as the weakly interacting massive particle (WIMP), suggest the halo lower mass limit as low as $10^{-11} M_\odot$ (Bringmann 2009; Cornell et al. 2013). However, low-mass subhalos have little to no stellar content, rendering direct detections ineffective. Determining whether these dark subhalos exist, hence, serves as an important test to the CDM. The non-existence of these dark subhalos would favor alternative models of dark matter where there is a suppression in the primodial matter power spectrum at the scales of the classical dwarf galaxies, such as the warm dark matter, which

predicts fewer low-mass subhalos (e.g. Knebe et al. (2003))

Long cold stellar streams, which are tidally disrupted dwarf galaxies or globular clusters, can be used to detect subhalos in the MW. The stream is sensitive to heating by repeated encounters with low-mass ($\sim 10^6 M_\odot$) dark subhalos (Ibata et al. 2002; Carlberg 2012, 2013; Erkal et al. 2016). Simple simulations of the interaction of such streams with a subhalo even of mass $< 10^7 M_\odot$ shows that it creates a discontinuity in the stream's orbital energy distribution, which will later evolve into a gap or fluctuation in surface density of stars along the stream (Yoon et al. 2011). Numerous cold streams have been observed in the Milky Way of which some, such as GD-1 and Pal 5, exhibit features that could be due to interactions with subhalos (Bonaca et al. 2014, 2019; Carlberg et al. 2012; Ibata et al. 2016; Bernard et al. 2016). These gaps, if really produced by stream-subhalo interaction, can possibly be modeled to estimate individual masses and mass profile of the interacting subhalo.

Formulating a robust stream-subhalo interaction model requires understanding dynamics of individual stars along the stream as well as dynamics of the entire stream within the main galaxy. These properties are directly related to the progenitor of each stream such as its initial mass, concentration, velocity dispersion and orbit. Many of these either have been assumed (such as inferring mass of the progenitor from velocity dispersion of the stream) or have only been studied in DM-only simulations.

In this paper, we study the physical and dynamical properties of stellar streams and their dwarf galaxy progenitors in MW-like galaxies in a suite of fully cosmological-hydrodynamical simulations. Section 2 describes this suite, the FIRE-2 hydrodynamical simulations. Section 3 presents the criteria used to distinguish stellar streams from bound satellites and phase-mixed accreted structures in these simulations at the present day. In Section 4 we present the properties of the progenitor galaxies that form these streams. In Section 5 we examine the orbital characteristics of the progenitors as they form streams. In Section 6 we look at the time-evolution of the velocity dispersion in the streams and evaluate its usefulness as an indicator of the progenitor mass and/or the age of the stream. In Section 7 we discuss how the stellar stream progenitors compare to satellite galaxies in the simulations that are still bound at the present day, and how they compare to observed satellite galaxies in the MW and M31. Section 8 summarizes our findings.

2. SIMULATIONS

For this work, we use cosmological zoom-in baryonic simulations of MW-mass galaxies from the Feed-

back In Realistic Environments (FIRE) project[1]. These simulations are run with Gizmo (Hopkins 2015), which uses an optimized TREE+PM gravity solver and a Lagrangian mesh-free, finite-mass method for accurate hydrodynamics. Star formation and stellar feedback are implemented using the state-of-the-art FIRE-2 physics model (Hopkins et al. 2018), which adopts a "bottom-up" approach for modeling the dense multiphase inter-stellar medium in galaxies and mono-age, mono-abundance stellar populations, and takes stellar feedback parameters directly from stellar evolution models like STARBURST99 (Leitherer et al. 1999).

The FIRE-2 simulations produce MW-mass galaxies with many properties that are broadly consistent with observations of the MW, including: MW-like thin stellar disks with transient bars and spiral structure (Ma et al. 2017; Hopkins et al. 2018; Debattista et al. 2018; Sanderson et al. 2018), a realistic distribution of massive GMCs (Guszejnov et al. 2020; Benincasa et al. 2020), stellar halos from disrupted satellite galaxies (Sanderson et al. 2017), and realistic populations of surviving satellite dwarf galaxies that do not suffer from the "missing satellites" or "too-big-to-fail" problems (Wetzel et al. 2016a, Garrison-Kimmel et al. 2018a, Samuel et al. 2019). Part of the reason why satellite dwarf galaxies in FIRE-2 simulations agree well with observations of the MW (and M31) is that these simulations self-consistently form a massive central MW-like galaxy, which efficiently destroys satellite galaxies *and low-mass subhalos* that orbit within the inner ~ 30 kpc via gravitational tidal forces (Garrison-Kimmel et al. 2017). The tidal destruction leads to satellite radial distributions that agree with MW and M31, especially in the inner ~ 150 kpc of the halo, as well as reasonably accurate planes of satellite dwarf galaxies for some of the hosts (Samuel et al. 2020b,a). The consistency of the phase-space volumes and orbits of the present-day satellites with observations, as well as the overall consistency of the stellar halo masses with observations, together support our expectation that a realistic population of streams should be created from tidally disrupted "building blocks" that are sufficiently resolved by the simulations. However, because of the particle resolution, stellar streams from globular clusters do not natively form in our simulations. In §4.5, we also explicitly compare these building blocks with the present day simulated satellites to determine whether this expectation is reasonable.

In this paper we study the tidal streams from dwarf galaxy progenitors around 13 MW-mass halos: 7 individual MW-mass halos and 3 Local-Group-like MW+M31 pairs. The isolated hosts are from the Latte suite, first presented in Wetzel et al. (2016b), while the paired hosts are from the ELVIS on FIRE suite, first presented in Garrison-Kimmel et al. (2019). All halos were

---

simulated in ΛCDM cosmology at particle mass resolution of 3500–7100 $M_\odot$ and spatial resolution of 1–4 pc for star/gas particles; 18,000–35,000 $M_\odot$ and 40 pc for DM particles. Each simulation stores 600 snapshots from $z = 99$ to 0, providing post-processing time resolution of ∼ 20 Myr—about 20 snapshots per orbit at 10 kpc, and more at larger galactocentric distances—allowing us to carry out dynamical analysis of the streams by post-processing.

The resolution of this suite of simulations allows both luminous and dark subhalos to be resolved well even near each MW-like galaxy, and follows the formation of tidal streams from dwarf galaxies down to slightly below the mass of the MW's "classical" dSphs: around 10^8 $M_\odot$ in total mass or 10^6 $M_\odot$ in stellar mass (at $z = 0$). We will discuss the effects of resolution on our results throughout the paper. For now we note that given the steep mass function of stellar halo building blocks, this lower limit in stellar mass for resolved building blocks corresponds to more than 99% of the total accreted mass in the stellar halo for a given simulation (Deason et al. 2015; Santistevan et al. 2020a).

The dark matter particles in each snapshot of the finished simulations are processed with `Rockstar` (Behroozi et al. 2013a) to produce halo catalogs. These are connected in time using `consistent-trees` (Behroozi et al. 2013b) to form a merger tree. Once the merger tree is constructed, we make a preliminary assignment of star particles to each halo and subhalo identified by `Rockstar` in each snapshot, by selecting star particles within the halo's virial radius and within twice the halo circular velocity relative to the halo's center. This is a conservative criterion that excludes nearly all stream stars; we use it as a starting point for collecting the full stellar distribution as described in §3.1. The post-processing is done using `gizmo_analysis` (Wetzel & Garrison-Kimmel 2020a) and `halo_analysis` (Wetzel & Garrison-Kimmel 2020b)

3. SELECTING STELLAR STREAMS IN THE SIMULATIONS

We select stream-like objects in each cosmological simulation using a standardized set of criteria. This section explains the steps that were taken to identify these streams, and their progenitors, for further study.

3.1. *Stream Candidate Identification*

We select candidates with non-zero stellar mass that were self-bound objects at some point in the past (and may or may not be bound at present day). Since it takes several dynamical times for a bound satellite to turn into a stream, we select structures that are bound at any time 2.7–6.5 Gyr ago (corresponding to $z \sim 0.25-0.75$) and whose star particles are within the virial radius of the main galaxy at present day. We keep track of each selected object within this time range to eliminate duplicates. Extending the time window to earlier times,

and increasing the tracking radius to larger than the present-day virial radius of the main galaxy, does not increase the total number of coherent streams. As we will show in §4.4, most satellite galaxies the are phase-mixed at present day become unbound before $z \sim 0.75$, given that the underlying potential of the main galaxy in most simulations is still changing non-adiabatically up until this point (see also Santistevan et al. 2020b). Moreover, at earlier times, there is not necessarily an unambiguous single host galaxy, so whether an object is accreted is not well-defined.

We track each substructure selected in this manner back in time to its formation, and forward to the point where it has merged with the main galaxy (or to the present day for objects that remain partially self-bound). We discard massive subhalos that contain more star particles than our upper-limit threshold stated in §3.2. A subhalo is considered "merged" when `consistent-trees` no longer continuously tracks its center. Since the code allows for individual subhalos to be intermittently missed by the halo finder, assigning a provisional placeholder or "phantom halo" to keep track of its expected position, this effectively means that `Rockstar`, a 6-D phase-space finder, must be unable to find the subhalo in > 3 consecutive snapshots. This is important for our study since halo finders often lose track of subhalos for a snapshot or two at pericenter; by bridging these gaps we can continue tracing the orbits of stream progenitors until a fairly advanced stage of tidal stripping.

To collect all the star particles that ever belonged to a given building block, we first determine the time in which the bound satellite obtains maximum stellar mass. We collect all star particles that were assigned to the same subhalo by our preliminary tracking (see §2) in any snapshot in a 200-Myr window around this time. This allows us to recover nearly all star particles associated with the subhalo, including those that may only be marginally bound. These star particles are then tracked up until present day to determine their current positions. We carry out this procedure for all the stream candidates identified in each MW-mass host galaxy in the 7 isolated and 3 paired systems in our suite for a total of 13 hosts. This process governs ∼ 15 − 75% of the accreted stellar halo since the time when the mass ratio of the main halo to the second most massive halo is 3:1. This wide range corresponds to a diverse accretion history among our simulations (Santistevan et al. 2020a).

3.2. *Stream Candidate Classification*

We use the following three criteria to determine whether star particles assigned to each accreted satellite galaxy in the previous section form a stellar stream at present day:
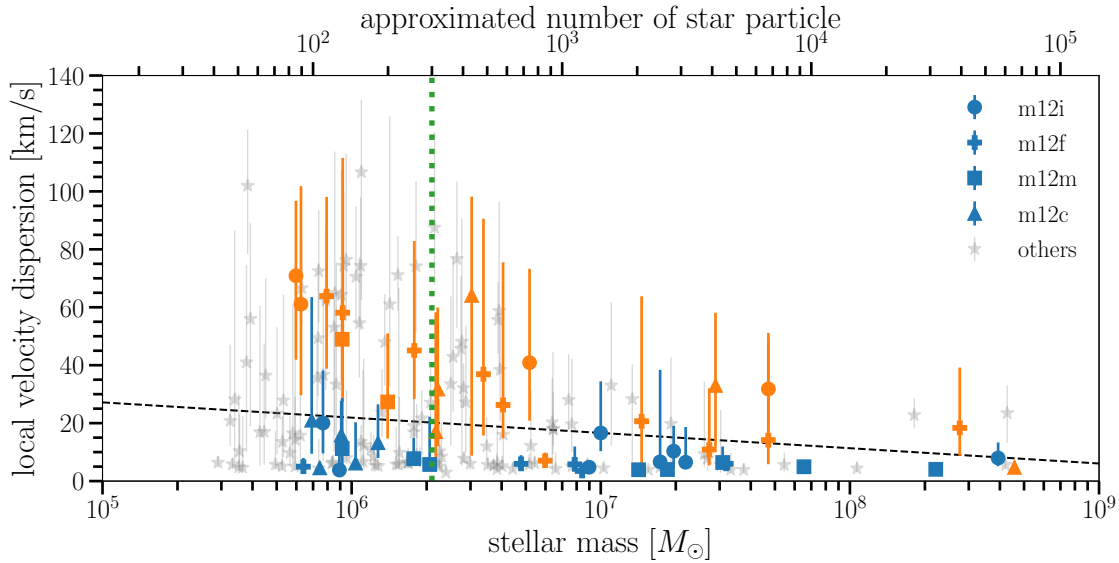
Figure 1. Local velocity dispersions of stream candidates that meet criteria (i) and (ii) of §3.2. The marker is the median value and the error bar spans the 16th to 83rd percentiles of particles across the stream. The orange points are identified by eye as phase-mixed, while the blue points are identified by eye as streams. The gray points are all of the stream candidates from other simulations. Stream candidates above the local velocity dispersion threshold determined by the SVM (black dashed line) are classified as phase-mixed. The green vertical line marks the transition between using 7 nearest neighbors (to the left) and using 20 nearest neighbors (to the right) to estimate local velocity dispersions.
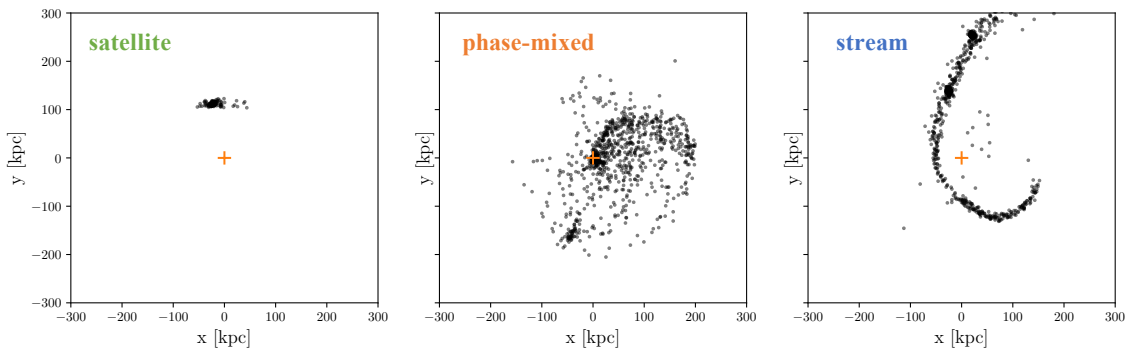
Figure 2. Example objects, showing satellite galaxy (left, `m12i`), phase-mixed (middle, `m12b`) and stream (right, `m12b`) structures as determined by the criteria listed in §3.2. The satellite fails criterion (ii) (on the distance between star particles), while the phase-mixed object fails criterion (iii) (on the local velocity dispersion). Note the that distance spread of this particular satellite galaxy is ~ 96 kpc, which is less than the distance threshold of 120 kpc in criterion (ii).

1. **Number of star particles.** The minimum number of star particles is greater than 120 and the maximum number of star particles is less than 10^5.

2. **Distance between star particles.** The maximum value of the pairwise separation between any two star particles in the group is greater than 120 kpc.

3. **Local velocity dispersion.** The median of the local velocity dispersion of star particles, $\langle\sigma\rangle$, and the total stellar mass of the object, $M_\star$, satisfies Equation 2.

We impose a lower bound on the number of star particles to ensure that all selected streams have enough star particles to be at least marginally resolved in the simulations. In the isolated simulations, since each evolved star particle has typical mass of $\sim 5000 M_\odot$, the streams and their progenitors have a lower mass limit of around $10^6 M_\odot$. This lower bound limits our study to simulated streams from dwarf galaxies at slightly above the mass of a classical dwarf satellite in the Milky Way. Stellar streams from globular clusters do not natively form in our simulations. An upper bound on the number of star particles is imposed to rule out objects that are even more massive than the Sagittarius stream, the most

massive coherent stream known to exist in the real MW. The star particle mass resolution is slightly better in the paired simulation with each evolved star particle having a mass of $\sim 3000 M_\odot$. This essentially decreases the lower mass limit that we can probe by half, compared to the isolated simulations. The upper bound on the number of star particles limits the stellar masses of the objects to $\lesssim 10^9 M_\odot$, similar to the stellar mass of the most massive streams such as Sagittarius. This is to ensure that we limit to objects that are minor mergers.

Requiring a large maximum separation between star particles serves to eliminate dwarf galaxies that are still bound at present day. Bound satellite galaxies are compact in both position and velocity, hence, their maximum pairwise distances are small. The distance separation threshold of 120 kpc is comparable to the size of the main galaxy's halo, which is much larger than characteristic sizes of bound dwarf galaxies. This criterion selects streams that have wrapped at least partway around the host galaxy by the present day, and is especially robust at eliminating dwarf galaxies that are slightly tidally deformed but without prominent tails. If we were to relax the distance separation threshold to 90 kpc, we would have let in 9 more objects across all simulations. We have checked by eye that most of these additional objects are bound dwarf galaxies that are slightly tidally deformed. Only one object looks like a coherent stream with a relatively small orbital diameter of ~ 100 kpc, hence, evading our original separation threshold. 1–2 additional streams with small orbital radii that we might obtain from lowering the distance separation threshold would not change any conclusions in this manuscript.

The purpose of the local velocity dispersion cut is to eliminate candidate streams that are phase-mixed, which happens at a later stage as a stream is fully absorbed by the main galaxy. Specifically, we refer to structures for which the *coarse-grained* phase-space distribution defined by the star particles associated with a particular stream is phase-mixed (i.e. approaches the overall phase-space distribution in the host halo). This criterion is thus resolution-dependent, as we discuss later in this section. Most known streams in the Milky Way and other galaxies are not phase-mixed, since phase-mixed objects are unlikely to be discovered and labeled outside of the local volume where 6 dimensional data (positions and velocities) are available. A few exceptions discovered in the Solar neighborhood using such data, such as the Helmi stream (Helmi et al. 1999) and the Gaia Enceladus stream (Helmi et al. 2018; Belokurov et al. 2018), are thought to be phase-mixed.

In our simulations, phase-mixed objects can be effectively ruled out by considering the median of the local velocity dispersion for all star particles along each candidate stream. The local velocity dispersion is computed using nearest neighbors in *phase space* around each star particle, not neighbors in real space where star particles in different orbital phases might be selected if the stream has multiple wraps. When the nearest neighbor star particles in phase space are no longer neighbors in orbital phase, the velocity dispersion should increase sharply and we can consider the stream phase-mixed for our purposes. The phase space distance between the ith particle and the selected particle is defined to be the Mahlanobis distance

$$d_i = \sqrt{\frac{(\vec{x}_i - \vec{x}_0)^2}{\sigma_x^2} + \frac{(\vec{v}_i - \vec{v}_0)^2}{\sigma_v^2}}, \qquad (1)$$

where $\vec{x}_i$ is the position of the ith particle, $\vec{v}_i$ is the velocity of the ith particle, and σ_x and σ_v are the standard deviations of positions and velocities of all star particles in the stream. d_i is thus a unitless quantity. The nearest neighbors to a given star, used to calculate its local velocity dispersion, are the stars with the 20 smallest d_i for stream candidates with more than 300 star particles, and the 7 smallest d_i otherwise. The local velocity dispersion is defined to be the velocity dispersion among these nearest neighbors.

Among the candidates that pass the number of star and the pairwise distance criterion, we identify phase-mixed candidates by eye in 4 of our 7 isolated simulations: `m12i`, `m12f`, `m12m` and `m12c` to calibrate this criterion. We compute the local velocity dispersion for all star particles in each candidate stream, estimating the values at 16th, 50th and 83rd percentile. Figure 1 shows these values for each candidate in all 4 simulations with respect to their total stellar mass. The marker represents the median value. The error bars represent the 16th and 83rd percentile. Phase-mixed candidates are shown in orange, while non-phase-mixed candidates are shown in blue. The phase-mixed candidates have systematically higher median local velocity dispersions compared to the non-phase-mixed ones. We use a linear kernel Support Vector Machine (SVM), implemented in the `scikit-learn` package, to determine the best line that separates the two groups (Cortes & Vapnik 1995; Pedregosa et al. 2011). The algorithm finds that the hyperplane that maximizes the width of the gap between the two groups is

$$\langle \sigma \rangle = -5.28 \log \left(\frac{M_\star}{M_\odot} \right) + 53.55, \qquad (2)$$

where $\langle \sigma \rangle$ is the median of the local velocity dispersion and $M_\star$ is the total stellar mass. Candidates that lie above this line are considered phase-mixed. The local velocity dispersion cut is stellar mass dependent since lower mass candidates have fewer star particles which result in higher estimated local velocity dispersion. This is shown by the negative slope of the cut, and is a reflection of the resolution dependence of this criterion. However, the resolution dependence is quite weak, as shown by the very shallow slope of the cutoff with stellar mass. In particular, it can still be applied to the

paired simulations we analyze, which have mass resolution about twice that of the isolated simulations (3500 versus 7100 $M_\odot$). We confirm by eye that the stream candidates in the paired simulations that lie above the threshold are phase-mixed or borderline phase-mixing.

Table 1 shows the number of streams that pass all three criteria for phase-coherent tidal streams in each simulation, along with the main halo's total mass, stellar mass, and virial radius. There are more streams in the paired simulations because of their better particle resolution. The number of streams in both sets of simulations are comparable if we only consider streams in the paired simulations that are more massive than the lower mass limit of the unpaired simulations at $\sim 10^6 M_\odot$. Example of objects that are classified by our criteria as satellite, stream, and phase-mixed are shown in Figure 2. To summarize our classification process: all objects that we consider pass criterion (i) (on the number of star particles). Satellites fail criterion (ii) (on the distance between star particles). Phase-mixed objects pass criterion (ii), but fail criterion (iii) (on the local velocity dispersion). Streams pass all the criteria. These important definitions are also summarized in Table 2.

3.3. *Selecting Recent/First Infall Satellites*

Since we only search for self-bound objects within the host halo's virial radius between 2.7–6.5 Gyr ago when we look for present-day coherent streams, our sample of satellite dwarf galaxies are incomplete, potentially missing some satellites that fell in more recently. To include these recent infall satellites, we first identify self-bound objects with star particles within the each host's virial radius at z=0. These luminous bound objects are then compared to our set of previously identified satellites and streams (with still visible bound part at $z = 0$). All phase-mixed objects no longer have identifiable bound parts. We select out extra, non-overlapping objects as recent infall satellites. We follow steps in §3.1 to collect most star particles that once belong to these additional satellites. 34 additional recent infall satellites are identified across all of the simulations.

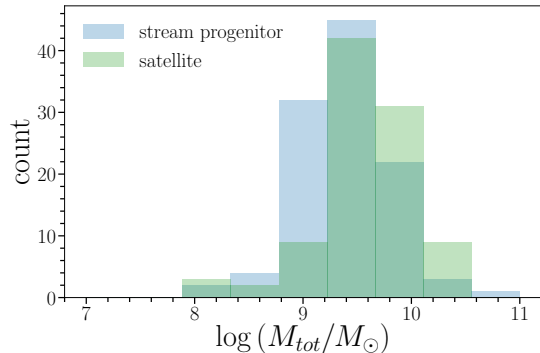## 4. PROPERTIES OF STELLAR STREAM PROGENITORS

Here we summarize the properties and statistics of the progenitors of all the streams that were selected as described in Section 3. These dwarf-galaxy progenitors of present-day ($z = 0$) tidal streams are characterized at the time when they have maximum stellar mass, unless otherwise stated.

4.1. *Mass Function*

The total and stellar mass functions of stream progenitors and satellites are alike, which show that there

sim name	$m_{200\mathrm{m}}[\mathrm{M}_\odot]$	$r_{200\mathrm{m}}[\mathrm{kpc}]$	N
m12i	1.18×10^{12}	336	9
m12f	1.71×10^{12}	380	8
m12m	1.58×10^{12}	371	8
m12c	1.35×10^{12}	351	7
m12b	1.43×10^{12}	358	8
m12r	1.10×10^{12}	321	3
m12w	1.08×10^{12}	319	3
Romeo	1.32×10^{12}	341	13 [10]
Juliet	1.10×10^{12}	321	12 [6]
Romulus	2.08×10^{12}	406	9 [6]
Remus	1.22×10^{12}	339	8 [5]
Thelma	1.43×10^{12}	358	10 [9]
Louise	1.15×10^{12}	333	8 [8]

Table 1. Total mass ($m_{200\mathrm{m}}$, DM+star+gas), virial radius ($r_{200\mathrm{m}}$) and number of stellar streams that pass all the criteria for phase-coherent tidal stream for each simulated Milky Way-like galaxies. The numbers in the square brackets are the effective numbers of stellar streams in the paired galaxy simulations that have mass greater than the lower bound mass we set for the unpaired simulations. Here, the total mass refers to the total mass enclosed within $r_{200\mathrm{m}}$, a radius containing 200 times the mean background matter density.
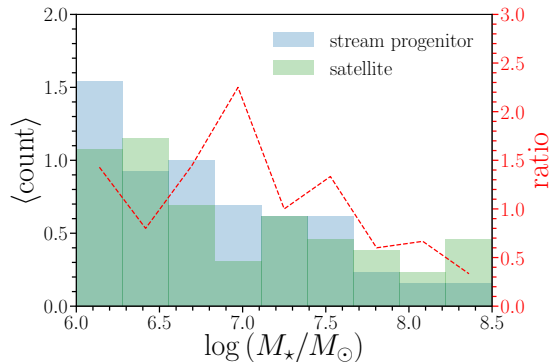

Figure 3. Total mass functions for satellite dwarf galaxies (green) and for stellar stream progenitors (blue), cumulative across all simulations. The total mass is evaluated at the peak stellar mass timescale of each satellite and progenitor.

is essentially no preferred range in total mass or stellar mass for which satellite galaxies are more likely to turn into stellar streams. Figure 3 compares the total mass (stars, gas, and DM) distribution of the stream progen-

Table 2. Important Definitions

Stream Candidate Identification (refer to §3.1)	
Stream candidate	An object that is bound between 2.7–6.5 Gyr ago and is within present day's virial radius of the host.
Candidate Classification (refer to §3.2)	
Satellite	A candidate that has between 120–10^5 star particles at present day, but has maximum pairwise distances between star particles less than 120 kpc.
Phase-mixed object	A candidate that has between 120–10^5 star particles at present day, has maximum pairwise distances between star particles greater than 120 kpc, and has median local velocity dispersion greater than Equation 2.
Stream	A candidate that has between 120–10^5 star particles at present day, has maximum pairwise distances between star particles greater than 120 kpc, and has median local velocity dispersion less than Equation 2.
Timescales	
τ_{infall}^i	First infall look-back time, measured at the *first* virial radius crossing.
τ_{infall}^f	Last infall look-back time, measured at the *last* virial radius crossing.
τ_{stream}	Stream-formation look-back time, measured when the stellar part of the progenitor starts to become unbound (refer to §4.2). This is also the age of each stream/phase-mixed object.
τ_{peak}	Peak stellar mass look-back time, measured when the bound part of the progenitor contains peak stellar mass throughout its evolution.

NOTE—For objects that only cross the virial radius once, $\tau_{\text{infall}}^i = \tau_{\text{infall}}^f$.

Figure 4. Stellar mass functions for satellite dwarf galaxies (green) and for stellar stream progenitors (blue). These are cumulative across all simulations, normalized by the total number of hosts; hence, ⟨count⟩ represents the average number of objects per host in any given stellar mass bin. The stellar mass is evaluated at the peak stellar mass timescale of each satellite and progenitor. The ratio of the progenitor to satellite galaxy is shown by the dashed line (red).

itors (blue) with the total mass distribution for satellite galaxies (green), both of which are evaluated at the peak stellar mass timescale of each object. These distributions are cumulative across all simulations. Both distributions show similar features. The majority of the stream progenitors and satellites have total mass of $\sim 10^9 - 10^{10} M_\odot$. However, note that these total mass distributions are localized and do not span the entire mass spectrum. This is primarily due to the lower bound and upper bound we set on the number of star particles in the object that we track.

Figure 4 is similar to Figure 3, but shows the distribution of stellar mass, evaluated at the peak stellar mass timescale of each object, instead of total mass and is only plotted over the span of stellar mass range that we have completed sample which corresponds to our criterion (ii) for the number of star particles. The stellar mass function of the stream progenitors is shown in blue, while the stellar mass function of satellites is shown in green. Both distributions have the most objects at low stellar mass, and the number decreases at higher stellar masses. The ratio of the number of stream progenitor to satellite at each stellar mass bin is shown by the dashed line (red). At stellar mass $\sim 10^{7.5} M_\odot$, there are comparable number of stream progenitors and satellites. There are more satellite galaxies at higher stellar mass (ratio < 1), while there are more stream progenitors at lower stellar mass (ratio > 1). The decrease in the relative numbers of stream progenitors to satellites at high stellar mass can be attributed to the shorter dynamical lifetimes for

massive streams due to experiencing higher dynamical friction.

4.2. *Infall and Stream-formation Timescales*

The infall and formation times of stellar streams are crucial in modeling and understanding the origin of real observed streams. In this work, we define the *infall time*, $\tau_{\rm infall}$, as the time before present day that each progenitor crosses the virial radius of the main galaxy. The virial radius of the main galaxy is determined independently for each snapshot in each simulation, and increases monotonically with time. We find that about half of stream progenitors cross the virial radius multiple times before complete disruption. In these cases, we distinguish between $\tau_{\rm infall}^i$, first infall, and $\tau_{\rm infall}^f$, final infall. For systems with a single virial radius crossing, $\tau_{\rm infall} = \tau_{\rm infall}^i = \tau_{\rm infall}^f$.

The *stream-formation time*, $\tau_{\rm stream}$, refers to the time when the stellar part of the dwarf galaxy progenitor starts to become unbound, and star particles that once belonged to the progenitor start to stretch out along the orbit. In this paper, we use the terms "stream-formation time" and "formation time" interchangeably. We determine $\tau_{\rm stream}$ by computing the moment-of-inertia tensor for the progenitor, I_{ij}:

$$I_{ij} = -\sum_{n=1}^{N} m_n x_{i,n} x_{j,n}, \qquad (3)$$

for $i \neq j$ and,

$$I_{11} = \sum_{n=1}^{N} m_n \left(x_{2,n}^2 + x_{3,n}^2 \right), \qquad (4)$$

$$I_{22} = \sum_{n=1}^{N} m_n \left(x_{1,n}^2 + x_{3,n}^2 \right), \qquad (5)$$

$$I_{33} = \sum_{n=1}^{N} m_n \left(x_{1,n}^2 + x_{2,n}^2 \right), \qquad (6)$$

for $i = j$, where $i, j = 1, 2, 3$ and N is the total number of star particles in each progenitor. $x_{i,n}$ is the position of the nth star particle along the ith direction, and m_n is the mass of the nth star particle.

We compute the three eigenvalues λ of I_{ij} and determine the maximum (λ_{max}) and minimum (λ_{min}) values. The ratio $\lambda_{max}/\lambda_{min}$ generally starts near 1 (while the object is self-bound), dips slightly during tidal compression, increases sharply as tidal disruption begins and eventually approaches 1 again as the stream wraps around the galaxy. A typical evolution of the eigenvalue ratio for an example stream progenitor is shown in Figure 5. The local minimum just before the maximum of $\lambda_{max}/\lambda_{min}$ is defined as $\tau_{\rm stream}$. We limit the search for $\tau_{\rm stream}$ to times after the first crossing of the virial

radius of the main galaxy, in order to ensure that the elongation is due to interaction with the main galaxy and not early assembly of the progenitor.

The top panel of Figure 6 shows histograms of the first infall look-back time $\tau_{\rm infall}^i$ (solid red), stream-formation look-back time $\tau_{\rm stream}$ (dashed black) and peak stellar mass look-back time (dotted blue) of stellar stream progenitors, while the bottom panel shows the virial radius of each main galaxy over the same timescale. Many stream progenitors fall into the main galaxy relatively early on: over 6 Gyr ago, when the virial radii of the main galaxies are approximately half of their present day values. However, most streams do not form as soon as their progenitors fall into the main galaxy, as shown by the significant shift between the two distributions. In some cases, it can take several Gyr before the progenitor is tidally disrupted, depending on its orbit. The stellar streams that have formed most recently are all either in very radial orbits or still have visible self-bound parts. Lastly, the distribution for the peak stellar mass timescale is also slightly shifted from both the infall and stream-formation timescales.

Our method of estimating the stream-formation time for each progenitor is purely geometrical; it does not use any dynamical information for each star particle. This strategy is intended to mimic observational identification of streams in large surveys, which generally uses only positional information. However, one drawback of our method is that it cannot reliably determine $\tau_{\rm stream}$ for members of an infalling group whose constituent objects are not fully merged after the first infall. Stream-formation times for such objects, which can tidally disrupt each other prior to infall into the host galaxy, are somewhat ambiguous. Out of the total of 106 streams, there are 13 objects that we mask as group infalls (roughly one per host). Table 2 contains definitions of these important timescales.

4.3. *Ordering of infall, quenching, and stream formation timescales*

Figure 7 shows the assembly history of present day dwarf galaxies, stellar streams and phase-mixed objects, identified by the selection criteria in §3.2, into the stellar halo of the isolated simulation m12i. This history is typical of our simulated systems without a relatively recent major merger. Phase-mixed components (orange lines) exclusively fall in very early and are all completely disrupted by 5 Gyr ago, while present-day dwarf satellite galaxies all have their last crossing of the virial radius after this time. The streams, meanwhile, probe basically the entire lifetime of the galaxy. Many of the more recently disrupted streams, as well as several of the satellite galaxies, have experienced multiple virial radius crossings, suggesting that their orbits could potentially be influenced by the local environment *beyond* the virial radius, especially in the paired environment. The transfer of an object from one host to the other does
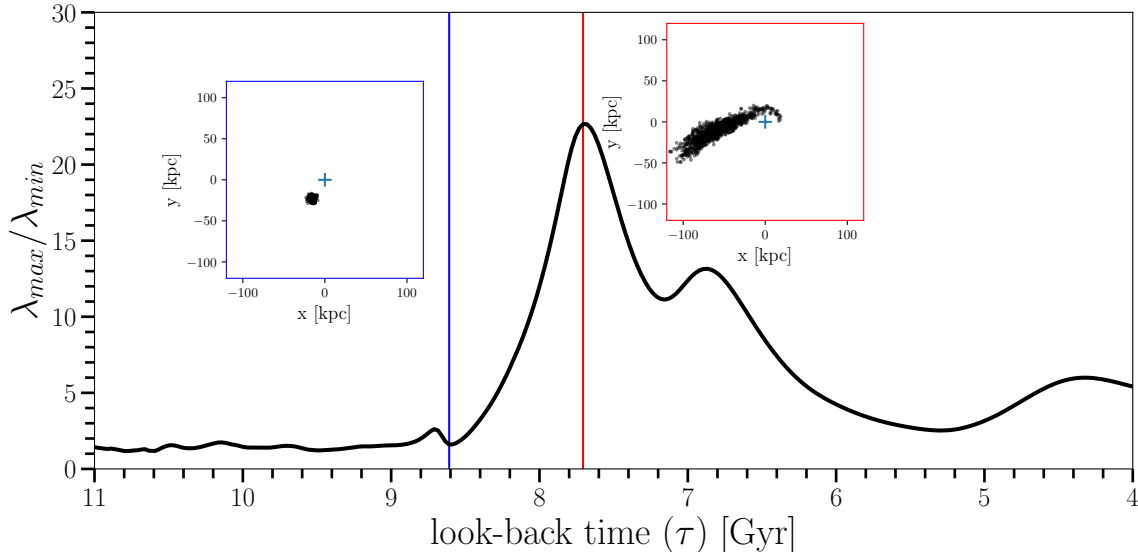
Figure 5. A typical evolution of the ratio of the maximum to minimum eigenvalue ($\lambda_{max}/\lambda_{min}$) of a stream progenitor's moment-of-inertia tensor as a function of look-back time. The graph shown is for an example stream progenitor from m12i. The red vertical line marks the global peak of the entire evolution. The blue vertical line marks the first local minimum just before the global peak and is the timescale that characterizes the stream-formation look-back time (τ_{stream}) for this example stream progenitor. The two panels show 2-D projections of the progenitor at these two timescales.

happen in the paired simulations, although they are not common ($\sim$1–2 objects per simulation). The dynamical properties of stream progenitors in the paired simulations are largely only confined within each host, similar to the isolated simulations.

The standard picture of hierarchical assembly in galactic stellar halos includes an implicit ordering of several important timescales in the life of each progenitor of a stellar stream. In this picture, each satellite galaxy crosses the virial radius of the host at τ_{infall}, has its star formation quenched (if it was not already) by the environment in the halo at τ_{peak}, and then is tidally disrupted to form a stream at τ_{stream}. From this picture, we would generally expect $\tau_{\text{infall}} > \tau_{\text{peak}} > \tau_{\text{stream}}$ for more massive satellites with sustained star formation, and $\tau_{\text{peak}} \geq \tau_{\text{infall}} > \tau_{\text{stream}}$ for less massive satellites quenched prior to infall by reionization or stellar feedback, with all ages measured back in time from the present day. Following this reasoning, the measured ages of stars in streams (which are $\geq \tau_{\text{peak}}$) are sometimes used as an upper bound on the stream age τ_{stream}. Figure 6 shows that statistically speaking, $\tau_{\text{peak}} \sim \tau_{\text{infall}}^i \geq \tau_{\text{infall}}^f \gtrsim \tau_{\text{stream}}$ when the entire sample is considered. However, the spread of all three timescales is extremely broad, and the size of the shift between τ_{peak} and τ_{infall} nearly vanishes if we consider τ_{infall}^i rather than τ_{infall}^f.

On closer examination, we find that this picture is not universally applicable to the progenitors of streams in our simulations. Figure 8 shows the relative times of these different events for all stream progenitors in our

sample, arranged by stellar mass. We bin the sample in increasing stellar mass bins such that each bin contains 7 progenitors. The median value in each bin are represented by lines and the shaded regions encapsulate minimum and maximum spreads. The infall time, τ_{infall}, is shown in red. The stream-formation time τ_{stream} and peak stellar mass time τ_{peak} are shown as black and blue, respectively. The median stellar mass of the entire sample $\sim 2.25 \times 10^6 M_\odot$ is shown by the dashed vertical line. The median stellar mass is used to divide the sample into two subgroups: the high-mass and low-mass progenitors. This stellar mass cut is supported by the star formation histories of the Local Group dwarf galaxies: the dwarf satellite Sculptor is the most massive satellite known for which models suggest early quenching of star formation (Weisz et al. 2014). Sculptor is estimated to have $M_\star \sim 2.3 \times 10^6 M_\odot$ (McConnachie 2012), consistent with our separation between high- and low-mass progenitors. Additionally, Figure 9 shows relative values of all the pairs between the three timescales. The three distributions are plotted: $\tau_{\text{infall}}^i - \tau_{\text{stream}}$, $\tau_{\text{infall}}^i - \tau_{\text{peak}}$ and $\tau_{\text{peak}} - \tau_{\text{stream}}$. The top row is the low-mass group and the bottom row is the high-mass group. Note that the bi-modalities in the distributions of $\tau_{\text{infall}}^i - \tau_{\text{stream}}$ is real and can be attributed to the fact that if the progenitor does not form a stream during the first pericenter, it is unlikely to form a stream until subsequent pericenters.

For these two groups in mass, we examine three different orderings of significant events:

1. $\tau_{\text{infall}}^i > \tau_{\text{peak}} > \tau_{\text{stream}}$. The progenitor first falls into the main halo, then has its star formation
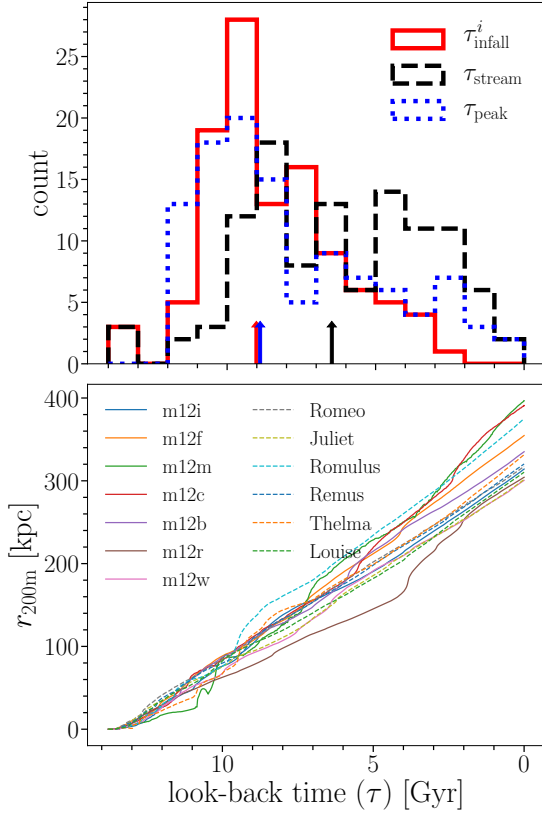
Figure 6. Top: first infall look-back time τ^i_{infall} (solid red), stream-formation look-back time τ_{stream} (dashed black) and peak stellar mass look-back time τ_{peak} (dotted blue) of stellar stream progenitors across all 13 simulations. These are shown in look-back time τ. The three distributions are not entirely identical, as these timescales do not happen at fixed ordering as discussed in section §4.3. Bottom: time evolution of the virial radii $r_{200\mathrm{m}}$ of the host galaxies in isolated simulations (solid lines) and paired simulations (dashed lines).

quenched, then forms a stream. This is the standard picture for high-mass progenitors.

2. $\tau_{\mathrm{peak}} > \tau^i_{\mathrm{infall}} > \tau_{\mathrm{stream}}$. The progenitor first stops forming stars, then falls into the main halo, then forms a stream. This is the standard picture for low-mass progenitors.

3. $\tau^i_{\mathrm{infall}} > \tau_{\mathrm{stream}} > \tau_{\mathrm{peak}}$. The progenitor first falls into the main halo, then begins to form a stream while still forming stars *at a higher rate than they are stripped by tides*, then reaches maximum stellar mass before being completely tidally disrupted.

In this subsection, we use τ^i_{infall} as the infall time, since this marks the time when the progenitor first enters the main halo environment, which is presumed to quench its star formation.

For the high-mass group, out of 47 objects, 26 are in the first scenario, which is expected for this group, ($\sim 55\%$), 9 are in the second scenario ($\sim 19\%$) and 10 are in the third scenario ($\sim 21\%$). Two objects have equal τ_{stream} and τ_{peak}, hence, are excluded. For the low-mass group, out of 46 objects, 20 are in the first scenario ($\sim 43\%$), 24 are in the second scenario, which is expected for this group, ($\sim 52\%$) and 2 are in the third scenario ($\sim 4\%$).

Our intuition for the ordering of timescales for high-mass and low-mass progenitors is thus approximately correct: around half of the members of each group follow our expected picture. However, there are large discrepancies within each group. A few progenitors in the high-mass group have their star formation quenched before their first infall. In the low-mass group, almost half of the progenitors still form star particles after their first infall. A significant fraction of the high-mass progenitors still have a high star formation rate even after undergoing tidal stripping, while this situation very rarely occurs for low-mass progenitors. In future work, Samuel et al. in prep. will examine the timescales and physical processes involved in satellite quenching in detail.

In Figure 10, we study the correlation between the pericenter distance of each progenitor and the timescale ordering that it experiences. Dividing the low-mass (light blue) and high-mass (dark blue) groups, the left plot shows the distribution of the pericenter distances. To remain coherent at present day, progenitors in the low-mass group overall have smaller pericenter distance with the peak at $\sim 15-20$ kpc, while the high-mass group peaks at $\sim 30-35$ kpc. This suggests that high-mass progenitors become phase-mixed with the host galaxy faster. Thus, high-mass progenitors with low pericenter distance are less likely to remain coherent at $z = 0$. The middle and right panels show the distribution of the pericenter distance and infall time of progenitors in three scenarios of the timescales ordering, combining both low-mass and high-mass groups. The first ordering scenario ($\tau^i_{\mathrm{infall}} > \tau_{\mathrm{peak}} > \tau_{\mathrm{stream}}$) is shown in red. The second ordering scenario ($\tau_{\mathrm{peak}} > \tau^i_{\mathrm{infall}} > \tau_{\mathrm{stream}}$) is shown in blue. The third ordering scenario ($\tau^i_{\mathrm{infall}} > \tau_{\mathrm{stream}} > \tau_{\mathrm{peak}}$) is shown in black. The median value of each distribution is illustrated by a dashed vertical line. The pericenter distributions are not clearly separable, with all three having similar median pericenter distance. However, they appear more distinct in the τ^i_{infall} projection. Progenitors in the second ordering fall in the earliest, followed by progenitors in the first ordering and the third ordering, respectively. This trend can be explained by the connection between τ^i_{inafll} and stellar mass of the progenitor. A small negative trend in Figure 8 shows a weak correlation between all the timescales and the stellar mass of the progenitor. This correlation can be explained by the hierarchical growth of structures: statistically, massive objects collapse at later times than lower mass objects. More mas-
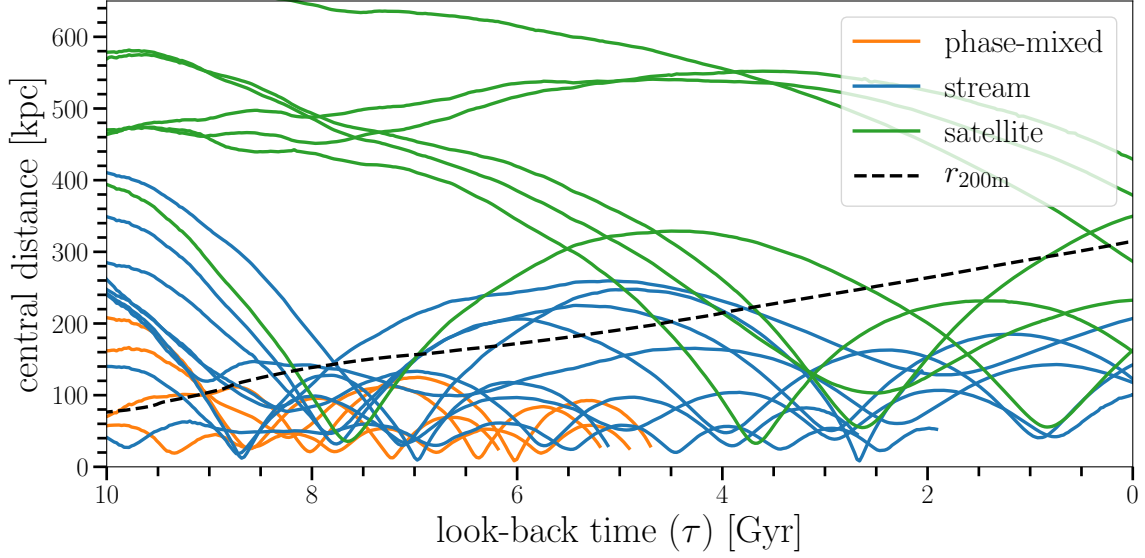
Figure 7. Assembly history of the stellar halo of `m12i` plotting object's central distance as a function of a look back time τ. Line colors represent object classifications at $z = 0$ based on criteria in §3.2. Each line terminates when the object no longer has an identifiable self-bound part, or is drawn up to $z = 0$ if the object is still partially self-bound at present day. The virial radius ($r_{200\mathrm{m}}$) of the main halo is shown as a black dashed line.
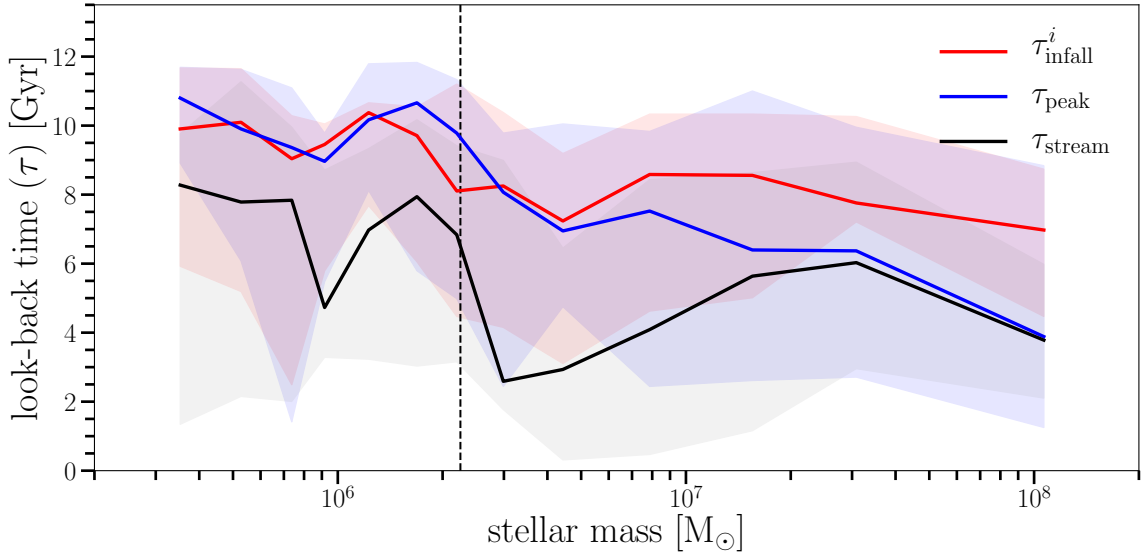
Figure 8. Ordering of different timescales in the life of stream progenitors. Each timescale is binned in increasing stellar mass bins such that there are 7 progenitors in each bin. The lines represent the median value and shaded region encapsulate minimum and maximum values witin a given bin. The red represents the first infall look-back time $\tau_{\mathrm{infall}}^{i}$. The blue represents the peak stellar mass time, τ_{peak}. The black represents the stream-formation look-back time, τ_{stream}, described in §4.2. The black dashed vertical line is the median stellar mass of all progenitors ($\sim 2.25 \times 10^{6} M_{\odot}$).
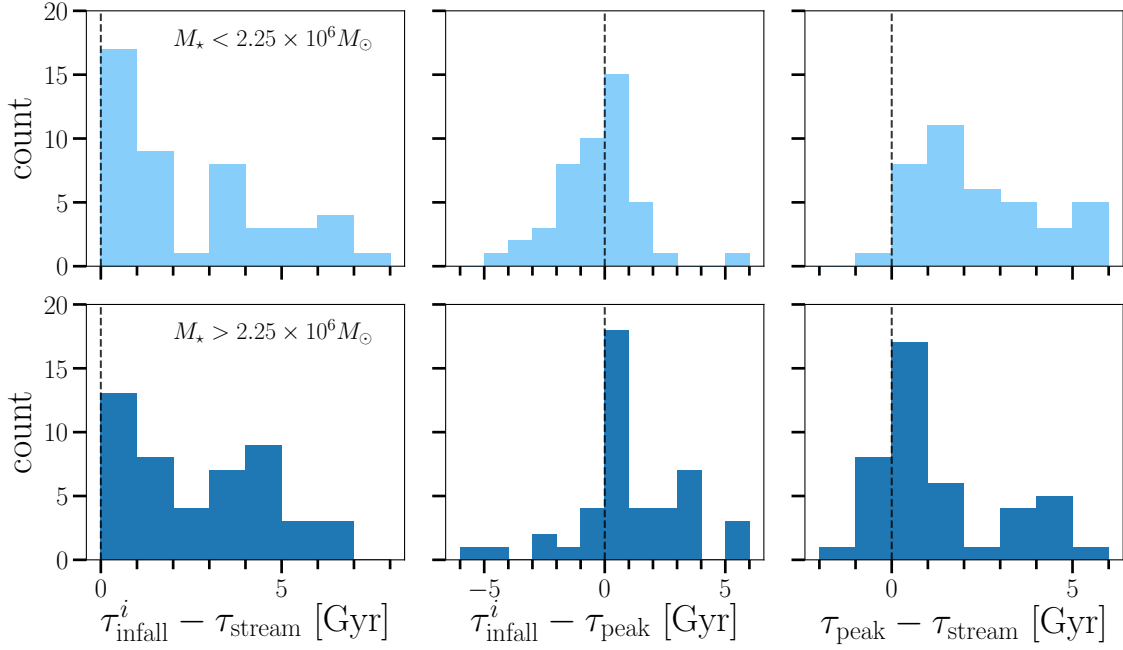
Figure 9. Relative values of different look-back timescales in the life of low-mass (top row; light blue) and high-mass (bottom row; dark blue) stream progenitors. The columns represent different pairs of timescales. Left is the infall look-back time relative to the stream-formation look-back time ($\tau_{\mathrm{infall}}^i - \tau_{\mathrm{stream}}$). Middle is the infall look-back time relative to the peak stellar mass look-back time ($\tau_{\mathrm{infall}}^i - \tau_{\mathrm{peak}}$). Right is the peak stellar mass look-back time relative to the stream-formation look-back time ($\tau_{\mathrm{peak}} - \tau_{\mathrm{stream}}$). The positive values translate to the first timescale occurring before the second timescale. The vertical dashed lines divide the positive and negative relative values.
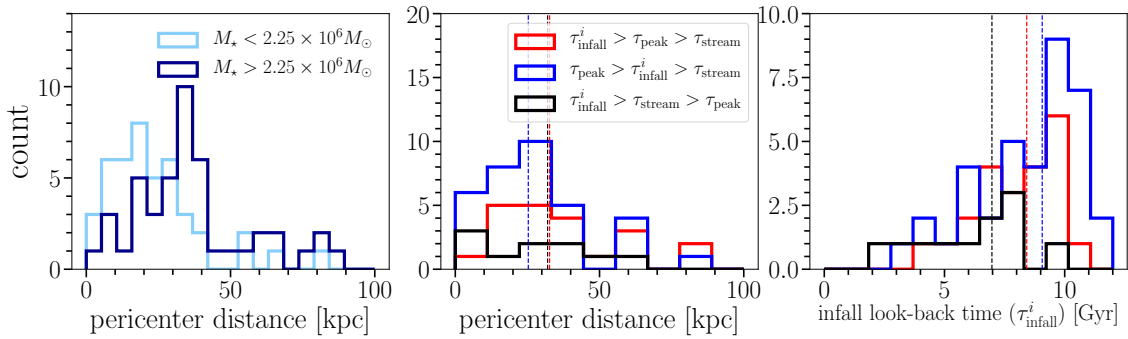
Figure 10. Left: Distribution of the pericenter distances for the low-mass progenitors (light blue) and high-mass progenitors (dark blue). Present day high-mass coherent streams are more likely to have larger pericentric distances compared to low-mass coherent streams. Middle: Distribution of the pericenter distances for progenitors (low-mass and high-mass combined) with three different ordering of timescales that they experience in their lifetime (see §4.3). The dashed vertical lines are the median of the distributions. Right: Distribution of the infall time (τ_{infall}^i) for progenitors with three different ordering of timescales. Note that there are 5 progenitors with pericenter distance > 100 kpc, which are not shown in the left and middle figures.

sive galaxies are thus expected to accrete onto the main halo slightly later than less massive ones (e.g. Wetzel et al. 2015). Additionally, note that there are 5 progenitors with pericenter distance > 100 kpc, which are not shown in the left and middle figures of Figure 10. The origin of these streams with large pericenter distances is expected to be due to interactions with other substructures within the hosts and will be studied in future work.

4.4. *Validating the Local Velocity Dispersion Criterion*

The stream formation time for each object can also be reinterpreted as the stream's age: specifically, how long ago each object began being tidally disrupted. This allows us to validate our classification of phase-mixed objects using their local velocity dispersions: since the phase-mixed objects were once stellar streams, we expect these objects to be older, as a population, than phase-coherent stellar streams.

Figure 11 plots each object's present-day local velocity dispersion, with uncertainties corresponding to the 16 and 83 percentile for all particles in the object, as a function of its stream-formation look-back time τ_{stream}. Each simulation is represented by a unique marker shape; the color of the marker signifies the stellar mass of the object. Objects that are considered phase-mixed by the local velocity dispersion criterion are shown with orange error bars, while streams that pass the local velocity dispersion criterion are shown with blue error bars. The shaded region corresponds to the span in the end of the star formation bursty phase/onset of the steady phase in all the hosts, determined in Yu et al. (2021), with the median value represented by the black vertical line. In the FIRE simulations, MW-mass galaxies generically have highly time-variable star formation histories at early times but eventually transition to more steady star formation rates after the disk settles (e.g. Muratov et al. 2015; Sparre et al. 2017; Faucher-Giguère 2018). This transition may be due to the virialization of the inner circumgalactic medium, which may stabilize the disk against disruption by stellar feedback (Stern et al. 2020, Gurvich et al. in prep.).

Phase-mixed objects that are ruled out by the local velocity dispersion criterion are indeed generally older than those that pass the criterion and are classified as streams. The former has a median age of ~ 9.98 Gyr, while the latter has a median age of ~ 6.45 Gyr. This validates our use of the local velocity dispersion criterion to separate phase-mixed objects from streams. An additional interesting feature of Figure 11 is the abrupt transition in the number of phase-mixed objects: nearly all began disruption more than 5 Gyr ago. This is partially a reflection of the time required for an object to phase-mix (further discussion in §5) but also reflects the transition between the early epoch of galaxy assembly, which is most chaotic, and the later establishment of a large, stable disk. In this suite of simulations this tran-

sition usually occurs at $z \sim 1$ (Garrison-Kimmel et al. 2018b), or about 6.5 Gyr ago. Another contributing factor is the end of the star formation bursty phase, which on average occurs after the disk orientation settlement, at ~ 5 Gyr ago (see Yu et al. 2021). Progenitors that fall in and disrupt before this time evolve in a potential with far less symmetry, changing on non-adiabatic timescales, which accelerates the process of phase-mixing substantially relative to the environment after the disk is established.

4.5. *Progenitors*

Here we compare properties of stellar stream progenitors with present-day simulated dwarf satellites, and with observational data for satellite galaxies and streams around the MW and M31. We consider three standard relationships: the stellar mass–velocity dispersion relation, the stellar mass–metallicity relation, and tracks in $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe/H}]$ chemical abundance space. In all cases we find that the present-day dwarf satellites in our simulations mostly resemble the progenitors of the simulated streams, suggesting that present-day dwarf satellites are useful proxies for stream searches in any of these spaces. We also find that the stellar mass–velocity dispersion relation of our simulated satellites and stream progenitors—and therefore the initial phase-space volume of each simulated stream—is consistent with observations of real dwarf satellites in the MW and M31. This agreement in velocity dispersion relation has been shown for a smaller sample (only satellites in `m12i`) in Wetzel et al. (2016b). Moreover, Garrison-Kimmel et al. (2019) shows that there is also an agreement between the circular velocity curves of surviving satellites.

4.5.1. *Stellar mass–velocity dispersion relation*

The 3-dimensional total velocity dispersions of stellar stream progenitors, simulated satellites and observed satellites are shown in Figure 12. For both stream progenitors and simulated satellites, their total velocity dispersion, σ_{tot}, agree well with the mass–velocity dispersion relation of observed satellite galaxies from McConnachie (2012, Table 4; black stars), with values between $10 - 30$ km/s for lower mass progenitors $(M_\star < 10^7 M_\odot)$, and as high as 50 km/s for higher mass progenitors. Wetzel et al. (2016b) and Garrison-Kimmel et al. (2019) present this comparison for surviving satellites and conclude that our simulations do not suffer from too-big-to-fail problem. Our analysis additionally extends this mass–velocity dispersion relation agreement to stream progenitors as well (disrupted satellites).

Compared to stellar stream progenitors, the observed satellite galaxies have slightly lower velocity dispersion for lower mass galaxies $(< 10^7 M_\odot)$, with the values between $5 - 25$ km/s. This might be due to a selection effect in our sample, as the velocity dispersion is overestimated for objects with fewer number of star particles. However, all observed satellites with velocity dispersion
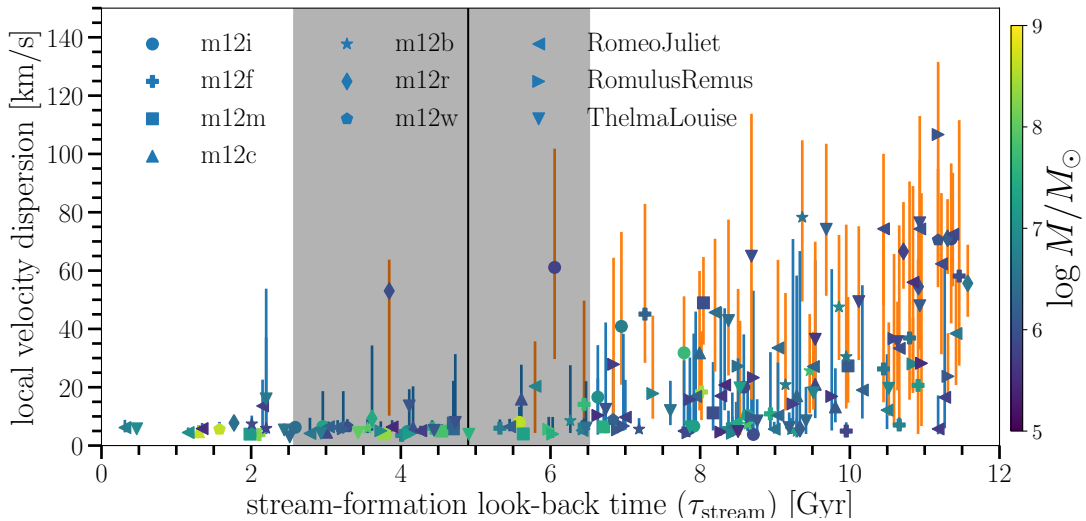
Figure 11. Validation of the phase-mixing criterion. Each object's local velocity dispersion σ_{local} is plotted against its age, which is defined to be its stream-formation time τ_{stream} (i.e. time since start of tidal disruption). The marker corresponds to the median local velocity dispersion of star particles in the stream; vertical error bars span 16th to 83rd percentile. The shapes and colors of the markers represent the simulation where each object is identified (see legend) and stellar mass (see colorbar), respectively. Objects that are considered phase-mixed by the local velocity dispersion criterion are shown with orange error bars, while streams that pass the local velocity dispersion criterion are shown with blue error bars. The shaded black region corresponds to the span in the end of the star formation bursty phase/onset of the steady phase in all the hosts. The solid black vertical line at $\tau = 4.90$ Gyr is the median of the span. These end of star formation bursty phase times are determined in Yu et al. (2021).

< 10 km/s also have stellar mass $< 10^6 M_\odot$. This is very close to the minimum possible stellar mass set by our particle number criterion, which shows our simulations overestimate dynamical masses of systems that have stellar mass $< 10^6 M_\odot$. In contrast, the velocity dispersions of objects that are $> 10^7 M_\odot$ agree well with observations.

4.5.2. *Stellar mass vs metallicity*

Figure 13 compares the mass-metallicity ([Fe/H]) relations for stellar stream progenitors (dots), present day simulated satellite galaxies (gray diamonds), observed satellite galaxies (black stars) and observed streams (red rectangles). Different colored dots represent streams from different simulations. Observational data is from Kirby et al. (2013). For the simulated systems, we evaluate [Fe/H] for stream progenitors at τ_{peak}, and for satellites at the present day. The solar abundance is adopted from Asplund et al. (2009).

For all groups except observed streams, there is a linear relationship between [Fe/H] and $M_\star$, with the more massive objects being more iron-rich. There is no significant difference between stellar stream progenitors from different simulations or between paired and isolated systems. However, there is a significant discrepancy in iron abundance between objects in the simulations and in observations (also noted in Escala et al. (2018)). At the

same stellar mass, observed satellite galaxies have higher metallicity compared to both stellar stream progenitors and present day simulated satellite galaxies. For the present day simulated satellite galaxies, a linear least-square fit yields

$$[\mathrm{Fe/H}]_{\mathrm{sim}} = 0.55 \times \log M_\star - 5.93, \qquad (7)$$

where $M_\star$ is the stellar mass. For stream progenitors, the fit is given by

$$[\mathrm{Fe/H}]_{\mathrm{stream}} = 0.52 \times \log M_\star - 5.76. \qquad (8)$$

The corresponding relationship from Kirby et al. (2013, Equation 4) for observed satellites is

$$[\mathrm{Fe/H}]_{\mathrm{obs}} = 0.3 \times \log M_\star - 7.69, \qquad (9)$$

or roughly half the log-slope of the simulated relation; discrepancies are larger for lower mass objects. There are several contributions to the underproduction of metals in these simulations. About 0.3–0.5 dex of the discrepancy appears to be resolution related, as illustrated by the additional points (blue and red crosses) from higher-resolution simulations of isolated dwarf galaxies (Wheeler et al. 2019). The rest of the discrepancy is probably attributable to the supernova delay-time distribution used in the simulations (Escala et al. (2018), Gandhi et al. in prep.), in which the delay before the
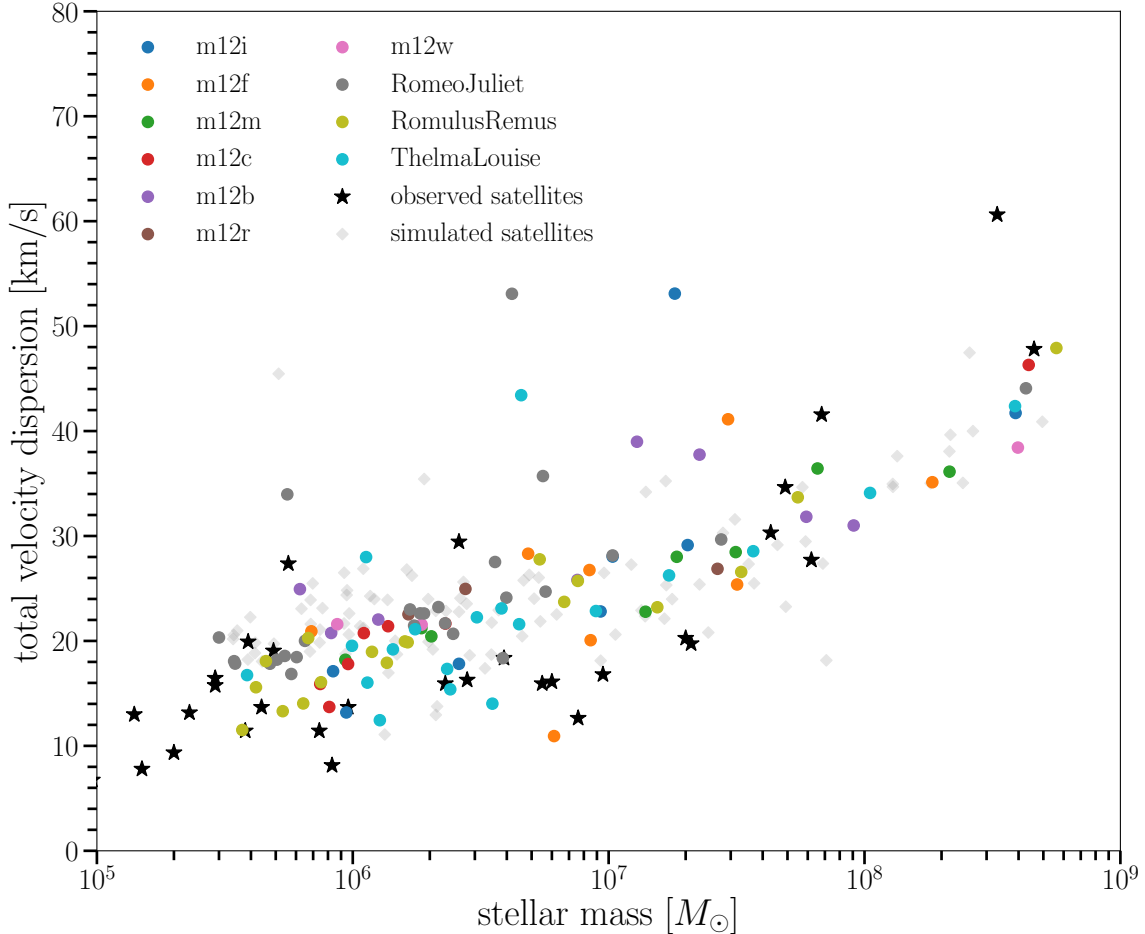
Figure 12. Comparison between velocity dispersion of stellar stream progenitors estimated at the peak stellar mass time (dots) and observed MW and M31 satellite galaxies (black stars) as a function of stellar mass. Data for simulated satellite galaxies are also shown (gray diamond). The colors represent individual simulations. Observational data are from McConnachie (2012).

onset of type Ia supernovae is likely too long and by the metallicity convergence tests in Hopkins et al. (2018). However, these properties are the same across all our simulations, so we can still draw some conclusions by comparing the *relative* metallicities of different groups of simulated objects. The discrepancies between simulations and observations only impact the metallicity normalization, but not the overall shape, spread and intrinsic scatter in the stellar metallicity distribution function (Escala et al. 2018). Relative to the present-day simulated satellites, the stream progenitors have very slightly lower [Fe/H], especially at the upper end of the mass range where the discrepancies are more evident, but are roughly consistent at lower masses. The trend can be explained by the fact that massive satellite galaxies are likely to form star particles long after their infalls (§4.3) and are likely to still be forming star particles at $z = 0$. We see the same trend in Figure 14. The small discrepancies as a function of stellar mass can be quantified by the differences between Equation 7 and 8. This suggests that present-day satellite galaxies are decent proxies to

use to estimate [Fe/H] of progenitors of coherent streams in MW-mass galaxies within ~ 0.10 dex for high-mass progenitors ($M_\star \sim 10^9 M_\odot$) or within ~ 0.02 dex for low-mass progenitors ($M_\star \sim 10^5 M_\odot$).

We also show data from a few MW streams with measured iron abundances. These data are not so well constrained, since for most streams we have only a rough estimate of the mass of their progenitor before tidal disruption. Interestingly, these measurements appear to be all over the place compared to the orderly mass-metallicity relation for satellite galaxies.

4.5.3. *[α/Fe]–[Fe/H] evolutionary tracks*

Chemical abundances of different elements give complementary information on a galaxy's evolution, especially if they are produced by different channels. Fe is mostly produced by Type Ia supernovae, on long timescales compared to α-elements produced by type II (core-collapse) SNe from massive stars. Hence, a galaxy's relative abundance of these two types of elements evolves with time, with [α/Fe] starting above the
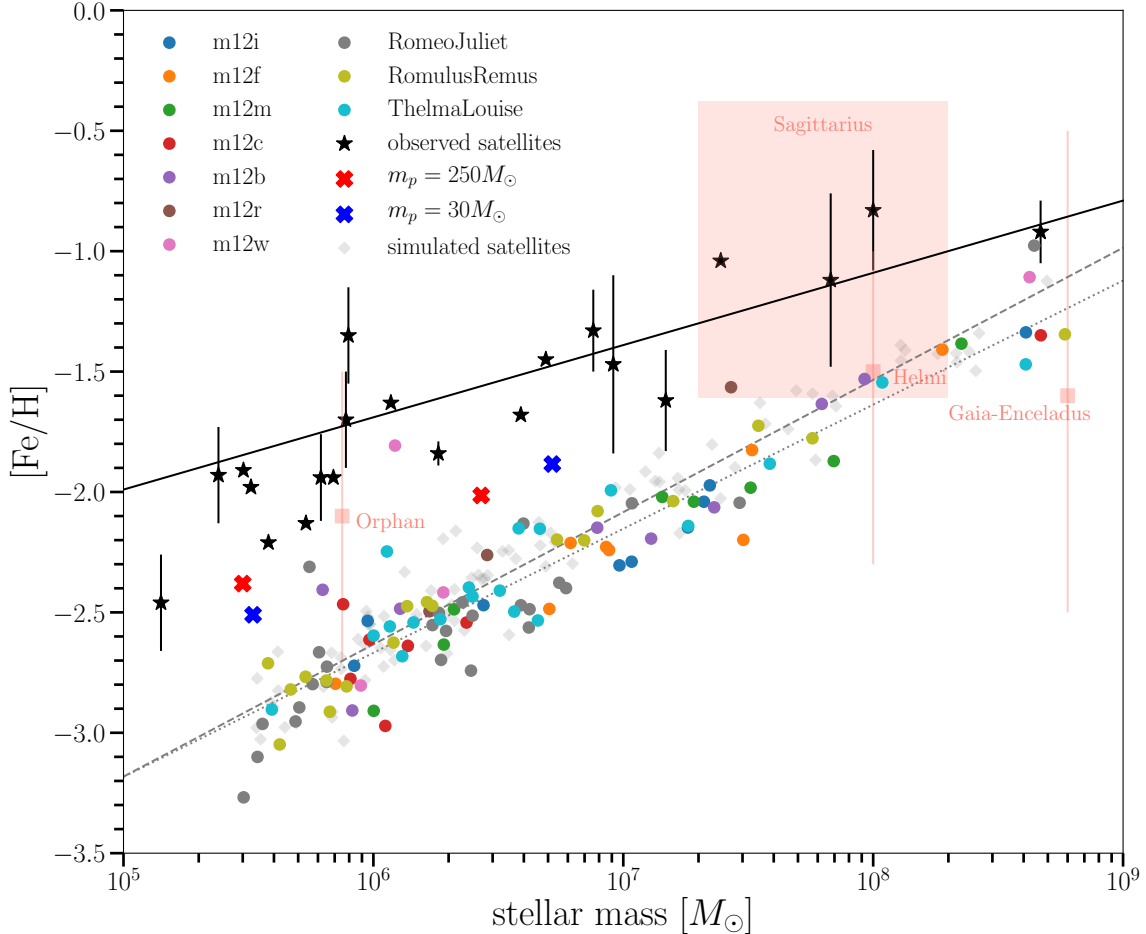
Figure 13. Stellar metallicity of stellar stream progenitors labeled by their simulations (colored dots) and present-day simulated satellite galaxies (gray diamonds), compared to observations of satellite galaxies (black stars) from Kirby et al. (2013) and streams (red rectangles). The solid line is Equation 4 of Kirby et al. (2013); the dashed and the dotted lines are the least-square fit for the simulated satellites and stream progenitors, respectively. High-resolution simulations of isolated dwarf galaxies from Wheeler et al. (2019) are also shown to approximate the size of resolution effects. These are the red and blue crosses, which are simulations with particle mass resolution of $250 M_\odot$ and $30 M_\odot$, respectively. For real streams (red rectangles), the data are from Mucciarelli et al. (2017); Gibbons et al. (2017) (Sagittarius), Helmi et al. (2018); Das et al. (2020) (Gaia-Enceladus), Sesar et al. (2013); Newberg et al. (2010) (Orphan) and Koppelman et al. (2019) (Helmi).

Solar value and decreasing over time (e.g. Tolstoy et al. 2009). This standard track in abundance space, and its variation with galaxy mass and duration of star formation, has been proposed as a basis for chemical decomposition of the stellar halo (Lee et al. 2015).

We study average abundance tracks for the three groups of objects (phase-mixed, streams, and dwarf galaxies) classified using our criteria in §3.2. The 2D [α/Fe] vs. [Fe/H] tracks and their respective 1D projections, stacked for all objects in each panel, are shown in Figure 14, for objects in different stellar mass ranges. Each 2D histogram is normalized such that the color in each bin represents the number of star particles in the bin normalized by the total number of star particles in that specific class and stellar mass range. For the 1D projections, each class of objects has a distinct

color, with lower mass objects using lighter shades of the same color. For phase-mixed objects, the low-mass, intermediate-mass and high-mass groups have the median values of [Fe/H] of -2.36, -1.86 and -1.53, respectively. For streams, the median values of [Fe/H] are -2.49, -2.00 and -1.27. For satellites, the median values of [Fe/H] are -2.32, -1.80 and -1.16. Similarly, for phase-mixed objects, the low-mass, intermediate-mass and high-mass groups have the median values of [α/Fe] of 0.27, 0.27 and 0.23, respectively. For satellites, the median values of [α/Fe] are 0.27, 0.27 and 0.25. Lastly, for satellites, the median values of [α/Fe] are 0.27, 0.24 and 0.22.

Within the same class of object, the highest stellar mass group is the most iron-enhanced, while the lowest stellar mass group is poor in iron. The [α/Fe] dis-

tributions are very similar across all stellar mass bins, with the highest stellar mass group having slightly lower [α/Fe]. Across different types of objects, the abundance distributions do not depend on mass, except for the high-mass group, in which surviving satellite galaxies have systematically higher [Fe/H] than the disrupted systems. This is in part, because high-mass dwarf galaxies likely have the longest duration of star formation, since they are least likely to be quenched by the present day relative to the progenitors of streams and phase-mixed components (see §4.3), unlike the smallest dark matter subhalos that can have their star-formation suppressed by a global outside influence, such as the reionization of the universe (Brown et al. 2014).

In addition to our abundance space analysis, Patel et al. in prep. will examine the origin of the visible features in this abundance space in FIRE dwarf galaxies.

5. ORBITS AND ALIGNMENT WITH THE GALACTIC DISK

In this section, we study the orbits of stream progenitors and their alignment with the disks of their host galaxies. The orbital plane of each progenitor is characterized by the direction of its total orbital angular momentum at the start of stream formation, while the disk plane is characterized as described in §5.1. We find a slight preference for streams to occupy orbits in the plane of the disk.

5.1. *Disk Angular Momentum Evolution*

The direction of the total angular momentum of the disk in each simulation is approximated by the direction of the principal axis vector of the moment-of-inertia tensor with the lowest eigenvalue. The principal axes are pre-computed for each snapshot using the youngest 25% of star particles within the distance that encloses 90% of the total stellar mass from the center of the host. We compute the disk's total angular momentum explicitly in three of the simulations (m12i, m12f, m12m) to confirm that the principal axes vectors are a good proxy for the disk angular momentum from ~ 4.6 Gyr ago onward, where the disks in those simulations are well-established and the direction of their angular momentum is relatively stable.

5.2. *Alignment of stream orbits with the disk plane*

We examine the alignment of the total orbital angular momentum of each stream's progenitor with the angular momentum of the disk, both evaluated at $\tau_{\rm stream}$ to approximate the initial orbits of tidally stripped stars. For this analysis we only include streams whose formation time is less than ~ 4.6 Gyr ago in the simulations, for three reasons. First, the principal axis vector with the smallest eigenvalue is a good proxy for the disk's total angular momentum only at later times (see §5.1). Second, the stream-formation time is generally later than both the infall time and the time where the progenitor

has peak stellar mass (see Figure 11 and §4.3). Between infall and stream-formation, there is often enough time for the direction of the disk to settle and the disk to become more stable. Finally, many orbital and dynamical properties of stellar streams, which will be studied later, are determined by the properties of their progenitors at stream-formation time rather than at infall or peak stellar mass time, since the orbit can undergo significant evolution, from dynamical friction and interactions with the disk, in the intervening period.

We compare orbits of bound subhalos (whether luminous or dark) at present day with those of satellites and stream progenitors across all simulations. We select all the bound subhalos at $z = 0$ within the virial radius of the main galaxy. For each bound object, we compute its total orbital angular momentum per unit mass with respect to the center of the main galaxy, and then compute the angle θ between this orbital angular momentum and the disk's angular momentum. Thus, subhalos that orbit in the disk plane have $\cos\theta = \pm 1$, while subhalos that have orbits perpendicular to the disk plane have $\cos\theta = 0$.

The left panel of Figure 15 shows the normalized distribution of $\cos\theta$ for all bound subhalos within the virial radius across all of the simulations (110,094 objects total). The presence of the stellar disk indeed has an impact on the overall orbital distribution of the subhalos: there is a peak in the otherwise almost flat distribution at $\cos\theta \sim \pm 1$. However, this distribution is cumulative across all simulations. The underlying distribution for each individual simulation varies quite significantly from this average profile, which is likely caused by variations in the assembly history of each host galaxy. In some individual cases the majority of subhalos are on retrograde orbits with respect to the galactic disk. To test its similarity with a uniform distribution, we randomly draw 1001 objects ($\sim 10\%$) from the distribution in the left panel of Figure 15 for 10,000 trials. In 55% of the draws, the majority of the objects have $|\cos\theta| > 0.5$. The quantity should be 50% if we draw from a perfectly uniform distribution, indicating a slight preference for subhalos to orbit closer to the disk plane.

Present day satellite galaxies with $10^5 M_\odot < M_\star < 10^9 M_\odot$ and stream progenitors with formation time later than ~ 4.6 Gyr (middle and right panel of Figure 15, respectively) are also slightly more likely to orbit on or close to the disk plane (these are evaluated at present-day for satellites, and at stream-formation time for stream progenitors). However, for satellite galaxies, this is only true for prograde orbits ($\cos\theta = +1$), but not for retrograde orbits ($\cos\theta = -1$). For stream progenitors, there are identifiable peaks in the orbital distribution at $\cos\theta \sim \pm 1$, similar to the overall subhalo orbital distribution. Interestingly, the prograde peak is not exactly at $\cos\theta = 1$. For each distribution, we perform random draws of 11 objects ($\sim 20\%$ of the sample) 10,000 times. The majority of the objects have
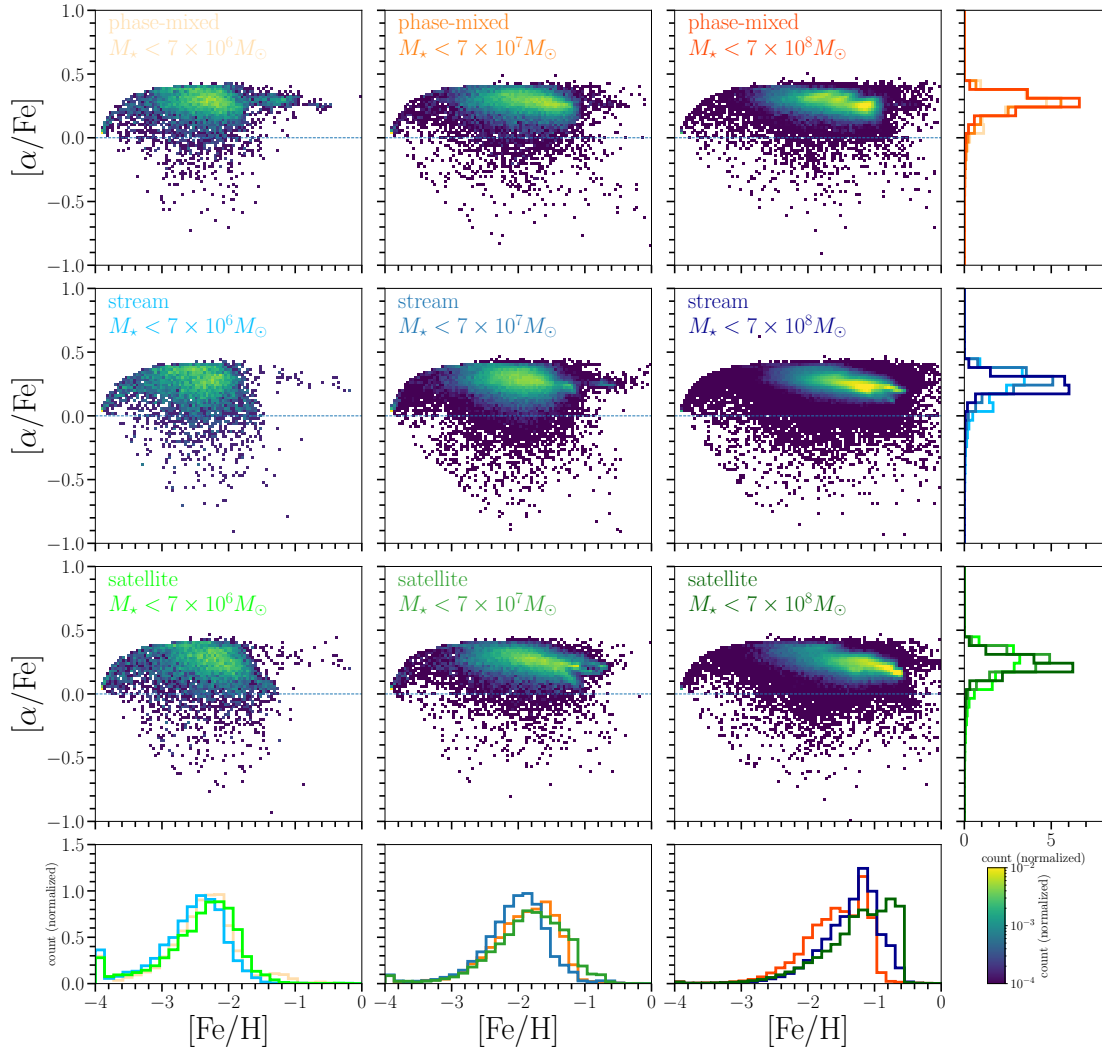
Figure 14. Histograms of [α/Fe] vs. [Fe/H] and their respective 1D projections for different classes of objects in three stellar mass bins for isolated simulations. Each bin is normalized by the total number of particles in each panel. Each class of objects has a distinct color, with lower mass objects using lighter shades of the color.

$|\cos\theta| > 0.5$ in 68% and 77% of the draws for the satellite galaxies and stream progenitors, respectively. The orbital distributions of these two groups of objects are thus non-uniform, and more skewed towards orbits in the disk plane than for the general subhalo population (see Samuel et al. (2020a) for a detailed analysis of satellite planarity in these simulations).

This finding has several implications. First, streams that orbit near the plane of the disk are more challenging to detect, and this "selection function" is likely to limit significantly our count of identified streams detected as overdensities. Streams in prograde orbits are also more likely to be affected by interactions with disk structures like bars, spiral arms, and molecular clouds which can mimic the effects of interactions with subha-

los (e.g. Pearson et al. 2017; Banik & Bovy 2018). Next, it implies that resonant interactions with the disk are slightly preferred, supporting interpretations of structures such as the "phase-space spiral" and related disk asymmetries (Widrow et al. 2012; Williams et al. 2013; Carlin et al. 2013; Antoja et al. 2018) as products of such interactions (Antoja et al. 2018; Laporte et al. 2018a,b, 2019a,b). Finally, the fact that the majority of the satellites are in prograde orbits underlines fluctuations in the orientation of the disk. In our simulations, big mergers can spark the formation of and/or shape the orientation of the stellar disk (Santistevan et al. 2021). D'Onghia et al. (2009) argues that resonant interactions on orbits prograde to the disk are most efficient at disrupting these satellites. If the orientation of the disk were fixed,
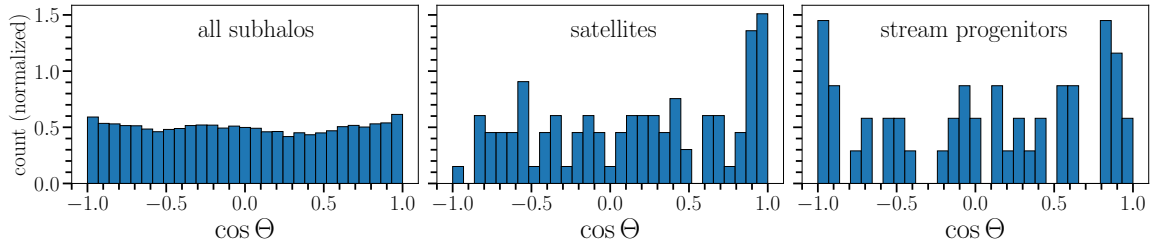
Figure 15. Alignment between orbital angular momentum and disk angular momentum for all bound subhalos (dark or luminous) within virial radii of the main host galaxies at present day (left); present-day satellite galaxies with $10^5 M_\odot < M_\star < 10^9 M_\odot$ (middle); and stream progenitors with stream-formation time $\tau_{\rm stream} < 4.6$ Gyr (right). The distributions for subhalos and satellites are evaluated at $z = 0$, while the distribution for stream progenitors is evaluated at each progenitor's $\tau_{\rm stream} <$. The angle θ between the subhalo's orbital angular momentum and the disk's angular momentum is evaluated at present day for the bound subhalos and satellite galaxies, and at $\tau_{\rm stream}$ for stream progenitors. $\cos\theta = 1$ corresponds to prograde orbits in the disk plane, while $\cos\theta = -1$ corresponds to retrograde orbits; $\cos\theta = 0$ corresponds to orbits perpendicular to the disk.

we would expect suppression in the number of prograde satellites, opposite from what we see.

6. MASS, ORBITAL CIRCULARITY, AND PHASE-MIXING TIME

Orbital circularity (the fraction of maximum allowed orbital angular momentum at a given energy) is one of the factors correlated with a stream's mixing time $\Delta\tau_{\rm mix}$, over which its initially gravitationally-bound stars become mixed with the equilibrium population of stars in the host galaxy. For constant orbital energy, a progenitor galaxy in a radial orbit (low circularity) experiences a steeper potential gradient along its motion; hence, positions of its stars in *phase-space* evolve more quickly compared to a progenitor in a more circular orbit. The increase in stream angular length per orbit can be approximated as (Yoon et al. 2011),

$$\Delta\Psi = \epsilon \left[\frac{2\pi}{T_\Psi} \frac{dT_\Psi}{dE} \right]_{L=L_{cir}}, \qquad (10)$$

where L_{cir} is the angular momentum of a circular orbit of energy E and azimuthal time-period $T_\Psi(E)$ at radius $r_{cir}(E)$. The prefactor $\epsilon \sim \Delta E$ is the spread in energy of the member star particles, which is approximately

$$\epsilon = \left(\frac{m}{M_p} \right)^{1/3} \frac{GM_p}{r_p}, \qquad (11)$$

where m is the mass of the dwarf galaxy with radius at the pericenter, r_p, and the mass of a host halo enclosed by the pericenter, M_p.

We consider objects in circular and radial orbits with the same mass m and energy E, assuming comparable T_Ψ and $\frac{dT_\Psi}{dE}$. These assumptions are true for the lowest order approximation of orbits around a point-like mass, where the energy determines the semi-major axis, which is then proportional to the azimuthal period of the orbit. The stream growth rate is

$$\frac{\Delta\Psi}{T_\Psi} \sim \frac{M_p^{2/3}}{r_p} \sim \frac{(\log(r_p/r_s))^{2/3}}{r_p}, \qquad (12)$$

assuming the standard Navarro-Frenk-White (NFW) profile with $r_p \sim r_s$, the scale radius. Since a circular orbit has the largest r_p, stellar streams that are in circular orbits should remain coherent longer than those in radial orbits with the same energy, which should lead to a bias toward circular orbits among older coherent streams. For phase-mixed objects, we thus expect a direct positive correlation between orbital circularity and mixing time. Since the energy spread is also a function of progenitor mass, we also expect more massive progenitors to mix faster.

### 6.1. *Determining Orbital Circularity*

Because the underlying global potential is time-dependent and not perfectly axisymmetric, orbits are not closed and evolve over cosmic time. Measuring the orbital circularity at different times will thus result in different outcomes. We determine orbital circularity of each object at stream-formation time, since tidal stripping "freezes" the orbital properties of stars in a progenitor galaxy with respect to their host. At any given point along the orbit, the orbital circularity η is defined as

$$\eta \equiv \frac{L(E)}{L_{cir}(E)}, \qquad (13)$$

where $L(E)$ and E are the angular momentum per mass and the total energy per mass, respectively, of the progenitor (streams or phase-mixed objects) at $t_{\rm stream}$. $L_{cir}(E)$ is the angular momentum per mass of a circular orbit with the same energy E. Thus $\eta = 1$ corresponds to a perfectly circular orbit at the time of measurement, while $\eta = 0$ corresponds to a perfectly radial orbit.

The total energy per mass, E, of the orbit is the mean kinetic and potential energy over all particles that belong to the object at stream-formation time. The kinetic energy per mass is the mean of the kinetic energy per mass of all star particles belonging to the progenitor. To estimate the potential energy due to the host galaxy without perturbations from large satellites, we use the AGAMA package (Vasiliev 2019) to model the smooth
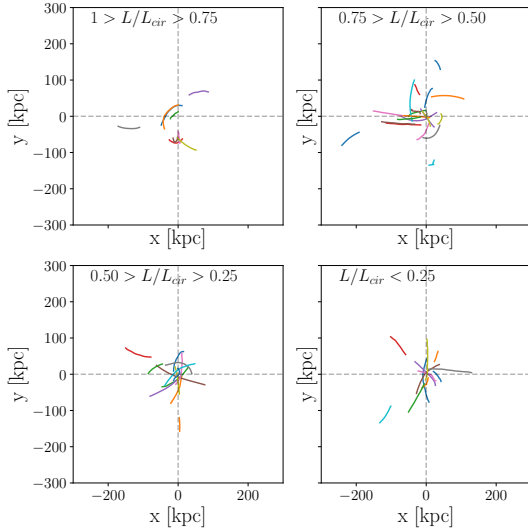
Figure 16. Orbital trajectories of stream progenitors and phase-mixed object progenitors in `m12i`, `m12f`, `m12m`, `m12b` and `m12r` , spanning ~ 0.5 Gyr around their stream-formation time, grouped by orbital circularity L/L_{cir}. The trajectories represent motion of the center of the bound part of these objects. The x-y plane here are in the coordinates of the simulations and hence are arbitrary with respect to the host or stream orbit.

component. The potential due to dark matter (within 500 kpc of the host) and gas (within 50 kpc of the host) is represented by an symmetric expansion in spherical harmonics up to $\ell = 4$, while the potential of the stars and cold gas within 50 kpc is approximated by an azimuthal harmonic expansion up to $m = 4$. The same radial cut offs are used in the paired simulations—since the paired galaxies are separated by ~ 800 kpc, the approximated potential around each host galaxy is only computed using particles within the same host. In all cases the potential is set to zero at infinity, so that particles with $E > 0$ can be considered unbound. We add the AGAMA potential at stream-formation time at the location (x, y, z) of the progenitor to the kinetic energy to get the total energy per mass, E.

For the circular orbit, the velocity of the object can be approximated as $v_{cir}^2(r_{cir}) = GM(< r_{cir})/r_{cir}$, while the potential energy per mass is from AGAMA evaluated at $x = r_{cir}$, $y = z = 0$. To estimate the radius of the circular orbit $r_{cir}(E)$, we compute the total energy of circular orbits every 2-kpc-interval within the virial radius of the host. The radius that yields the closest value to E is appointed as $r_{cir}(E)$. The angular momentum per mass of the progenitor is given by $L(E) = |\vec{r} \times \vec{v}|$, while $L_{cir}(E) = |r_{cir} \times v_{cir}(r_{cir})|$.

For isolated simulations, 19% have $L/L_{cir} > 1$ (indicating a problem), 29% have $1 > L/L_{cir} > 0.5$ (relatively circular) and 52% have $L/L_{cir} < 0.5$ (relatively radial). For paired simulations, over 60% of the objects

have $L/L_{cir} > 1$. If the global potential is static and spherically symmetric, the orbital shape, $L(E)/L_{cir}(E)$, should yield a value between 0 and 1 since the circular orbit has maximum angular momentum across all family of orbits with energy E. However, the potential in the simulation is time-dependent and not spherically symmetric, especially at earlier times, where the main halo accretes other substructures. The stream-formation time, especially for phase-mixed objects, tends to happen very early on in the simulations. Half of the phase-mixed objects start to form streams over 8.8 Gyr ago (see §4.4). The early stages of evolution in the paired simulations are also much more chaotic, compared to the isolated simulations, and the most-massive and second-most-massive hosts are not well-determined. Thus, the paired simulations and objects with $L/L_{cir} > 1$ are excluded in the analysis for the rest of this section. *The large number of objects that must be excluded for having nonsensical circularities underlines the limitations of this simple model for the orbital evolution of streams in realistic cosmological potentials.*

In figure 16, we show the orbital trajectories of streams and phase-mixed objects in `m12i`, `m12f`, `m12m`, `m12b` and `m12r` , spanning ~ 0.5 Gyr around their stream-formation time. These objects are grouped based on the circularity, L/L_{cir}, of their orbits, and we only consider objects with $L/L_{cir} < 1$. The orbital trajectories of the objects visibly become more radial as the computed circularity L/L_{cir} goes from higher to lower values, confirming that the circularity is at least somewhat correlated with orbital shape.

6.2. *Calculating Mixing Time and Dynamical Time*

In our analysis, the mixing time $\Delta\tau_{\mathrm{mix}}$ is defined to be the difference between the stream formation time and the time when the object is first considered phase-mixed according to Equation 2. The evolution of the velocity dispersion within each object is not a monotonically increasing function. The period of fluctuations in the local velocity dispersion is equal to the *radial* orbital period, with maximum at pericenter. Hence, the velocity dispersion for some objects crosses the phase mixing criterion threshold multiple times. Our definition of $\Delta\tau_{\mathrm{mix}}$ refers specifically to the difference between the stream formation time and the first up-crossing of the phase mixing criterion threshold. We verified by eye that a subset of objects indeed look phase-mixed after the first up-crossing.

The dynamical time $\Delta\tau_{\mathrm{dyn}}$ relevant for phase-mixing corresponds to the *azimuthal* period, T_ϕ, of the progenitor around stream formation time (see Equation 12). Rather than computing $\Delta\tau_{\mathrm{dyn}}$ from the approximate potential model, we do so directly from the simulation. We transform the position of a randomly chosen star particle from the simulation Cartesian coordinates into galactocentric spherical coordinates (r, θ, ϕ) in every snapshot using the instantaneous principal-axis frame
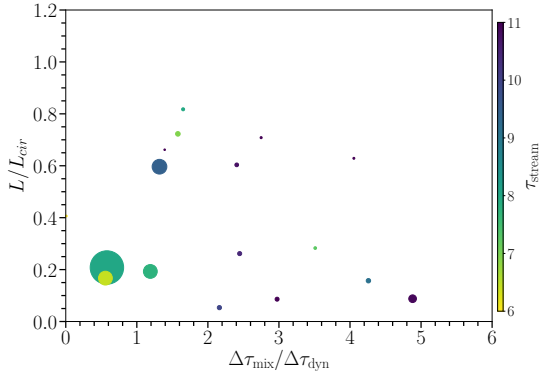
Figure 17. Orbital circularity of phase-mixed objects in isolated simulations as a function of the ratio between mixing time and dynamical time, $\Delta\tau_{\mathrm{mix}}/\Delta\tau_{\mathrm{dyn}}$. The stellar mass of each object is represented by the size, while the stream-formation time τ_{stream} is represented by the color of each marker. There is no apparent correlation between the quantities, although no streams on very circular orbits ($L/L_{\mathrm{circ}} \gtrsim 0.5$) mix in less than about 1 dynamical time. Streams that mix the fastest tend to be of higher mass, consistent with Equation 12.

determined as described in §5.1, consistent with the assumption of approximate spherical symmetry that underlies the derivation of Equation 12. We begin tracking the particle around 0.5 Gyr before stream formation time and follow the particle until ϕ comes back to its starting value.

6.3. *What determines how fast streams phase-mix?*

Figure 17 shows the orbital circularity of phase-mixed objects in isolated simulations as a function of the ratio between their mixing time and dynamical time, $\Delta\tau_{\mathrm{mix}}/\Delta\tau_{\mathrm{dyn}}$, described in §6.2. The color of each point represents the stream-formation time, while the size represents the stellar mass of the object. Contrary to the standard picture described in §6, we do not see a correlation between the orbital circularity with $\Delta\tau_{\mathrm{mix}}/\Delta\tau_{\mathrm{dyn}}$. We argue that this lack of correlation is physical, and not caused by the resolution limitation of our simulations. If the resolution were too low, we would underestimate the mixing time of the objects, especially the low-mass objects with fewer number of star particles. As a result, there would be a trend with stellar mass such that low-mass objects would have the lowest $\Delta\tau_{\mathrm{mix}}/\Delta\tau_{\mathrm{dyn}}$. Instead, we see that high mass objects, for which we can most confidently estimate $\Delta\tau_{\mathrm{mix}}$, have the lowest $\Delta\tau_{\mathrm{mix}}/\Delta\tau_{\mathrm{dyn}}$.

The assumption that the global potential is smooth, static and axisymmetric is clearly not good enough to model the time-evolution of a stream from formation to mixing, especially for objects with very early stream-formation times, when the halo's potential is lumpy and its time-evolution is non-adiabatic. During the early

stages of galaxy formation, this chaotic environment accelerates the phase-mixing of accreted satellite galaxies, as illustrated in Figure 11. *Since the typical time for a galaxy to become a satellite, begin tidally disrupting, and spread out into a stream is long compared to the age of the host galaxy, the non-adiabatic evolution of the galactic potential at early times is potentially one of the main contributors to phase-mixing in the stellar halo.*

7. GRAVITATIONAL COOLING OF STREAMS

This section studies the evolution of the local velocity dispersion as dwarf galaxies evolve into stellar streams. Under an adiabatic potential of the host halo, the conservation of phase-space volume in collisionless systems predicts that streams should grow kinematically colder as they grow in length, albeit in a phase-dependent way (Helmi & White 1999). The evolution of the local velocity dispersion is thus an important test of whether we are sufficiently resolving our simulated streams. We can also determine to what extent the local velocity dispersion in a stream is a good proxy for the mass of its progenitor.
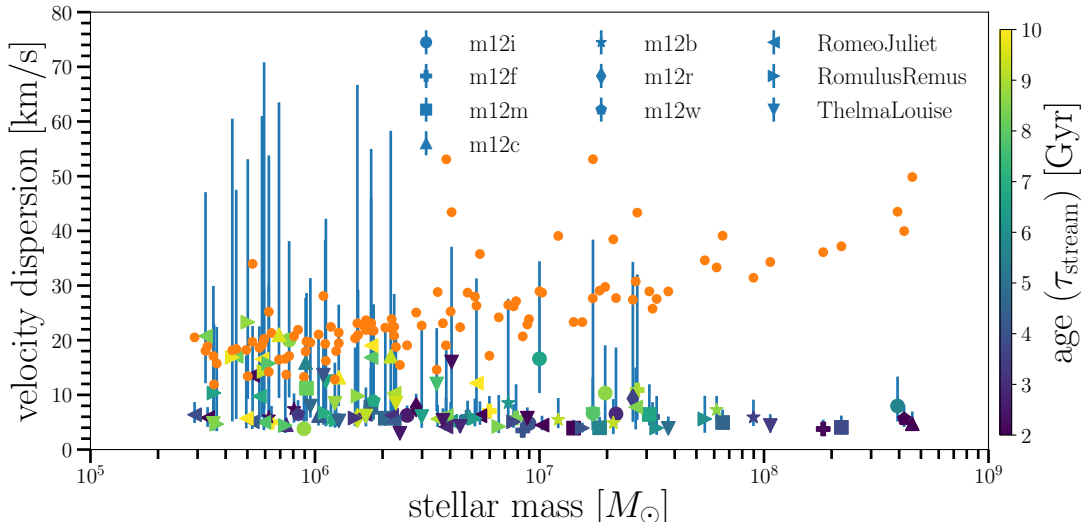
7.1. *Comparison of Progenitors' and Streams' Velocity Dispersions*

Figure 18 compares the *global* velocity dispersions of the stellar stream progenitors (standard deviation over all star particles; orange) with the local velocity dispersion when those progenitors become streams (blue). The local velocity dispersions are computed as shown in §3.2, where we use 20 nearest neighbors for streams with more than 300 star particles and 7 nearest neighbors otherwise. The median local velocity dispersions along the streams are $\sim 4 - 20$ km/s, while the global velocity dispersion of the stream progenitors are $\sim 15 - 50$ km/s. The global velocity dispersion σ increases as a function of the total progenitor mass M as predicted for a dispersion-supported system: $\sigma \propto M^{1/3}$.

The wide range of the local velocity dispersion along a single stream suggests that we should refer to a median value of the local velocity dispersion when we discuss the velocity dispersion of a stellar stream. Our stellar streams indeed get colder as a function of time, as predicted by the conservation of phase-space volume.

7.2. *Evolution of local velocity dispersion*

We select a sample stellar stream from m12i to study how its local velocity dispersion evolves with time. This sample stream has about 2000 particles in total, which is roughly the median number of particles for our sample of streams. We track the local velocity dispersion of one randomly-chosen star particle in this stream from the time when the progenitor is still bound until the stream forms. This is shown in Figure 19 where the blue line represents the local velocity dispersion of this star particle as a function of the simulation time t. The start of the simulation corresponds to $t = 0$ Gyr, while the

Figure 18. Comparison between local velocity dispersions $\sigma_{\rm local}$ along stellar streams (blue error bars) and total velocity dispersion $\sigma_{\rm total}$ of their progenitors (orange), evaluated at the time when they have maximum stellar mass. The error bars for streams span the 16th to 83rd percentile of $\sigma_{\rm local}$, with the median values shown by the markers. Their colors represent their ages, which are defined to be the stream-formation times $\tau_{\rm stream}$ (look-back) described in §4.2.

present day corresponds to $t \sim 13.7$ Gyr. The orange line represents the distance between this star particle and the center of the main host galaxy. The time elapsed between adjacent pericentric passages, or the radial period T_r, is ~ 1 Gyr.

The object has $\tau_{\rm stream} \approx 8$ Gyr, but this particular star particle leaves the bound part at $t \approx 11$ Gyr (3 Gyr after tidal disruption commences). The local velocity dispersion of this star particle is roughly time-independent before the star particle is stripped from the bound part, even after $t_{\rm stream}$. After the star particle leaves the bound part ($t > 11$ Gyr), its local velocity dispersion fluctuates between ~ 5 and 20 km/s after the progenitor is tidally disrupted, and is anti-correlated with the star particle's distance to the center of the main galaxy. The orbital phase of the stream governs the velocity dispersion, as the value peaks when the star particle approaches the pericenter of the orbit. This behavior, pointed out in Helmi & White (1999), indicates that we are successfully approximating the phase-space evolution of our simulated stellar streams.

It is common to infer the mass of a stellar stream's progenitor by using the velocity dispersion of stars along the stream as a proxy for the dispersion in the progenitor, corrected by a "gravitational cooling factor" that scales with the stream age. This example illustrates that this method is complicated by the phase-dependent nature of the velocity dispersion, especially in real observations where we can only observe the brightest fraction of the stellar streams. Fortunately, these brightest parts likely include the remnant of the bound part of the progenitor, which is likely to have a higher stellar density than the stream itself. Using the local velocity dispersion of stars
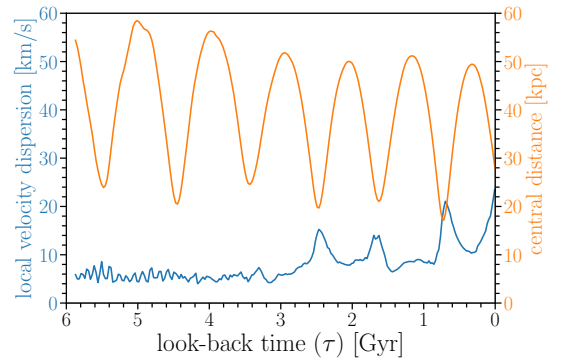
Figure 19. Evolution of local velocity dispersion (blue; left y-axis label) and distance to the center of the main galaxy (orange; right y-axis label) as a function of look-back time for one random star particle in a stellar stream selected from `m12i`.

in the bound remnant, in which the fluctuation with orbital phase is small, we can possibly model the mass of the progenitor with higher accuracy. Otherwise, the fluctuation with orbital phase will likely be the dominant source of uncertainty in any mass estimate based on stellar velocity dispersions in streams.

8. SUMMARY

In this paper, we present a set of stellar streams, dwarf satellite galaxies and phase-mixed objects selected from the FIRE-2 cosmological hydrodynamical simulations, including 7 isolated MW-mass systems and 3 paired

MW-Andromeda-like systems[2]. Objects are classified using a set of phase-space criteria (§3.2): satellite galaxies have a small size in both position and velocity, streams have a large extent in position and small local extent in velocity, and phase-mixed structures have a large range in both position and velocity. We show that these stellar streams are dynamically cold, and that the median value of the local velocity dispersion within an object, $\langle\sigma\rangle$, can be used to separate phase-mixed objects from stellar streams, with only a weak dependence on stellar mass. Applying these criteria, we find a total of 106 simulated coherent stellar streams with stellar masses $10^5 \lesssim M_* \lesssim 10^9 M_\odot$ across all 13 MW-mass hosts (§1). We use this sample to study the origin, chemistry, and dynamical properties of simulated stream progenitors. The important findings are the following:

1. Present-day satellite galaxies are good proxies for stellar stream progenitors (§4). They have similar stellar mass and total mass functions (§4.1), implying that satellite galaxies are equally likely to turn into stellar streams independent of mass. The stellar mass–velocity dispersion relation for stream progenitors, evaluated at $t_{\rm peak}$, resembles that for real dwarf satellites (§4.5). The velocity dispersions of stellar stream progenitors range from 10 km/s for low-mass objects ($M_\star \approx 10^6 M_\odot$) to 40 km/s for high-mass objects ($M_\star \approx 10^9 M_\odot$). The [Fe/H]–[$\alpha$/Fe] evolution tracks for simulated stream progenitors, dwarf galaxies and phase-mixed objects are similar for objects with $M_\star < 7 \times 10^7 M_\odot$, while the high-mass dwarf satellites are slightly more iron-rich compared to stream progenitors and phase-mixed objects with similar masses (§4.5.3).

2. The order in which infall, quenching, and tidal disruption occur for progenitors of stellar streams varies with stellar mass (§4.3). Low-mass progenitors ($M_\star < 2.25 \times 10^6 M_\odot$) are likely to have their star formation quenched before their first infall, while most high-mass progenitors ($M_\star > 2.25 \times 10^6 M_\odot$) have their star formation quenched by the host environment, with many continuing to form star particles and reaching peak stellar mass *after* stream-formation time.

3. The orientation of the galactic disk affects the orbital distributions of all surviving subhalos (luminious or dark), dwarf galaxies and stream progenitors (§5). All substructures slightly prefer orbits that align with the galactic disk plane; the degree of bias is higher for dwarf satellites than subhalos in general, and higher for streams than for dwarf satellites. Subhalos and stream progenitors appear to equally pre-

fer prograde and retrograde orbits, while dwarf satellites are apparently biased towards prograde orbits.

4. For streams that form more than 6–8 Gyr ago, the non-adiabatic evolution of the global potential during the early, chaotic phase of the formation of the host galaxy determines how quickly streams become phase-mixed with the host, regardless of their orbital circularity or mass (§6). For more than half of the streams in our simulations, a smooth, static, and axisymmetric potential is insufficient to model their time evolution after tidal disruption.

5. Orbital-phase-dependent fluctuations complicate using the velocity dispersion to estimate the mass of a stream progenitor (§7). Most stream progenitors have a total velocity dispersion > 20 km/s, with a strong mass-dependence consistent with observations and the theory of dispersion-supported systems. Most of the streams have a median local velocity dispersion $\langle\sigma\rangle < 10$ km/s, but this fluctuates by up to a factor 4 with orbital phase (highest at pericenter), which translates to more than a factor 50 in mass.

The use of cosmological-hydrodynamical simulations allows us to study streams and their progenitors in a realistic system without relying on simplifying theoretical assumptions, and avoiding the strong biases in observational measurements (findings 1, 2 and 3). Simulated streams also serve as a tool to test our theoretical models of streams in realistic MW-mass host environment where many of these assumptions no longer hold exactly (findings 4 and 5). The non-adiabatic time-evolution of the host galaxy (especially in its early stages), the non-smoothness in the global potential and the resolution-dependence in the simulations (see Figures 1 and 13) all contribute to the divergence from theoretical predictions and semi-analytic models. However, we have shown that the resolution limit alone is not the major source or attribute to all of differences (see Figures 12, 13, 14 and 17). These discrepancies are thus mainly physical in nature, and call for more realistic models in the post-*Gaia* era.

---

[2] full dataset at: https://flathub.flatironinstitute.org/sapfire.

REFERENCES

Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, Nature, 561, 360, doi: 10.1038/s41586-018-0510-7

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481, doi: 10.1146/annurev.astro.46.060407.145222

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Banik, N., & Bovy, J. 2018, ArXiv e-prints, arXiv:1809.09640

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013a, ApJ, 762, 109, doi: 10.1088/0004-637X/762/2/109

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., et al. 2013b, ApJ, 763, 18, doi: 10.1088/0004-637X/763/1/18

Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611, doi: 10.1093/mnras/sty982

Benincasa, S. M., Loebman, S. R., Wetzel, A., et al. 2020, MNRAS, 497, 3993, doi: 10.1093/mnras/staa2116

Bernard, E. J., Ferguson, A. M. N., Schlafly, E. F., et al. 2016, MNRAS, 463, 1759, doi: 10.1093/mnras/stw2134

Bonaca, A., Geha, M., Küpper, A. H. W., et al. 2014, ApJ, 795, 94, doi: 10.1088/0004-637X/795/1/94

Bonaca, A., Hogg, D. W., Price-Whelan, A. M., & Conroy, C. 2019, ApJ, 880, 38, doi: 10.3847/1538-4357/ab2873

Bringmann, T. 2009, New Journal of Physics, 11, 105027, doi: 10.1088/1367-2630/11/10/105027

Brown, T. M., Tumlinson, J., Geha, M., et al. 2014, ApJ, 796, 91, doi: 10.1088/0004-637X/796/2/91

Carlberg, R. G. 2012, ApJ, 748, 20, doi: 10.1088/0004-637X/748/1/20

—. 2013, ApJ, 775, 90, doi: 10.1088/0004-637X/775/2/90

Carlberg, R. G., Grillmair, C. J., & Hetherington, N. 2012, ApJ, 760, 75, doi: 10.1088/0004-637X/760/1/75

Carlin, J. L., DeLaunay, J., Newberg, H. J., et al. 2013, ApJL, 777, L5, doi: 10.1088/2041-8205/777/1/L5

Cornell, J. M., Profumo, S., & Shepherd, W. 2013, Phys. Rev. D, 88, 015027, doi: 10.1103/PhysRevD.88.015027

Cortes, C., & Vapnik, V. 1995, Mach. Learn., 20, 273, doi: 10.1023/A:1022627411411

Das, P., Hawkins, K., & Jofré, P. 2020, MNRAS, 493, 5195, doi: 10.1093/mnras/stz3537

Deason, A. J., Belokurov, V., & Weisz, D. R. 2015, MNRAS, 448, L77, doi: 10.1093/mnrasl/slv001

Debattista, V. P., Gonzalez, O. A., Sanderson, R. E., et al. 2018, ArXiv e-prints, arXiv:1805.12199

D'Onghia, E., Besla, G., Cox, T. J., & Hernquist, L. 2009, Nature, 460, 605, doi: 10.1038/nature08215

Erkal, D., Belokurov, V., Bovy, J., & Sand ers, J. L. 2016, MNRAS, 463, 102, doi: 10.1093/mnras/stw1957

Escala, I., Wetzel, A., Kirby, E. N., et al. 2018, MNRAS, 474, 2194, doi: 10.1093/mnras/stx2858

Faucher-Giguère, C.-A. 2018, MNRAS, 473, 3717, doi: 10.1093/mnras/stx2595

Garrison-Kimmel, S., Wetzel, A., Bullock, J. S., et al. 2017, MNRAS, 471, 1709

Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2018a, ArXiv e-prints, arXiv:1806.04143

—. 2018b, MNRAS, 481, 4133

—. 2019, MNRAS, 487, 1380, doi: 10.1093/mnras/stz1317

Gibbons, S. L. J., Belokurov, V., & Evans, N. W. 2017, MNRAS, 464, 794, doi: 10.1093/mnras/stw2328

Guszejnov, D., Grudić, M. Y., Offner, S. S. R., et al. 2020, MNRAS, 492, 488, doi: 10.1093/mnras/stz3527

Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85, doi: 10.1038/s41586-018-0625-x

Helmi, A., & White, S. D. M. 1999, MNRAS, 307, 495, doi: 10.1046/j.1365-8711.1999.02616.x

Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Nature, 402, 53, doi: 10.1038/46980

Hopkins, P. F. 2015, MNRAS, 450, 53

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800, doi: 10.1093/mnras/sty1690

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Ibata, R. A., Lewis, G. F., Irwin, M. J., & Quinn, T. 2002, MNRAS, 332, 915, doi: 10.1046/j.1365-8711.2002.05358.x

Ibata, R. A., Lewis, G. F., & Martin, N. F. 2016, ApJ, 819, 1, doi: 10.3847/0004-637X/819/1/1

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102, doi: 10.1088/0004-637X/779/2/102

Knebe, A., Devriendt, J. E. G., Gibson, B. K., & Silk, J. 2003, MNRAS, 345, 1285, doi: 10.1046/j.1365-2966.2003.07044.x

Koppelman, H. H., Helmi, A., Massari, D., Roelenga, S., & Bastian, U. 2019, A&A, 625, A5, doi: 10.1051/0004-6361/201834769

Laporte, C. F. P., Gómez, F. A., Besla, G., Johnston, K. V., & Garavito-Camargo, N. 2018a, MNRAS, 473, 1218, doi: 10.1093/mnras/stx2146

Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018b, MNRAS, 481, 286, doi: 10.1093/mnras/sty1574

Laporte, C. F. P., Johnston, K. V., & Tzanidakis, A. 2019a, MNRAS, 483, 1427, doi: 10.1093/mnras/sty2362

Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019b, MNRAS, 485, 3134, doi: 10.1093/mnras/stz583

Lee, D. M., Johnston, K. V., Sen, B., & Jessop, W. 2015, ApJ, 802, 48, doi: 10.1088/0004-637X/802/1/48

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3

Ma, X., Hopkins, P. F., Wetzel, A. R., et al. 2017, MNRAS, 467, 2430

McConnachie, A. W. 2012, AJ, 144, 4, doi: 10.1088/0004-6256/144/1/4

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51 – 56

Mucciarelli, A., Bellazzini, M., Ibata, R., et al. 2017, A&A, 605, A46, doi: 10.1051/0004-6361/201730707

Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, MNRAS, 454, 2691, doi: 10.1093/mnras/stv2126

Newberg, H., Willett, B., Yanny, B., & Xu, a. 2010, The Astrophysical Journal, 711, 32, doi: 10.1088/0004-637X/711/1/32

Oliphant, T. E. 2006, A guide to NumPy, Vol. 1 (Trelgol Publishing USA)

Pearson, S., Price-Whelan, A. M., & Johnston, K. V. 2017, Nature Astronomy, 1, 633, doi: 10.1038/s41550-017-0220-3

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Perez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21, doi: 10.1109/MCSE.2007.53

Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Samuel, J., Wetzel, A., Chapman, S., et al. 2020a, arXiv e-prints, arXiv:2010.08571. https://arxiv.org/abs/2010.08571

Samuel, J., Wetzel, A., Tollerud, E., et al. 2019, arXiv e-prints, arXiv:1904.11508. https://arxiv.org/abs/1904.11508

—. 2020b, MNRAS, 491, 1471, doi: 10.1093/mnras/stz3054

Sanderson, R. E., Garrison-Kimmel, S., Wetzel, A., et al. 2017, ApJ, in press, arXiv:1712.05808

Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2018, ArXiv e-prints, arXiv:1806.10564

Santistevan, I. B., Wetzel, A., El-Badry, K., et al. 2020a, MNRAS, 497, 747, doi: 10.1093/mnras/staa1923

—. 2020b, arXiv e-prints, arXiv:2001.03178. https://arxiv.org/abs/2001.03178

Santistevan, I. B., Wetzel, A., Sanderson, R. E., et al. 2021, arXiv e-prints, arXiv:2102.03369. https://arxiv.org/abs/2102.03369

Sesar, B., Grillmair, C. J., Cohen, J. G., et al. 2013, ApJ, 776, 26, doi: 10.1088/0004-637X/776/1/26

Sparre, M., Hayward, C. C., Feldmann, R., et al. 2017, MNRAS, 466, 88, doi: 10.1093/mnras/stw3011

Stern, J., Faucher-Giguère, C.-A., Fielding, D., et al. 2020, arXiv e-prints, arXiv:2006.13976. https://arxiv.org/abs/2006.13976

Tolstoy, E., Hill, V., & Tosi, M. 2009, ARA&A, 47, 371, doi: 10.1146/annurev-astro-082708-101650

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22, doi: 10.1109/MCSE.2011.37

Vasiliev, E. 2019, MNRAS, 482, 1525, doi: 10.1093/mnras/sty2672

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, doi: https://doi.org/10.1038/s41592-019-0686-2

Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014, ApJ, 789, 148, doi: 10.1088/0004-637X/789/2/148

Wetzel, A., & Garrison-Kimmel, S. 2020a, GizmoAnalysis: Read and analyze Gizmo simulations. http://ascl.net/2002.015

—. 2020b, HaloAnalysis: Read and analyze halo catalogs and merger trees. http://ascl.net/2002.014

Wetzel, A. R., Deason, A. J., & Garrison-Kimmel, S. 2015, ApJ, 807, 49, doi: 10.1088/0004-637X/807/1/49

Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016a, ApJL, 827, L23

—. 2016b, ApJ, 827, L23, doi: 10.3847/2041-8205/827/2/L23

Wheeler, C., Hopkins, P. F., Pace, A. B., et al. 2019, MNRAS, 490, 4447, doi: 10.1093/mnras/stz2887

Widrow, L. M., Gardner, S., Yanny, B., Dodelson, S., & Chen, H.-Y. 2012, ApJL, 750, L41, doi: 10.1088/2041-8205/750/2/L41

Williams, M. E. K., Steinmetz, M., Binney, J., et al. 2013, MNRAS, 436, 101, doi: 10.1093/mnras/stt1522

Yoon, J. H., Johnston, K. V., & Hogg, D. W. 2011, ApJ, 731, 58, doi: 10.1088/0004-637X/731/1/58

Yu, S., Bullock, J. S., Klein, C., et al. 2021, arXiv e-prints, arXiv:2103.03888. https://arxiv.org/abs/2103.03888