

# Reconciling Observed and Simulated Stellar Halo Masses

Robyn E. Sanderson<sup>1,2,3</sup> , Shea Garrison-Kimmel<sup>2,9</sup> , Andrew Wetzel<sup>2,4,5</sup> , Tsang Keung Chan<sup>6</sup> , Philip F. Hopkins<sup>2</sup> , Dušan Kereš<sup>6</sup> , Ivanna Escala<sup>7</sup> , Claude-André Faucher-Giguère<sup>8</sup> , and Xiangcheng Ma<sup>2</sup> 

<sup>1</sup> Department of Physics & Astronomy, University of Pennsylvania, 209 S. 33rd St., Philadelphia, PA 19104, USA; [robyn@caltech.edu](mailto:robyn@caltech.edu)

<sup>2</sup> TAPIR, MC 350-17, California Institute of Technology, Pasadena, CA 91125, USA<sup>10</sup>

<sup>3</sup> Columbia University, Department of Astronomy, 550 W. 120th St., Mail Code 5246, New York, NY 10027, USA

<sup>4</sup> Department of Physics, University of California, Davis, CA 95616, USA

<sup>5</sup> The Observatories of the Carnegie Institution for Science, Pasadena, CA 91101, USA

<sup>6</sup> Department of Physics, Center for Astrophysics and Space Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

<sup>7</sup> Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

<sup>8</sup> Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

Received 2017 December 13; revised 2018 September 26; accepted 2018 October 22; published 2018 December 6

## Abstract

We use cosmological hydrodynamical simulations of Milky Way–mass galaxies from the FIRE project to evaluate various strategies for estimating the mass of a galaxy’s stellar halo from deep, integrated-light images. We find good agreement with integrated-light observations if we mimic observational methods to measure the mass of the stellar halo by selecting regions of an image via projected radius relative to the disk scale length or by their surface density in stellar mass. However, these observational methods systematically underestimate the accreted stellar component, defined in our (and most) simulations as the mass of stars formed outside of the host galaxy, by up to a factor of 10, since the accreted component is centrally concentrated and therefore substantially obscured by the galactic disk. Furthermore, these observational methods introduce spurious dependencies of the estimated accreted stellar component on the stellar mass and size of galaxies that can obscure the trends in accreted stellar mass predicted by cosmological simulations, since we find that in our simulations, the size and shape of the central galaxy are not strongly correlated with the assembly history of the accreted stellar halo. This effect persists whether galaxies are viewed edge-on or face-on. We show that metallicity or color information may provide a way to more cleanly delineate in observations the regions dominated by accreted stars. Absent additional data, we caution that estimates of the mass of the accreted stellar component from single-band images alone should be taken as lower limits.

**Key words:** dark matter – galaxies: halos – galaxies: structure – methods: numerical – methods: observational

## 1. Introduction

Recent advances in observational astronomy have begun to reveal the faint stellar halos surrounding Milky Way (MW)–mass galaxies in integrated light (Martínez-Delgado et al. 2010; Bakos & Trujillo 2012; D’Souza et al. 2014; Merritt et al. 2016). Cosmological simulations (Bullock & Johnston 2005; Purcell et al. 2007; De Lucia & Helmi 2008; Font et al. 2008; Cooper et al. 2010, 2013; Font et al. 2011; Tissera et al. 2013, 2014; Pillepich et al. 2014) have long predicted that a significant fraction of the light in these halos, especially at large distances from the main galaxy, should come from the remains of smaller galaxies that were accreted onto the main galaxy and tidally disrupted. Throughout this work, we will refer to this component, which is a reflection of the hierarchical nature of structure formation in a cold dark matter (CDM) universe, as the “accreted stellar component.” In the CDM picture, the process of hierarchical accretion thus ties the variation in the mass fraction of accreted stars to the accretion history of the host galaxy (Bullock & Johnston 2005; Tissera

et al. 2012, 2013; Rodriguez-Gomez et al. 2016; Amorisco 2017). In principle, this component holds some of the few memories of the accretion process. Identifying regions of galaxies that are primarily made up of accreted material is thus the first step toward testing these predictions.

To verify the predictions of CDM, there have been recent attempts to compare the amount of stellar mass observed in the outskirts of galaxies with the mass fraction in accreted stars predicted by simulations (e.g., Font et al. 2008; Pillepich et al. 2014; Merritt et al. 2016; D’Souza & Bell 2018; Elias et al. 2018; Huang et al. 2018). For the purposes of this paper, we will use the term “stellar halo” in an observational sense, referring to the faint structure in the outskirts of galaxies beyond the central concentration of stellar mass. In practice, there are many observational definitions of this term that can include radial profile, surface brightness, or metallicity characteristics. Most recent observational attempts to characterize the stellar halos of galaxies, including the MW (e.g., Carollo et al. 2010) and the Andromeda galaxy (M31; e.g., Courteau et al. 2011), have used analysis of resolved stellar populations to identify the accreted component, usually by searching for an old, metal-poor population extending far from the central galaxy (e.g., Seth et al. 2007; Cockcroft et al. 2013). However, this method requires extremely deep images, mainly obtained using the *Hubble Space Telescope* (with the exception of Greggio et al. 2014), and has therefore been limited to a small handful of galaxies so far. Observing stellar halos in

<sup>9</sup> Einstein Fellow.

<sup>10</sup> NSF Astronomy & Astrophysics Postdoctoral Fellow.

integrated light is, in principle, more easily scalable to the sample sizes needed to explore the wide variation in the accreted component that is predicted by simulations (e.g., Bakos & Trujillo 2012; D’Souza et al. 2014; Duc et al. 2015; Merritt et al. 2016; Huang et al. 2018), presuming that it is possible to account for the contribution of scattered light (de Jong 2008; Slater et al. 2009; Sandin 2014). However, the lack of resolved stellar population information makes it far more challenging to identify the regions of an image dominated by accreted material. So far, no work has attempted to account for how the method used to select the stellar halo from a galaxy observed in integrated light may bias the comparison to simulations, where the provenance of material is perfectly known and a variety of definitions of “stellar halo” are imposed. Despite efforts such as that in Rodriguez-Gomez et al. (2016) to understand whether the mass in the stellar halo comes primarily from accreted material or from stars formed in the central galaxy (which in this work we call “formed in situ”) and expelled to the halo, and its dependence with separation from the central galaxy, it is not straightforward to apply these results to the spatial selections in projection that are commonly used in integrated-light images. In fact, most prior work has focused on comparisons between observed galaxies and predictions for the stellar halo based on dark matter-only (DM-only) simulations tagged with stars, where the stellar halo is by definition 100% accreted. However, in simulations that include baryonic physics, both of these distinct channels are observed to contribute to the stellar halos of galaxies (Font et al. 2011; Cooper et al. 2013; Tissera et al. 2013; Pillepich et al. 2015; Anglés-Alcázar et al. 2017; Gómez et al. 2017a), and both are interesting for what they tell us about the process of galaxy formation, as well as the cosmology in which galaxies are formed.

Recently, Merritt et al. (2016, hereafter M16) presented a sample of eight MW-mass galaxies with stellar halos observed in integrated light, from which they estimated stellar halo masses and mass fractions and compared them to predictions from cosmological simulations. Although still small, this is the first such sample to exist in the literature at this mass scale (D’Souza et al. 2014 stacked many galaxies together, and Duc et al. 2015 looked at more massive early-type galaxies) and is a promising step toward placing the MW’s stellar halo into a cosmological context. Interestingly, most of their measured stellar halo mass fractions lie systematically lower when compared to simulations by nearly an order of magnitude, but both the simulated predictions and the observational methods could potentially have systematic offsets in mass. On the simulation side, selections in present-day, three-dimensional radial distance from the host galaxy are often used to define the halo component, sometimes scaled to the size of the central galaxy in the case of simulations that include baryons (e.g., Pillepich et al. 2014, 2015) or with a fixed radial range in the case of tagged DM-only (and hence accretion-only) simulations (e.g., Cooper et al. 2010, 2013). On the observational side, M16 used a spatial selection in projected radius, based on the best-fit disk scale length for the central galaxy, to define the stellar halo component. These differing definitions could be responsible for at least part of the apparent discrepancy between the measured and simulated halo mass fractions reported in M16.

In this work, we use a set of high-resolution cosmological zoom simulations from the FIRE-2 suite<sup>11</sup> (Hopkins et al.

2018) to directly test the consistency and accuracy of the methods used in simulations and observations for measuring the mass fraction in stellar halos of MW-mass galaxies and to explore which observational methods might be effective for separating *in situ* from accreted stars in images. In Section 2, we describe the simulated galaxies in our sample. In Section 3, we describe how we define the accreted component of the stellar halo and discuss the general properties of the simulated halos. In Section 4, we describe how we produce mock images from the simulations. In Section 5, we reproduce methods of measuring the halo mass using our mock data and discuss how the different measurement methods can introduce unwanted biases in halo masses. In Section 6, we explore possible ways to reduce these biases, and in Section 7, we summarize.

## 2. Simulations

In this work, we study the stellar halos of a suite of simulated MW-mass galaxies at different resolutions and with different environments and accretion histories. Their basic properties are summarized in Table 1. All are cosmological zoom-in, hydrodynamical  $N$ -body simulations carried out with the GIZMO meshless hydrodynamic simulation code (Hopkins 2015) and FIRE-2 model for star formation and stellar feedback (Hopkins et al. 2018). In some cases, an explicit implementation of subgrid turbulent metal diffusion (Hopkins 2016) is included, which reduces artificial numerical noise in the metallicity distributions (Escala et al. 2018) but has almost no effect on the large-scale properties of the simulated galaxy (Su et al. 2017). For this study, we use the highest-resolution simulation available for each set of initial conditions; however, for many of these simulations, we also have variations at lower resolution. Appendix A illustrates that we do not expect these differences to significantly affect the results.

The suite includes three basic groups of simulations.

1. Six isolated halos with a resolution of  $57,000 M_{\odot}$  or better per star particle, with a variety of different accretion histories and stellar masses.
2. Four pairs of halos, with similar resolution to the six isolated halos ( $32,000$ – $57,000 M_{\odot}$  per star particle), selected to roughly resemble the MW–M31 configuration, for a total of eight halos in a Local Group-like environment. Two of these pairs are hydrodynamical resimulations of pairs in the ELVIS suite (Garrison-Kimmel et al. 2014a), while the other two are new additions (see S. Garrison-Kimmel et al. 2018, in preparation, for more details).
3. Three isolated halos simulated at the standard FIRE-2 resolution for  $M_h(z=0) \sim 10^{12} M_{\odot}$  halos ( $7070 M_{\odot}$  per star particle). These simulations are part of the “Latte” suite first described in Wetzel et al. (2016). Two of the three (m12i and m12f in Table 1) have been resimulated to include subgrid turbulent metal diffusion.

The simulations we consider are selected from large cosmological boxes containing thousands of dark matter halos in the mass range  $10^{11}$ – $10^{12} M_{\odot}$  (see Hopkins et al. 2014 and Kim et al. 2014 for details of how m11 and m12 halos are chosen and Garrison-Kimmel et al. 2014a for a description of how the paired halos were chosen). Although the main halos were chosen to fall in specific mass ranges, and in the case of the ELVIS suite, to be in pairs with a separation and orbit similar to the MW–M31 system, in all cases, these selections

<sup>11</sup> See the FIRE project website: <http://fire.northwestern.edu>.

**Table 1**  
Simulations Used in This Work

Name	$m_p (M_\odot)$	$M_{\text{vir}} (10^{12} M_\odot)$	$r_{\text{vir}} (\text{kpc})$	$r_{-2} (\text{kpc})$	$M_{*,90} (10^{10} M_\odot)$	$r_{*,90} (\text{kpc})$	References
m11f	17000	0.50	207.7	8.5	2.5	8.3	C
m11g	17000	0.64	225.3	8.1	4.8	8.1	B
m12b	56500	1.37	290.6	5.4	14.3	10.0	F
m12c	56500	1.30	285.9	4.7	8.9	5.5	F
m12q	56500	1.71	313.3	3.2	16.7	6.3	F
m12z	33000	0.87	249.9	10.2	4.0	13.0	H
m12m	7070	1.47	297.8	10.7	12.6	15.7	L
m12f	7070	1.40	293.2	14.1	8.4	16.1	T
m12i	7070	1.07	268.0	12.3	6.4	10.5	T
Romeo	28000	1.29	285.7	8.5	7.9	18.6	E
Juliet	28000	1.06	268.0	8.5	6.0	15.2	E
Romulus	31900	1.95	325.5	9.8	15.5	13.6	E
Remus	31900	1.25	280.4	5.1	11.5	9.0	E
Thelma	31900	1.44	294.4	10.7	13.2	12.8	E
Louise	31900	1.10	269.4	5.4	7.4	14.1	E
Batman	57000	1.90	325.2	4.9	12.6	4.3	E
Robin	57000	1.58	305.9	8.5	7.1	12.2	E

**Notes.**  $m_p$ : baryonic particle mass.  $M_{\text{vir}}$ ,  $r_{\text{vir}}$ : Bryan & Norman (1998) virial quantities.  $r_{-2}$ : radius where the log-slope of the dark matter density profile is  $-2$ .  $M_{*,90}$ ,  $r_{*,90}$ : mass and radius of 90% of stellar mass within 30 kpc of the central galaxy at  $z = 0$ .

**References.** (C) Chan et al. (2018). (B) El-Badry et al. (2018). (F) Part of the FIRE-2 suite (Hopkins et al. 2018). (H) Hafen et al. (2017). (L) Part of the Latte simulation series (Wetzel et al. 2016). (E) Part of the ELVIS simulation series (Garrison-Kimmel et al. 2014a; S. Garrison-Kimmel et al. 2018, in preparation). (T) This work.

were agnostic to the formation histories of the halos (Garrison-Kimmel et al. 2014a; Wetzel et al. 2016; El-Badry et al. 2018). Garrison-Kimmel et al. (2014a) further found (in DM-only simulations) no statistical difference between the subhalo properties (counts and kinematics) and host halo properties (formation times and concentrations) of the paired halos in ELVIS compared to similar isolated halos, suggesting that selecting paired halos does not itself result in a selection on the distribution of satellites that form the accreted stellar halo. Our sample, although too small to span the full range, can therefore be considered an unbiased and representative sampling of accretion histories for host galaxies in this mass range.

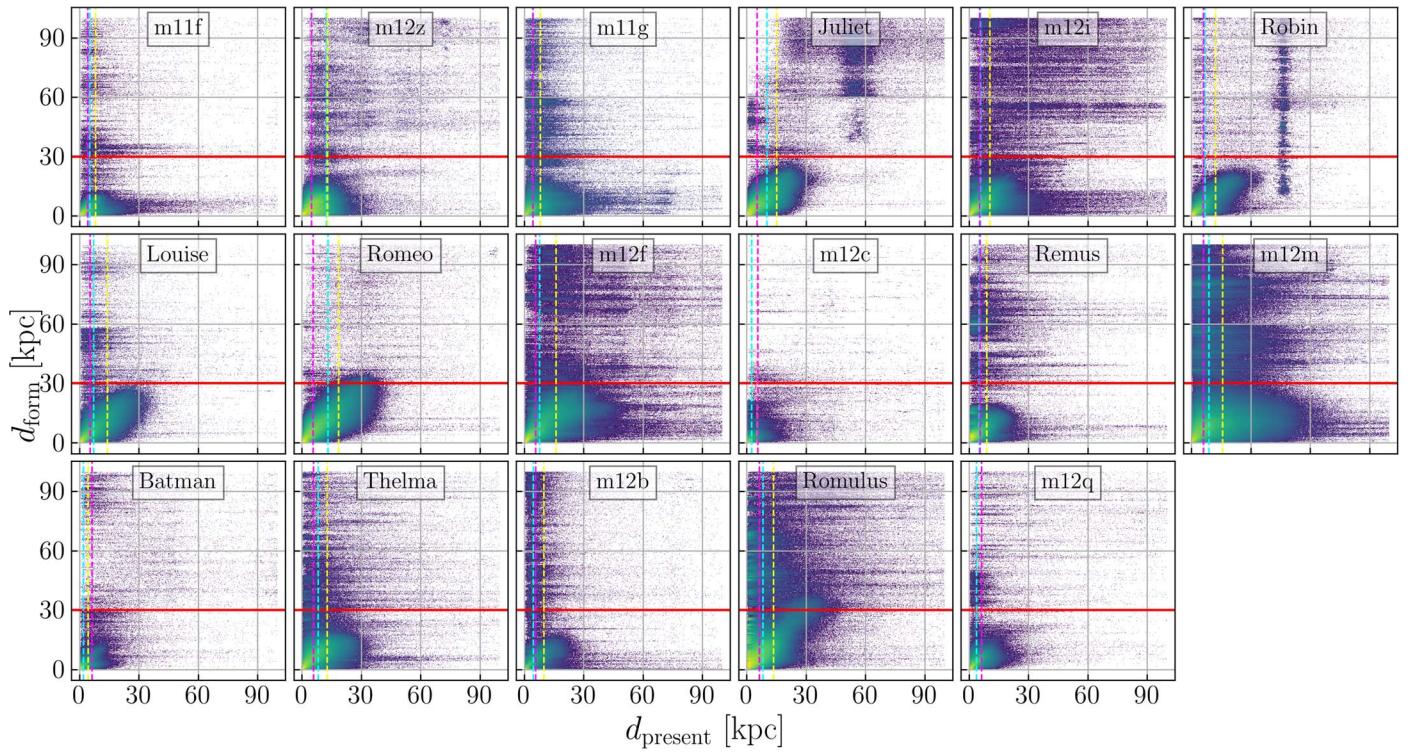
FIRE-2 galaxies have already been shown to match the observed stellar mass–dark matter halo mass relation over cosmic time (Hopkins et al. 2018). This is critical to present-day comparisons of the stellar halo mass fraction in order to ensure that the simulated galaxies falling in this mass range at  $z = 0$  have had accretion histories consistent with expectations for this mass scale. The lowest-resolution simulations within the group used for this work may slightly overestimate the stellar mass in the main galaxy (see Appendix A). The FIRE-2 galaxies also feature systems of satellite galaxies at  $z = 0$  that are consistent with the mass, size, and number distribution of present-day satellites around the MW and M31 (Wetzel et al. 2016); although these are not necessarily identical to the building blocks of the stellar halo, this agreement supports the assumption that previously accreted galaxies also had realistic properties and that the disruption rate of satellites by tides is plausible. Finally, the main galaxies in the FIRE-2 suite at this mass scale also match the distribution of observed galaxy sizes (Ma et al. 2017; S. Garrison-Kimmel et al. 2018, in preparation), which is crucial given that the observational method for selecting the stellar halo involves a spatial cut proportional to the disk scale length.

### 3. The Stellar Halos of Simulated Galaxies

#### 3.1. Distinguishing Accreted from *In Situ* Stars

In our simulated galaxies, we distinguish star particles that were accreted from those formed *in situ* using the distance of each star particle from the center of the main progenitor halo of the  $z = 0$  galaxy of interest in the first snapshot after it is formed ( $d_{\text{form}}$  for short), as in Bonaca et al. (2017). The time interval between snapshots is approximately 24 Myr over most of the star formation history, so we consider this to be a suitable approximation to a star particle’s “birth distance.” This diagnostic is far simpler but also far less computationally intensive than the detailed particle-tracking analysis applied to the FIRE-1 simulations by Anglés-Alcázar et al. (2017, hereafter AA17). In terms of their hierarchy of definitions (see Figure 1 of AA17), this criterion selects stars formed outside the galaxy, whether from externally processed or nonexternally processed material, that are then incorporated into the galaxy either by merger or intergalactic transfer. Our selection also includes star particles that may have been gas particles at first accretion but form into stars within the accreting galaxy before being tidally stripped, which AA17 would formally consider *in situ* star formation from externally processed material. Although it fails to capture the nuances in the origin of the star particles, we consider that our definition based on formation distance adequately captures the accreted stellar component of each simulated galaxy for the purpose of this work, which is to broadly investigate what is being selected using common observational definitions of “stellar halo.”

As with any such investigation, there is necessarily some dependence on the specifics of our criterion for which stars are accreted. To better understand this dependence, we consider Figure 1, which shows the distribution of  $d_{\text{form}}$  relative to the present-day distance ( $d_{\text{present}}$ : the distance of that star particle to the galaxy center at  $z = 0$ ) for all star particles in each of the



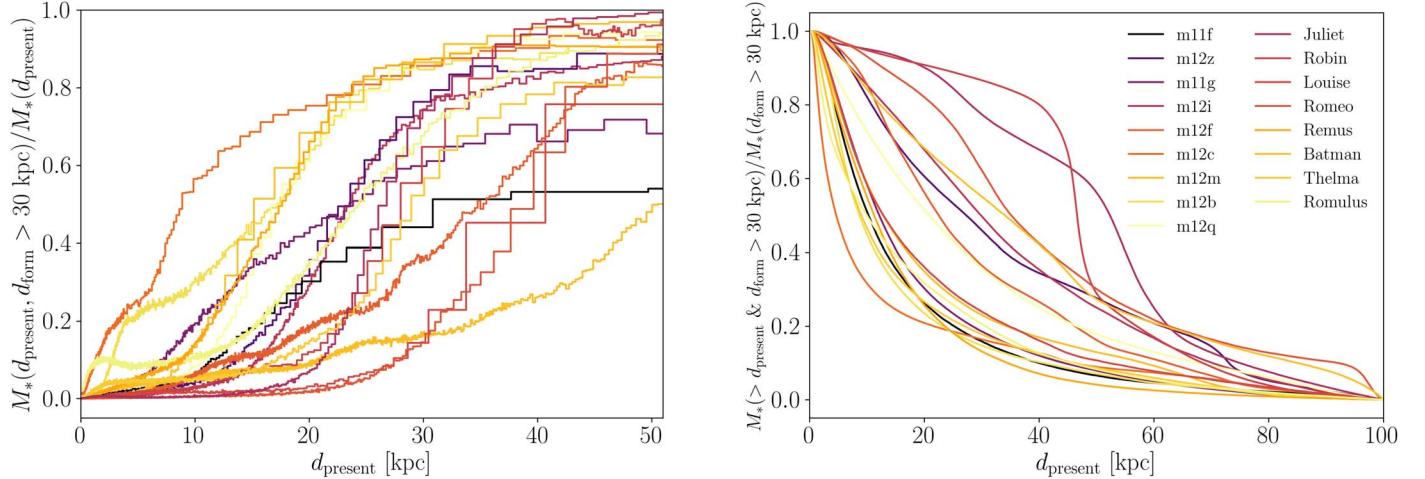
**Figure 1.** Maps of the formation distance  $d_{\text{form}}$  as a function of present-day distance for star particles within 100 kpc of the centers of the simulated galaxies at the present day, in order from lowest (top left) to highest (bottom right) stellar mass  $M_{90}$ . The color scale varies as the base-10 logarithm of the stellar mass density per pixel from low (blue) to high (yellow). Star particles formed beyond 30 kpc from the main galaxy (above the red horizontal line) are considered accreted. Dashed vertical lines indicate  $2r_{*,50}$  (cyan; twice the present-day half-mass radius in stars),  $r_{*,90}$  (yellow; radius enclosing 90% of the stellar mass at the present day, as in Table 1), and  $R_{\text{vir}}/50$  (magenta).

simulated galaxies out to a present-day galactocentric radius of 100 kpc. In each galaxy, the in situ population is visible as a concentration in the lower left-hand corner of the plot that has a roughly 1:1 relationship with large scatter as a function of radius, indicating that the majority of stars in this population tend to stay near where they are born but may have migrated somewhat (especially the older stars; see Ma et al. 2017). Material near either the  $x$  or  $y$  axis in this diagram is also interesting: along the  $y$  axis, there is often a fairly significant amount of stellar mass that was formed at large distances from the main halo but is now located in the inner galaxy (though not necessarily in the disk plane), while along the  $x$  axis, there is material that was formed relatively close in and then scattered out of the inner galaxy. Some tidal streams from disrupted satellites are evident in this diagram as horizontal streaks (that is, they formed as a bound object at some well-defined distance, perhaps in a burst of star formation, and their stars now span a range of radii). Star-forming satellite galaxies are evident as vertical streaks: they are forming new stars at a variety of radii while changing their distance from the main galaxy, but they retain all their stars as a bound object with a small present-day distance range. We do not exclude star particles within satellite galaxies from our estimate of the accreted stellar mass, and in fact, there are very few galaxies in the inner 50 kpc that have not been tidally disrupted.

Based on the view in Figure 1, we will define  $d_{\text{form}} > 30$  kpc as the separation between in situ and accreted populations. This cut is sufficient to exclude the in situ population for all the galaxies in our sample. We emphasize that this is a conservative choice meant to ensure that very little in situ material is included rather than attempting to capture all the

accreted material: clearly some galaxies in Figure 1 have material with  $d_{\text{form}} < 30$  kpc that looks like it was accreted. We did check that changing the criterion to  $d_{\text{form}} > 20$  kpc does not appreciably change the results. This can be at least partially attributed to the fact that about half of the material designated as “accreted” and ending up with  $d_{\text{present}} < 100$  kpc has  $d_{\text{form}} > 100$  kpc and so does not show up on this figure at all (although its mass is indeed counted as part of the accreted component). This choice also excludes the small amount of material that formed in the disk and is scattered to large radii by the present day (such as the horizontal streaks near the  $y$  axis in, for example, m12f). We therefore expect that our quoted accreted stellar masses are probably underestimated; Figure 1 gives an idea of the degree of underestimation for each halo.

Figure 1 also shows, for reference, several length scales commonly used to describe the extent of simulated galaxies or define the region considered the stellar halo in simulations:  $2r_{*,50}$  (twice the three-dimensional spherical radius enclosing 50% of the stellar mass within 30 kpc of the central galaxy, shown in cyan; Elias et al. 2018),  $r_{*,90}$  (the radius enclosing 90% of the stellar mass within 30 kpc of the central galaxy, in yellow; S. Garrison-Kimmel et al. 2018, in preparation), and  $R_{\text{vir}}/50$  (the fraction of the virial radius used in Pillepich et al. 2014 to define the stellar halos of Illustris galaxies, shown in magenta). For most of our galaxies, all three of these length scales are fairly similar and even overlap in some cases, but in most galaxies, the in situ sequence (where  $d_{\text{form}} \sim d_{\text{present}}$ ) extends further than any of these distances. For our simulated sample at least, using any of these criteria to select the accreted stellar halo component would overestimate the mass fairly significantly in most cases. Using a criterion related to a



**Figure 2.** Transition from in situ to accreted material occurs at a wide range of radii for our simulated galaxies. In both panels, lines are colored by the present-day total stellar mass of the galaxy ( $M_{90}$  in Table 1) from lowest (dark purple) to highest (yellow). Left: fraction of star particles that were accreted (with  $d_{\text{form}} > 30 \text{ kpc}$ ) per bin of 5000 star particles in present-day distance  $d_{\text{present}}$  for each simulated galaxy. At 25 kpc from the center, the fraction of material that is accreted ranges from a few to 80%. Right: fraction of stellar mass outside a given present-day distance of the galaxy’s center that was accreted (i.e., formed beyond 30 kpc), relative to the total accreted stellar mass presently within 100 kpc of the central galaxy.

fraction of the stellar mass would also set a maximum value on the stellar halo mass fraction. We therefore avoid using any of these length scales to define the stellar halo.

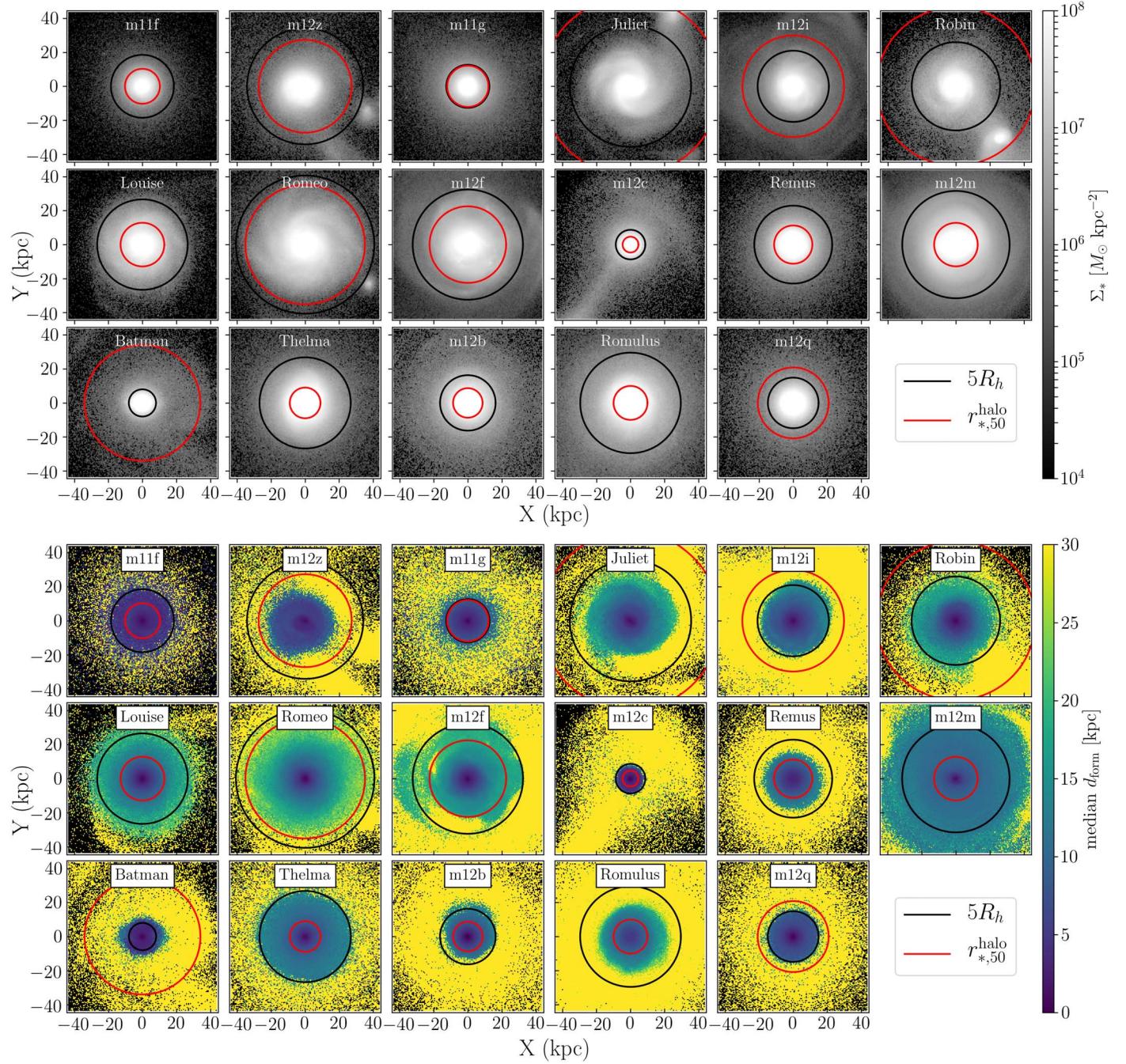
### 3.2. Variation of Simulated Stellar Halo Properties

The galaxies in our sample have a wide variety of accretion histories and differ substantially in the present-day morphologies of both their stellar halos and their disk–bulge systems (S. Garrison-Kimmel et al. 2018, in preparation). Figure 2 shows how the wide variation in the accretion histories of the galaxies is reflected in the radial distribution of their accreted material. The left-hand panel shows a broad diversity in the fraction of accreted material, relative to the total stellar mass within 50 kpc, as a function of radius (computed using bins of 5000 star particles). The range of radii where the transition from mostly in situ to mostly accreted material occurs is quite broad, consistent with the results of Rodriguez-Gomez et al. (2016) for a much larger sample of simulated galaxies in this stellar mass range. In the right-hand panel of the figure, the y axis shows the fraction of accreted material (defined as  $d_{\text{form}} > 30 \text{ kpc}$ ) presently at distances larger than  $x$  relative to the total accreted mass within 100 kpc for each simulated galaxy. The degree of central concentration of the accreted material also varies widely from halo to halo: the half-mass radius of the accreted component can be anywhere from 5 to 50 kpc, varying by an order of magnitude for a range of about a factor of 7 in host stellar mass. For comparison, the ratio of largest to smallest  $r_{90}$  within our sample, which measures the extent of the total stellar mass, is 4.3. No trend with host stellar mass is apparent in Figure 2. An investigation of the Spearman rank coefficient, which was used to test for correlations between the half-mass radius of the accreted material and the host stellar mass, bears this out: the value of  $r_{\text{Sp}} = -0.24$  is in the 16th percentile of correlation values computed for a bootstrapped sample of 1000 shuffled versions of the same data. These are approximately normally distributed; thus, the computed value is roughly  $1\sigma$  from the median of the decorrelated samples, indicating a statistically insignificant degree of correlation.

Examining the outliers in Figure 2, we find that some of them would be readily apparent from images of the galaxy, while others are less so. In cases where the accreted component is unusually extended (like Robin, Romeo, and Juliet), a relatively massive companion contributes a substantial fraction of accreted material at relatively large separation, while at the opposite extreme (like m12c), nearly all the accreted material is within 20 kpc of the galactic center at the present day. Tellingly, these two extremes are in different stages of a relatively major interaction whose signatures are still present in images of the galaxy (see Figure 3). Even excluding such cases, however, one is left with the difference between galaxies like m12f and m12b, neither of which show visible signs of a recent merger with a large companion but that have substantial differences in the concentration of their accreted material: only 20% of m12b’s accreted stars lie beyond 20 kpc, while 70% of m12f’s do. In fact, m12f did have a recent merger with a galaxy of roughly the original mass of the Sagittarius dwarf that contributes to its large and extended halo and has substantially disrupted its outer disk, but this is not immediately apparent from the mock image alone.

The face-on projections shown in the top panel of Figure 3 show how this maps onto the present-day appearance of the different simulated galaxies. The sample considered in M16 tends toward spiral galaxies with extended disks (like m12i and m12f) but also includes a few with prominent bulges (like m12b and Batman).

The masses and mass fractions of the stellar halos in our suite are broadly similar to the general trends found in studies that used semi-analytic modeling (Bullock & Johnston 2005) or particle tagging in DM-only simulations (Cooper et al. 2010, 2013), or in larger cosmological volumes simulated at lower resolution using different physical models than those in FIRE (Font et al. 2011; Tissera et al. 2013; Pillepich et al. 2014, 2015; Rodriguez-Gomez et al. 2016). Because (as we will show in Section 5.2) the strategies used to define halo mass make an important difference, we will not make more detailed comparisons between our results and those of simulations by other groups using different codes. Instead, we will focus in the rest of this work on something that the high-resolution galaxies



**Figure 3.** Galaxies in our simulated sample show distinct transitions from primarily in situ to primarily accreted stars that are not always accompanied by a distinct transition in the surface mass density. Top: simulated stellar mass surface density maps of all the galaxies in our sample, created as described in Section 4, in order from lowest to highest stellar mass  $M_{*,90}$  (Table 1). The lognormalized gray scale shows surface densities between  $10^4$  (black; roughly one particle in a pixel in most simulations) and  $10^8 M_\odot \text{ kpc}^{-2}$  (white) to emphasize fainter outer features. The graininess at the lowest surface densities in the outskirts of each map is due to fluctuations in the number of particles per pixel; no additional observational noise sources were simulated. The pixels used here and in the bottom panels are  $12''$  on a side, corresponding to 0.6 kpc at the fiducial distance of 10 Mpc used to create these maps. For comparison, central surface densities tend to fall in the range  $10^9\text{--}10^{10} M_\odot \text{ kpc}^{-2}$ , as shown in Table 2. Bottom: median mass-weighted formation distance of star particles in each pixel of the simulated images in the top panel. Yellow pixels have average  $d_{\text{form}} > 30$  kpc, our cutoff for material considered accreted (see Section 2). Black pixels contain no star particles. The degree to which the region inside  $5R_d$  (black circle; see Section 4.2) corresponds to in situ material varies widely from galaxy to galaxy and does not systematically correlate with the half-mass radius of the accreted stellar halo within 100 kpc,  $r_{*,50}^{\text{halo}}$  (red circle; see right panel of Figure 2).

here can do uniquely well: testing the fidelity of observational methods for measuring the mass in accreted stars.

#### 4. Mock Observations of Simulated Stellar Halos

In order to compare different observational methods for determining stellar halo mass, we first need to create a mock

“observation” of each galaxy to which we can apply the method. Given that the goal of this paper is not to explore the validity of assumptions about color corrections, stellar population modeling, or correction for background light, we chose to produce and compare maps of stellar mass surface density, rather than modeling the spectral energy distribution (SED) to produce mock images in different bands. This approach

necessarily assumes that the process of converting from flux in a set of filters produces an accurate estimate of the stellar mass surface density in each pixel.

We attempt to reproduce as closely as possible the method used in M16 to determine  $f_{\text{halo}}$  for our simulated galaxies, which comprises the following steps.

1. Obtain a map of the stellar mass surface density for each system. M16 uses a color transformation on images; we bin the star particles in the simulation as described in Section 4.1.
2. Fit a double Sérsic profile to the disk and bulge (see Section 4.2).
3. Sum the stellar mass in pixels with ellipsoidal radius larger than 5 times the scale length of the disk component in the Sérsic fit to obtain the mass in the stellar halo.
4. Divide by the total stellar mass in the image to obtain the stellar halo mass fraction,  $f_{\text{halo}}$ .

#### 4.1. Maps of Stellar Mass Surface Density

We place each simulated galaxy at a distance of 10 Mpc, roughly the median of the M16 sample, and rotate it to a face-on orientation defined by calculating the principal axes of the stellar mass distribution within 10 kpc of the Galactic center. We then create  $60' \times 60'$  maps of stellar mass surface density (corresponding to about  $85 \times 85$  kpc) made up of pixels  $12''$  on a side (about 0.6 kpc) by binning the star particles of the simulated galaxy in projection. The mock images are shown as thumbnails in the top panel of Figure 3. Consistent with the approach in M16, we do not expressly attempt to remove satellite galaxies from the mock images.

Spiral arms are faint but evident in the disks of many of our simulated galaxies. In the outskirts, faint structure is seen in the stellar halos that in most cases extends beyond the frame of the map. This is expected, since the maps only probe to about 50 kpc in projected radius, while the typical virial radii of the dark matter halos for our simulated galaxies are about 300 kpc (see Table 1).

#### 4.2. Galaxy Profile Fitting

As in M16, we describe the disk and bulge of each simulated galaxy using a double Sérsic profile as a function of ellipsoidal projected radius  $\tilde{R}$ :

$$\Sigma_*(\tilde{R}) = I_b \exp \left\{ -b_n \left[ \left( \frac{\tilde{R}}{R_b} \right)^{1/n} - 1 \right] \right\} + I_d \exp \left\{ -b_1 \left[ \left( \frac{\tilde{R}}{R_d} \right) - 1 \right] \right\}, \quad (1)$$

with  $b_n$  defined such that the effective radius ( $R_b$  for the bulge or  $R_d$  for the disk) contains half the total luminosity. We fit this function to the region of each simulated stellar mass surface density map out to approximately the last spiral features or drop-off in surface density. For most galaxies, this region is within 20–30 kpc in  $\tilde{R}$ . We restrict the scale radius  $R_b$  of the bulge component to be less than or equal to the scale radius  $R_d$  of the disk component, fix the index of the disk component to 1 (i.e., we assume an exponential disk), and allow the index  $n$  of the bulge component to vary between 0 and 5. We set a broad

prior on the logarithm of the two normalizations  $I_b$  and  $I_d$  of the bulge and disk components, respectively.

In M16, the outermost spiral features in the galaxy images were used to select the region used to fit a Sérsic profile to the disk and bulge, but since we use stellar mass density directly rather than synthesizing an SED, spiral features are often less apparent than in an image, since young stars are much brighter in luminosity than they are massive. If no clear spiral features are present, we use the radius corresponding to a drop-off in surface mass density. For all of our galaxies, this is roughly 20–40 kpc; the results of the Sérsic fit are not sensitive to the exact value.

In order to account for the fact that galaxies are not perfectly round, the mass and light profiles of galaxies are commonly described as a function of ellipsoidal projected radius  $\tilde{R}$ . Instead of determining the ellipsoidal radius of each pixel by fitting ellipsoids in annuli, as was done in M16, we directly compute them from the same principal-axis vectors used to rotate the disks into a face-on projection, such that

$$\tilde{R}^2 \equiv x^2 + \left( \frac{a}{b} y \right)^2, \quad (2)$$

where  $a$  and  $b$  are the axis ratios of the principal axes defining the directions  $x$  and  $y$ , respectively. This choice is not as sensitive to possible twists in the orientation of the disk as a function of projected radius, but since most of our simulated galaxies are fairly regular in shape, the resulting scatter in the density profile tends to be fairly small; hence, we do not expect it to significantly affect the fitted parameters. We also checked that the by-eye choice of radial range for the fit does not significantly affect the results.

We perform a least-squares fit to the surface mass density of the individual pixels for each galaxy using Levenberg–Marquardt minimization with a fit-by-eye first guess. However, since this algorithm finds only the nearest local minimum, we also did a more complete Monte Carlo sampling of parameter space using emcee (Foreman-Mackey et al. 2013) for one galaxy to understand how the well-known degeneracies inherent in Sérsic fitting affected the determination of the disk scale length  $R_d$ , which is used to delineate between disk and halo in M16. The details of this test are described in Appendix B; we conclude that even when other parameters are degenerate, the value of  $R_d$  is sufficiently robust that we can use the Levenberg–Marquardt minimization results (Table 2) with confidence.

#### 5. Performance of Typical Methods to Measure Stellar Halo Mass

We examine two existing observational methods for measuring the stellar halo mass fraction  $f_{\text{halo}}$ : the disk-fitting method employed by M16 and the surface mass density cutoff motivated by the Cooper et al. (2013) retagging of DM-only halos in the Millennium-II simulation at the low end of their mass range (see their Figure 6). We also look at the performance of three selections used in simulations for separating stellar halo from disk populations to determine whether the simulated and observed methods agree on stellar halo mass when applied to the same simulated galaxy. All these methods rely on a selection in either present-day spatial location (3D or projected) or surface density (which is

**Table 2**  
Results of Profile Fitting for Simulated Galaxies

Name	$M_*$ ( $10^{10} M_\odot$ )	$I_d$ ( $10^9 M_\odot \text{ kpc}^{-2}$ )	$R_d$ (kpc)	$I_b$ ( $10^9 M_\odot \text{ kpc}^{-2}$ )	$R_b$ (kpc)	$n$
m11f	2.8	0.1	3.7	0.8	1.0	0.9
m11g	5.4	0.5	2.5	3.3	0.7	0.6
m12b	16.1	0.7	3.3	7.4	0.8	0.7
m12c	10.1	1.5	1.7	11.0	0.6	0.5
m12q	18.7	1.1	3.0	7.7	1.0	0.6
m12z	4.6	0.1	6.8	0.3	1.0	0.6
m12m	14.3	0.2	6.3	2.3	1.3	1.2
m12f	9.5	0.1	6.5	2.2	1.2	0.8
m12i	7.2	0.2	4.2	2.0	1.0	0.9
Romeo	9.0	0.1	8.1	0.7	1.5	0.9
Juliet	6.9	0.1	7.1	0.6	1.6	1.2
Romulus	17.7	0.3	5.9	3.0	1.3	0.8
Remus	12.9	0.3	4.6	3.4	1.3	0.8
Thelma	14.8	0.3	5.4	3.6	1.0	0.7
Louise	8.2	0.2	5.3	2.3	1.0	0.7
Batman	14.1	2.5	1.6	23.9	0.5	0.6
Robin	8.9	0.2	5.2	1.9	1.1	0.7

**Notes.** See Equation (1) for explicit definitions of fit parameters.  $M_*$ : total stellar mass in frame of mock image (see Section 4).  $I_d$ : central surface density of disk component.  $R_d$ : scale radius of disk component.  $I_b$ : central surface density of bulge component.  $R_b$ : scale radius of bulge component.  $n$ : Sérsic index of bulge component.

**Table 3**  
Estimates of Stellar Halo Mass

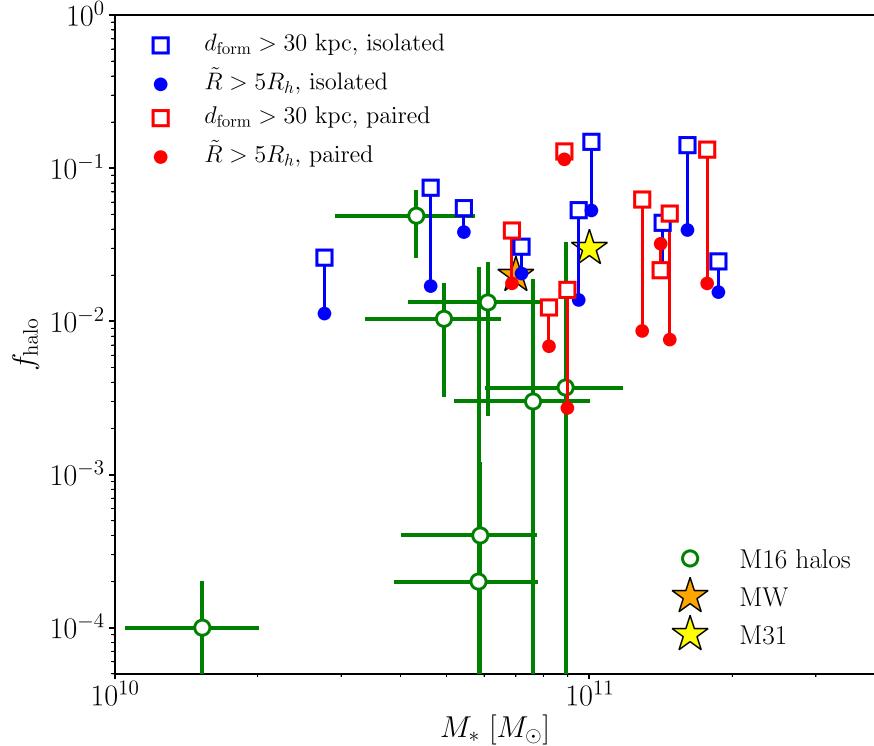
Name	$M_*$	$5R_d$ (kpc)	$M_*(d_{\text{form}} > 30 \text{ kpc})$	$M_*(\tilde{R} > 5R_d)$	$M_*(\Sigma < 10^6)$	$f_{\text{halo}}^{d_{\text{form}}}$	$f_{\text{halo}}^{5R_d}$	$f_{\text{halo}}^{\Sigma}$
m11f	2.8	18.5	0.07	0.04	0.06	0.026	0.011	0.021
m11g	5.4	12.7	0.30	0.22	0.11	0.055	0.038	0.020
m12b	16.1	16.3	2.28	0.66	0.19	0.142	0.040	0.012
m12c	10.1	8.7	1.51	0.56	0.15	0.149	0.053	0.015
m12q	18.7	14.9	0.46	0.31	0.16	0.025	0.016	0.009
m12z	4.6	34.2	0.35	0.09	0.14	0.075	0.017	0.029
m12m	14.3	31.6	0.63	0.32	0.16	0.044	0.021	0.011
m12f	9.5	32.4	0.51	0.14	0.16	0.053	0.015	0.017
m12i	7.2	20.9	0.22	0.16	0.16	0.031	0.021	0.022
Romeo	9.0	40.7	0.14	0.04	0.10	0.016	0.003	0.011
Juliet	6.9	35.6	0.27	0.13	0.11	0.039	0.018	0.017
Romulus	17.7	29.6	2.34	0.34	0.17	0.132	0.018	0.010
Remus	12.9	22.9	0.81	0.13	0.13	0.062	0.009	0.010
Thelma	14.8	26.8	0.75	0.14	0.13	0.051	0.008	0.009
Louise	8.2	26.6	0.10	0.07	0.07	0.012	0.007	0.009
Batman	14.1	8.0	0.31	0.47	0.17	0.022	0.032	0.012
Robin	8.9	26.0	1.14	1.02	0.11	0.129	0.114	0.013

**Notes.** All masses are in units of  $10^{10} M_\odot$ .  $M_*$ : total stellar mass in field of view (see Figure 3).  $5R_d$ : projected distance threshold used in M16 to distinguish disk and halo (see Section 4.2 and Table 2).  $M_*(d_{\text{form}} > 30 \text{ kpc})$  ( $f_{\text{halo}}^{d_{\text{form}}}$ ): mass (mass fraction) in stellar halo determined by selecting material with  $d_{\text{form}} > 30 \text{ kpc}$  (see Section 3).  $M_*(\tilde{R} > 5R_d)$  ( $f_{\text{halo}}^{5R_d}$ ): mass (mass fraction) in stellar halo determined by selecting pixels with ellipsoidal radius  $\tilde{R} > 5R_d$  (see Section 5.1).  $M_*(\Sigma < 10^6)$  ( $f_{\text{halo}}^{\Sigma}$ ): mass (mass fraction) in stellar halo determined by selecting pixels with  $\Sigma < 10^6 M_\odot \text{ kpc}^{-2}$  (see Section 5.3).

generally correlated with position) to define the accreted stellar halo, but we see in Figure 1 that while  $d_{\text{form}}$  and  $d_{\text{present}}$  are tightly correlated for most stars formed in situ, this is not the case for accreted material. Accreted stars can have large  $d_{\text{form}}$  but small  $d_{\text{present}}$  due to dynamical friction; adding to the confusion, some material formed in situ can be removed to large  $d_{\text{present}}$ , where a spatial cut on  $d_{\text{present}}$  would mistake it for accreted stars. The accreted material at small  $d_{\text{present}}$  could make up only a small fraction of the total mass at these distances but a large part of the total halo mass. In essence, therefore, we are asking whether any of these spatial selections are successful as a relatively unbiased proxy of the total mass in accreted stars.

### 5.1. Stellar Halo Mass Fraction Estimates from Profile Fitting

We calculate  $f_{\text{halo}}$  for each of our simulated galaxies in a manner analogous to M16 by summing the stellar mass in pixels with ellipsoidal radius  $\tilde{R} > 5R_d$ , subtracting the extrapolated mass of the disk and bulge in the fitted profiles, and dividing by the total stellar mass  $M_*$  within the field of view. The results of this calculation are shown in Table 3 and as filled symbols in Figure 4, along with the results from M16 and estimates for the MW (Carollo et al. 2010) and M31 (Courteau et al. 2011). To determine the accuracy of this method for estimating  $f_{\text{halo}}$ , we also calculated the fraction of stellar mass in each simulated galaxy within the field of view of the mass map that was formed more than 30 kpc from the



**Figure 4.** Stellar halo mass fractions  $f_{\text{halo}}$  calculated for the simulated galaxies in Table 1 (red and blue points) compared with the observed galaxies in M16 (green). Estimates from resolved stellar populations for the MW (Carollo et al. 2010; orange) and M31 (Courteau et al. 2011; yellow) are given for context. Filled red and blue symbols show estimates using fitted values of  $R_d$  given in Table 2 to distinguish accreted from in situ stars (the same method as in M16); open symbols show the actual accreted mass fraction in our simulations using  $d_{\text{form}} > 30$  kpc. Pairs of estimates for the same simulated galaxy are connected with vertical lines. Data for the simulations are given in Table 3. For simulated galaxies with similar total stellar mass to the observed sample, we obtain halo fractions that are statistically consistent with those measured by M16, with no systematic difference between paired and isolated halos. The true accreted fraction in our simulated galaxies is systematically higher, as seen in previous work.

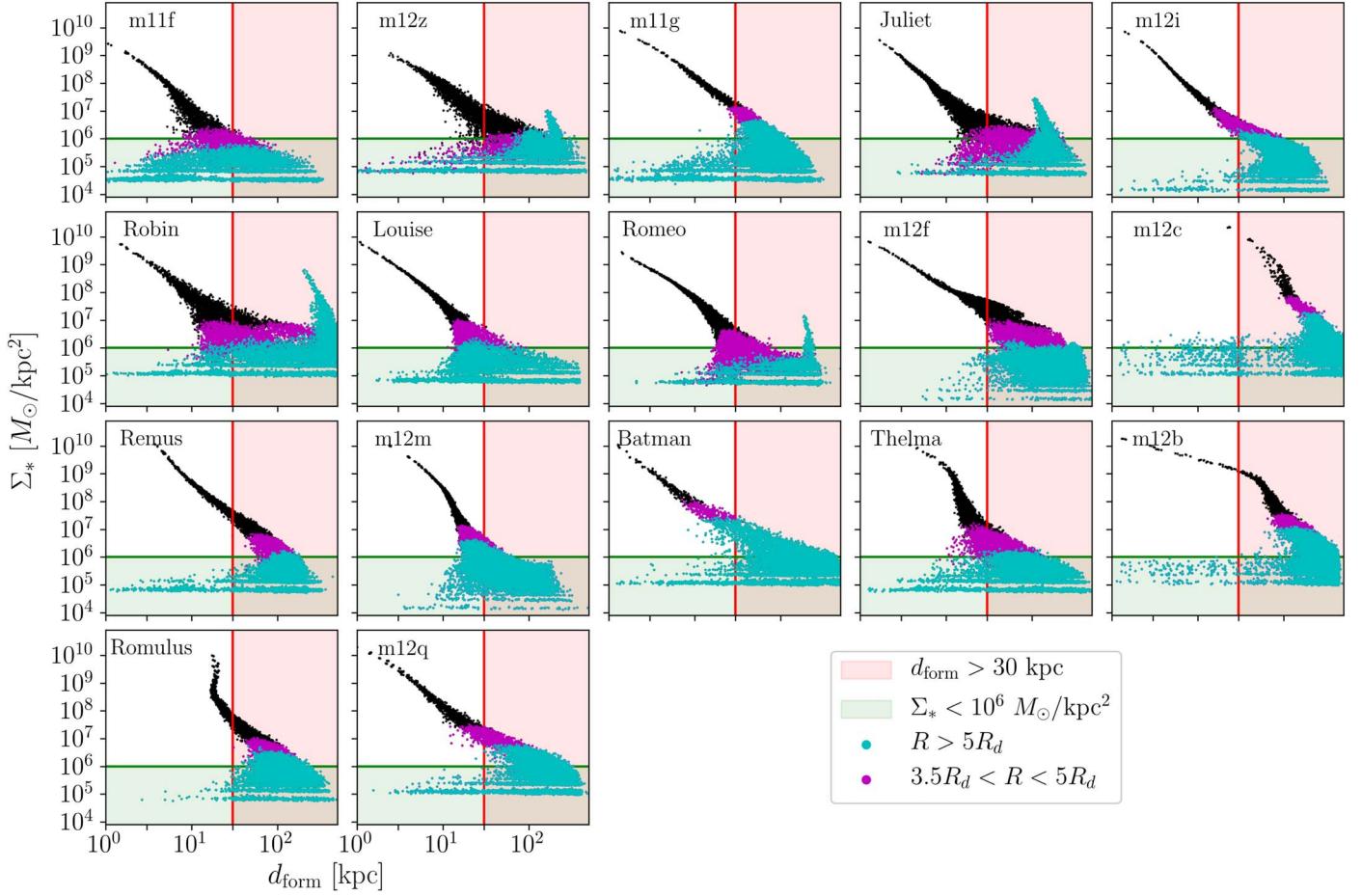
center of the galaxy. This quantity (i.e., the true accreted fraction) is also given in Table 3 and is plotted with open symbols in Figure 4. When we reproduce the observational method on our simulated galaxies, we obtain halo fractions that agree with those measured by M16, but the accreted mass fraction in our simulated galaxies is systematically higher than this.

Compared to the galaxies analyzed by M16, our simulated galaxies tend to have slightly higher total stellar mass. Remarkably, some of the simulated galaxies have stellar masses and  $f_{\text{halo}}$  values quite close to those estimated for the MW (m12i and Juliet) and M31 (m12f); these methods use resolved stellar populations and so, in principle, should not be subject to the same biases as the integrated-light estimates. The weak trend toward lower  $f_{\text{halo}}$  at lower stellar mass apparent in the M16 galaxies is also present in our simulations and appears to connect smoothly with their lower-mass galaxies in terms of “observed” stellar halo mass fraction (filled symbols). The large variation in  $f_{\text{halo}}$  at a given stellar mass is also apparent in our simulated galaxies, although our sample is too small to assess the full extent of the scatter. Galaxies simulated in pairs resembling the MW and M31 (red points) do not appear to have systematically different  $f_{\text{halo}}$  than those evolved in isolation; we expect this is because our mock observations probe a relatively small region around each galaxy (out to 50 kpc) relative to the separation of the pairs (about 1 Mpc).

The lowest mass fraction estimated in M16, which occurs for the lowest-mass galaxy in the sample, is still significantly below the range of stellar halo mass fractions in our suite of

simulations. However, we note that this galaxy (M101) has one of the most extended disks in the sample, which, as we will discuss shortly, corresponds to the most severe underestimation scenario in our simulated sample. We also expect the scatter in halo fractions to increase significantly at lower total stellar mass (Cooper et al. 2013; D’Souza et al. 2014; Pillepich et al. 2015): the building blocks of accreted stellar halos in galaxies of lower mass are themselves lower in mass and therefore have a larger scatter in stellar mass that may be at least partially responsible for this trend (e.g., Cooper et al. 2010). Although our sample is unbiased with regard to formation history (see Section 2), it is too small to understand whether M101 is consistent with the tails of the distribution; Elias et al. (2018) has explored this question using the Illustris simulations.

It is also evident from Figure 4 that the choice of  $5R_d$  as the dividing line between in situ and accreted populations underestimates  $f_{\text{halo}}$  in nearly every galaxy. The extremely steep falloff of most stellar halos with distance means that the estimate of the mass in the stellar halo is extremely sensitive to the radius at which this selection is made, and for most of our halos,  $5R_d$  is well outside the region where most of the stars have  $d_{\text{form}} > 30$  kpc. To illustrate this, in the bottom panel of Figure 3, we show the average formation distance of the star particles in each pixel of the simulated face-on images shown in the top panel, with a black circle indicating  $5R_d$  and a red circle indicating the half-mass radius of the stellar halo within 100 kpc (see the right-hand panel of Figure 2), superposed for each galaxy. It is clear that selecting material outside  $5R_d$  to represent the halo is a far better assumption for some galaxies



**Figure 5.** Distribution of stellar mass surface densities in individual pixels (black points) as a function of the mass-averaged formation distance  $d_{\text{form}}$  of star particles in each pixel. The red shaded region is dominated by “accreted stellar halo,” that is, pixels where stars have a mass-weighted median  $d_{\text{form}} > 30$  kpc. Magenta (cyan) points have  $\tilde{R} > 3.5R_d$  ( $\tilde{R} > 5R_d$ ). The green shaded region shows pixels with  $\Sigma_* < 10^6 M_\odot \text{ kpc}^{-2}$  (see Section 5.3). The typical surface density at the transition from stars formed in situ to accreted material ranges over three orders of magnitude, from  $10^6$  to  $10^9 M_\odot \text{ kpc}^{-2}$ . The wide variation in the relationship between surface density and stellar origin underlines the difficulty of separating accreted material using a surface density criterion.

than others. It is also clear that while for some galaxies, the accreted material is quite centrally concentrated, in others, it is much less so, with the half-mass radius well outside  $5R_d$ . In these cases (such as Juliet), one could indeed hope to identify a region that includes most of the accreted stellar mass while still excluding the in situ material, whereas observationally disentangling the majority of the stellar halo mass in galaxies like m12b may prove impossible.

Comparing the top and bottom panels of Figure 3, it is not clear immediately whether the morphology of the galaxy could be used to diagnose whether selection by projected radius is likely to give a good estimate of the halo-dominated region. Nevertheless, we can ask whether simply using a smaller characteristic radius can give a more consistent estimate of  $f_{\text{halo}}$ . We find that using  $\tilde{R} > 3.5R_d$  as the cutoff (plotted as orange pentagons in the right-hand panel of Figure 6) gives a relatively unbiased result across the limited mass range of stellar halos in our sample, with over- and underestimates all less than 0.5 dex.

Given the wide variety of distributions of accreted material apparent in Figure 3, it is not clear whether a superficial adjustment to the cutoff radius used to select the halo is really a complete solution to the problem. To understand better how using a radial selection may be biasing stellar halo mass estimates, and to investigate further whether a galaxy’s

morphology could be a clue to calibrate estimates of the stellar halo mass, we looked for correlations between the true or estimated stellar halo mass and three basic parameters: the total stellar mass of the galaxy, the extent of the disk (represented by  $R_d$ ), and the bulge-to-disk mass ratio  $M_b/M_d$  (computed by integrating the best-fit Sérsic profile to 5 effective radii for each component). The true accreted halo masses have no apparent dependence on the extent of the stellar disk, but as expected due to the steep falloff in the mass profile of most stellar halos, a larger stellar disk will induce a greater underestimate of the stellar halo mass when  $\tilde{R} > 5R_d$  is used as the cutoff. Because of the correlation between  $R_d$  and  $M_*$ , the bias from using  $\tilde{R} > 5R_d$  can artificially suppress the dependence of  $f_{\text{halo}}$  on  $M_*$ . Adopting  $\tilde{R} > 3.5R_d$  seems to mitigate this issue somewhat. We see no dependence of stellar halo mass on the Sérsic index of the bulge or the bulge-to-disk ratio.

Figure 5 plots the distribution of surface densities and average formation distances for all the pixels in the face-on galaxy images. The pixels selected by the wider radial cut  $5R_d$  (cyan points) often do not include the pixels with an average formation distance beyond 30 kpc (the red shaded region) at the highest stellar mass surface densities: in other words, any black point falling in the red shaded area represents stellar mass that should be counted as accreted stellar halo but is excluded by the radial selection. Changing to  $3.5R_d$  rather than  $5R_d$

(magenta points) appears to help by including primarily brighter pixels; in some cases, this also agrees better with the selection based on  $d_{\text{form}}$ , but in others, it includes material formed closer to the main galaxy than our criterion considers accreted (although many of the galaxies shown here and in Figure 1 have disks that do not extend all the way to  $d_{\text{form}} = 30$  kpc). Reassuringly, however, using a smaller disk exclusion region tends to incorporate material that was still formed at appreciably large distances from the main galaxy.

For many of the galaxies in our sample, there is a significant transition between a relatively well-correlated sequence in formation distance at high surface densities, driven in part by star formation *in situ*, to a more scattered picture at lower surface density. Material at small  $d_{\text{form}}$  is not part of the accreted stellar halo under our definition, but it nonetheless exists at large separation from the main galaxy and low stellar mass surface density in many cases and could fairly be considered part of the stellar halo by many definitions. However, this makes up a very small portion of the stellar halo mass (usually less than 1%) for most of our simulated galaxies.

Using a less conservative selection of  $\tilde{R} > 3.5R_d$  (shown as magenta points) seems to result in less bias over the entire sample, mainly because this selection appears to do a slightly better job getting most of the pixels with average  $d_{\text{form}} > 30$  kpc at the cost of occasionally including some material at lower  $d_{\text{form}}$ . Given the strong correlation between surface density and  $d_{\text{form}}$ , and the relatively conservative nature of this criterion to begin with, most of the overshoot is still material that was formed relatively far from the galaxy or is at very low surface densities, as discussed above. Using a smaller radius also means that the differences between fitted profiles are less exaggerated, especially given that some of the visible disks in the mock images of Figure 3 seem to fade out before  $5R_d$ , and that the bottom panel of Figure 3 shows that material that looked like part of an extended disk could actually be considered accreted for many of the galaxies with the largest  $R_d$ . Practically speaking, one could envision optimizing a selection criterion over a larger sample of simulated galaxies (especially spanning a wider range in total stellar mass), but Figure 5 underlines how the correspondence between stellar mass surface density and stellar origin differs so widely between galaxies that any selection based on projected radius should be used with caution.

## 5.2. Estimation Methods Used in Simulations

In simulations of galaxy halos, a spatial cut is sometimes used to try to separate accreted from *in situ* material, defined in terms of the three-dimensional rather than projected radius. In Figure 1, which shows three examples of cutoff radii used in other works ( $r_{\text{vir}}/50$ ,  $2r_*$ , 50, and  $r_{*,90}$ ; see Section 3), we showed that these choices tend to fall within the outskirts of the disk sequence for many of the FIRE simulated galaxies, which would indicate that these cuts are likely to overestimate the mass in the stellar halos by mistakenly including stars formed *in situ*. The left panel of Figure 6 illustrates that this is indeed the case: all of these selection criteria tend to overestimate the stellar mass in the halo, and the bias is worse for lower-mass halos. These results may change for simulations in which the stellar-to-dark matter halo mass ratio (and hence the distribution of  $r_{\text{vir}}$  at a given stellar mass) or the size distribution of galaxies at a given stellar mass differs substantially from our

sample, illustrating the importance of matching the observed distributions of these properties with samples of simulated galaxies.

## 5.3. Stellar Halo Mass Fraction Estimates Based on Surface Density Criteria

We also considered using a surface density criterion to define the stellar halo as an alternative to the profile-based method used in M16. To compare with simulations, M16 defined the simulated tagged stellar halos of the Aquarius simulations (Cooper et al. 2010) as the region where  $\Sigma_* < 10^6 M_\odot \text{ kpc}^{-2}$  (shown as the green shaded region in Figure 5). For most galaxies in our sample, this selection picks out the same approximate set of pixels as one of the selections on projected radius, with a few exceptions where it is more conservative. We apply this criterion to our simulated stellar mass surface density maps and compare it to the value obtained by selecting stars with  $d_{\text{form}} > 30$  kpc, as we did for the  $R_d$ -based estimate in Section 5.1. The yellow circles in the right-hand panel of Figure 6 show how this strategy corresponds to the mass in accreted stars. The values obtained are given in Table 3.

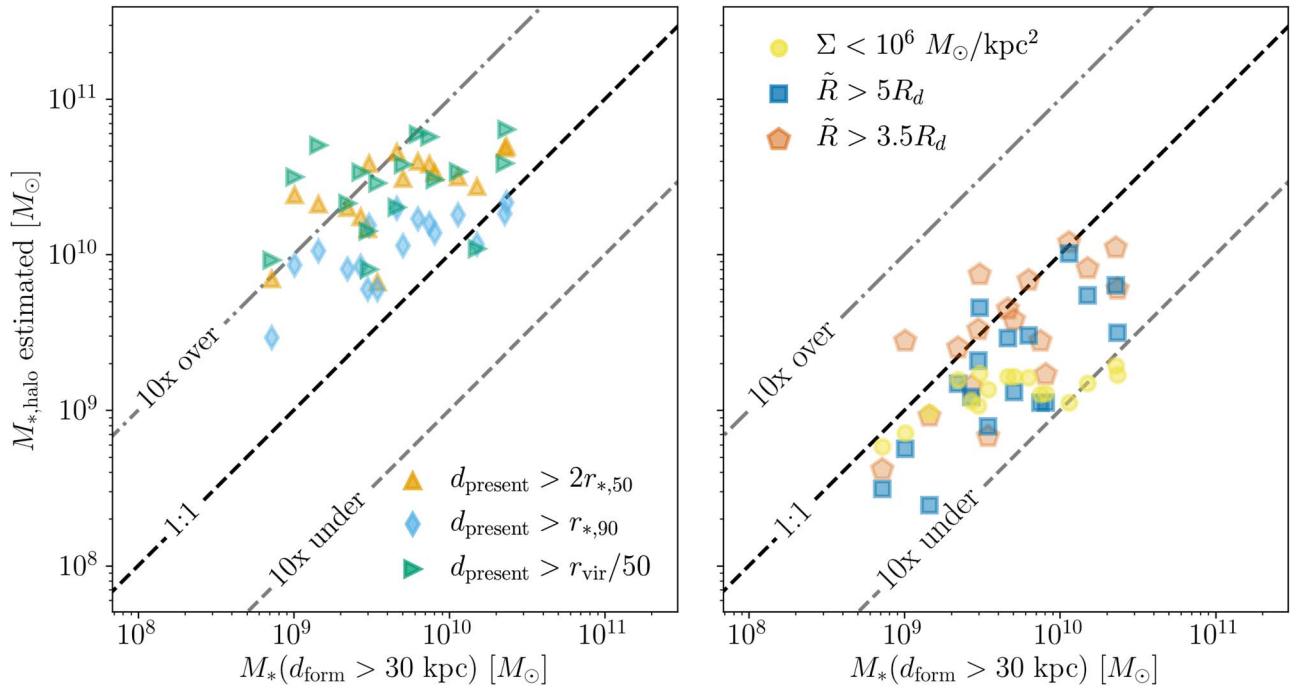
Although both of the observational methods shown in Figure 6 (excluding the criterion specifically calibrated on this data set) systematically underestimate the mass of the stellar halo, the  $R_d$ -based selection does better than the surface density criterion. Remarkably, although the masses of the stellar halos in our simulated galaxies (as defined by formation distance) range over more than an order of magnitude, the mass in pixels with surface densities less than  $10^6 M_\odot \text{ kpc}^{-2}$  is nearly constant over this entire range.

## 6. Improved Accreted Halo Mass Fractions

Given that the observational methods we have tried so far tend to underestimate the accreted halo mass, while the methods used in simulations tend to overestimate it, we now consider possible improvements in separating the accreted component from stars formed *in situ* using observational proxies. Given that a large part of the mass being missed is at relatively small radii, we first consider whether targeting edge-on rather than face-on galaxies helps mitigate confusion with the disk. Second, we consider whether there is a nonparametric way to determine which regions of a surface density map correspond to halo rather than disk stars by looking for inflection points in the cumulative distribution function (CDF) of the surface densities of the pixels. Third, we look at the distribution of metallicities in the field of view to determine whether this additional information could help determine which regions of a galaxy are more likely to be made up of accreted stars.

### 6.1. Edge-on Galaxies

Our simulated galaxies pose a different but equally frustrating set of problems for estimating  $f_{\text{halo}}$  when viewed edge-on, as illustrated in Figure 7. In most of the galaxies, there is a clear delineation, especially in the direction perpendicular to the disk, between material with  $d_{\text{form}} > 30$  and  $d_{\text{form}} < 30$  kpc. However, there is still a fairly wide variety of distributions of *in situ* material. In some cases, like m12i, there is substantial *in situ* material at higher latitudes above the disk plane (i.e., an extended thick disk) and in a spheroid, while in others, like m12b, there is virtually no region outside of the



**Figure 6.** Comparison of the extent to which different definitions of “stellar halo” correspond to the true accreted stellar mass (stars formed beyond 30 kpc from the main galaxy). Left: criteria commonly used to select stellar halos of simulated galaxies (Section 5.2). All stars with present-day distance  $d_{\text{present}} > x$ , for each of the listed thresholds  $x$ , are considered part of the stellar halo; this tends to overestimate the mass in accreted stars because the cuts include a significant amount of material formed in situ (see Figure 1). Right: definitions based on observed galaxy images (Sections 5.1–5.3), where all stars below a threshold surface density or with projected ellipsoidal radius  $\tilde{R}$  (Section 4.2) larger than some value are considered part of the stellar halo. This tends to underestimate the accreted stellar mass when applied to our simulated galaxies viewed face-on. Blue squares show the criterion used in M16, yellow circles show the criterion proposed in Cooper et al. (2013), and orange pentagons show a recalibrated version of the M16 criterion chosen to produce an unbiased estimate for this sample (see extended discussion in the text). The dashed lines in both panels show unbiased estimates (black) and a factor of 10 under- or overestimate (gray) for reference.

thin disk that is dominated by stars formed in situ. In still other cases (like Robin), an interaction with a massive satellite (apparent in the image) has pulled in situ material out of the disk plane, as has been recently predicted for the Sagittarius dwarf galaxy in the MW, which is of a similar mass (Laporte et al. 2016; Gómez et al. 2017a, 2017b). While in this particular case, it is easy to tell this is happening because the interacting satellite is clearly visible, such an effect is more difficult to pinpoint at later stages of such a merger, as is probably the case in m12c. In yet other cases (like m11f), there is little accreted material visible in this view at all. Therefore, although in general the in situ material appears to dominate a lower fraction of pixels in the edge-on case, the prospect of fitting a profile and choosing a threshold beyond which stars are part of the halo, as in M16, seems to be equally problematic here. On the other hand, the dominance of the accreted component in this view validates the strategy of the GHOSTS team (Harmsen et al. 2017) to sample along the minor axis of edge-on galaxies in order to optimize for the highest possible fraction of stars that were unambiguously formed outside the galaxy and then accreted, with the caveat that for small  $z \sim R$ , there is still a substantial fraction of in situ scattered material present in many of our simulated galaxies.

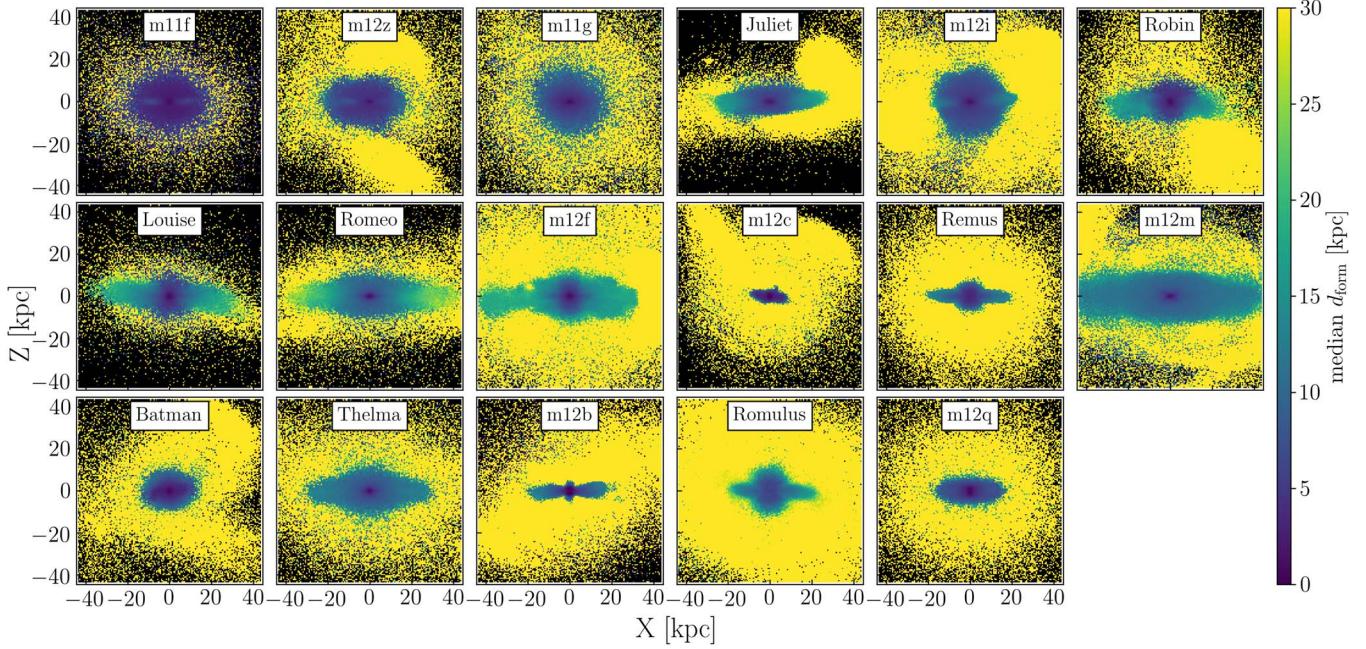
## 6.2. Stellar Mass Density Distributions

We next examine the full distribution of stellar mass densities in the pixels of each mock image, to explore (following Cooper et al. 2013, who studied more massive galaxies) whether an inflection point in this distribution will naturally differentiate between the concentrated stars in the disk

and bulge and the more diffuse halo. Figure 8 shows the CDF of the surface mass densities in the pixels of each mock image for both face-on (left) and edge-on (right) views. The line color is proportional to the mass-weighted average formation distance for material in each pixel. Unfortunately, no obvious criterion, such as a characteristic change in the CDF slope, appears to consistently delineate pixels dominated by accreted or in situ material: some galaxies are dominated by in situ material down to very low surface mass densities, while others are dominated by accreted material to surface densities as high as  $10^7 M_\odot \text{ kpc}^{-2}$ . Some CDFs do appear to exhibit an inflection point around the transition between in situ and accreted material, but most do not. This is true whether the galaxies are viewed edge-on or face-on and suggests not only that a single surface mass density cut is not suitable but that the surface density tends to smoothly connect between pixels containing mostly accreted halo and those containing mostly in situ material, with no “break” or other obvious feature in the light profile or CDF.

## 6.3. Metallicity

We next consider whether spatial variations in metallicity can indicate the region where most of the stars are accreted, inspired by works like Tissera et al. (2013), D’Souza et al. (2014), Harmsen et al. (2017), and D’Souza & Bell (2018). For this analysis, we consider only simulations from our suite that include subgrid metal diffusion, to reduce artificial numerical noise in the distribution of [Fe/H]. This subset has similar properties to the full sample (see Appendix A). The range of metallicities in the halos of the different simulated galaxies is



**Figure 7.** Edge-on images of mass-weighted median  $d_{\text{form}}$ , as in Figure 3. In some galaxies (m12i), an in situ component extends well above the thin disk plane, while in other cases (m12b), in situ material is closely confined to the thin disk plane. In other galaxies (Robin, m12f), interactions with massive satellites warp the in situ component; in later stages of mergers, the accreted material dominates the in situ component entirely (m12c). Although the disk generally subtends less area than in a face-on orientation, the great variety of accretion histories still complicates the determination of halo stellar mass in edge-on galaxies.

substantial, reflecting the wide variety of accretion histories in the sample: some (like m12i) have very well-defined, metal-poor halo regions, while in others (like m12f), the halo is much closer in metallicity to the disk.

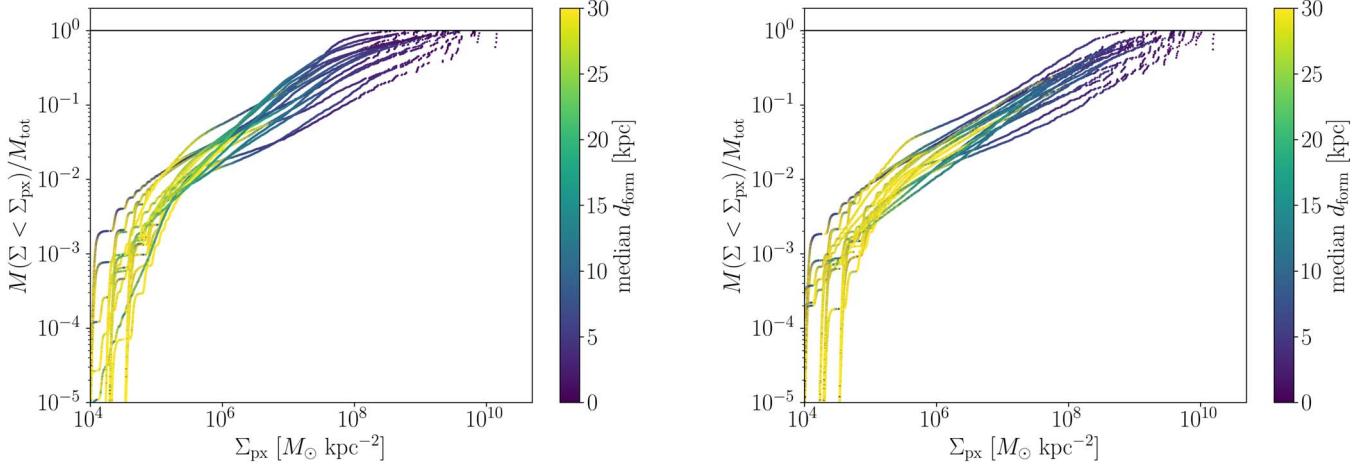
Figure 9 suggests that the metallicity distribution can indeed be diagnostic, since in general the disk (i.e., in situ) regions have higher metallicity than the accreted outskirts regardless of their absolute metallicity, which varies substantially. Structures clearly due to mergers stand out as higher-metallicity regions on a lower-metallicity background. Information on metallicity is therefore a valuable complement in understanding which regions of a galaxy were accreted (see also Bonaca et al. 2017). The obvious next step of testing whether photometric metallicities are sufficient would require modeling the SEDs of the simulated galaxies, a task we defer to future work.

## 7. Conclusions

In this paper, we compared the stellar halos of high-resolution simulated galaxies from the FIRE-2 suite (Sections 2–3) with recent measurements by M16 of stellar halo mass fractions for eight MW-like galaxies. Because of the high resolution of these simulations, and thanks to the clear and detailed description in M16 of their analysis, we were able to reproduce the same steps they used to measure the stellar halo mass fractions of our simulated galaxies (Sections 4.1–4.2) and found that these had similar magnitude and scatter to the observations (Section 5.1).

In the simulations, we can record whether each star particle was formed in situ or accreted. Thus, we can consider how well the stellar halo mass estimated from deep images agrees with the mass of accreted stars, an important theoretical quantity that is not directly accessible observationally. Inspired by M16 and other recent works, we considered both spatial selection based on modeling and subtracting the disk and bulge regions (Section 5.1) and a selection based on surface mass density as

proposed in Cooper et al. (2010; Section 5.3). We found that these methods can underestimate the accreted stellar mass, usually by a factor of 2–3 and up to a factor of 10. In selections based on disk fitting and subtraction, this can introduce spurious correlations into the mass fraction that obscure the true dependence of the accreted mass fraction on stellar mass. Although Sérsic profiles can be notoriously degenerate, we found that determining the disk scale radius  $R_d$  was fairly robust and not the main source of trouble. Instead, the main issues with using  $R_d$  to select a region of a galaxy containing mostly accreted stars are the confusion of accreted and in situ material in the outskirts of the disk and the obscuration of accreted material by the disk at small projected radii. The contribution of each of these two complications varies substantially from galaxy to galaxy based on the individual accretion history and the morphology of the central galaxy. Using edge-on rather than face-on systems (Section 6.1) helps in the sense that the disk obscures less of the halo, but it does not address the problems posed by the variety of accretion histories in selecting the accretion-dominated region. Because of the wide variety of accretion histories spanned by our set of 17 simulated galaxies, there is no characteristic surface mass density or inflection point in the distribution of stellar mass surface densities that could help systematically diagnose the presence of mostly accreted material (Section 6.2). Given the narrow mass range (less than an order of magnitude) of this set of simulations, we emphasize that this significant variation in halo properties (over three orders of magnitude in surface mass density) is not simply an effect of mass dependence but reflects the wide variety of accretion histories across galaxies of similar masses. If only integrated-light images are available, we suggest that examining the spatial variation in the metallicity of the galaxy, accessed perhaps through colors, is a promising approach to identifying the accreted component using minimal additional information, since the simulated galaxies show a



**Figure 8.** The CDF of the surface mass densities in the pixels of each mock image for both face-on (left) and edge-on (right) views of the same sample of galaxies. The appearance of step features at low surface density indicates the resolution limit (pixels containing only one or a few particles). The surface density corresponding to the transition to an accreted component varies substantially from galaxy to galaxy, with no obvious transition in the CDF corresponding to a transition in the origin of the material.

strong contrast in metallicity between the accreted stars and those formed in situ (Section 6.3). Quantifying this relationship will be the subject of future work.

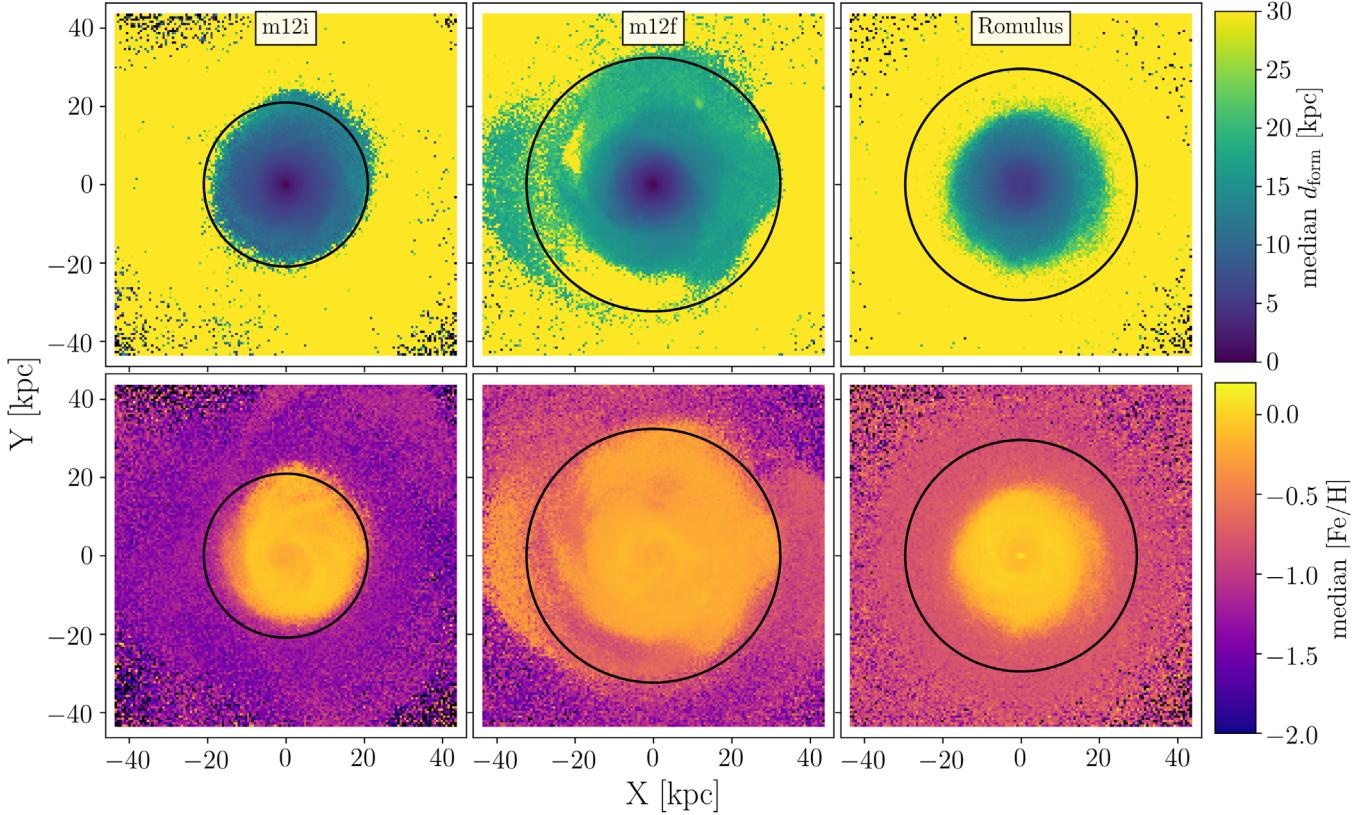
Given the limitations of any suite of simulated galaxies, these conclusions come with a few caveats. Even in our highest-resolution simulations, we can only satisfactorily resolve star formation in satellite galaxies down to the mass scale of the classical dwarfs ( $M_{\text{halo}} \gtrsim 10^9 M_{\odot}$ ,  $M_{\text{star}} \gtrsim 10^5 M_{\odot}$ ; Wetzel et al. 2016), so we are still missing some of the mass in the accreted stars, but we expect this fraction to be small given that the mass-to-light ratio of galaxies increases so sharply below this level (e.g., Moster et al. 2010; Behroozi et al. 2013; Brook et al. 2014; Garrison-Kimmel et al. 2014b, 2017a) and the total stellar halo masses are  $10^9 M_{\odot}$  or above for all the galaxies. On the other hand, unlike Pillepich et al. (2014), we do not remove stars in bound subhalos, so this likely raises the mass fractions somewhat for a few of our simulated galaxies, notably those where a companion is clearly visible. There are not too many of these cases in our suite, which is reasonable given that the field of view shows the inner  $\sim 50$  kpc where most structures will have been tidally disrupted (Garrison-Kimmel et al. 2017b). M16 did not attempt to fit and remove any satellites in their measurements, but neither are any companion galaxies clearly apparent in their images. The main point is to maintain consistency in the treatment of satellites between analyses of the observations and simulations, so where some of our simulated galaxies have companions, we consider this part of the scatter.

Most of our simulated galaxies have somewhat higher stellar masses than the bulk of the observed sample in M16, a limitation shared by many of the other simulations to which those observations were compared (most of which, like many of our simulated galaxies, were originally matched to the MW's properties). The most massive galaxy in the M16 sample has a stellar mass of  $9 \times 10^{10} M_{\odot}$ , which is less massive than half the simulated galaxies we consider, while their lowest-mass galaxy, at about  $1.5 \times 10^{10} M_{\odot}$ , is only half as massive as our least massive simulated galaxy (see Figure 4). The simulated systems in the lower end of the mass range differ quite widely in terms of their present-day appearance (bulge- or disk-dominated) and accretion histories, and clearly a larger sample than ours is needed to get a real sense of the scatter

inherent at this mass scale. Work with the Illustris simulations (Pillepich et al. 2014; Elias et al. 2018) is complementary to this study: thanks to the necessary trade-off between simulation box size and resolution, such work can more systematically explore the scatter and mass dependence of stellar halo masses, but the star particle masses are too high to produce images, like those in Figure 3, with resolved structure at the surface densities reported in M16.

Finally, in determining whether stars were formed in situ or accreted, we used a constant cutoff in the distance from the main galaxy where each star particle was formed, as in Bonaca et al. (2017), rather than taking a more detailed particle-tracking approach, as in previous studies (Font et al. 2011; Pillepich et al. 2015; Rodriguez-Gomez et al. 2016; Anglés-Alcázar et al. 2017). Some star particles may therefore be mistaken for accreted rather than in situ or vice versa if they came in very early and/or formed during the infall of a gas-rich galaxy. This includes only a small fraction of stars for most of our galaxies, and in some cases, it is genuinely debatable what should be considered accreted or formed in situ; our method does not allow for any nuance in this area. An example in Figure 1 is the infalling satellite in Robin, in the lower left-hand corner, where star formation clearly proceeded along with the merger. Our simplistic 30 kpc distance cutoff is clearly not the right choice in this case, but one could argue for interpreting those stars either as formed in situ (since the stars were formed when the satellite galaxy was clearly within and influenced by the halo of the main galaxy) or accreted (since they formed in a significantly different environment than the bulk of the stars in the main galaxy). However, given that we are comparing mainly total masses, and given the relative insensitivity of our results to the exact value of  $d_{\text{form}}$  that distinguishes accreted stars, we consider that this approximation is likely sufficient for the present study. In future work, we plan to extend the current analysis by using particle tracking to take into account the full time history of the star particles that end up in the outskirts of our simulated galaxies.

Our results underline the importance of parallel analysis of observations and simulations of galaxies as the way forward in robustly comparing the two to construct tests of cosmological models of galaxy formation. In this study, such an approach revealed the difficulty in choosing a single scale or scales for



**Figure 9.** Median formation distance (top) and median metallicity (bottom) for three simulated galaxies viewed face-on. Black pixels contain no star particles. As in Figure 3, the black circles indicate  $5R_d$ . There is a clear contrast between accreted and in situ material apparent in the metallicity map, which suggests that it may be a useful diagnostic for distinguishing between in situ and accreted components in MW-mass galaxies.

apertures around galaxies that are most sensitive to a given stellar population of theoretical interest (i.e., accreted or formed in situ), at least for galaxies in the mass range of the MW (and hence the targets observed in M16). For higher-mass galaxies, this may be a more viable solution: Huang et al. (2018) explored the use of physically constant-size apertures (10 and 100 kpc) around massive elliptical galaxies in a related strategy, based on indications from simulations of elliptical galaxies in, e.g., Rodriguez-Gomez et al. (2016). However, such a method is unlikely to be viable for lower-mass galaxies, where morphologies and sizes differ quite widely: as seen in Table 1, the radius enclosing 90% of the stellar mass in our sample varies over an order of magnitude even for our relatively narrow mass range due in part to the variety of formation histories (Garrison-Kimmel et al. 2018), so choosing a single set of physical apertures is likely to yield even worse results than scaling the aperture to a fitted disk scale length. Simultaneous analysis of real and simulated observations of stellar halos can thus also indicate which proxies are most appropriate for separating accreted stars from those formed in situ at different mass scales, since galaxies certainly vary widely with stellar mass. These valuable insights, enabled by the match between state-of-the-art observation and simulation techniques, are best gained by the type of in-depth conversation between simulators and observers that inspired this work.

The authors thank Laura Sales, Lydia Elias, and Kyle Stewart for helpful suggestions during the GalFRESCA 2017 meeting, organized by Shea Garrison-Kimmel and Coral

Wheeler, and Alison Merritt for helpful background on her measurement techniques.

R.E.S. was supported by an NSF Astronomy & Astrophysics Postdoctoral Fellowship under grant AST-1400989.

Support for S.G.K. was provided by NASA through Einstein Postdoctoral Fellowship grant number PF5-160136 awarded by the *Chandra* X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060.

A.W. was supported by a Caltech-Carnegie Fellowship, in part through the Moore Center for Theoretical Cosmology and Physics at Caltech, and by NASA through grants *HST*-GO-14734 and *HST*-AR-15057 from STScI.

T.K.C. was supported by NSF grant AST-1412153.

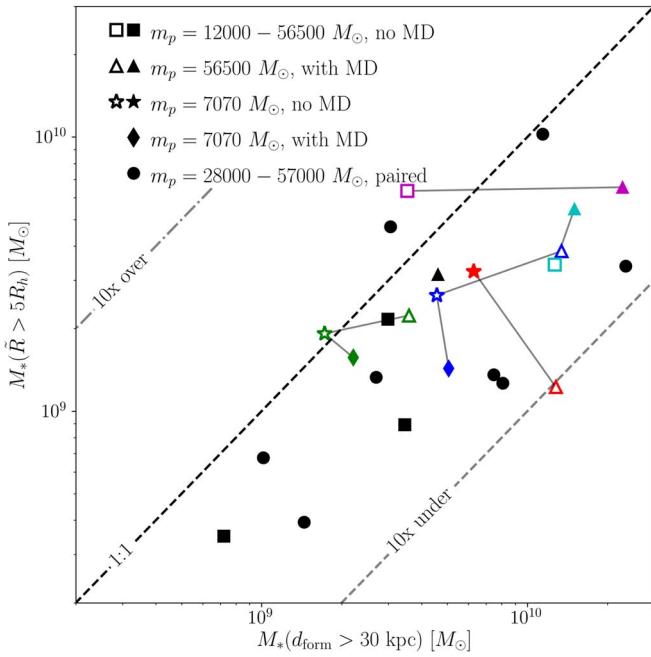
Support for P.F.H. was provided by an Alfred P. Sloan Research Fellowship, NSF Collaborative Research Grant #1715847, and CAREER grant #1455342.

D.K. acknowledges support from NSF grants AST-1412153 and AST-1715101 and the Cottrell Scholar Award from the Research Corporation for Science Advancement.

E.E. was supported by a Ford Foundation Predoctoral Fellowship.

C.-A.F.-G. was supported by the NSF through grants AST-1412836, AST-1517491 and AST-1715216 and CAREER award AST-1652522 and by NASA through grant NNX15AB22G.

Numerical calculations were run on the Caltech computer cluster “Wheeler,” allocations from XSEDE TG-AST130039 and PRAC NSF.1713353 supported by the NSF, NASA HEC



**Figure 10.** Estimated mass (as in the right-hand panel of Figure 6 but using the  $\tilde{R} > 5R_h$  criterion only) relative to the “true” accreted stellar mass (formed beyond 30 kpc) for the simulated galaxies in Table 1 (filled symbols) and companion simulations with the same initial conditions but different resolution and/or physics (open symbols). Points in the same color have the same initial conditions and are connected by gray lines. In the legend, “MD” indicates whether a numerical implementation of subgrid turbulent metal diffusion is included in the simulations. Different resolutions and physics produce no significant trends in the relative accuracy of estimates of the accreted stellar mass using this criterion.

SMD-16-7592, and the High Performance Computing at Los Alamos National Labs.

## Appendix A Numerical Effects

Figure 10, similar to the right-hand panel of Figure 6, shows the estimated mass using the  $\tilde{R} > 5R_h$  criterion relative to the mass formed beyond 30 kpc for different resimulations of the galaxies in Table 1 that use different resolution and/or subgrid physics. The filled symbols are the set used in the main paper and listed in the table. As was pointed out in Hopkins et al. (2018), higher-resolution simulations (stars and diamonds) tend to have slightly lower total stellar masses than lower-resolution versions (squares and triangles) with the same initial conditions, and this also appears to be true for the mass in accreted stars. However, this trend is quite noisy and probably complicated by differences in the treatment of turbulent metal diffusion between runs at different resolution. To illustrate, we consider the two sets of simulations shown in the figure, m12i and m12f, that have been run with resolutions lower by a factor of 8 relative to the fiducial run with an otherwise identical setup, permitting a controlled comparison. In one case, lowering the resolution by this factor increases the mass formed beyond 30 kpc by 60%, while in the other, it is lower by 40% relative to the fiducial run used in our analysis. Resolution does not appear to significantly affect the degree to which the selection criterion over- or underestimates the mass in accreted stars relative to its true value.

Simulations including subgrid turbulent metal diffusion (triangles and diamonds) look to perhaps have slightly higher

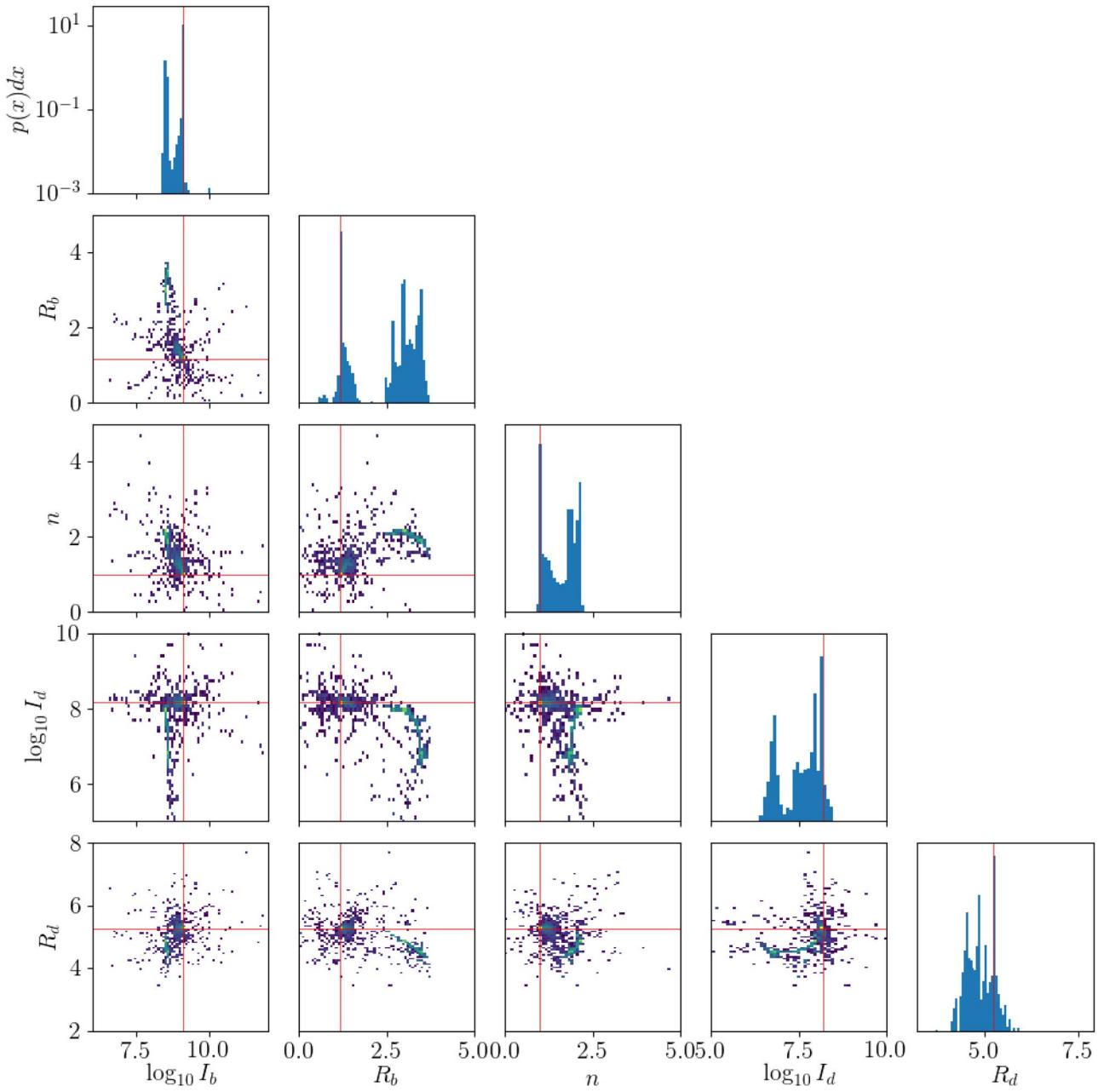
accreted masses than their counterparts without this additional physics (squares and stars). Comparing different versions of m12i and m12f again, but now selecting those that have been run at the same (highest) resolution while varying only the treatment of metal diffusion, the mass formed beyond 30 kpc is 22% and 10% lower in the runs without metal diffusion relative to fiducial, respectively. Varying the diffusion prescription likewise shows no trend in terms of estimator accuracy, so we can safely assume that the results of the subset discussed in Section 6.3 extend to our entire sample.

Finally, we find that paired and isolated halos are also similarly distributed within the scatter (which is again likely to be partially attributable to the variation in resolution and metal diffusion treatments, as discussed above). Based on these findings, we do not expect that our results are highly affected by numerics.

## Appendix B Monte Carlo Sérsic Fit Test

In order to understand how degeneracies between the five Sérsic parameters could affect the fit results, we carried out a more complete five-dimensional exploration of parameter space using the affine-invariant sampler `emcee`, starting from a broad distribution in a ball around the best fit reported by the local Levenberg–Marquardt minimization. We used 50 walkers to sample the parameter space 10,000 times per walker, discarding the first 200 steps per walker as burn-in. The distribution of the remaining samples is shown in Figure 11. There are often multiple peaks in most projections of the parameter space, but one peak (at the intersection of the red lines) is much higher and narrower than the other and displays less degeneracy between parameters. It is also clear that the more-degenerate region spans to much less sensible values of some of the parameters, notably the disk normalization  $I_d$  and the bulge scale radius  $R_b$ . The normalization of the disk profile is of special concern, since it could potentially bias the disk scale radius  $R_d$ , but in fact this parameter is relatively robust over a broad range of values, which gives us some confidence that a more localized minimization algorithm will not obtain wildly different values for  $R_d$  depending on which local minimum it ends up in, even if the other parameters are more strongly affected.

The taller peak is extremely narrow and not very degenerate between different parameters, except for in the extreme tails of the  $I_d$ – $R_d$  projection. The narrowness is probably due in part to the lack of a proper treatment of errors in the least-squares likelihood, where we simply use the square root of the mass surface density. However, the well-confined nature of this peak shows that it is a robust solution to minimizing least-squares differences, and its relative prominence and symmetrical shape suggest that one can tell when a fit by a local minimization algorithm has obtained it by starting with a good guess, considering the results of some perturbations to the initial location of the best fit, and making sure that the normalizations to the two components have reasonable best-fit values. Additionally, the most important parameter for this analysis, the disk scale length  $R_d$ , is not very susceptible to variations in the other best-fit parameters. We therefore use the results from the Levenberg–Marquardt minimization, shown in Table 2, with confidence.



**Figure 11.** MCMC exploration of the five-dimensional parameter space of the double Sérsic profile for the simulated galaxy m12i. Details of the sampler setup are given in Section 4.2. The two-dimensional projections show the binned samples on a logarithmic color scale from low (blue) to high (yellow) density. The histograms show one-dimensional projections on a logarithmic scale with identical limits on the y axis. The median of all the samples along each parameter is marked by red lines. The distribution of samples is multivalued in most projections, with the most probable of the peaks also displaying the least degeneracy.

### ORCID iDs

- Robyn E. Sanderson  <https://orcid.org/0000-0003-3939-3297>  
 Shea Garrison-Kimmel  <https://orcid.org/0000-0002-4655-8128>  
 Andrew Wetzel  <https://orcid.org/0000-0003-0603-8942>  
 Tsang Keung Chan  <https://orcid.org/0000-0003-2544-054X>  
 Philip F. Hopkins  <https://orcid.org/0000-0003-3729-1684>  
 Dušan Kereš  <https://orcid.org/0000-0002-1666-7067>  
 Claude-André Faucher-Giguère  <https://orcid.org/0000-0002-4900-6628>  
 Xiangcheng Ma  <https://orcid.org/0000-0001-8091-2349>

### References

- Amorisco, N. C. 2017, *MNRAS*, **469**, L48  
 Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, *MNRAS*, **470**, 4698  
 Bakos, J., & Trujillo, I. 2012, *ApJ*, submitted (arXiv:1204.3082)  
 Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJL*, **762**, L31  
 Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Kereš, D. 2017, *ApJ*, **845**, 101  
 Brook, C. B., Di Cintio, A., Knebe, A., et al. 2014, *ApJL*, **784**, L14  
 Bryan, G. L., & Norman, M. L. 1998, *ApJ*, **495**, 80  
 Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, **635**, 931  
 Carollo, D., Beers, T. C., Chiba, M., et al. 2010, *ApJ*, **712**, 692  
 Chan, T. K., Kereš, D., Wetzel, A., et al. 2018, *MNRAS*, **478**, 906  
 Cockercroft, R., McConnachie, A. W., Harris, W. E., et al. 2013, *MNRAS*, **428**, 1248

- Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, *MNRAS*, **406**, 744
- Cooper, A. P., D’Souza, R., Kauffmann, G., et al. 2013, *MNRAS*, **434**, 3348
- Courteau, S., Widrow, L. M., McDonald, M., et al. 2011, *ApJ*, **739**, 20
- D’Souza, R., & Bell, E. 2018, *MNRAS*, **474**, 5300
- D’Souza, R., Kauffmann, G., Wang, J., & Vegetti, S. 2014, *MNRAS*, **443**, 1433
- De Lucia, G., & Helmi, A. 2008, *MNRAS*, **391**, 14
- de Jong, R. S. 2008, *MNRAS*, **388**, 1521
- Duc, P.-A., Cuillandre, J.-C., Karabal, E., et al. 2015, *MNRAS*, **446**, 120
- El-Badry, K., Quataert, E., Wetzel, A., et al. 2018, *MNRAS*, **473**, 1930
- Elias, L. M., Sales, L. V., Creasey, P., et al. 2018, *MNRAS*, **479**, 4004
- Escala, I., Wetzel, A., Kirby, E. N., et al. 2018, *MNRAS*, **474**, 2194
- Font, A. S., Johnston, K. V., Ferguson, A. M. N., et al. 2008, *ApJ*, **673**, 215
- Font, A. S., McCarthy, I. G., Crain, R. A., et al. 2011, *MNRAS*, **416**, 2802
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Kirby, E. N. 2014b, *MNRAS*, **444**, 222
- Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Lee, K. 2014a, *MNRAS*, **438**, 2578
- Garrison-Kimmel, S., Bullock, J. S., Boylan-Kolchin, M., & Bardwell, E. 2017a, *MNRAS*, **464**, 3108
- Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2018, *MNRAS*, **481**, 4133
- Garrison-Kimmel, S., Wetzel, A., Bullock, J. S., et al. 2017b, *MNRAS*, **471**, 1709
- Gómez, F. A., Grand, R. J. J., Monachesi, A., et al. 2017a, *MNRAS*, **472**, 3722
- Gómez, F. A., White, S. D. M., Grand, R. J. J., et al. 2017b, *MNRAS*, **465**, 3446
- Greggio, L., Rejkuba, M., Gonzalez, O. A., et al. 2014, *A&A*, **562**, A73
- Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2017, *MNRAS*, **469**, 2292
- Harmsen, B., Monachesi, A., Bell, E. F., et al. 2017, *MNRAS*, **466**, 1491
- Hopkins, P. F. 2015, *MNRAS*, **450**, 53
- Hopkins, P. F. 2016, *MNRAS*, **455**, 89
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, *MNRAS*, **445**, 581
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, **480**, 800
- Huang, S., Leauthaud, A., Greene, J. E., et al. 2018, *MNRAS*, **475**, 3348
- Kim, J.-h., Abel, T., Agertz, O., et al. 2014, *ApJS*, **210**, 14
- Laporte, C. F. P., Gómez, F. A., Besla, G., Johnston, K. V., & Garavito-Camargo, N. 2016, *MNRAS*, submitted (arXiv:1608.04743)
- Ma, X., Hopkins, P. F., Wetzel, A. R., et al. 2017, *MNRAS*, **467**, 2430
- Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, *AJ*, **140**, 962
- Merritt, A., van Dokkum, P., Abraham, R., & Zhang, J. 2016, *ApJ*, **830**, 62
- Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, *ApJ*, **710**, 903
- Pillepich, A., Madau, P., & Mayer, L. 2015, *ApJ*, **799**, 184
- Pillepich, A., Vogelsberger, M., Deason, A., et al. 2014, *MNRAS*, **444**, 237
- Purcell, C. W., Bullock, J. S., & Zentner, A. R. 2007, *ApJ*, **666**, 20
- Rodríguez-Gómez, V., Pillepich, A., Sales, L. V., et al. 2016, *MNRAS*, **458**, 2371
- Sandin, C. 2014, *A&A*, **567**, A97
- Seth, A., de Jong, R., Dalcanton, J. & GHOSTS Team 2007, in IAU Symp. Ser. 241, Stellar Populations as Building Blocks of Galaxies, ed. A. Vazdekis & R. Peletier (Cambridge: Cambridge Univ. Press), 523
- Slater, C. T., Harding, P., & Mihos, J. C. 2009, *PASP*, **121**, 1267
- Su, K.-Y., Hopkins, P. F., Hayward, C. C., et al. 2017, *MNRAS*, **471**, 144
- Tissera, P. B., Beers, T. C., Carollo, D., & Scannapieco, C. 2014, *MNRAS*, **439**, 3128
- Tissera, P. B., Scannapieco, C., Beers, T. C., & Carollo, D. 2013, *MNRAS*, **432**, 3391
- Tissera, P. B., White, S. D. M., & Scannapieco, C. 2012, *MNRAS*, **420**, 255
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, **827**, L23