Probing Local Group Galactic Substructure with Cosmological Simulations

by

Gregory Dooley

A.B., Princeton University (2011)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2017

Author . Signature redacted .....
Department of Physics
January 13, 2017

Certified by . Signature redacted .....
Anna Frebel
Professor of Physics
Thesis Supervisor

Accepted by Signature redacted .....
Nergis Mavalvala
Professor of Physics
Associate Head, Department of Physics

Probing Local Group Galactic Substructure with Cosmological Simulations

by

Gregory Dooley

Submitted to the Department of Physics
on January 13, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The Lambda cold dark matter (ΛCDM) model is enormously successful at predicting large scale structure in the Universe. However, some tensions still remain on small scales, specifically regarding observed satellites of the Milky Way (MW) and Andromeda. Foremost among the problems have been the missing satellite, too big to fail, and cusp/core problems, which concern the expected abundance of satellites and their inner structure. This Ph.D. thesis consists of a series of studies using dark matter only cosmological N-body simulations of MW-mass galaxies to address topics related to these issues.

In light of the recent *Planck* mission, I investigate how changes to cosmological parameters affect dark matter halo substructure. I find that the process of continuous subhalo accretion and destruction leads to a steady state description of most subhalo properties in a given host, unchanged by small fluctuations in cosmological parameters. Subhalo concentration, maximum circular velocity, and formation times, however, are somewhat affected. One way to reduce the central density of satellites, as needed to solve the cusp/core and too big to fail problems, is through self-interacting dark matter (SIDM). I search for new implications of SIDM and find that stars in satellites spread out to larger radii and are tidally stripped at a higher rate in SIDM than CDM, even though the mass loss rate of dark matter is unchanged. These signatures should be particularly prominent in ultrafaint dwarf galaxies for the class of otherwise difficult to constrain velocity-dependent SIDM models.

I also helped carry out the *Caterpillar* project, a suite of 36 high mass resolution ($\sim 10^4\,\mathrm{M}_\odot$/particle) simulations of MW-like galaxies used to study diversity in halo substructure. To these, I apply abundance matching and reionization models to make novel predictions about the abundance of satellites in isolated dwarf galaxies out to 8 Mpc to help guide future searches. Applying the same techniques to predict satellites within 50 kpc of the LMC, I discover large discrepancies with the observed stellar mass function, which may lead to new constraints on the galaxy stellar mass-halo mass relationship, and the ability of reionization to leave dark matter halos entirely dark.

Thesis Supervisor: Anna Frebel
Title: Professor of Physics

Contents

List of Figures

List of Tables

Acknowledgments

I would first like to give a big thank to my advisor Anna Frebel. I met her as a first year graduate student unsure about the research project I had just begun. After a meeting where she showed me cosmological simulation videos and explained her research goals, I was hooked on the topic and very grateful that she was soon to be starting at MIT. Five years later I am still very grateful that she took me on as a student and set me on a path of rewarding research. I learned a tremendous amount from her in graduate school, from pure science, to time management, to strategic decision making, to presenting work effectively, and writing and editing scientific papers. She knew when to connect me with other collaborators, when to put the pressure on, and how to adjust when I invariably missed self-imposed deadlines. She also built a close-knit research group which made my experience at MIT both more productive and more personally rewarding. I look forward to hearing about the work of future generations of graduate students, and little Philip's growth amid the unique and special world of MIT.

I would next like to give a big thank you to Annika Peter, with whom I have collaborated for the past three years and on three of my four primary papers. While under no obligation to meet with me so frequently and consistently, Annika always found a way to have weekly Skype meetings with me discussing my latest progress, whether big or small. These meetings were invaluable to my progress, as I was able to vet ideas with her and take suggestions on the next path forward. I greatly enjoyed and benefited from being able to have technical discussions, and to catch up on weekly life. I would also like to thank her again for orchestrating a productive visit to Ohio State.

Mark Vogelsberger brought an incredible boon of scientific and technical knowledge to our *Caterpillar* research group when he joined MIT, and I am very thankful to him for that. He also set up my project on self-interacting dark matter, a topic of great interest to me which I would not have been able to break into without his guidance and simulations.

I also want to give a huge thanks to my Caterpillar project partners Brendan Griffen and Alex Ji. Brendan and Alex both came to MIT in my second year, and dramatically changed and accelerated my work environment. Brendan's expertise fresh out of his PhD was extremely welcome, as he immediately took on responsibility in pushing my first paper to completion and gave me a friendly companion to ask basic questions. Our constant collaboration since then, trips to Germany, regular chats, and his humor have been equally fun, memorable, and scientifically valuable. As my graduate student peer, Alex and I "grew-up" together in Anna's group, starting with our long diligent hours setting up the backbone of reading, running, and analyzing simulations, to eventually using those tools to explore interesting scientific questions. It has been a pleasure working together, and discussing all aspects of graduate student life.

I also thank the other half of Anna's research group, Heather, Andy, Ani, and Rana, for teaching me about metal poor stars, how to interpret spectra, and all of our fun times on the group retreats. I would especially like to thank Ani for the times we worked together on MOSTEC and for our regular conversations, and to Rana for her critical help in improving my thesis defense.

Much of my enjoyment of graduate school must be attributed to all of the astro grad students, post-docs, and staff, and to each of them I am very thankful. To Jess and Peter in particular I am thankful for bonding quickly after arrival and the countless fun times working on psets late into the night (Pablo!), parties at the "Morrison Mansion" late into the night, and growing up together as young adults. For all of the conversations, foosball games, dinners, parties, sports games, etc., I am thankful for David, Mike, Kat, Dylan, Phil, Bryce, Josh, Cameron, Aaron, Abraham, Ryan, Reed, Keaton, Tom, Nancy, Aaron, Liang, Ani, Alex, Sherry, Ryan, Uchupol,

Phyllis and Al.

My life was also shaped for the better by my friends outside of astrophysics. I am grateful for my wonderful roommates Tony and Amira (and the kitties Batcat and Molly) - I will miss our after work conversations, workouts, and videos. A huge part of my time at MIT was shaped by Mariana and I am grateful for her, and all of our time navigating the place together. Another big thank you goes to my first MIT roommate and dear friend Jon Goh with whom I signed the paperwork to go to MIT in the first place. Our exploits in cooking, cardboard boat building, obelisk costumes, and so on were entirely expected and as fun as ever. Thank you also to my many other friends from Princeton who lived or spent significant time in the Boston area including Andra, Krishnan, Jen, and my previous roommate Katherine.

Between curling and intramurals, sports were an integral part of my time during the PhD. A thank you goes out to everyone from MIT curling over the years including Phil, Andrea, Dave, Nish, Sean, Alex, Cody and Nate. Another one goes out to the many, many physics IM teams and players including Dan, Brian, Axel, Mike, Nikki, Ryan, Reed, Keaton, and Graeme.

Since I met her, I am enormously thankful for my girlfriend Sara who stole my heart and made me chase her to Mongolia to earn hers. The final stretch of finishing my thesis was so much easier having your support as you also worked endlessly on yours.

Lastly, my family deserves an incredible thank you for providing the foundation upon which my entire journey began. You provided a household where satisfying my curiosity through learning was always encouraged, where excelling was rewarded but never forced upon me, and where the idea of pursuing a PhD was planted and welcomed since middle school. I finally made it to the end of that dream. Thank you tremendously Mom, Dad, Kate, Brian and Tim.

Chapter 1

Introduction

1.1 Galaxies and their Satellites

1.1.1 The Milky Way and the Local Volume

The Milky Way galaxy in which our solar system resides has been observed for millennia as a streak of light amidst stars on the night sky. Not until Galileo's observations in 1610 did anyone identify that it consists of individual stars, and not until pioneering work of Harlow Shapley and Edwin Hubble in the 1910s and 1920s did anyone begin properly understanding the Milky Way (MW) as a rotating spiral galaxy (van den Bergh, 2011). In the past 100 years, sophisticated telescopes, satellites, and analysis have revealed a more complete description of the MW . Our galactic system consists not only of the spiral arms which are visible on the night sky, but a large encompassing gravitationally bound volume, the "halo", that consists of dark matter, globular clusters, halo stars, and smaller satellite galaxies. A diagram of this is shown in Fig. 1-1. At the center of the MW is a supermassive black hole estimated to be $4 \times 10^6\,\mathrm{M}_\odot$ (Reid, 2009; Falcke and Markoff, 2013). Surrounding that is the bulge, a roughly spherical distribution of predominately old stars orbiting the Galactic Center with a $\sim$ 1.5 kpc radius (Gonzalez and Gadotti, 2016). Extending from the bulge is a flattened disk of stars roughly 15 kpc in radius, in which our Sun is situated 8.5 kpc from the Galactic Center (Carraro, 2015). It is divided into two structurally different

layers, the thin disk which is 0.3 kpc in scale height, and the thick disk which is 0.9 kpc in scale height (Jurić et al., 2008; Chang et al., 2011; Carraro, 2015; van Vledder et al., 2016). From a top down view the disk is organized into spiral arms, and in total has a stellar mass of $6 \times 10^{10}\,M_{\odot}$ (Licquia and Newman, 2015), or about 10^{11} stars.

In the 1930s, Fritz Zwicky measured velocity dispersions of galaxies in the Coma cluster, finding that the velocity of stars indicated more gravitational mass than could be explained by visible stars (van den Bergh, 1999). One early hypothesis to explain this "dark matter" was that it is in the form of massive astrophysical compact halo objects (MACHOs), such as black holes, neutron stars, and brown dwarfs. However, microlensing studies (Schmidt and Wambsganss, 1998; Tisserand et al., 2007; Wyrzykowski et al., 2011a,b), black hole-neutron star merger rates Capela et al. (2013), and Planck CMB data Chen et al. (2016) collectively rule out nearly all mass ranges for MACHOs. A more promising proposal and still the leading theory today is that dark matter is in the form of weakly interacting massive particles (WIMPs) (Drees and Gerbier, 2012; Bergström, 2012). These are particles beyond the standard model which hypothetically formed and self-annihilated in equilibrium with the hot, dense early universe. As the universe expanded and cooled, the rate of production and destruction of particles dropped to negligible amounts, causing the density of dark matter to "freeze out" at the correct density observed if they interact at the scale of the electroweak force. Hence WIMPs are proposed to interact via the weak force (weakly interacting), have non-zero mass in order to interact gravitationally (massive), but do not interact with the electromagnetic force in order to remain dark. Along with galaxy rotation curves, evidence for a collisionless dark matter like WIMPs has independently come in the form of explaining mass distributions measured by gravitational lensing in the Bullet Cluster (Clowe et al., 2006; Angus et al., 2007), and is an essential ingredient in explaining the cosmic microwave background radiation Komatsu et al. (2011); Planck Collaboration et al. (2016b).

The MW galaxy sits at the center of a so-called dark matter halo, which is a roughly spherical distribution of dark matter whose density is highest at its center,

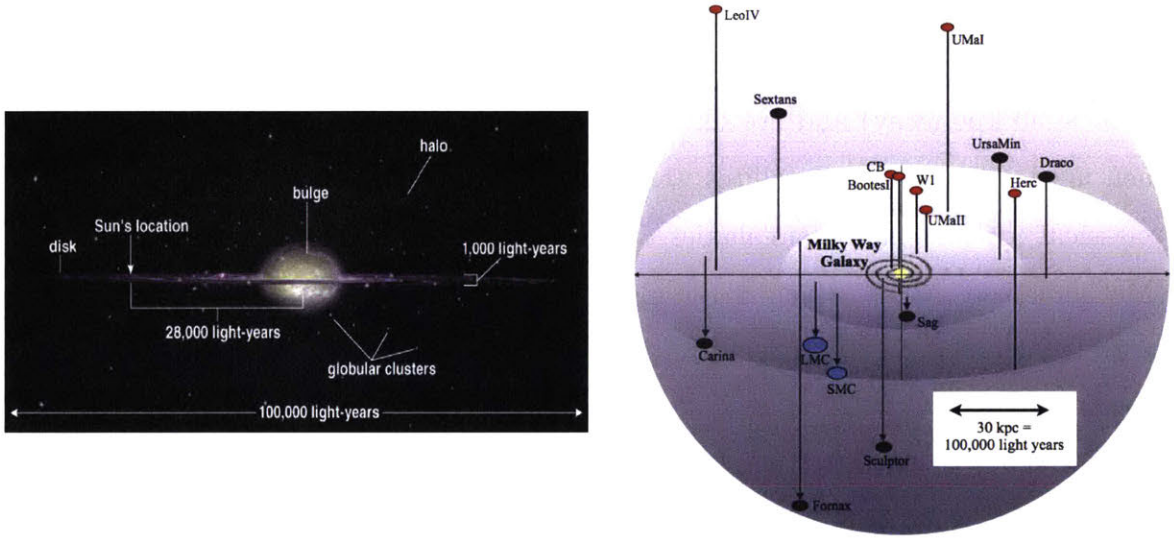

Figure 1-1 *Left:* The Milky Way galaxy consists of a central supermassive black hole, a spheroidal bulge of orbiting stars, and a disk organized into spiral arms. The disk has a radius of 15 kpc, or $\sim 50,000$ light years. Our solar system is located 8.5 kpc (or 28,000 light years) into the disk. The galaxy is centered in a large encompassing dark matter halo which contains stars, globular clusters, and satellites. The halo has a radius of $\sim$ 300 kpc. Image credit: Addison Wesley. *Right:* The MW has many satellite galaxies orbiting it within the halo. This diagram shows a selection of MW satellites within 250 kpc of the galactic center. In blue are the largest satellites, the Large and Small Magellanic Clouds. In purple are 7/9 of the classical dwarfs. Not included are Leo I and Leo II which are located beyond 250 kpc. In red are a few of the earliest discovered ultrafaint dwarfs (UFDs). Figure based on image by M. Geha, 2007.

and drops as $\rho(r) \propto r^{-3}$ beyond some characteristic scale radius Hernquist (1990); Navarro et al. (1997). The scale radius of the MW is around 40 kpc, and the full radius about 300 kpc. The halo has a combined dark matter and baryonic mass of $0.5 - 2.5 \times 10^{12}\,\mathrm{M}_\odot$ Wang et al. (2015). Within this halo are halo stars and globular clusters that make up the stellar halo (Helmi, 2008). There are also galaxies orbiting the MW, the satellite galaxies. The largest of these, the Large Magellanic Cloud (LMC) and the Small Magellanic Cloud (SMC), are visible to the naked eye. In the 1920s, Harlow Shapley identified that they are close to the MW center (known now to be $\sim$ 50 kpc away) and are likely orbiting the MW (Shapley, 1922, 1924). From then until 1994, 9 more satellites were discovered that are collectively known as the "classical" dwarf spheroidal galaxies. These all have stellar masses with $M_* > 10^5\,\mathrm{M}_\odot$ (McConnachie, 2012). Seven of the nine (Leo I and Leo II are excluded due to their larger galactocentric distances) are depicted in the right-hand panel of Fig. 1-1 as purple ellipses.

Since the classical dwarfs, more faint satellites were discovered with stellar masses between $10^3\,\mathrm{M}_\odot$ and $10^5\,\mathrm{M}_\odot$, the ultrafaint dwarfs (Willman et al., 2005b; Zucker et al., 2006b; Belokurov et al., 2006, 2007, 2008, 2010; Irwin et al., 2007; Walsh et al., 2007; Bechtol et al., 2015; Drlica-Wagner et al., 2015; Kim et al., 2015; Kim and Jerjen, 2015; Koposov et al., 2015a; Laevens et al., 2015; Martin et al., 2015; Luque et al., 2016; Torrealba et al., 2016b; Drlica-Wagner et al., 2016). Even fainter galaxies below $10^3\,\mathrm{M}_\odot$ have been found and are named the hyper-faint dwarfs (Willman et al., 2005a; Zucker et al., 2006a; Belokurov et al., 2009). Satellite galaxies are not exclusive to the MW, with more satellite galaxies known in the MW's similar sized neighbour, Andromeda (M31) (Zucker et al., 2004, 2007; McConnachie et al., 2008, 2009; Majewski et al., 2007; Irwin et al., 2008; Martin et al., 2009; Bell et al., 2011; Slater et al., 2011; Richardson et al., 2011), and several nearby galaxies and clusters (Jang and Lee, 2014; Sand et al., 2014; Crnojević et al., 2016). Explaining the abundance and properties of satellite galaxies, both for the purpose of theoretically understanding galaxy formation and for understanding the fundamental physics involved in galaxy formation (such as the properties of dark matter) are the core motivators of this

thesis.

1.1.2 Milky Way's Place in Cosmology

1.1.2.1 The Early Universe

The formation of galaxies, including their dark matter halos and satellite galaxies, is explained well within the Lambda cold dark matter (ΛCDM) model of cosmology (see Springel et al. (2006); Yoshida (2010) and Frenk and White (2012) for reviews). This explains how the universe and its constituent components evolved over time since the Big Bang. Λ refers to dark energy, a mysterious energy that is driving the expansion of the universe. "Cold" refers to the non-relativistic thermal velocities of dark matter particles in the early universe (Armendariz-Picon and Neelakanta, 2014). The very beginning fractions of a second of the universe remain an area of active research and even philosophical discussion. However, just 10^{-36} seconds after the Big Bang, the hot, dense universe likely underwent a period of exponential expansion in spacetime. Between 10^{-36} and $\sim 10^{-32}$ seconds, the theory of inflation posits that the linear scale of the universe increased by a factor of $\geq 10^{26}$. The particle associated with the driving energy behind this expansion, the inflaton, then decayed into particles including the standard model particles, dark matter, and perhaps others that make up the universe today. At the time though, temperatures were too high for hadrons (composite quark particles) to form, meaning most standard model mass was distributed in quarks, anti-quarks and gluons.

The matter distribution in this early universe was isotropic, and nearly fully homogeneous (uniform density). However, owing likely to quantum vacuum fluctuations in the early universe, the matter density contained small inhomogeneities. The primordial power spectrum of these density fluctuations is described by a power law, $P_0(k) \propto k^{n_s-1}$, where the wavenumber k describes different physical size scales, and the exponent n_s is named the scalar spectral index. Observations of light from the early universe indicate a value of $n_s = 0.965 \pm 0.006$ Planck Collaboration et al. (2016b). While favored, the theory of inflation is not verified. What is known is the

primordial matter power spectrum at the beginning of the universe and its isotropy. Possible alternate explanations for this density distribution, as well as inflation, are described in Brandenberger (2011).

Due to the driving force of dark energy, the universe continued to expand and cool. Between 10 and 10^3 seconds, quarks coalesced into protons and neutrons, and protons and neutrons formed bound states to make the first atoms (73% hydrogen, 25% helium, 2% heavier elements by mass). At this point the universe consisted of a hot, dense ionized plasma of protons, neutrons, electrons and photons. Due to high number densities, photons would travel a short mean free path between Thompson scattering off one electron and the next. During this phase, the matter distribution changed under the influence of the expanding universe, gravity, and radiation pressure, thus beginning the first steps of galaxy formation. Dark matter and baryons spread out along with the expansion of the universe. In time though, the mutual gravitational pull of masses began pulling clumps of matter together in overdense regions (regions where the local density is higher than the mean density of the universe). Due to the finite speed of gravity (equal to that of light), there is a delay from the Big Bang until distant masses can feel each other's mutual gravitational pull. Thus matter started to collapse on small distance scales first, and later on larger distance scales. Being collisionless and not interacting with photons, dark matter collapsed with no outward pressure to withstand its collapse.

Baryonic matter (astronomers' term for ordinary standard model matter) on the other hand, experienced an outward radiation pressure. As baryons collapsed they became hotter, releasing more photons, which exerted an outward pressure through Thompson scattering on electrons. Consequently, baryons alternately collapsed and heated, expanded and cooled, then collapsed again in what are called baryonic acoustic oscillations (Bassett and Hlozek, 2010). This process for baryons stopped once the universe expanded and cooled below 3000 K, enough for electrons to become bound to protons thereby diminishing Thompson scattering and the radiation pressure needed to counteract gravitational collapse. This moment of the formation of the first atoms is called recombination, and occurred $\sim 380,000$ years after the Big Bang.

The process of dark matter overdensities growing in their density contrast relative to mean of the universe, and baryonic matter overdensities growing and oscillating is mathematically described by a transfer function, $T(k)$. This transforms the primordial power spectrum as a function of wavenumber k to a new power spectrum. The power spectrum at recombination is thus $P(k) = P_0(k)T(k)^2$. The full functional form is dependent on the cosmological parameters, of which we focus on Ω_m, Ω_b, Ω_γ, Ω_Λ, n_s, σ_8, and H_0. Ω_m is the normalized present day matter energy density of the universe (dark matter + baryons), Ω_b is the baryonic matter energy density, Ω_γ is the energy density from radiation, Ω_Λ is the dark energy density, n_s is the scalar spectral index already introduced, σ_8 is a measure of the amplitude of the power spectrum, and H_0 is the present day expansion factor for the universe. Without a high density of free electrons for light to scatter off at recombination, the universe became optically transparent and photons were able to travel unimpeded. That light from recombination can still be observed today on earth as the cosmic microwave background radiation, and has been measured in detail with the COBE (Smoot et al., 1992), WMAP (Bennett et al., 2012), and Planck (Planck Collaboration et al., 2016b) satellite missions.

Measuring the power spectrum of the CMB at recombination and fitting it with theoretical transfer functions has allowed these missions to pinpoint the cosmological parameters. It also confirmed the early universe physics as described above, revolutionizing the field of cosmology. Current measurements indicate the energy density of the universe is $\Omega_\Lambda = 0.69$, $\Omega_\mathrm{m} = 0.31$, $\Omega_b = 0.05$, and $\Omega_\gamma \approx 10^{-4}$. This means 69% of the energy of the universe is in the form of dark energy, and the remaining 31% is in the form of matter. Radiation makes up a negligible fraction today. Of the matter fraction, only 16% is composed of ordinary baryonic matter, and the rest of the 84% is dark matter. The spectral index value of $n_s = 0.965$ is close to 1, which would generate equal power at all scales. Instead, being slightly < 1 means power decreases for smaller scales in the primordial universe. The Hubble factor is $H_0 = 68\,\mathrm{km/s\ Mpc^{-1}}$, meaning from the Hubble Law, $v = HD$, that objects 1 Mpc away from each other are moving away from each other at 68 km/s. The factor is sometimes expressed as $H = 100h\,\mathrm{km/s\,Mpc^{-1}}$, in which case $h = 0.68$. Finally, the

amplitude of the power spectrum is measured in terms of the standard deviation in the matter density field within spheres of radius $8/h$ Mpc. Its value measured from the CMB then linearly extrapolated to present day is $\sigma_8 = 0.8$. The influence of the cosmological parameters on the number of dark matter halos as a function of mass (the halo mass function) and in particular the population of halos within halos (subhalos) is the topic of Chapter 2.

1.1.2.2 Halo Virialization

As the universe continued to expand after the recombination, overdensities of dark matter and baryons continued to grow in density relative to the diminishing mean background density of the universe. Eventually, self-gravitation wins against the initial outward flow, and mass begins retracting, leading to the collapse of matter into a self-bound halo. Since dark matter makes up 84% of the mass in the universe, its overdensities dominate those of baryons, which make up the other 16%. Hence, as dark matter collapses into denser clumps, baryonic matter follows. Rather than collapsing on radial trajectories to the center of mass, gravitationally bound clumps of dark matter undergo a process of virialization. Random thermal tangential velocities provide angular momentum which prevents total collapse. An equilibrium state is reached when the time averaged kinetic energy is half the potential energy, $2T = U$, by the virial theorem.

A schematic of the virialization process is shown in Fig. 1-2. The overdensity of matter is approximated as a sphere (or spherical tophat), whose overdensity is $\delta = (\rho_\mathrm{p} - \rho_\mathrm{b})/\rho_\mathrm{b}$ where ρ_p is the density of the evolving density perturbation, and ρ_b is the mean density of the "background" expanding universe in which it resides. Solving equations for the motion of the outer shell of halo mass as it expands and then turns around yields a solution for when the virialization condition $2T = U$ is satisfied. At this point $\delta = 18\pi^2 = 178$ Peebles (1980). It is therefore common to define the somewhat arbitrary radius of a dark matter halo as the "virial radius," R_Δ, which is the radius in which the mean density of the halo is Δ times that of the background density of the universe. For a geometrically flat universe, which describes our universe,

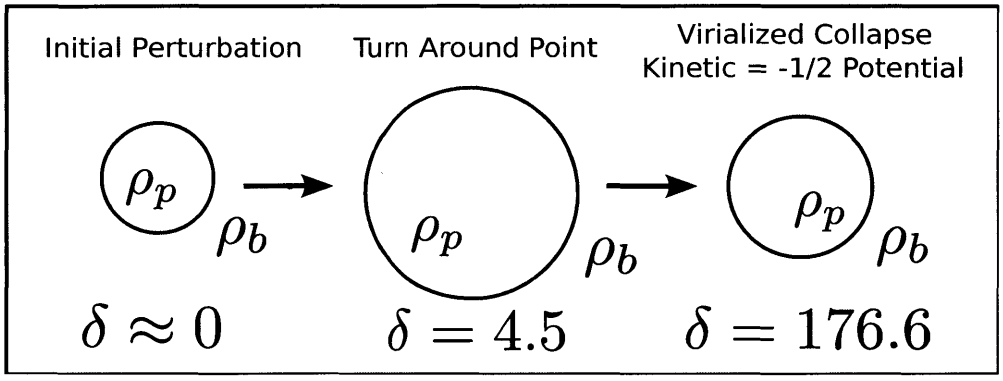

Figure 1-2 Spherical tophat model of dark matter halo growth, turnaround, and collapse until it is stabilized by virialization. $\delta \equiv (\rho_{\rm p} - \rho_{\rm b})/\rho_{\rm b}$ where $\rho_{\rm p}$ is the density of the evolving halo, and $\rho_{\rm b}$ is the mean density of the expanding background universe.

the background density equals the critical density of $\rho_{\rm crit} = 3H_0^2/8\pi G$, which today is $\approx 10^{-26}\,{\rm kg/m^3}$ or $148\,{\rm M_\odot/kpc^3}$. Rather than directly using $\Delta = 18\pi^2$, it is common to use a rounded number of $\Delta = 200$, although other values are motivated and used as well. Another common way to define the radius is where Δ follows a fitting function from Bryan and Norman (1998) which accounts in detail for the dependence of Δ on cosmology while solving the spherical top-hat perturbation problem.

Over time, more dark matter is pulled into a virializing halo, and the radius at which it is virialized grows. Since dark matter is collisionless (although small rates of collisions are not ruled out, see Chapter 3) and it does not emit photons, it has no efficient mechanism to lose its thermal energy. Baryonic matter, mostly in the form of Hydrogen and Helium, however, is able to cool. Collisions excite molecular vibrational states like those of H_2 and excite electrons into higher energy shells which then fall back, emitting photons and dissipating energy. Collisions also transfer angular momentum and leave some particles with less angular momentum. Consequently, baryonic matter becomes more concentrated at the center of the growing dark matter halo and increases in density. If, in the process of radiating away energy and collapsing, it reaches a large enough critical mass, the Jean's mass (Jeans, 1902), it will undergo unstable gravitational collapse and eventually become hot and dense enough for fusion to begin. In this way stars and then galaxies are born (see (Larson, 2003) for more details).

The clumps of dark matter which evolve into halos do not form in isolation. Near a large halo for instance are smaller clumps of dark matter which also virialize. The gravitational pull of the larger halo can drag the smaller ones into it. Once they pass within the virial radius of the larger "host" galaxy they are called subhalos. This process of entering the halo of a host is called "accretion" or "infall". If the dark matter subhalo hosts stars, it is called a satellite galaxy. Over time, two processes, dynamical friction and tidal stripping, can cause subhalos to be entirely disrupted. Dynamical friction refers to an effect in which particles are attracted to a massive object's center of mass, but since that object is moving, it will pull ahead of the location where mass was just dragged. Thus it creates a slight overdensity in its wake, which in turn exerts a backward gravitational force on the halo, slowing it down and reducing its angular momentum (Binney and Tremaine, 2008). The radius of its orbit around the host is therefore diminished. Tidal stripping refers to a process where the gravitational force of a host halo exceeds the force of a subhalo at a point within the subhalo's radius, thus removing mass from the subhalo (Binney and Tremaine, 2008). As a subhalo orbits closer and closer to the center of its host, tidal stripping increases and can completely disrupt the previously self-bound halo. This results in a stream of stars and dark matter, or a complete merger of the galaxies. The process of halo formation in which smaller halos are accreted onto larger halos and merge into them is called hierarchical galaxy formation.

The Milky Way and its satellite galaxies are a product of this long 13.8 billion year journey of growing density perturbations, halo virialization, and hierarchical galaxy formation.

1.1.3 Outstanding Problems with Local Group Satellites

The ΛCDM paradigm is impressively successful at explaining and predicting properties of structure formation in the universe. It can explain the number of galaxies, the mass function of these galaxies, their spatial correlation functions, shapes of large galaxies, and the occurrence of filaments, voids, clusters, and halos (Frenk and White, 2012). For instance, modeling the early distribution of dark matter in the universe and

its subsequent gravitational evolution in simulations naturally produces large scale structure that looks like the distribution of galaxies in our actual universe. Fig. 1-3 shows filamentary structure of galaxies in the Sloan Digital Sky Survey, the second CfA survey, and the 2dF Galaxy Redshift Survey in blue and purple. In comparison, mock observations of the equal sized volumes were made from the Millennium simulation (Springel et al., 2005) and are shown in red. They reproduce the same patterns of filaments and voids, and are visually strikingly similar.

In spite of so many successes on large scales, tensions still exist on the small scales. These tensions drive much of the research on low mass galaxies and satellite galaxies of the MW. The most prominent of them are the "missing satellite" problem, the "Too Big to Fail" problem, and the "cusp/core" problem.

1.1.3.1 Missing Satellite Problem

The missing satellite problem (Moore et al., 1999; Klypin et al., 1999b), refers to the fact that the number of satellites known to be orbiting the MW is vastly smaller than the number of dark matter subhalos in simulations of Milky Way-sized galaxies. Moore et al. (1999) initially estimated that the Milky Way should contain $\sim$ 500 satellites equal to the size of the Draco dwarf spheroidal satellite galaxy or larger, whereas only 11 such satellites are actually known in the MW. The magnitude of this problem has garnered much attention and effort to resolve it. In the past 17 years, the problem has been largely solved from two avenues. First, more satellite galaxies have been found, suggesting that dark matter subhalos host smaller galaxies than initially proposed, and bringing the known number of galaxies closer to the predicted value. Whereas only 10 satellites were known in 1999 when the missing satellite problem was identified, a proliferation of discoveries of ultrafaint dwarf satellites by the Dark Energy Survey, Pan-STARRS, ATLAS, and MagLiteS surveys has increased the count to 53 total candidates at the time of writing (see online version of McConnachie (2012) for a regularly updated list and for references within `http://www.astro.uvic.ca/~alan/Nearby_Dwarf_Database.html`). The names and galactic positions of the majority of the UFDs are included in Fig. 1-4.

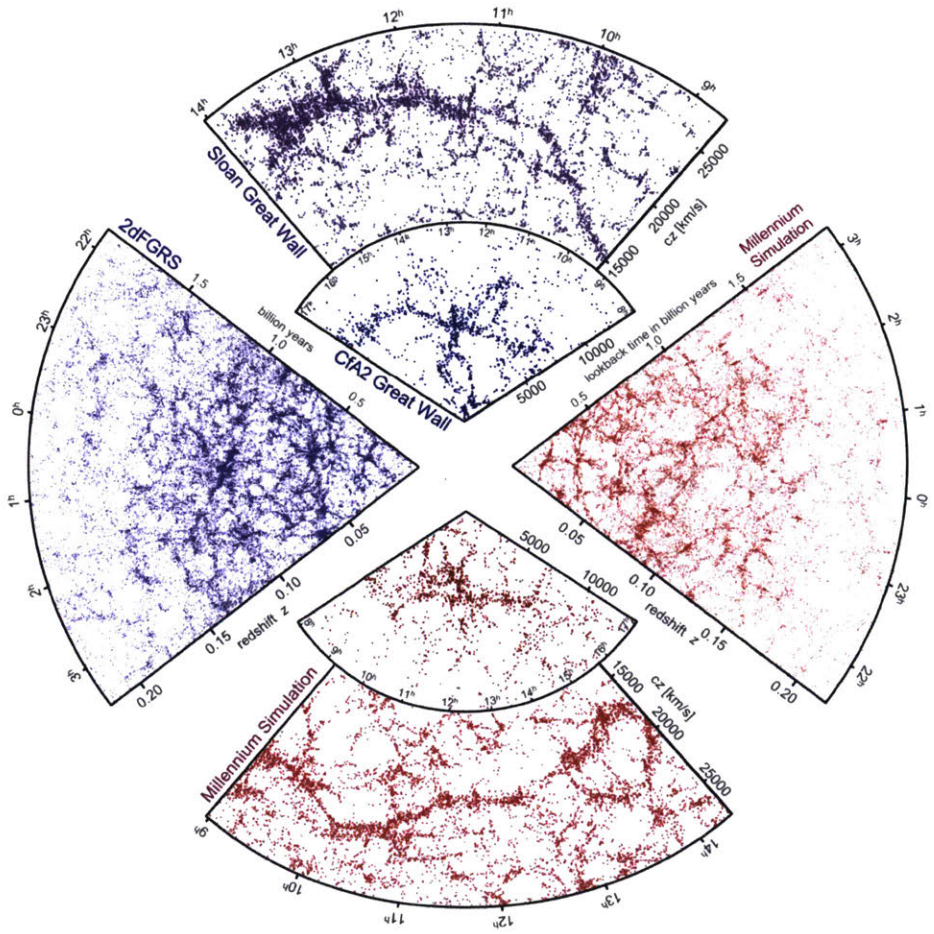

Figure 1-3 Mock surveys conducted on the cosmological galaxy simulation "Millennium", shown in red, are able to reproduce the filamentary structure and density of galaxies seen in the CfA2, SDSS, and 2dFGRS surveys, shown in blue. This is an indicator of the remarkable success of the ΛCDM model of the universe on which the simulations were based. The simulations begin with a realization of the early universe that matches the CMB power spectrum, meaning the mock universe should be statistical similar to the real universe, but not identical. Figure from Springel et al. (2006).

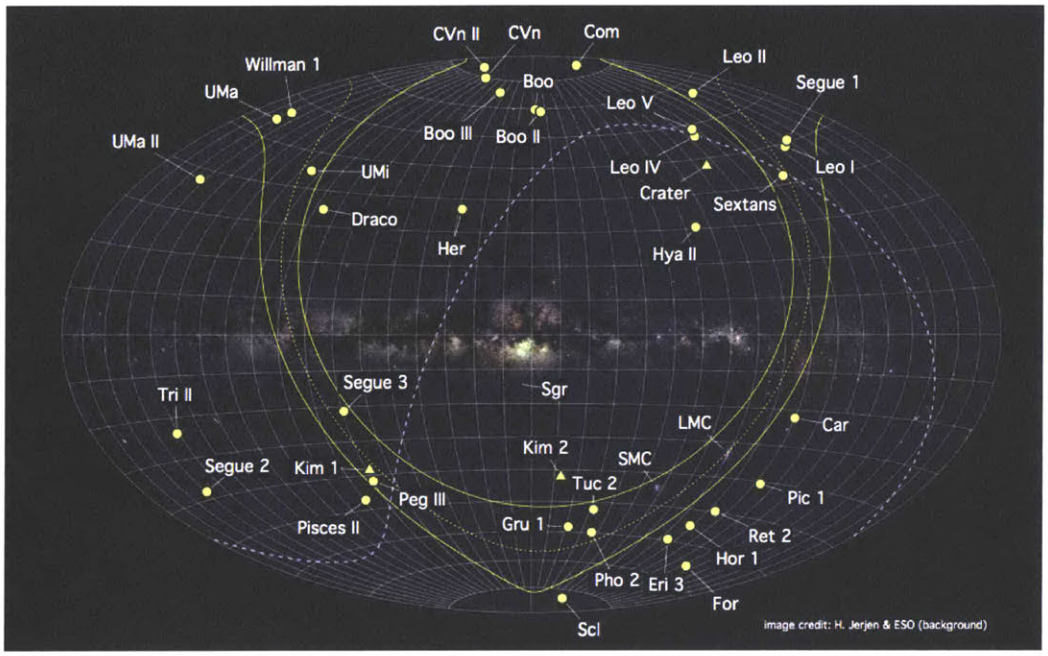

Figure 1-4 Known ultrafaint dwarf satellites in early 2015 in galactic coordinates on the sky. Image from Helmut Jerjen 2015.

Second, mechanisms to suppress star formation in dark matter halos have been proposed and theoretically verified. After the very first stars in the universe were born, between a redshift of $6 < z < 14$, they began emitting UV photons. These photons are of the right energy to dissociate molecular hydrogen, and ionize neutral hydrogen, making the the period of time known as the "Epoch of Reionization" (Zaroubi, 2013). Gas clouds that would otherwise cool and collapse to form stars lose H_2 and H cooling mechanisms and are heated by UV photons, preventing star formation (Efstathiou, 1992; Thoul and Weinberg, 1996; Gnedin, 2000; Wiersma et al., 2009; Pawlik and Schaye, 2009). In low mass galaxies, the radiation pressure can be strong enough and the gravitational pull weak enough to lead to photoevaporation, where gas is expelled from the halo. This process halts star formation, leaving low mass galaxies with a small ratio of stellar mass to dark matter mass. Additionally, halos that were not sufficiently large to form stars before reionization lose their gas or cannot cool it enough to ever form stars (Barkana and Loeb, 1999; Shapiro et al., 2004; Okamoto et al., 2008). This can leave dark matter halos entirely devoid of stars, an effect simulated in Sawala et al. (2013, 2015, 2016), and initially proposed in Bullock et al.

(2000); Somerville (2002); Benson et al. (2002).

Yet another consequence of weaker gravitational energy in low mass galaxies is that supernova from the first generation of stars can expel surrounding gas, preventing it from ever collapsing and forming stars in the halo (Wyithe and Loeb, 2013; Power et al., 2014). This contributes to very low and stochastic stellar mass to halo mass ratios. Thus the missing satellite problem can be explained by allowing for hundreds of dark matter subhalos to exist, but stipulating that only a small fraction of them host luminous galaxies due to reionization, and that most of the luminous galaxies are rendered very faint by reionization and supernova feedback.

While the number of MW satellites is no longer a serious challenge to the ΛCDM paradigm, the simple question of how many luminous satellites exist remains an uncertain and debated question in theory. The mass of the MW, which affects the number of satellites it should host, is unknown up to a factor of a few. Various estimates place its mass between $0.5 - 2.5 \times 10^{12}\,M_{\odot}$ Wang et al. (2015). Factors that control the fraction of galaxies which do and do not form stars, including the redshift of reionization, the halo size necessary for the first stars to form, the duration of reionization, the number of ionizing photons that escape galaxies, etc. are all too poorly constrained to enable precise estimates (Tegmark et al., 1997; Madau et al., 2008; Okamoto and Frenk, 2009; Bovill and Ricotti, 2009; Peter and Benson, 2010; Bovill and Ricotti, 2011a,b; Bromm and Yoshida, 2011; Lunnan et al., 2012; Power et al., 2014; Garrison-Kimmel et al., 2014b; Oñorbe et al., 2016; Griffen et al., 2016b). Furthermore, there are open questions about the number of stars that occupy each galaxy as a function of galaxy mass (Ural et al., 2015; Sawala et al., 2016; Garrison-Kimmel et al., 2016; Read et al., 2016b; Dooley et al., 2016b). Gas accretion rates, reionization, the rate of supernovae, the coupling of supernovae winds to gas, the geometry of supernovae explosions, etc. all effect the rate of star formation. Further complicating the picture is the fact that satellite galaxies are tidally stripped and have gas removed by ram pressure stripping (Emerick et al., 2016). Uncertainty also remains over quenching mechanisms for satellites, i.e. mechanisms that stop star formation van den Bosch et al. (2008); Lu and Mo (2015); Fillingham et al. (2016).

These questions must all be answered and understood before accurate theoretical predictions can be made about the number of satellites that ought to exist.

Even without modelling the physics with sufficient accuracy, simplified models calibrated to observations can provide constraining estimates of the abundance of satellite galaxies (Moster et al., 2013; Behroozi et al., 2013a; Brook et al., 2014; Garrison-Kimmel et al., 2014b, 2016). Chapters 4 and 5 make predictions about the number of satellite galaxies in the MW and around isolated dwarfs throughout the Local Volume.

1.1.3.2 Too Big to Fail and Cusp/Core

In addition to the abundance of satellites, questions regarding their internal density structure exist. When considering dark matter halos produced in simulations, a plot of its density as a function of radius (averaging over spherical shells) results in a profile known as a "cusp" (Navarro et al., 1997; Bullock et al., 2001; Wechsler et al., 2002). Specifically, at small radii, $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln(\mathrm{r}) \approx -1$ for a cuspy density profile. In contrast, observations of Milky Way satellite galaxies (Walker and Peñarrubia, 2011; Salucci et al., 2012; Breddels and Helmi, 2013) as well as more distant low mass galaxies (de Blok and McGaugh, 1997; de Blok et al., 2001; Gentile et al., 2004; Simon et al., 2005; Kuzio de Naray and Spekkens, 2011; Oh et al., 2011; Castignani et al., 2012; Adams et al., 2014) indicate a density profile known as a "core". In a cored profile, the central density is reduced relative to a cusp, and $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln(\mathrm{r}) \approx 0$. A diagram of a "cuspy" compared to a "cored" profile is presented in Fig. 1-5. This discrepancy is referred to as the "cusp/core" problem (Flores and Primack, 1994; Moore, 1994).

An additional related issue concerns the circular velocity profiles of dark matter halos in simulations. The circular velocity profile measures the typical velocity of stars (or dark matter particles in simulations) as a function of radius. For instance, in a spherically symmetric halo with circular orbits, the circular velocity would be $v_{\mathrm{circ}} = \sqrt{GM(<r)/r}$. In simulations of Milky Way mass galaxies, it was identified that a population of subhalos exists which have circular velocity profiles that are

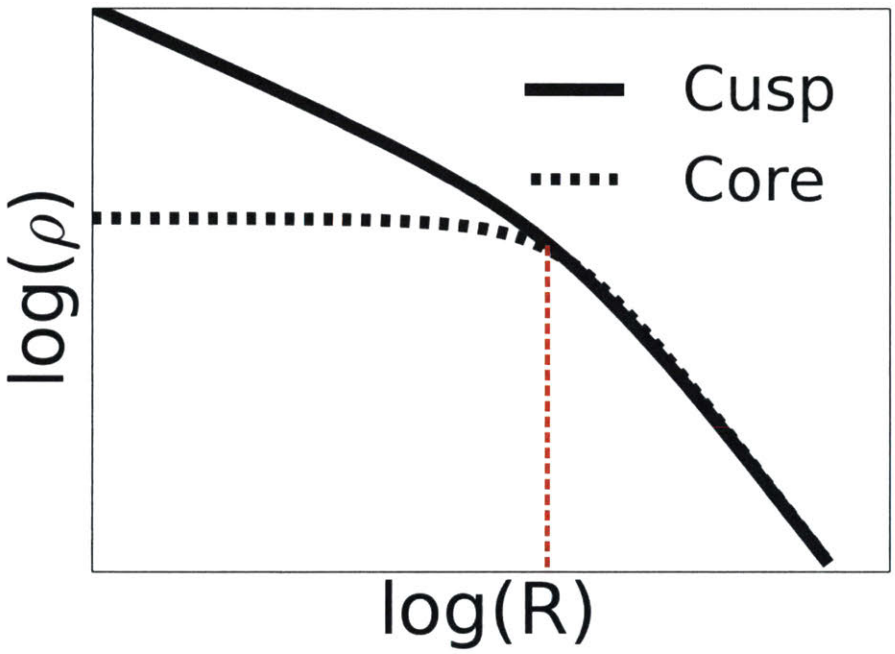

Figure 1-5 Diagram contrasting a dark matter halo "cusp" and "core." These refer to the shape of the density profile of a halo at low radii. In a cuspy halo, the central density rises steeply as $r \to 0$, whereas in a cored profile, the density remains nearly constant. More precisely, $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln(\mathrm{r}) \approx -1$ at low radii for a cuspy density profile, but is ≈ 0 for a cored profile.

larger in magnitude than any known Milky Way satellite. If such subhalos did exist as simulated, they would be too big to fail to form stars, and as such there should be large visible satellite galaxies observed that match the simulated circular velocity profiles. This discrepancy was dubbed the "too big to fail" problem (Boylan-Kolchin et al., 2011, 2012), in a timely reference to the banks of the 2008 financial crisis. Fig. 1-6 highlights this problem. It plots the circular velocity profiles of the 15 largest subhalos in a simulation of a Milky Way mass galaxy, along with the measured circular velocity at specific radii of the largest Milky Way satellites. As seen, none of the largest simulated subhalos are consistent with the largest Milky Way satellites.

Solutions to the cusp/core problem simultaneously help resolve the too big to fail problem. By lowering the central density of halos, the mass enclosed $M(< r)$ functions are reduced, and hence so are the circular velocity profiles. The too big to fail problem, however, could be solved without requiring cores in halos. A lower mass Milky Way for instance, would contain in general smaller subhalos with diminished circular velocity profiles (González et al., 2014).

Commonly proposed solutions to these two issues involve either creating cores through baryonic means, or through adjusting the properties of dark matter. For baryonic matter, mechanisms include the reduction of central mass via supernovae (Navarro et al., 1996b; Read and Gilmore, 2005; Governato et al., 2010; Oh et al., 2011; Governato et al., 2012; Pontzen and Governato, 2012; Teyssier et al., 2013; Madau et al., 2014; Amorisco et al., 2014; Pontzen et al., 2015), tidal stripping (Zolotov et al., 2012; Brooks and Zolotov, 2014), bursty star formation (Oñorbe et al., 2015), and baryon loss due to reionization (Sawala et al., 2014). A potential non baryonic solution is that of relaxing the assumption that dark matter is collisionless, and allowing for a low rate of scattering between dark matter particles. This "self-interacting dark matter" and its signatures are discussed in more detail in Chapter 3, as are details of the cusp/core and too big to fail problems.

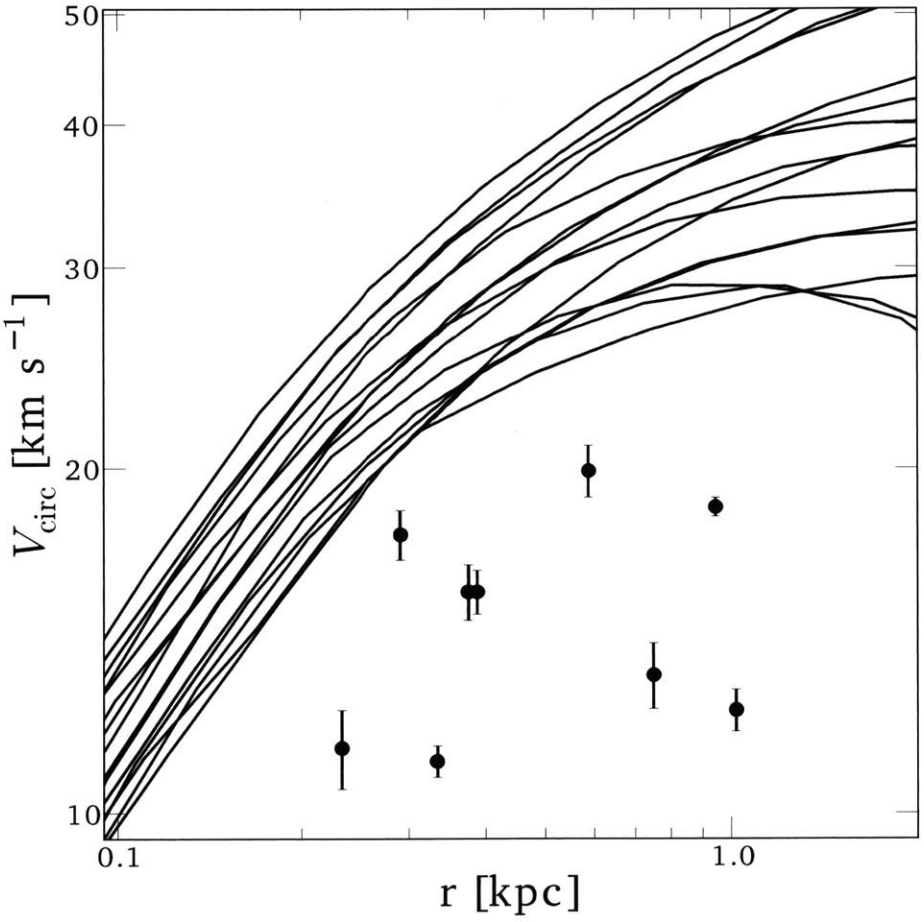

Figure 1-6 The circular velocity profile of the 15 largest subhalos in a simulation of a Milky Way-like galaxy (black curves) are all too high to be consistent with any of the Milky Way satellites. Values of measured circular velocity at specific radii for the largest Milky Way satellites are indicated by black dots with error bars. This discrepancy is known as the "too big to fail" problem. Figure from Vogelsberger et al. (2012).

1.2 Simulating Galaxies

1.2.1 N-body Galaxy Simulations

Understanding galaxies in our Local Group is advanced by an interplay of observations and interpretation of those observations in light of theory. In the early universe, the theory of structure formation can be understood analytically with the primordial power spectrum, the transfer function, and linear growth of overdense regions of matter, as discussed in Section 1.1.2.1. However, by $z \approx 127$, or 11 Myr after the Big Bang, nonlinear growth of halos and halo-halo interactions become important. These are chaotic processes that cannot be modeled analytically, making simulations necessary for further study.

Since dark matter makes up 84% of the mass of the universe and galaxies form at the center of dark matter halos, the first step towards understanding late-epoch structure formation is to follow the evolution of dark matter in the universe. Thus, there is considerable value in running simulations with only dark matter ("dark matter only simulations") and no baryons. Semi-analytic methods can then be applied to estimate how stars and baryons in general are distributed on the dark matter backbone. Simulations that include baryons and solve the hydrodynamic equations governing their motion also exist, but are more computationally expensive and require implementing uncertain physics of star formation, reionization, supernovae, black hole accretion, etc. (O'Shea et al., 2004; Hopkins et al., 2014; Vogelsberger et al., 2014a; Muratov et al., 2015; Schaye et al., 2015). Which type of simulation is best to use depends on the science questions being asked, and computational resources available (see Somerville and Davé (2015) for a review of hydrodynamic and semi-analytic codes). This thesis focuses on the dark matter only simulations, as described further below.

The basic idea of cosmological simulations is to distribute discrete particles in a volume that mimics the early universe, and then to evolve them under the influence of gravity. More precisely, the initial conditions of the simulation are a list of (x, y, z) positions and (v_x, v_y, v_z) velocities of massive particles generated in such a way to satisfy the known early universe power spectrum at a chosen time (or cosmologi-

cal redshift). The velocities of the particles are determined by the gradient of the gravitational potential, ϕ, which is in turn determined by solving Poisson's equation, $\nabla^2\phi = 4\pi G\rho$, on the discretized matter density field (Hahn and Abel, 2011). There is no one unique set of initial conditions that satisfy the power spectrum. An initial density field is generated randomly with white noise, and then transformed so as to meet the necessary requirements. There are an infinite set of possible initial white noise fields, and thus initial conditions.

After establishing the initial conditions, numerical integration schemes evolve all of the particles in time steps, determining their next position and velocity according to the gravity exerted on each and every particle. A brute force implementation of computing the force of gravity on every particle from every other particle would be an $O(N^2)$ algorithm, which does not scale well for large numbers of particles. Instead, more sophisticated $O(N \log N)$ algorithms are used to accurately compute short range forces, and quickly estimate long-range forces. For instance, particles can be organized into quad-trees, reducing the effective number of particles needed to compute short range forces. For long range forces, they can be placed onto the nearest nodes of a mesh which allows for a Fourier transform to solve Poisson's equation for gravity. Additionally, care must be taken in close encounters to use small enough time steps so as to conserve energy and angular momentum. More details on the algorithms of N-body codes can be found in Springel (2005); Dolag et al. (2008).

The mass of each particle is determined by the volume of the mock universe simulated, the mean density of the universe, and the number of particles used. Due to regular improvements in computational ability, the power of cosmological simulations has evolved tremendously. While the earliest simulation of a ΛCDM universe had 32,768 particles in 1984 (Davis et al., 1985), it is now common for simulations to use over one billion particles.

1.2.2 Caterpillar Simulation Suite

Since initial conditions are generated randomly, the locations and sizes of galaxies that will form in a simulation are not known ahead of time. Therefore, a special

technique known as a "zoom-in" simulation (Katz et al., 1994; Navarro and White, 1994; Diemand et al., 2007; Springel et al., 2008) is required to simulate galaxies similar to the Milky Way (MW). In this method, a large cosmological volume is simulated first (e.g. a box with 100/h Mpc side length at $z = 0$) and then halos are selected within according to their desired $z = 0$ mass and formation history. The particles of interest in and around the chosen halo are then traced back to their initial conditions, and a bounding box, sphere, or ellipse is made around them with some extra spatial padding allocated. The initial conditions in that region are then re-sampled with higher resolution particles, and a new simulation is run. This time high resolution particles make up the selected halo, and increasingly lower resolution particles are used at larger radii to include the effect of long range forces. This technique allows for an efficient simulation of a MW-like galaxy with high particle resolution, while still preserving the full cosmological setting.

A set of zoom-in simulations run and analyzed as part of this thesis is the *Caterpillar* simulation suite (Griffen et al., 2016b), a set of 36 and growing dark matter only simulations of Milky Way-like galaxies. This suite provides the means to tackle a diverse set of questions about the formation history and structure of the Milky Way, Andromeda, and other Local Group galaxies. Uniting all questions is the need to understand the diversity and variance of characteristics of Milky Way-like halos. For instance, the occurrence of two large satellites as big as the LMC and SMC is expected in only $\sim 5\%$ of MW mass galaxies (Boylan-Kolchin et al., 2009). The stellar halo, satellite mass functions, number of UFDs, spatial distribution of satellites, and reionization, formation, accretion, and merger history of component galaxies are all subject to similar variance which can only be properly understood with a large sample of simulated MW-like galaxies.

A guideline to quantify how many a "large sample" needs to be is to produce at least three halos as rare as one which contains and LMC/SMC pair. This results in a number of ~ 70. An additional key design requirement is to have a small enough particle mass to study the stellar halo and UFDs. Since part of the stellar halo is composed of stars that were tidally removed from infalling galaxies(Pillepich et al.,

2015), the motion of different mass elements representative of the stars in satellites need to be traceable. A way to do this in dark matter only simulations is via "particle tagging" in which the innermost dark matter particles of a halo are followed as tracers for the motion of stars (Bullock and Johnston, 2005; De Lucia and Helmi, 2008; Le Bret et al., 2015). A sample of $\sim$ 100 tagged particles per halo is roughly the number needed to begin tracing the diverse motion of stars from a single halo, and the smallest halos that need to be tagged have a total mass of $\sim 10^8\,M_\odot$. Following a the lower limit recommendation to tag 1% of the innermost particles (Cooper et al., 2010; Rashkov et al., 2012; Bailin et al., 2014), a mass resolution of $\sim 10^4\,M_\odot$ per particle is therefore required. Furthermore, another goal of the *Caterpillar* project is to identify the formation sites of the very first stars and track their subsequent evolution. The first stars are estimated to form in halos when they reach a total mass of 10^6 $M_\odot$ (Tegmark et al., 1997; Madau et al., 2008; Power et al., 2014). Since the existence of a self-bound halo can only be reliably identified in simulations if it contains at least 100 particles (Onions et al., 2012), a particle mass of $\sim 10^4\,M_\odot$ is once again needed. Having a low particle mass is valuable for many other studies as well, such as measuring density profiles and circular velocity profiles of halos at small radii, and research goals not yet identified. High particle resolution is thus important for the immediate goals of the *Caterpillar* project, usefulness to other researches, and the longevity of the project.

Running a large simulation suite requires a significant investment of resources. If possible, it can be advantageous to use existing simulations to answer scientific questions. Fig. 1-7 shows the state of the art of zoom-in simulation suites at the time of writing along two axes: the number of simulations (x-axis) and the particle resolution (y-axis). To satisfy the goal of studying the diversity and variance of MW-like halos down to the few % level, a suite of least > 30 simulations is needed. Before the start of *Caterpillar* in 2012, no suite existed with so many halos. Since then, two suites, those of Mao et al.(Mao et al., 2015) and ELVIS(Garrison-Kimmel et al., 2014b), have surpassed 40 halos. However, these two sets of simulations have a particle mass that is a factor of 10 times larger than needed for the *Caterpillar*

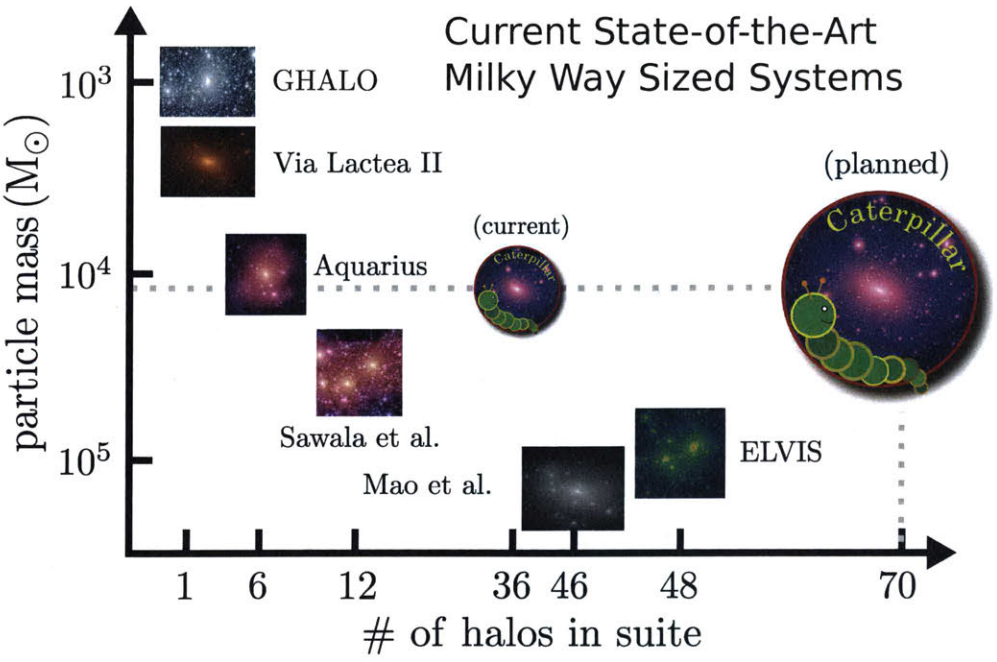

Figure 1-7 State of the art zoom-in simulations of Milky Way-mass galaxies, shown in terms of their particle mass and the number of simulations in each project. No existing simulations satisfied our research team's desire for 70 halos with a particle mass resolution of $\sim 10^4\,M_\odot$ per particle. Consequently, the *Caterpillar* simulation suite was developed to fill this gap. Indicated is the present day progress of the suite at 36 halos, and the final goal of 70 halos. Figure adapted from an image by Brendan Griffen, 2016.

project goals. Of the simulations that do have a mass resolution $\leq 10^4\,M_\odot$, the largest number of halos is only 6 from the Aquarius project (Springel et al., 2008). Thus there was a need for a new simulation suite, and *Caterpillar* was born. At the time of writing, there are 36 halos fully completed, as indicated in Fig 1-7, 8 more still running, and initial conditions generated for a full 70. The highest resolution particle mass in each simulation is $2.99 \times 10^4\,M_\odot$. A density projection of the first four halos at $z = 0$ is presented in Fig. 1-8.

The flagship publication announcing the *Caterpillar* project and presenting initial science results is included as Chapter 6. More technical details of the project can be found there. Additionally, a website with more information, images, and videos can be found at `http://www.caterpillarproject.org/`. In addition to making up a chapter of this thesis, the *Caterpillar* project is connected to the work presented in

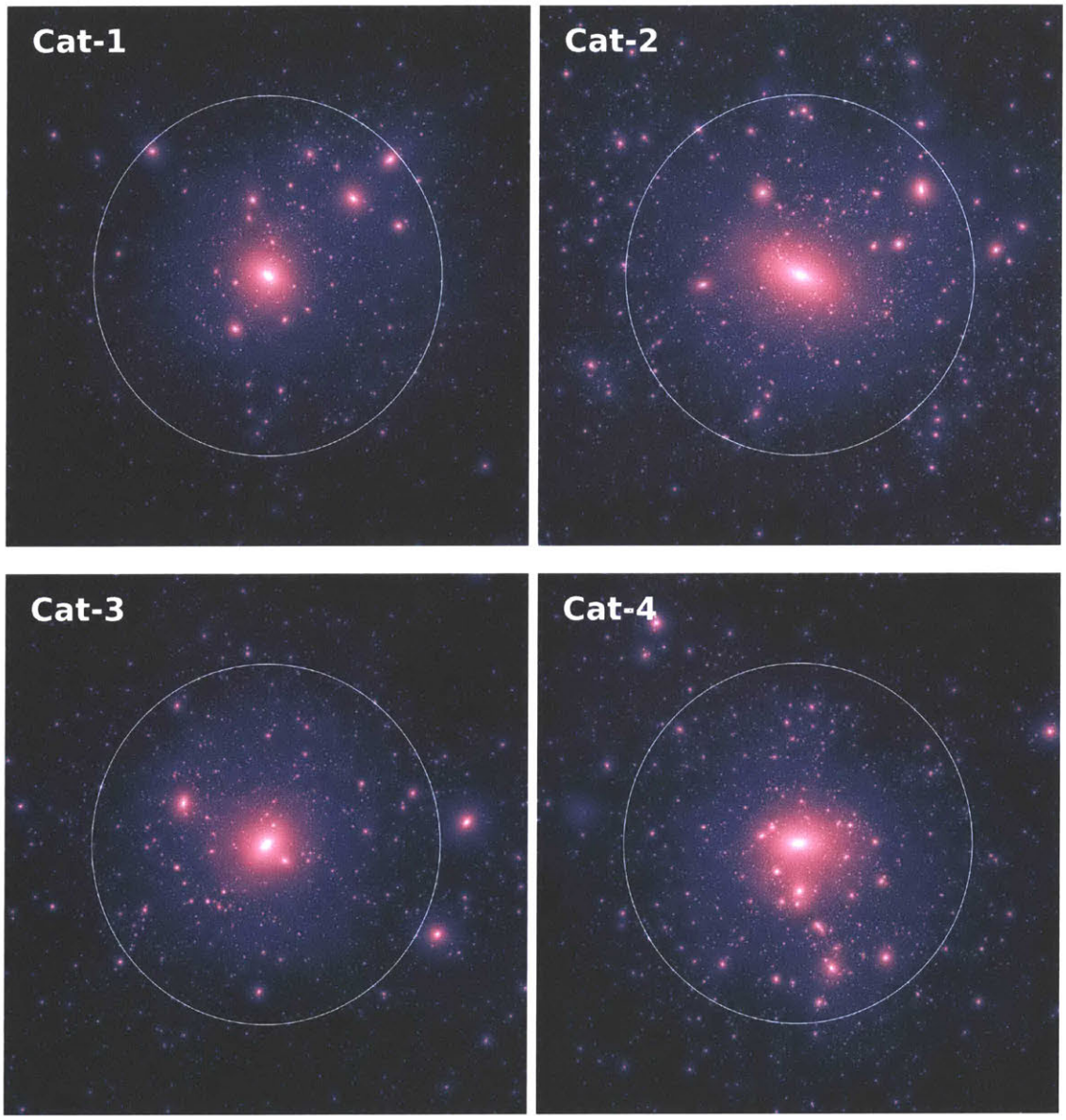

Figure 1-8 Density projections of the first four Caterpillar simulations at present day. Mass is integrated in the viewing direction and color coded according to the $\log_{10}$ of its density. Virial radii are indicated by white circles, and are typically between 250 and 350 kpc. The high density center of the Milky Way-like halo is where the spiral arm disk galaxy in which our solar system resides would be located. Surrounding this, clumps of dark matter can be seen within the virial radius. These are the subhalos, of which a subset host luminous satellite galaxies. The diversity of substructure in just four simulations is visually apparent.

all other chapters. The code developed to analyze results from the *Caterpillar* suite branched off of the work presented in Chapter 2. Additions to the *Caterpillar* analysis code, particularly regarding the merger history of halos, were developed in tandem with the work in Chapter 3. In Chapters 4 and 5, 33 of the *Caterpillar* simulations were used directly for science results. Finally, in Chapter 7 30 simulations were used to study the formation sites and evolution of the first galaxies.

Chapter 2

The Effects of Varying Cosmological Parameters on Halo Substructure

The content of this chapter was published in Astrophysical Journal *Volume 786, Issue 1 in May 2014. It has an arXiv location of* `https://arxiv.org/abs/1403.6828`. *The authors are: Gregory A. Dooley, Brendan F. Griffen, Phillip Zukin, Alexander P. Ji, Mark Vogelsberger, Lars E. Hernquist, and Anna Frebel.*

Disclaimer

This chapter was predominately a product of my own work. The simulations analyzed were run by Phillip Zukin. Brendan Griffen contributed a majority of the text in the introduction and describing the halo finders. He also carried out the analysis fitting functions to the host halo concentration-mass relationship and wrote the corresponding text. All authors provided comments and feedback on the paper.

Acknowledgements

We thank Peter Behroozi for help with using ROCKSTAR and Volker Springel for giving access to GADGET3. G.A.D. acknowledges support from an NSF Graduate Research Fellowship under Grant No. 1122374.

Abstract

We investigate how different cosmological parameters, such as those delivered by the *WMAP* and *Planck* missions, affect the nature and evolution of dark matter halo substructure. We use a series of flat Λ cold dark matter (ΛCDM) cosmological N-body simulations of structure formation, each with a different power spectrum but the same initial white noise field. Our fiducial simulation is based on parameters from the *WMAP* 7th year cosmology. We then systematically vary the spectral index, n_s, matter density, Ω_M, and normalization of the power spectrum, σ_8, for 7 unique simulations. Across these, we study variations in the subhalo mass function, mass fraction, maximum circular velocity function, spatial distribution, concentration, formation times, accretion times, and peak mass. We eliminate dependence of subhalo properties on host halo mass and average over many hosts to reduce variance. While the "same" subhalos from identical initial overdensity peaks in higher σ_8, n_s, and Ω_m simulations accrete earlier and end up less massive and closer to the halo center at $z = 0$, the process of continuous subhalo accretion and destruction leads to a steady state distribution of these properties across all subhalos in a given host. This steady state mechanism eliminates cosmological dependence on all properties listed above except subhalo concentration and V_{max}, which remain greater for higher σ_8, n_s and Ω_m simulations, and subhalo formation time, which remains earlier. We also find that the numerical technique for computing scale radius and the halo finder used can significantly affect the concentration-mass relationship computed for a simulation.

2.1 Introduction

The cold dark matter (ΛCDM) model of our universe has been well constrained to be flat, dark energy dominated, and filled predominantly with cold, collisionless dark matter (Bennett et al., 2012; Planck Collaboration et al., 2014). It is partly parametrized by four quantities: the matter fraction of the universe at present day, Ω_m, the primordial power spectrum scalar spectral index n_s, the Hubble constant at present day $H_0 = 100h$ km s^{-1} Mpc^{-1} , and the amplitude of the linear power

spectrum at the scale of 8 h^{-1} Mpc, σ_8. In a flat universe the dark energy content Ω_Λ, is constrained by $\Omega_\Lambda + \Omega_M = 1$. While adherence to ΛCDM has not changed, the best estimates of these parameters have varied between recent *Planck* and *WMAP* measurements from *WMAP* values of $\Omega_m = 0.27$, $n_s = 0.96$, $\sigma_8 = 0.80$, and $h = 0.71$ to *Planck* values of $\Omega_m = 0.32$, $n_s = 0.96$, $\sigma_8 = 0.83$, and $h = 0.67$ (Spergel et al., 2003, 2007; Dunkley et al., 2009; Larson et al., 2011; Bennett et al., 2012; Planck Collaboration et al., 2014).

For over a decade now, numerical simulations adopting the ΛCDM paradigm have shown that large dark matter halos contain substructures or *subhalos* which survive to the present day (e.g., Moore et al. 1999; Tormen et al. 1998; Diemand et al. 2007; Springel et al. 2008; Giocoli et al. 2010). Various studies which attempt to connect the properties of dark matter halos to present day observables have yielded some conflicting results. There remain two key problems: the so-called "missing satellite problem" where there is a dearth of observed low $V_{\rm max}$ subhalos when compared to simulations (Moore et al. 1999), and the "too big to fail" problem where there is a lack of dark ($L_V < 10^5 L_\odot$), dense, high $V_{\rm max}$ subhalos when compared to simulations (Boylan-Kolchin et al., 2011). While these issues depend critically on the influence of baryonic and radiative processes (Brooks et al., 2013), as well as the possible warm (Lovell et al., 2013) or self-interacting nature of dark matter (Vogelsberger et al., 2012; Zavala et al., 2013; Vogelsberger and Zavala, 2013), a more complete understanding of the characteristics of the dominant dark matter structures is also required, including the effect of changes in cosmological parameters on halo properties.

Recently, Angulo and White (2010) developed a technique, whose accuracy was tested by Ruiz et al. (2011), to transform large simulation results into those of a slightly modified cosmology by adjusting length, mass, and velocity units as well as changing time time scale and amplitudes of large scale fluctuations to successfully reproduce the mass power spectrum of a given target cosmology to better than 0.5 per cent on large scales (k $< 0.1\ h^{-1}$ Mpc). Guo et al. (2013) recently used this technique to compare *WMAP*1 and *WMAP*7 cosmologies on the *Millennium* and *Millenium-II* simulations (Springel et al. 2005, Boylan-Kolchin et al. 2009). Assuming a halo

mass-luminosity relationship, they determined that the differences were not significant enough to be found observationally. Wang et al. (2008) conducted a similar study between *WMAP*1 and *WMAP*3 cosmologies, but used two distinct simulations instead of transforming one into another cosmology. While the cosmologies produced minimal observable differences at low redshifts, they had to use significantly different star formation and feedback efficiencies in their models to match results to observational data. Furthermore, they concluded observational differences should be noticeable at high redshifts ($z > 3$). A number of studies have also been made of the concentration, spin and shape of dark matter halos across various WMAP cosmologies (Duffy et al. 2008, Macciò et al. 2008) and within the larger hierarchical framework (e.g., Zhao et al. 2009). These studies all focused on host halos and have largely ignored subhalo properties across cosmologies. Guo et al. (2013) did examine subhalos, but only their global abundances, not as a function of their host halos.

The dependence of host halo abundance on cosmological parameters is well understood within the Press-Schechter (Press and Schechter, 1974) and improved Extended Press-Schechter (see Sheth and Tormen (2002) for example) formalism of halo mass functions. Subhalo mass functions (SHMFs), however, are not as well predicted analytically and have not been studied extensively with respect to changing cosmologies. Zentner and Bullock (2003) did partially investigate the effect of input cosmology on substructure using their merger history, destruction rate and survival probability as a function of spectral index, n_s, as well as a running spectral index model. They did not, however, investigate how these properties depend on other cosmological parameters (e.g., σ_8 or Ω_m).

Several studies have investigated more general subhalo properties. Oguri and Lee (2004) developed an analytic model for the subhalo mass function based on the Extended Press-Schechter formalism and took account of the effects of tidal disruption and dynamical friction to estimate that the subhalo mass function is virtually independent of host halo mass. But since they used the host mass at the present day to calculate the impact of dynamical friction they inaccurately predicted the SHMF. In turn, van den Bosch et al. (2005) examined the SHMF, mass fraction, and ac-

cretion history and found that the SHMF may not be universal, arguing that the slope and normalization depend on the ratio of the parent mass to that of the characteristic non-linear mass, M*. M* indicates the typical mass scale of halos that are collapsing as a function of redshift and is defined by $\sigma(M^*, z) = \delta_c(z)$ where $\sigma(M, z)$ is the rms density fluctuation for a spherical volume of mass M at redshift z, and $\delta_c(z) = 0.15(12\pi)^{2/3}[\Omega_m(z)]^{0.0055} \approx 1.68$ is the critical threshold for spherical collapse (Navarro et al., 1997). The value of M* depends on cosmology and is larger for cosmologies where objects collapse sooner, i.e., higher σ_8, n_s and Ω_m. van den Bosch et al. (2005) along with Giocoli et al. (2010) and Gao et al. (2004) however, further found that the SMHF per unit host halo mass at z = 0 is universal.

Whilst these studies have determined many of the fundamental properties of substructures, they ultimately do not systematically investigate the effect a varied cosmology has on their properties. Those that attempt to, only focus on the variance of their properties using *one* cosmological parameter (e.g., Zentner and Bullock 2003). In this work, we vary three key cosmological parameters (σ_8, n_s, Ω_m) systematically and quantify the effect it has on the substructure population using simulations. In this manner, we help quantify what effect these variations have on the subhalo mass function, mass fraction, maximum circular velocity function, spatial distribution, concentration, formation times, accretion times, and peak mass in a self-consistent manner.

This paper is organized as follows. In Section 2.2, the simulations and halo finders used in this work are presented. Section 2.3.1 presents the known major effects of cosmology on host halos. In Sections 2.3.2 and 2.3.3, we discuss the averaged statistical properties of substructure as a function of cosmology. In Section 2.3.4 we discuss the differences in substructure that is directly matched between each of our cosmological simulations. A summary of the effects of cosmology on subhalo properties is given in Section 2.4 and conclusions are given in Section 2.5.

2.2 Numerical Methods

2.2.1 Simulations

For our fiducial simulation we adopt an approximately *WMAP7* cosmology characterized by the present-day matter density parameter: $\Omega_m = 0.27$; a cosmological constant contribution, $\Omega_\Lambda = 0.73$; and Hubble parameter: $h = 0.7$ ($H_0 = 100\ h$ km s^{-1} Mpc^{-1}). The mass perturbation spectrum has a spectral index, $n_s = 0.95$, and is normalized by the linear rms fluctuation on 8 Mpc h^{-1} radius spheres, $\sigma_8 = 0.8$. Six of our seven simulations adopt cosmologies which are identical to our fiducial run except for individual variations in n_s, Ω_m and σ_8.

Cosmological initial conditions were generated at redshift z = 127 using the public code GRAPHIC (Bertschinger 2001) with an Eisenstein-Hu transfer function (Eisenstein and Hu 1998). All seven simulations employed $N = 512^3$ dark matter particles, a Plummer-equivalent comoving softening length of 1.22 h^{-1} kpc, the same comoving box size, $L_{\rm box} = 25\ h^{-1}$ Mpc, and were evolved using an unreleased version of GADGET3 (Springel 2005). Our particle masses across the seven runs range from 6.14 to 11.3 $\times\ 10^6\ h^{-1}\ M_\odot$, and all simulations were written out at 64 snapshots. Our entire suite with all pertinent parameters are listed in Table 2.1.

Rather than performing convergence test runs, we only investigate halos above an appropriate minimum number of particles as determined by previous studies for each halo property. We also note that our small simulation volume could lead to different absolute results than those obtained from larger simulation volumes that contain larger wavelengths of the power spectrum, particularly for halo mass functions (Sirko, 2005; Power and Knebe, 2006; Bagla et al., 2009). However, we are only concerned with the relative differences in results between cosmologies for halo substructure as a function of its host halo. In light of the conclusion by Power and Knebe (2006) that the internal properties of CDM halos are relatively unaffected by a finite box size, it is reasonable to assume that our box's lack of larger wavelength modes does not affect any of our conclusions.

Table 2.1 Summary of the cosmological simulations

Run	Ω_m	σ_8	n_s	h	m_p (h^{-1} $M_\odot$)
$\Omega_m = 0.35$	**0.35**	0.8	0.95	0.7	$\mathbf{11.3 \times 10^6}$
$\Omega_m = 0.19$	**0.19**	0.8	0.95	0.7	$\mathbf{6.14 \times 10^6}$
$\sigma_8 = 0.9$	0.27	**0.9**	0.95	0.7	8.72×10^6
$\sigma_8 = 0.7$	0.27	**0.7**	0.95	0.7	8.72×10^6
$n_s = 1.0$	0.27	0.8	**1.0**	0.7	8.72×10^6
$n_s = 0.9$	0.27	0.8	**0.9**	0.7	8.72×10^6

NOTE. — $\Omega_b = 0.045$ in all runs. $\Omega_\Lambda + \Omega_m = 1$. The cosmological parameters used in the seven simulations are shown above. Each uses *WMAP7* values with one parameter (highlighted in boldface) varied at a time.

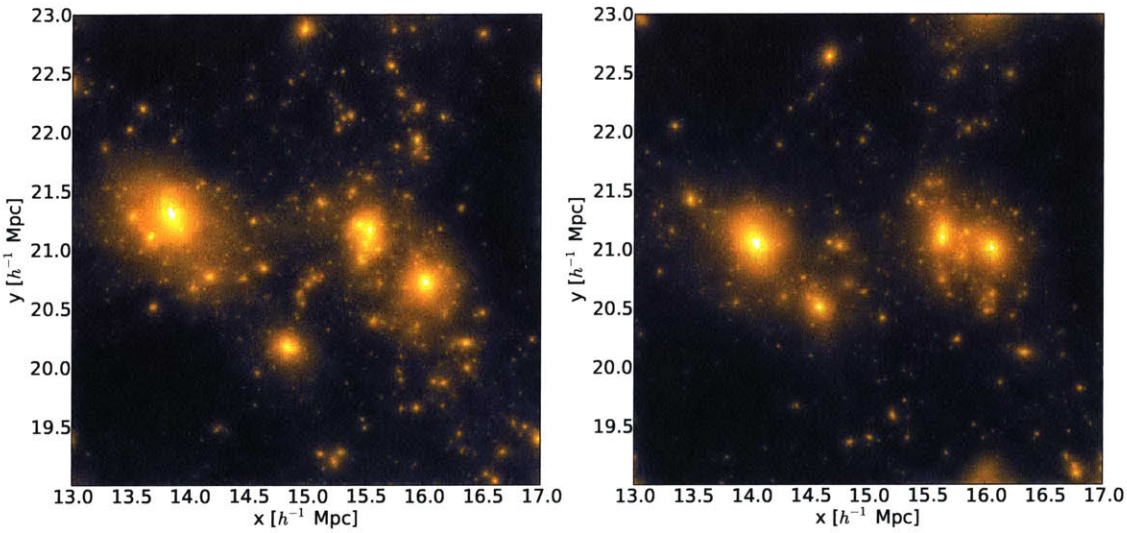

Figure 2-1 Projected dark matter column density ($\log \rho$) of corresponding halo objects in the *WMAP7* cosmology simulation with $\sigma_8 = 0.8$ at z = 0 (left panel) and the $\sigma_8 = 0.9$ simulation (right panel). The largest four halos in each visualization are matched pairs, as explained in Section 2.2.4. The cosmology with higher σ_8 results in the "same" halos forming earlier and merging earlier. This is visually exemplified by the pair of halos in the right side of each panel that are in the process of merging in the $\sigma_8 = 0.9$ simulation, but not in the *WMAP7*, $\sigma_8 = 0.8$ simulation.

2.2.2 Halo Finders

Two different halo finders are used throughout our analysis so as to not bias our results by using a particular algorithm: ROCKSTAR (Behroozi et al. 2013b) and SUBFIND (Springel et al. 2001). As each halo finder produced the same conclusions for the relative effect of cosmology, all figures present data from ROCKSTAR only. Any important systematic differences due to the halo finder are discussed in the text of the results.

• ROCKSTAR (Robust Overdensity Calculation using K-Space Topologically Adaptive Refinement) is a phase-space halo finder which considers the position *and* velocity of all particles to determine the location of halos in an N-body simulation. The algorithm initially selects groups of particles based on a 3D Friends-of-Friends (FOF) algorithm with a large linking length (b = 0.28). Within each FOF group, ROCKSTAR builds more FOF groups in a hierarchical fashion by adapting the linking length such that a user-specified percentage of the particles contained within a subgroup is also contained within the parent group. Once complete, these FOF groups are used to generate seed halos from the innermost level of the tree. This process is repeated until all particles at every level of the FOF group have been assigned to a halo. A further unbinding procedure is carried out to determine the final list of particles contained within each halo. For more details, see Behroozi et al. (2013b).

• SUBFIND defines halos as locally over-dense, gravitationally bound groups of particles and begins by conducting a FOF search of the simulation volume. The local density around each particle is calculated using a smoothing Kernel over its nearest neighbors. Whenever a saddle point in a density contour which bridges two regions is reached, the smallest of the two is considered a subhalo candidate. As with the other halo finders, these candidates then undergo an unbinding procedure to produce a list of halos and subhalos. For more details, see Springel et al. (2001).

It is now well established that different halo finders have various strengths and weaknesses depending on the environment in which they are used (Knebe et al. 2011, Onions et al. 2012, Knebe et al. 2013). For example, since SUBFIND is based upon an

over-density criterion, its ability to identify substructure is strongly dependent on the radial position of the structure from the host (Muldrew et al. 2011). ROCKSTAR is based on a phase space algorithm and thus it does not suffer from the same problem. Onions et al. (2012) found that for properties which rely on particles near the outer edge of the subhalo (e.g., total halo mass), the majority of available halo finders agree to within 10 per cent. Basic properties such as mass or the maximum circular velocity can also be reliably recovered if the subhalo contains more than 100 particles. To ensure we are not resolution limited, we only include halos and subhalos which have 300 particles or more, unless otherwise specified.

The mass and radius of host halos used throughout this work are r_{200} and M_{200}. r_{200} is defined such that the average density of a halo within r_{200} is $\bar{\rho} = \Delta_h \rho_{crit}$ where $\Delta_h = 200$, independent of cosmology as in Jenkins et al. (2001), and $\rho_{crit} = 3{H_0}^2/8\pi G$. The mass M_{200} is then simply found as $M_{200} = \frac{4}{3}\pi r_{200}^3 \Delta_h \rho_{crit}$.

For subhalos, which experience tidal stripping, m_{200} is a less meaningful quantity. We therefore define subhalo mass, $M_{\rm sub}$, as the total bound mass. This is computed as an extra parameter within ROCKSTAR's source code. Defining the radius of a subhalo is even more troublesome. The distance to the furthest bound particle depends on the random kinematics of a few particles, and the tidal radius is not a spherically symmetric value. Since defining a useful measure of the outer edge of a subhalo is difficult, we avoid using it. In the one analysis where it could be used, in finding subhalo concentration, we instead use a concentration definition that is independent of outer radius.

2.2.3 Merger Trees

For the merger tree component of this work (Section 2.3.3), we use only the merger tree generated by ROCKSTAR CONSISTENT TREES on the 64 snapshots per simulation. Behroozi et al. (2013c) have shown that the merger tree catalogues produced by ROCKSTAR are consistent in determining halo masses, positions, and velocities when compared to merger trees constructed via different methods, e.g., BDM (Klypin et al., 1999a, 2011) and SUBFIND.

2.2.4 Halo Matching Across Simulations

Since the same white noise field was used in each simulation, particles tagged with the same ID in each simulation have a low mean displacement from each other, ~ 4 h^{-1} kpc at $z = 0$. The aggregate of these similarly distributed particles are halos of similar size, position, and formation history between simulations. These "same" halos can be matched to each other, enabling a direct assessment of how cosmological parameters affect properties of individual halos and subhalos. This analysis is carried out in Section 2.3.4. Figure 2-1 shows a visual example of matched halos in two different simulations.

Halos are matched by their particle content. In a given set of two simulations, one halo from each – a halo pair – is compared and given a matching strength value based on the number of particles they have in common (same ID), weighted by the gravitational boundedness of each particle. For a pair of (sub)halos A and B, there are two match values, one where particles are weighted by boundedness within A, and one where particles are weighted by boundedness within B. In order for a pair to qualify as a match, (sub)halo B must be the best possible match out of all (sub)halos in its simulation to (sub)halo A, and (sub)halo A must be the best possible match to (sub)halo B. This dual requirement eliminates cases when a small fragment of a halo matches to a larger encompassing halo.

The particle weighting of the i^{th} most bound particle (starting from 0) in a (sub)halo is

$$W_i \propto \frac{n-i}{n} \tag{2.1}$$

where n is the number of particles in the halo. The total match value is given as $\sum W_i$ for all particles in common to the (sub)halo pair. In order to sum to unity for a perfect match, the normalization constant is chosen as $\frac{2n}{n+1}$. A perfect match is when all particles in (sub)halo A are found in (sub)halo B or vice-versa. Additional constraints of $\sum W_i > 0.2$ and $n > 40$ are imposed to eliminate uncertain matches. Host halos are matched first, then subhalos are matched within matched host halo pairs.

2.3 Results

This section presents results on the differences and lack of differences induced by varying cosmology on halo and subhalo properties. Effects on host halos are presented in Section 2.3.1. Section 2.3.2 shows results for subhalo characteristics at $z = 0$, Section 2.3.3 presents subhalo characteristics from merger tree analysis, and Section 2.3.4 compares subhalo characteristics for matched subhalos.

2.3.1 Averaged Properties of Host Halos

While the focus of this paper is on how cosmological parameters affect substructure, we first verify and summarize for reference how cosmological parameters affect the host halo mass function and concentration.

2.3.1.1 Mass Function

In accordance with Jenkins et al. (2001), we compute $\mathrm{d}n/\mathrm{d}\log M$ where $\mathrm{d}n(M)$ is the number of host halos on an infinitesimal mass interval centered at M, and $\mathrm{d}\log M$ is the logarithm of the width of the mass interval. This function shows a characteristic monotonic trend of higher abundances of low mass objects. Rather than presenting the mass function directly, we compute the ratio of the mass function at redshift $z = 0$ in each cosmology to our fiducial simulation mass function in order to accentuate any differences. This is shown in Figure 2-2. We also use the Extended Press-Schechter formalism developed by Sheth and Tormen (2002) with the Eisenstein-Hu transfer function (Eisenstein and Hu, 1998) to compute analytic estimates of these mass function ratios. We compute error bars using a jackknife method (Tinker et al., 2008). We create five sub-volumes, each with $\frac{4}{5}$ of the original volume, by removing a different fifth of the original volume for each sub-volume. Error bars are then the $1\,\sigma$ standard deviation of the mass function ratio as computed in each sub-volume. Our data agree with analytic estimates except on the high mass end above $10^{12}\ h^{-1}\mathrm{M}_{\odot}$, where there are fewer halos and Poisson noise becomes important.

In the top panel of Figure 2-2, we show that higher σ_8, which means higher ampli-

tude for primordial density fluctuations, produces more halos above the characteristic mass scale of collapse at $z = 0$, $M^* \approx 10^{13}\, h^{-1}\mathrm{M}_\odot$, and fewer halos below. Intuitively, a higher initial density amplitude causes more mass to end up in large halos after hierarchical merging, leaving less mass available for low mass halos. This agrees with a study by Guo et al. (2013). The reverse is true for lower σ_8.

In the middle panel, we plot the scalar spectral index, n_s, which controls the slope of the initial power spectrum, $P_0(k) \propto k^{n_s}$. Larger values of n_s correspond to more initial power on small scales (and less on large scales), and thus more collapsed objects early on. The earlier collapsing small objects then hierarchically merge into larger objects. As evidenced by Figure 2-2 larger n_s continues to correspond to more collapsed objects by $z = 0$ for the mass range considered.

We also show the effect of more exaggerated variations in Ω_m. More matter content in the universe translates into objects of higher mass, and more objects, as demonstrated in the bottom panel.

Since $z = 0$ subhalos form independently of their hosts at high redshifts, these qualitative mass function differences apply just as well to subhalos before accretion. Any changes in mass and abundance are due to the dynamics of subhalo-host interactions, including tidal stripping, dynamical friction, ad collisions with other subhalos.

2.3.1.2 Host Concentration

The density profile of dark matter halos are well approximated by the spherically symmetric Navarro, Frenk, & White (NFW) Profile (Navarro et al., 1996b). The profile is defined by

$$\rho(r) = \rho_{crit} \frac{\delta_c}{(r/r_s)(1 + r/r_s)^2}, \tag{2.2}$$

where the dimensionless δ_c is the characteristic overdensity, r_s is the scale radius, and ρ_{crit} is the critical density of the universe. Concentration, defined by $c_{200} \equiv r_{200}/r_s$, gives a measure of how centrally concentrated the particles in the halo are. Concentration could also be defined in terms of the spherically symmetric Einasto

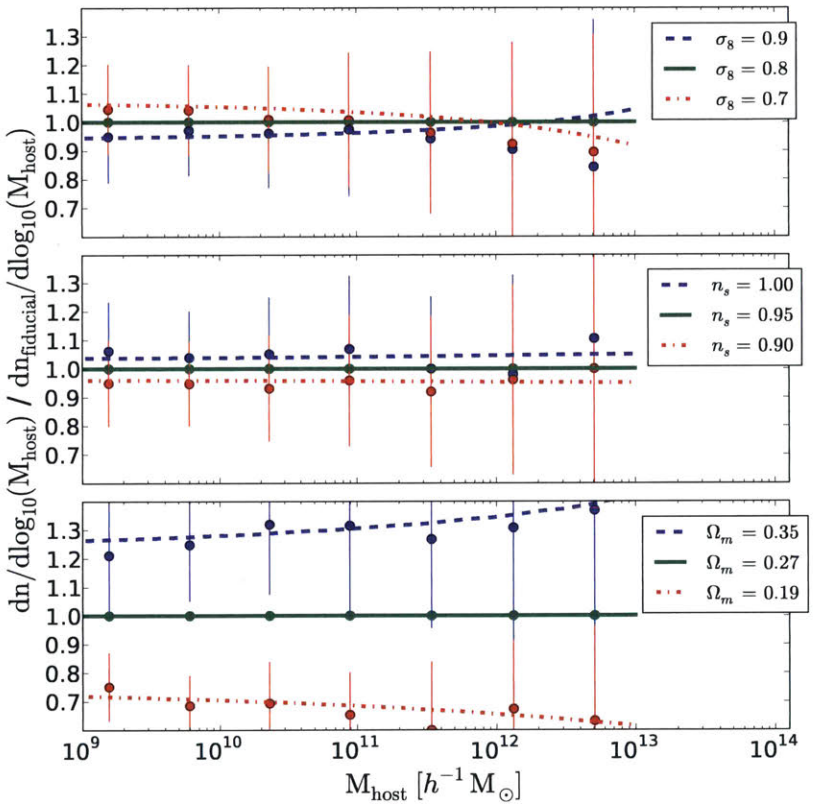

Figure 2-2 Ratio of host halo mass functions to our fiducial simulation for different cosmological parameters. Dots correspond to data points computed from ROCKSTAR halo finder data. Lines correspond to the Sheth & Tormen analytical prediction. Error bars are computed using a jackknife method. As anticipated, in cosmologies where structure forms earlier there are more halos at $z = 0$

density profile, given as

$$\rho(r) = \rho(r_{-2}) \exp\left[-\frac{2}{\alpha}\left(\left(\frac{r}{r_{-2}}\right)^{\alpha} - 1\right)\right], \tag{2.3}$$

where r_{-2} is the radius at which the logarithmic slope of the profile is -2, the shape parameter α is an extra free parameter to fit to, and $\rho(r_{-2})$ is the density at r_{-2} (Einasto, 1965). In this case the concentration would be $c_{200} \equiv r_{200}/r_{-2}$. Furthermore, many studies use outer radii other than r_{200}. Since each definition quantifies the same qualitative idea, we consider only the first definition in terms of the NFW profile. We use the default r_{200} as determined by the halo finder, but compute r_s directly from the particle data of the halo. To facilitate a direct comparison to past studies, we find r_s by fitting an NFW profile to the density profile of each halo, determined by binning particles in 16 bins equally spaced in log space from $\log_{10}(\frac{r}{r_{200}}) = -1.25$ to 0, as in Duffy et al. (2008).

Host halo concentrations have been studied extensively (e.g., Zhao et al. 2009). It is now well established that halos which collapse earlier have higher concentrations (Neto et al. 2007, Duffy et al. 2008, Macciò et al. 2008, Giocoli et al. 2010). This is simply because the universe was of higher density at early times. One would then expect cosmologies with a higher M^* to have higher concentrations. Similarly, halos with a higher mass should have smaller concentrations since they formed later. In particular, for increasing values of σ_8 we expect the concentration to increase since it reflects the background density of the universe at the time when the halo formed.

To test this, we consider all "relaxed" halos whose hosts satisfy the following criteria similar to Neto et al. (2007):

- The host must contain at least 600 particles.
- The substructure mass fraction of a given host must be, $f_s < 0.1$.
- The offset parameter or center of mass displacement must be s < 0.07 where $s = |r_c - r_{cm}|/r_{200}$. r_c and r_{com} are the center of halo peak density and mass respectively.

- The virial ratio, $2T/|U|$, must be less than a pre-set value of 1.35, where T is the total kinetic energy and U is the gravitational potential self-energy of the host halo.

This combination of mass and kinematic information encoded in the substructure mass fraction, offset parameter, and virial ratio ensures that the hosts in our relaxed sample are well fitted by NFW profiles (Navarro et al., 1996b). Unrelaxed halos are more difficult to accurately fit a scale radius to. Figure 2-3 shows the concentration-mass ($c - M$) relation for each of our simulations. Using the mean and rms deviation of $\log_{10}(c_{200})$, we fit the binned $c - M$ relation using,

$$\log_{10}(c_{200}) = A \log_{10}(M_{vir}) + B, \tag{2.4}$$

where $\log_{10}(c_{200})$ and M_{vir} are the mean values in each bin. For the error of the mean in each bin, σ_c, we use the rms deviation of $\log_{10}(c_{200})$ divided by the square root of the number of objects in each bin. For each fit, we define

$$\chi^2 = \sum_{j=1}^{N_b} \left(\frac{\log_{10}(c_{200}) - \log_{10}(c_{200fit_j})}{\sigma_{\bar{c}_j}} \right)^2, \tag{2.5}$$

where N_b is the number of bins over which the fit is performed and c_{200fit_j} is obtained from the best fit of Equation (2.4), and find the values of A and B such that the reduced chi-squared, $\bar{\chi}^2$ (χ^2 divided by the number of degrees of freedom), is minimized. All error bars and best fit functions in future figures are found using this same method.

Our data and best fit functions for the relaxed sample are shown in Figure 2-3. We also add for comparison the relation for c_{200} derived by Duffy et al. (2008) who used *WMAP5* data ($\sigma_8 = 0.796$, $n_s = 0.963$, $\Omega_m = 0.258$). The data confirms that higher M^* cosmologies and smaller host halos do in fact have higher concentrations. The slopes of each $c - M$ relation range from -0.060 to -0.097 with a typical value of -0.083, but do not vary systematically with cosmological parameters. Furthermore, the slope is sensitive to within ± 0.008 to how the host halos are binned and the halo

mass range considered. Therefore we can only definitively confirm how the amplitude of the $c - M$ relation is affected by cosmological parameters, and not the slope. The slopes obtained do agree, within error, with the slope of -0.092 obtained by Duffy et al. (2008) for relaxed halos.

We further note that the slope and magnitude of the $c-M$ relation is very sensitive to the method for finding r_s and the halo finder. By default, ROCKSTAR computes r_s by fitting an NFW profile to a density profile found by dividing particles into 50 radial bins such that each bin contains the same mass (Behroozi et al., 2013b). This results in a much steeper slope for all cosmologies, with a typical slope of -0.15. Furthermore, the concentrations computed using ROCKSTAR's r_{200} and particle assignments are systematically higher than those found using SUBFIND's parameters. Using SUBFIND, our fiducial simulation's $c - M$ relation is very similar to the Duffy et al. (2008) relationship in magnitude and slope, which was also computed using SUBFIND. The $c - M$ relationships found using ROCKSTAR, however, are on average 12% greater in overall magnitude. Both halo finders do, however, agree on the relative differences between cosmologies. The large dependencies of concentration on its method of computation are examined in greater detail by Meneghetti and Rasia (2013).

2.3.2 Averaged Properties of Subhalos

Subhalo abundances and properties depend dramatically on the size of their host halo. Any useful comparison between simulated and observed subhalo distributions thus requires comparing distributions within hosts of the same size. In this section, we explore the effects of cosmology on averaged subhalo properties as a function of their host halo mass.

2.3.2.1 Subhalo Mass Function

Similar to the host halo mass function, the subhalo mass function, SHMF, counts subhalo abundance within a chosen host halo as a function of mass. Whereas the

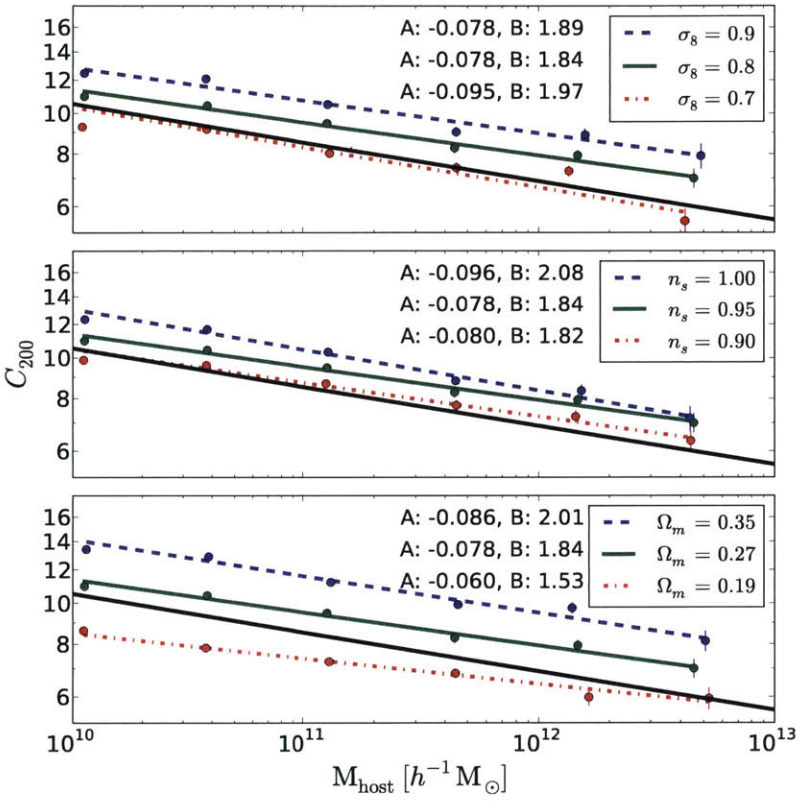

Figure 2-3 The averaged host halo concentration as a function of host halo mass and cosmology. Best fit lines and best fit parameters as described in Eq. 2.4, and error of the mean bars are shown. The $c - M$ relation derived by Duffy et al. (2008) (solid black line) is consistent in its slope to our findings. Its overall magnitude, however, is lower primarily due to its computation using SUBFIND and our data using ROCKSTAR. Cosmologies with higher M^* lead to more concentrated halos, but do not affect the slope of the $c - M$ relation.

host halo mass function has been studied extensively by numerical simulations and agrees well with fully analytic predictions, see Figure 2-2, only semi-analytic models exist for the SHMF (Oguri and Lee, 2004; van den Bosch et al., 2005; Springel et al., 2008; Zentner and Bullock, 2003; Gao et al., 2004). This is due to the more complex effects of collisions, dynamical friction, and tidal stripping changing subhalo masses over time.

We investigate the SHMF, $\mathrm{d}n/dM_{sub}$ vs M_{sub}, where $\mathrm{d}n$ is the number of subhalos in a mass interval $\mathrm{d}M_{sub}$. Within each particular cosmology, we find that the SHMF per unit host halo mass (i.e., dividing the differential abundance by the mass of the host halo), yields a universal function. This confirms the same result found in Gao et al. (2004) and van den Bosch et al. (2005). We use this trait to average the normalized subhalo mass functions for all hosts halos with $M_{host} > 10^{12.5}\ h^{-1}\mathrm{M}_{\odot}$ to account for halo-to-halo variance of the SHMF. Lowering the M_{host} mass threshold simply increases variance, due to resolution effects, without changing the value of the average mass function. The top panel of Figure 2-4 shows the SHMF per unit host halo mass scaled for a Milky Way sized host of $M_{host} = 0.84 \times 10^{12}\ h^{-1}\mathrm{M}_{\odot}$ (1.2×10^{12} $\mathrm{M}_{\odot}$). We also indicate the magnitude of $1\,\sigma$ halo-to-halo variation with bars on each point to compare changes due to cosmology with inherent variance between halos. This strategy of handling variance by finding characterizations that are independent of host halo mass, then averaging the characterizations over all host halos, is used throughout the remainder of Section 2.3.2.

Varying the parameters σ_8 and Ω_m has no noticeable effect on the subhalo abundance at $z = 0$. Our simulations with higher values of n_s have a greater abundance of subhalos in each bin. However, the error on this overall trend is equal to the magnitude of the trend so it may have arisen due to chance. Additionally, variations from halo-to-halo in the SHMF exceed any possible small cosmological effect. The best fit to the data produced from ROCKSTAR data is given by

$$\frac{\mathrm{d}n}{\mathrm{d}M_{sub}} = K \times \left(\frac{M_{sub}}{h^{-1}\mathrm{M}_{\odot}}\right)^{-\alpha} \frac{M_{host}}{h^{-1}\mathrm{M}_{\odot}} \tag{2.6}$$

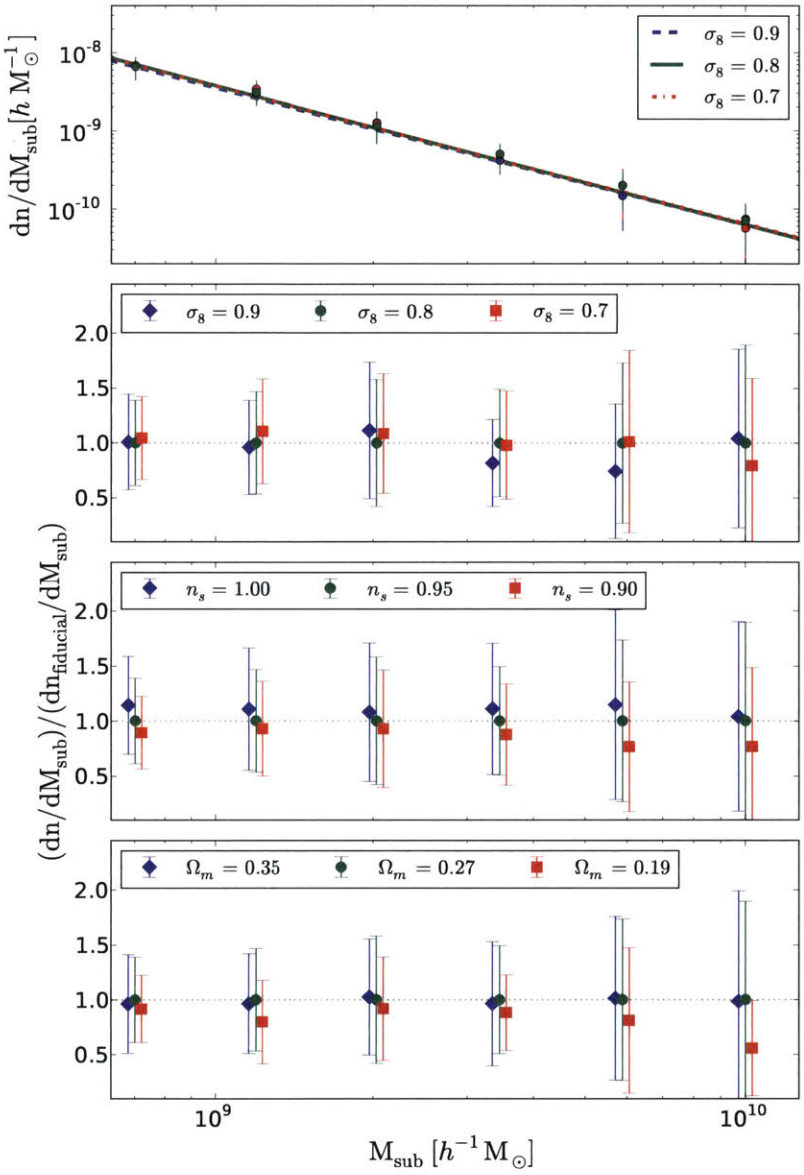

Figure 2-4 *Top Panel:* Subhalo mass function normalized to host halo mass and scaled to a Milky Way sized host of $0.84 \times 10^{12}\ h^{-1}\mathrm{M}_\odot$ ($1.2 \times 10^{12}\ \mathrm{M}_\odot$), averaged over all host halos above $10^{12.5}\ h^{-1}\mathrm{M}_\odot$ for cosmologies with different σ_8. Resolution constraints on low masses and large Poisson errors for high masses constrain the subhalo mass range considered. Best fit lines and $1\,\sigma$ halo-to-halo variation bars, not error bars, are shown. *Bottom Panels:* Ratio of subhalo mass function to the fiducial *WMAP7* subhalo mass function. Data points for the non *WMAP7* cosmologies are shifted slightly left and right of their true values to help distinguish them. The effects of varying σ_8 and Ω_m are consistent with no change in the SHMF. A small trend of higher abundance with higher n_s exists but is also consistent with no change in the SHMF within error. Variation on a halo-to-halo basis dominates any effects of cosmology on the SHMF as seen in the $1\,\sigma$ variation bars.

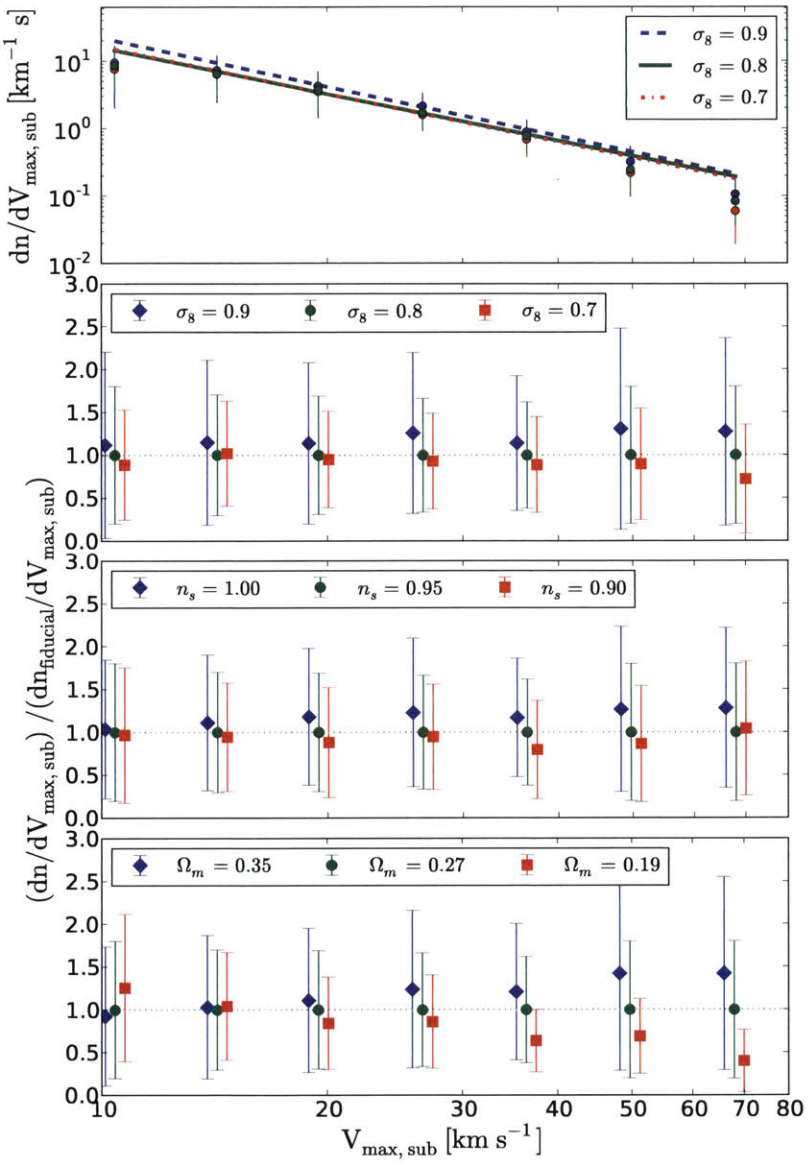

Figure 2-5 *Top Panel:* Subhalo V_{max} function normalized to host halo V_{max} and scaled to a Milky Way sized host of 220 km s^{-1}, averaged over all host halos above $10^{12.5}$ $h^{-1}\mathrm{M}_{\odot}$ for cosmologies with different σ_8. Best fit lines and 1 σ halo-to-halo variation bars are shown. *Bottom Panels:* Ratio of subhalo V_{max} function to the fiducial *WMAP7* subhalo V_{max} function. Data points for the non *WMAP7* cosmologies are shifted slightly left and right of their true values to help distinguish them. Higher values of each of the cosmological parameter (higher M^*) result in higher abundances of V_{max} subhalos over the mass range considered. The cosmologies chosen correspond to $\sim 18\%$ greater abundance for the higher M^* cosmologies and $\sim 10\%$ less abundance for the lower M^* cosmologies. Variation on a halo-to-halo basis dominates any effects of cosmology on the V_{max} function as seen in the 1 σ variation bars.

Where $K = 4.5 \pm 0.3 \times 10^{-5} h\ \mathrm{M}_{\odot}^{-1}$ and $\alpha = 1.78 \pm 0.04$. The same results and fit values were reproduced within error using SUBFIND, the halo finder used in most previous studies. The power law exponent of -1.78 falls between the values of -1.73 reported by Helmi et al. (2002) and -1.9 reported by Springel et al. (2008), van den Bosch et al. (2005) and Gao et al. (2004). Individual halos in this study have SHMFs with α ranging from 1.65 to 1.95, encompassing the range of values reported in the literature.

Following from an unchanging SHMF per host halo mass, the subhalo mass fraction is also not changed by variations in cosmological parameters. This was directly confirmed by simulation data.

2.3.2.2 Maximum Circular Velocity

While mass may be the most intuitive description of a subhalo's size, the maximum circular velocity, V_{max}, is a related measure that is easier to ascertain observationally and more robust to measure from simulations. It is defined as the maximum velocity of an orbiting body in the potential of a halo:

$$V_{max} = \max\left(\frac{GM(<r)}{r}\right)^{\frac{1}{2}}. \tag{2.7}$$

Unlike mass, V_{max} is not sensitive to the poorly defined boundary between the halo and the background (Kravtsov, 2010). Additionally, it avoids the arbitrariness that plagues any definition of mass. While closely related, V_{max} does not just trace mass, but also has a weak dependence on concentration: higher concentration leads to higher V_{max} (Bullock et al., 2001). Thus, if there is no cosmology dependence on the SHMF, there should in principle still be a weak dependence on V_{max} due to the dependence of subhalo concentration on cosmology (see Figure 2-9). For these reasons, we repeat the subhalo abundance analysis done for subhalo mass with subhalo V_{max} instead. The subhalo V_{max} function is computed for each host halo, normalized to the host's V_{max}, averaged over all host halos with $M_{host} > 10^{12.5}\ h^{-1}\mathrm{M}_{\odot}$, then scaled to a host with $V_{max} = 220$ km s^{-1}. The resulting function and the ratio of it to our fiducial

simulation's function for each cosmology are shown in Figure 2-5. We indeed find that this V_{max} function depends weakly on the cosmological parameters considered. The greater value of each cosmological parameter increases the abundance of subhalos as a function of V_{max} by $\sim$ 18% and the lower values of each parameter decreases abundance by $\sim$ 10%. Still, the error on this trend is substantial at $\sim$ 10%, and variation from halo to halo exceeds this effect.

Just as for the SHMF, we fit an exponential function to the data. This equation is

$$\frac{\mathrm{d}n}{\mathrm{d}V_{max,\ sub}} = K_v \times \left(\frac{V_{max,\ sub}}{\mathrm{km\ s^{-1}}}\right)^{-\alpha_v} \frac{V_{max,\ host}}{\mathrm{km\ s^{-1}}} \tag{2.8}$$

and the best fit parameters to the fiducial simulation are $K_v = 12.5 \pm 0.9$ km^{-1} s and $\alpha_v = -2.3 \pm 0.2$.

V_{max} is of particular interest in relation to the Missing Satellite Problem. Wang et al. (2012) discuss the disparity between subhalo abundance above $V_{max} = 30$ km s^{-1} observed in the Milky Way and the abundance predicted in simulations of a 10^{12} $h^{-1}\mathrm{M}_\odot$ halo. The Aquarius simulations predict eight subhalos, whereas the Milky Way has only three with V_{max} above 30 km s^{-1} (the Large Magellanic Cloud, the Small Magellanic cloud, and the Sagittarius Dwarf). They discover that this number is very sensitive to the host halo mass and very insensitive to the cosmology. With the data to test this more explicitly, we show in Figure 2-6 the average number of subhalos with $V_{max} > 30$ km s^{-1} as a function of host halo mass. We find that this V_{max} function is consistent with being independent of the variation in cosmological parameters considered, confirming the result of Wang et al. (2012).

2.3.2.3 Spatial Distribution

The subhalo spatial distribution in terms of its number density as a function of radius is also studied. Following Springel et al. (2008), we first investigate the number density profile, $n(r)/<n>$, where $n(r)$ is the local number density of subhalos, and $<n>$ is the average subhalo number density within the virial radius. With the radial distance normalized to r_{200}, we discover that this distribution takes on a characteristic form for

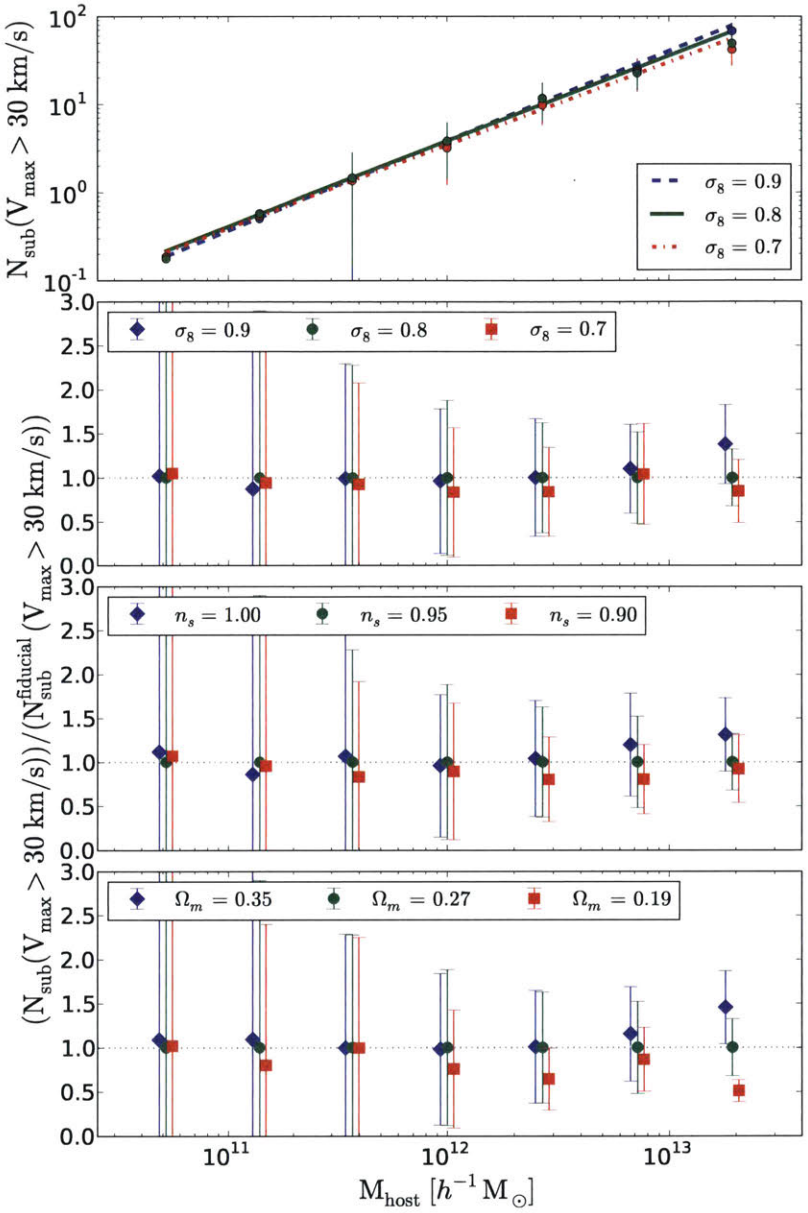

Figure 2-6 *Top Panel:* Average number of subhalos with $V_{max} > 30$ km/s as a function of host halo mass and cosmology for variations in σ_8. The number of subhalos above this threshold is not affected by cosmology, but is substantially affected by the mass of the host halo. Halo-to-halo variation is shown in the 1 σ variation bars. *Bottom Panels:* Ratio of the number of subhalos with $V_{max} > 30$ km/s in a given cosmology to that of the fiducial *WMAP7* cosmology. Data points for the non *WMAP7* cosmologies are shifted slightly left and right of their true values to help distinguish them. All changes in cosmology tested are consistent with having no effect on this function. Variation on a halo-to-halo basis dominates any effects of cosmology as seen in the 1 σ variation bars.

all host halos, regardless of mass, for $r/r_{200} > 0.4$. We tested for this independence of mass by computing the average distribution over several host halo mass intervals in the range $10^{10}\ h^{-1}\mathrm{M}_{\odot} < M_{host} < 10^{13.6}\ h^{-1}\mathrm{M}_{\odot}$. No mass dependent difference was found in the profiles above $r/r_{200} > 0.4$. Below this threshold, corresponding to only 6.4% of the halo volume, there is a clear mass dependent trend with smaller host halos having higher number densities than larger host halos. However, this region is also one where halo finders have difficulty identifying subhalos, so it is unclear whether the effect is real, or a halo finding artifact. We resolve to exclude this range from the cosmology comparison analysis.

Figure 2-7 shows the characteristic subhalo radial distribution averaged over all host halos with $M_{host} > 10^{12.5}\ h^{-1}\mathrm{M}_{\odot}$ for each cosmology. Once again, this distribution is found to be independent of the cosmological parameters studied.

2.3.2.4 Subhalo Concentration

Using similar methodology presented in Section 2.3.1.2, we examine the subhalo $c-M$ relation for subhalos within relaxed hosts at redshift $z = 0$. Due to difficulty in defining subhalo mass, we also explored the subhalo $c - V_{max}$ relation and found all of the same results. Since both options lead to the same conclusions, we choose to present only the $c - M$ relation in order to compare to previous literature and the host halo $c - M$ relation. The concentration, c, is once again defined by the ratio $c \equiv r_{200}/r_s$. However, directly computing r_{200} and r_s for subhalos is troublesome. As noted in Section 2.3.1.2, the $c - M$ relation is very sensitive to how r_s in computed. Whereas for host halos the method used by Duffy et al. (2008) of dividing particles into 16 logarithmic bins to make a density profile, then fitting an NFW profile to find r_s works consistently well, this method does not work reliably for subhalos which typically have too few particles to populate all 16 bins. The default method from ROCKSTAR overcomes this issue by choosing 50 bins such that each contains an equal number of particles. However, we find that this method results in erratically varying subhalo concentrations with standard deviations of concentration per subhalo mass interval exceeding 80.

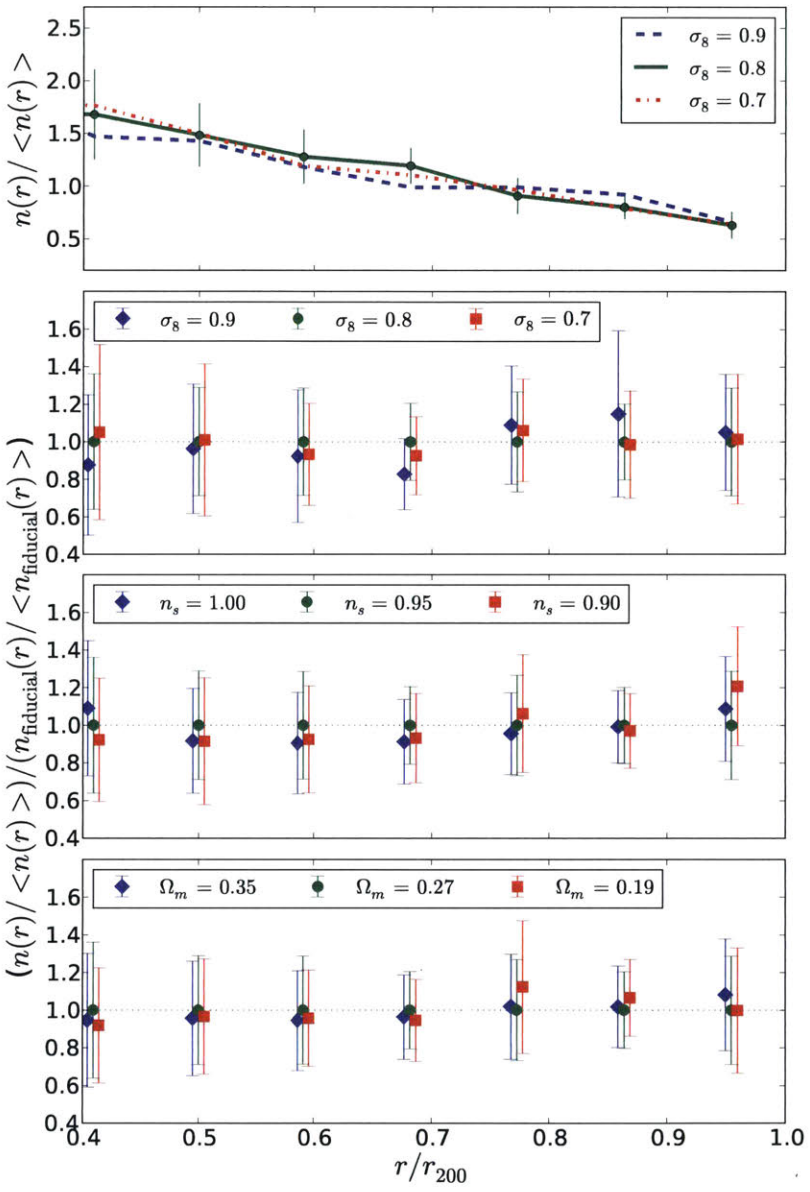

Figure 2-7 *Top Panel:* The averaged subhalo number density as a function of r/r_{vir} normalized to the mean subhalo number density for cosmologies with varied σ_8. This function is found to not vary with the mass of the host halo nor the cosmological parameters considered for $r/r_{vir} > 0.4$. At smaller radii, halo finding becomes too uncertain to make any claim. Variation bars show 1σ variations in the density profiles of individual halos. *Bottom Panels:* Ratio of the subhalo spatial distribution as in the top panel to that of the to the fiducial *WMAP7* cosmology for all cosmologies. Data points for the non *WMAP7* cosmologies are shifted slightly left and right of their true values to help distinguish them. All changes in cosmology tested are consistent with having no effect on this function. Variation on a halo-to-halo basis dominates any effects of cosmology on subhalo spatial distribution as seen in the $1\,\sigma$ variation bars.

We therefore avoid computing r_s directly and instead use other parameters and properties of an NFW profile to infer c. (Klypin et al., 2011) uses V_{max} and M_{200} to numerically solve for c assuming an NFW profile. While we did test out this method, M_{200} is nonsensical for very tidally stripped subhalos where r_{200} exceeds the furthest bound particle. This occurs in $\approx 5\%$ of subhalos above $10^9\ h^{-1}\mathrm{M}_\odot$. We thus finally turn to the method used by Springel et al. (2008) which also assumes an NFW profile, but uses V_{max} and the radius where V_{max} occurs, r_{vmax}, to compute c. The concentration is found numerically from the equation

$$\frac{c^3}{\ln(1+c) - c/(1+c)} = .21639 \left(\frac{V_{max}}{H_0 r_{vmax}}\right)^2 . \tag{2.9}$$

The $c - M$ relation from this method for all subhalos with greater than 500 particles is shown in Figure 2-8. As a function of cosmology, we discover that the differences in concentration demonstrated in Figure 2-3 still last in the subhalo population. Cosmologies with higher M^* (higher σ_8, n_s, and Ω_m) lead to subhalos with higher concentration. We further fit the form of Eq. 2.4 to the data and find concentrations similar to the host halos but with more shallow negative slopes. Due to a more limited sample of subhalos than host halos after the selection criteria, ~ 900 vs. ~ 6000 in the $10^{9.64} - 10^{12}\ h^{-1}\mathrm{M}_\odot$ interval, and density profiles that deviate further from NFW, the strength of each fit is worse and the variation in slopes between simulations is much greater for subhalos. Furthermore, using the Klypin method of computing subhalo concentration yields steeper negative slopes than the host halo $c - M$ relation (it yields a median slope of -0.122, in agreement with the slope of -0.12 found by Klypin et al. (2011)), and a $70 - 95\%$ increase in concentration relative to hosts for the mass range considered. We therefore cannot definitively claim a characteristic slope value or magnitude for subhalo concentrations and caution that results are very sensitive to the method used to compute halo concentration. The Klypin method does nonetheless agree with the relative differences between cosmologies, and we conclude that varying cosmology does have a real effect on subhalo concentrations.

With a consistent computation of concentration for host halos and subhalos ac-

cording to *Eq.* 2.9, we find that the median slope of the $c - M$ relation for hosts is -0.06, which is consistent with the slopes found for the subhalos. The concentration of subhalos in the same mass intervals is typically $\approx 15\%$ greater. Merger tree histories of subhalos demonstrate that their concentrations do in fact continually rise after accretion, and at a faster rate than host halos. This can be explained by a change in subhalo density profiles due to tidal stripping and tidal heating. Hayashi et al. (2003) models the density profiles as a modified NFW function that changes over time as a function of the ratio of current subhalo mass to its mass at infall. The modified NFW profile and supporting fitting functions are given in Hayashi et al. (2003) as equations 8, 9, and 10. Following this model, both V_{max} and r_{vmax} decrease as a function of mass loss, and thus time since accretion. However, since r_{vmax} decreases faster than V_{max}, subhalo concentration as computed in Eq 2.9 increases monotonically. These trends are confirmed in the merger tree histories of the subhalos, but with significant variance. Since the subhalo profile is no longer NFW in this model, Eq 2.9 only approximates the original definition of concentration, $c = r_{200}/r_s$. We therefore also study the poorly defined r_s and r_{200}. This similarly shows r_s decreasing at a faster rate than r_{200}. Below the limit of 500 particles, the concentration of subhalos and host halos systematically and unrealistically decreases as a function of mass, but the relative differences between cosmologies remains.

We further investigate the subhalo concentration as a function of host mass in Figure 2-9. We find that the average subhalo concentration is weakly dependent on the host mass, with larger hosts having more concentrated subhalos. We account for the bias of larger hosts having a different distribution of subhalos by only considering subhalos in $d\log_{10}(M_{sub}) = 0.5$ mass intervals from 10^8 to 10^{11} $h^{-1}\mathrm{M}_\odot$. In every case we observe the same positive slope. We also rule out the possibility that larger hosts tend to have halos which were accreted earlier through studies of the merger tree history, as presented in Sections 2.3.3.1 and 2.3.3.2. Without a strong theoretical motivation for any functional form, we abstain from making a best fit. We speculate that the trend may be due to stronger tidal forces in larger hosts. This would indicate that the subhalo profile models from Hayashi et al. (2003) should depend weakly on

the host halo's size in addition to the fraction of remaining bound subhalo mass. A full explanation is beyond the scope of this paper.

2.3.3 Merger Tree Analysis

While the previous subsection focused on static distributions at $z = 0$, we now investigate the evolution of subhalos as a function of cosmology through merger tree analysis. Halos that become subhalos go through a trajectory of forming, gaining mass, reaching a maximal mass, losing mass due to tidal stripping from a nearby halo, entering r_{200} of their eventual host (accretion), and ultimately merging with their host through stripping and disruption. Subhalos detected at $z = 0$ are ones that have not yet been fully torn apart by their host. The following sections explore these steps with the approach of finding average distributions for subhalos within a host. Rather than presenting results for all cosmologies, we highlight only simulations with varied σ_8, which has historically been the least well constrained parameter of the three studied in this paper.

2.3.3.1 Formation Time of Subhalos

Numerous definitions exist for the formation time of dark matter halos. As discussed in Li et al. (2008), these definitions generally fall into two classes:

1. When a halo reaches a fraction of its final mass or V_{max}.

2. When a halo first reaches a threshold mass or V_{max}.

While both types of definitions can be applied to host halos, the first class does not apply well to subhalos. Subhalos do not monotonically grow in mass, but rather reach a peak before the effects of tidal stripping from a nearby host removes mass. Thus, the time when a subhalo reaches a fraction of its final or maximal mass intertwines its formation and post-accretion history. We therefore use the definition that a halo forms once it first surpasses a threshold mass of $M_{200} = 3 \times 10^8\ h^{-1}\mathrm{M}_\odot$, corresponding

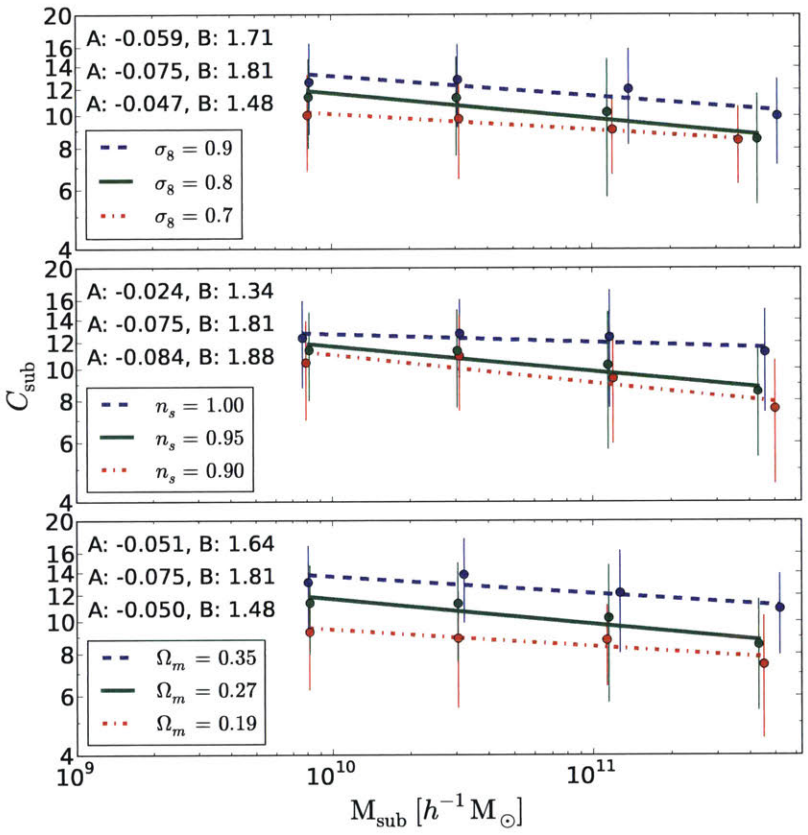

Figure 2-8 The averaged subhalo concentration as a function of subhalo mass and cosmology. Concentration is computed using r_{vmax} and V_{max} according to Eq. 2.9 which assumes an NFW profile. The differences in concentration due to cosmology seen in the host halos remain in the subhalos. Cosmologies with higher M^* lead to subhalos with higher concentration. The same trend holds for the $c - V_{max}$ relationship. The magnitude of the subhalo concentrations cannot be reliably compared to the host halo concentrations of Figure 2-3 since they were computed differently. Variation bars show large $1\,\sigma$ variations in the concentration of individual subhalos. Best fit lines and best fit parameters as described in Eq. 2.4 are also shown. Difficulty in computing concentration reliably below 500 particles per subhalo sets the low mass cut-off in the plot.

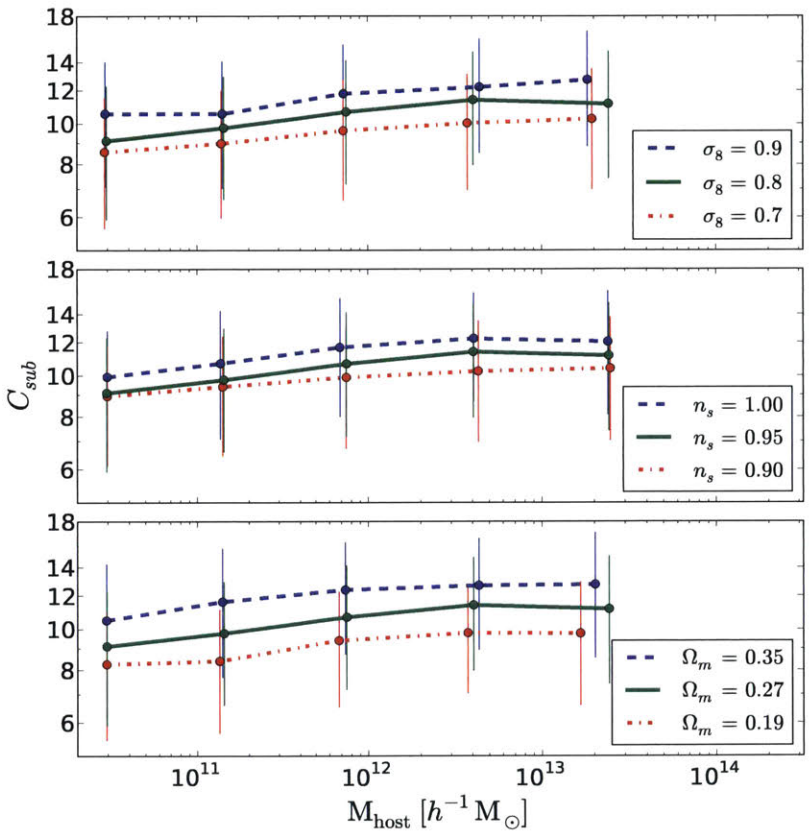

Figure 2-9 The averaged subhalo concentration as a function of host halo mass and cosmology for subhalos between 10^9 and 10^{10} $h^{-1}\mathrm{M}_\odot$ in relaxed hosts. The slight positive trend suggests that larger host halos cause a greater increase in subhalo concentration than smaller host halos, possibly due to stronger tidal forces. Without a theory to predict a form of this trend, no function is fit to the data. Variation bars show large 1 σ variations in the concentration of individual subhalos.

to ~ 35 particles. Changing the mass threshold shifts when halos have "formed", but does not influence how cosmology affects formation times. This mass was chosen to minimize the number of halos which were first detected above the threshold mass, and those which form but never reach the threshold mass. It succeeds in characterizing $\sim 55\%$ of subhalos at $z = 0$. The remaining 45% of subhalos typically have incomplete mass histories, so their formation time would be hard to deduce under any definition of formation time. The threshold mass definition has the useful property that the formation time of subhalos is a tracer of formation time of the oldest stars in each subhalo (Li et al., 2008).

We use the merger trees to study the probability density distribution of subhalo formation times for all surviving subhalos at $z = 0$ within a host. This distribution is similar for all host halos in a simulation, and thus averaged over all $z = 0$ host halos within the mass range $10^{12} < M < 10^{13.7}\ h^{-1}M_{\odot}$ to characterize a particular cosmology. A comparison of the averaged distribution is seen in the top panel of Figure 2-10. Note that the distribution is computed as a function of look-back time for the given cosmology, not redshift. A small effect is seen demonstrating that in cosmologies with higher σ_8, $z = 0$ subhalos tend to be older. The mean age of subhalos in the $\sigma_8 = 0.9, 0.8$, and 0.7 simulations are 11.1, 11.0, and 10.6 billion years respectively. Higher σ_8 indicates that halos of the same mass should collapse sooner, so this result is expected for host halos. For subhalos, the age distribution can change due to their destruction over time, but evidently this effect is not enough to erase the difference in ages between lasting $z = 0$ subhalo populations in different cosmologies.

2.3.3.2 Accretion Time of $z = 0$ Subhalos

In addition to formation time, the same analysis is applied to subhalo accretion time. The probability density distribution of subhalo accretion times is similar for all host halos and thus averaged to reduce scatter. The accretion time is defined by when the center of an eventual $z = 0$ subhalo crosses its host's r_{200}, defined at the time of crossing, and remains inside. This characterization is relevant in that it roughly indicates the time when ram pressure stripping and tidal stripping become important

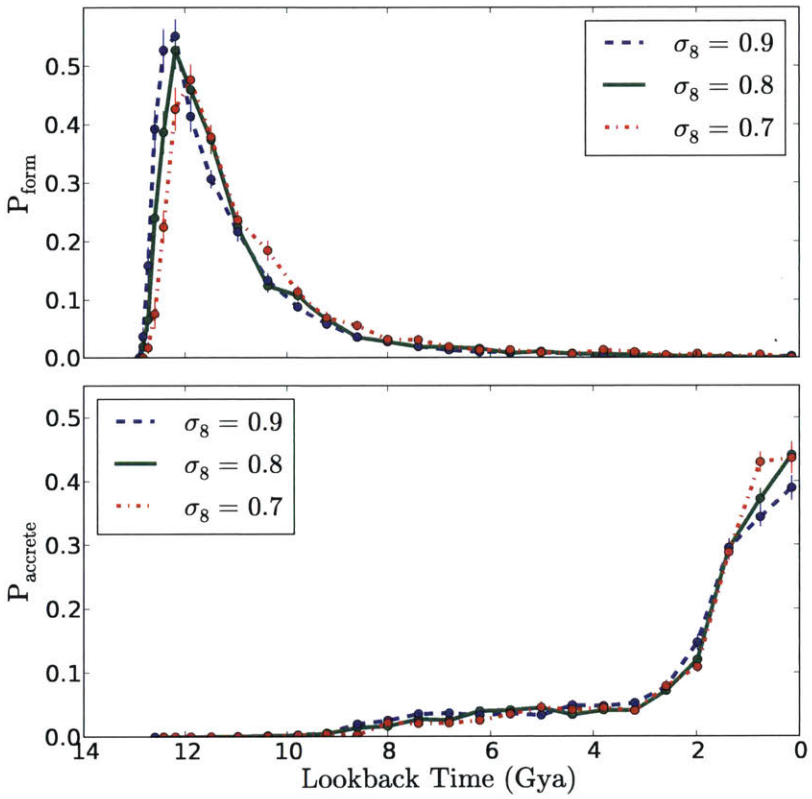

Figure 2-10 *Top Panel*: Average formation time probability density distribution of $z = 0$ subhalos within a host for cosmologies with varied σ_8. Formation time defined by when halos first reach a mass of $M_{200} = 3 \times 10^8\ h^{-1}\mathrm{M}_\odot$. Subhalos lasting to $z = 0$ typically formed earlier in higher σ_8 cosmologies by .1 to .4 billion years for a 0.1 increment in σ_8. *Bottom Panel*: Average accretion time probability density distribution of subhalos within a host for cosmologies with varied σ_8. No significant differences are found between the cosmologies. Error bars show 1 σ uncertainty of each data point.

mechanisms of gas loss, which suppresses star formation (Mayer et al., 2006; Nickerson et al., 2011). The accretion time probability density distribution is shown in the bottom panel of Figure 2-10. Unlike formation time, no difference inconsistent with scatter is observed between cosmologies. This indicates that subhalos are destroyed as a function of when they were accreted in a process that is independent of small variations in cosmology.

2.3.3.3 Maximal Mass

Finally, we study the averaged mass function per unit $z = 0$ host halo mass for the maximal masses of eventual $z = 0$ subhalos. As in Figure 2-4, this mass function is scaled to a host of mass $0.84 \times 10^{12}\ h^{-1}\mathrm{M}_{\odot}$. The total stellar mass/luminosity is related to the maximum mass obtained by the subhalo. Since the more concentrated luminous matter is less easily stripped than dark matter, the maximal mass is a much better indicator of stellar mass than the $z = 0$ mass of a subhalo (Springel et al., 2001; Gao et al., 2004; Guo et al., 2013).

Figure 2-11 compares the M_{max} mass function for two different cosmologies, showing no dependence on σ_8. This suggests varying σ_8 has no effect on the stellar mass/-luminosity functions.

2.3.4 Properties of Matched Subhalos

In order to ascertain the difference between specific subhalos due to cosmology, host halos are matched to each other by a procedure described in Section 2.2.4. About 90% of hosts above $10^{10}\ h^{-1}\mathrm{M}_{\odot}$ are successfully matched. Subhalos within matched hosts are then matched to each other, and their history is compared. The fraction of subhalos from one host successfully matched to subhalos from the other matched host are shown in Fig. 2-12. The remaining unmatched subhalos either are the result of an imperfect matching system, or are ones that correspond to subhalos whose counterpart in the other simulation have already merged with the host, or have not yet been accreted into the host, and are thus not identified as subhalos. The follow-

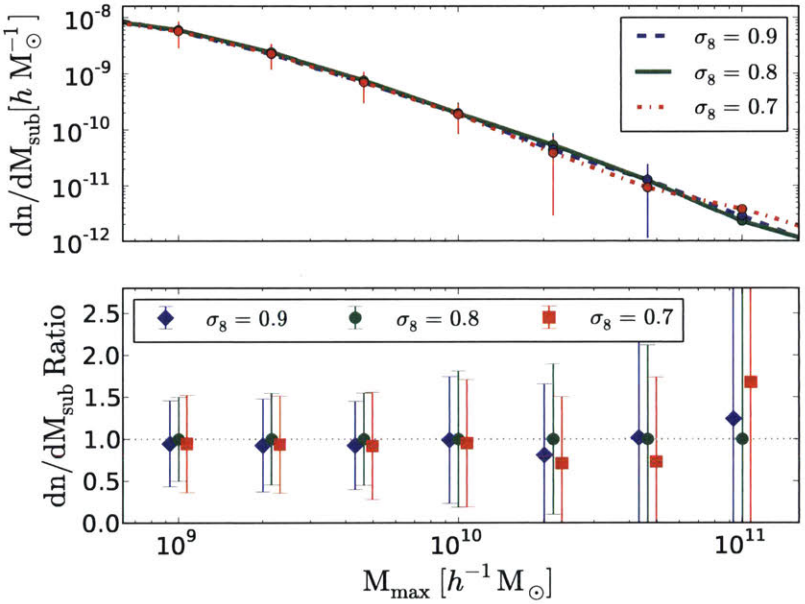

Figure 2-11 *Top Panel:* Average mass function of M_{max} for subhalos that survive until $z = 0$, normalized by $z = 0$ host halo mass, and scaled to a host of 0.84×10^{12} $h^{-1}\mathrm{M}_\odot$ for cosmologies with different σ_8. $1\,\sigma$ halo-to-halo variation bars are shown. *Bottom Panel:* Ratio of the subhalo M_{max} function for each cosmology to the fiducial *WMAP*7 cosmology. Data points for the non *WMAP*7 cosmologies are shifted slightly left and right of their true values to help distinguish them. The subhalo M_{max} function is found to not vary with σ_8. Large variation on a halo-to-halo basis is seen in the $1\,\sigma$ variation bars.

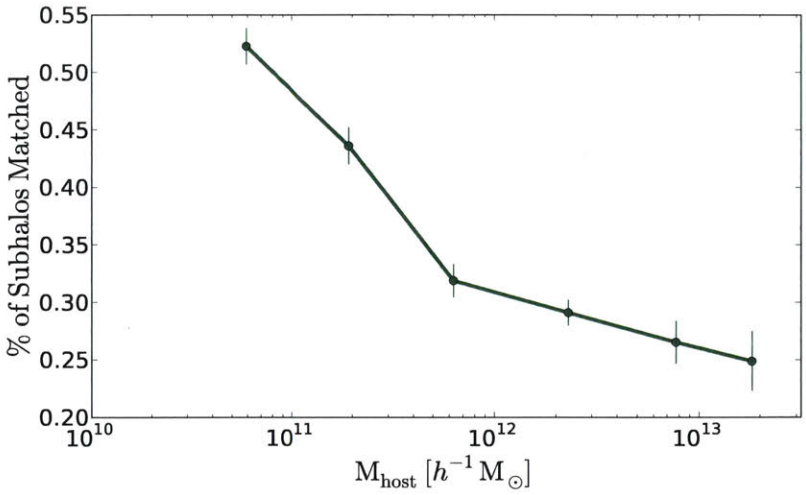

Figure 2-12 Fraction of subhalos in a matched pair of hosts between simulations of different cosmology that are successfully matched to each other as a function of their halo mass. Stronger dynamical processes in the larger halos inhibit subhalo matches more than in smaller host halos. 1σ error bars also shown.

ing subsections investigate the same averaged distributions as done in Section 2.3.3, except with only the subset of subhalos that have been matched to each other at $z = 0$. Additionally, all figures show only the data matched between $\sigma_8 = 0.8$ and $\sigma_8 = 0.9$ boxes for conciseness. All of the same trends are prevalent in the $\sigma_8 = 0.7$ and $\sigma_8 = 0.8$ matches.

2.3.4.1 Formation and Accretion Time

Figure 2-13 shows the probability density distributions of matched subhalos for subhalo formation time and accretion time for two different cosmologies. As expected, matched subhalos from the higher M^* simulation, $\sigma_8 = 0.9$, form earlier on average as seen in the top panel of the figure. The difference in the mean ages of the matched subhalos between $\sigma_8 = 0.9$ and $\sigma_8 = 0.8$ is 0.2 billion years, slightly greater than the 0.1 billion year difference found in the general sample.

In addition to forming earlier, the matched subhalos in the higher σ_8 simulation are accreted earlier. This is seen in the bottom panel of Figure 2-13. Comparison with Figure 2-10 shows that the matching process preferentially selects subhalos in the $\sigma_8 = 0.9$ simulation that were accreted ~ 2 billion years earlier than the typical

subhalo.

Thus, in a direct subhalo to subhalo comparison, varying cosmological parameters do have significant effects on the life of a subhalo, even though there are small or no differences found when averaging over all subhalos. Reconciliation of Figure 2-13 and Figure 2-10 indicates that the unmatched subhalos in the $\sigma_8 = 0.9$ simulation must be ones that were very recently accreted. Following the trend of earlier formation and earlier accretion with higher σ_8, the counterparts of these unmatched subhalos must be ones that have not yet accreted in the $\sigma_8 = 0.8$ simulation. Similarly, the unmatched subhalos in the $\sigma_8 = 0.8$ simulation are ones that are close to the end of their life. Their counterparts in the $\sigma_8 = 0.9$ simulation are ones that were already tidally destroyed. Therefore, the majority of unmatched subhalos are not stragglers from a faulty matching scheme, but rather the necessary result of different cosmologies leading to a shift in the formation, accretion, and thus destruction of subhalos. Evidently, when studying the accretion times of all subhalos in a host, this accretion and destruction process erases hints of the cosmology the subhalos formed in.

2.3.4.2 Maximal Mass and $z = 0$ Mass

Comparison of the mass function of the maximal mass of matched subhalos, as seen in the top panel of Figure 2-14, shows a weak dependence on cosmology. Subhalos in the $\sigma_8 = 0.9$ simulation have their life cycle shifted earlier, forming and accreting earlier. The earlier accretion leads to less mass at its peak. Likewise subhalos in the $\sigma_8 = 0.7$ simulation peak at a higher mass than in the $\sigma_8 = 0.8$ simulation. Knowing that such subhalos in the higher σ_8 simulations accrete earlier, one can expect them to reach $z = 0$ at an even smaller mass due to more time spent being tidally stripped. This is confirmed in the mass function displayed in the bottom panel of Figure 2-14. Matched subhalos are on average 25% less massive in the $\sigma_8 = 0.9$ simulation than the $\sigma_8 = 0.8$ simulation, and 50% more massive in the $\sigma_8 = 0.7$ simulation than the $\sigma_8 = 0.8$ simulation.

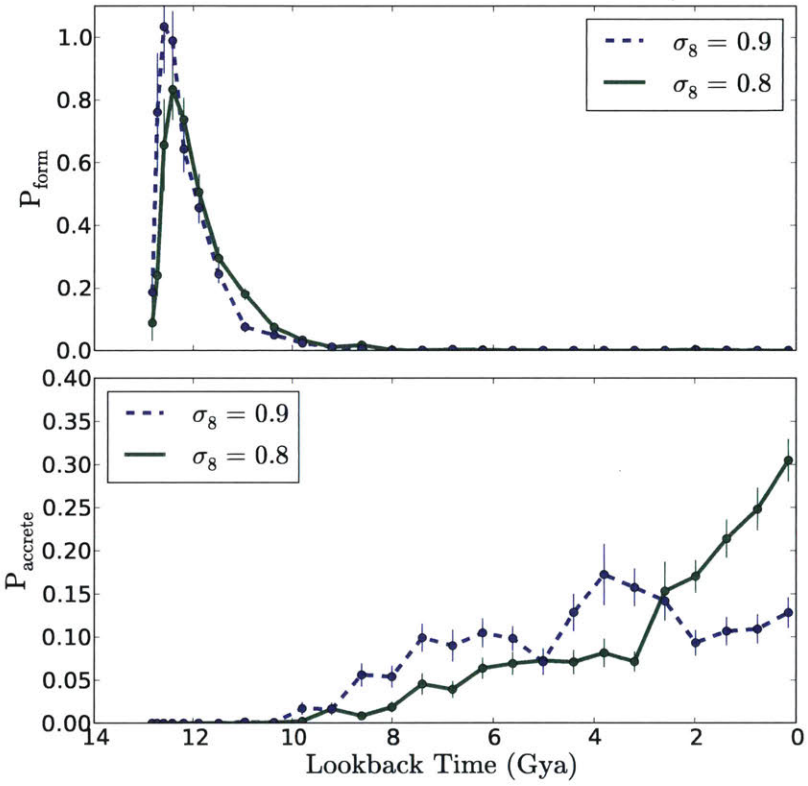

Figure 2-13 *Top Panel*: Average formation time probability density distribution of subhalos matched between $\sigma_8 = 0.9$ and $\sigma_8 = 0.8$ simulations. Matched subhalos in the higher σ_8 simulation form earlier on average, just as in the general case. *Bottom Panel*: Average accretion time probability density distribution of subhalos matched between $\sigma_8 = 0.9$ and $\sigma_8 = 0.8$ simulations. Unlike the general case, matched subhalos in the higher σ_8 simulation are accreted significantly earlier. Error bars show 1 σ uncertainty of each data point.

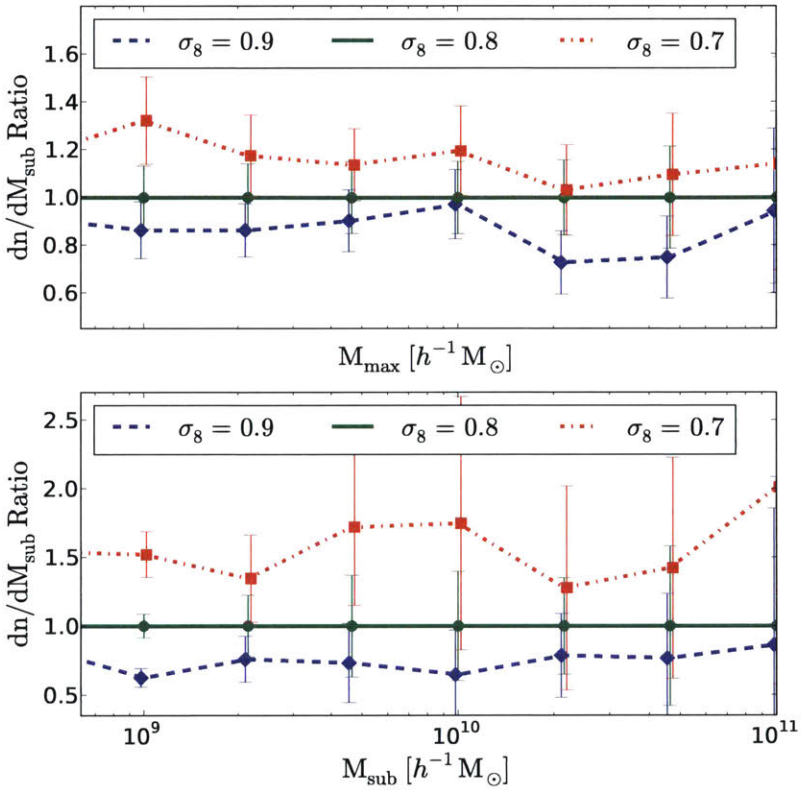

Figure 2-14 *Top Panel*: Ratio of the average mass function of M_{max} normalized by the $z = 0$ host halo mass for matched subhalos that survive until $z = 0$ to the M_{max} function of the fiducial *WMAP7* simulation for cosmologies with varied σ_8. Higher σ_8 results in lower peak masses in each pair of cosmologies tested. Error bars show $1\,\sigma$ uncertainty of each data point. See Fig. 2-11 for the true M_{max} function. *Bottom Panel*: Ratio of the average SHMF normalized to host halo mass for matched subhalos to the SHMF of the fiducial *WMAP7* simulation for cosmologies with varied σ_8. Unlike the general SHMF, the matched SHMF is significantly affected by cosmology. Subhalos in higher σ_8 simulations end up less massive. Error bars show 1σ uncertainty of each data point. See Fig. 2-4 for the true SHMF.

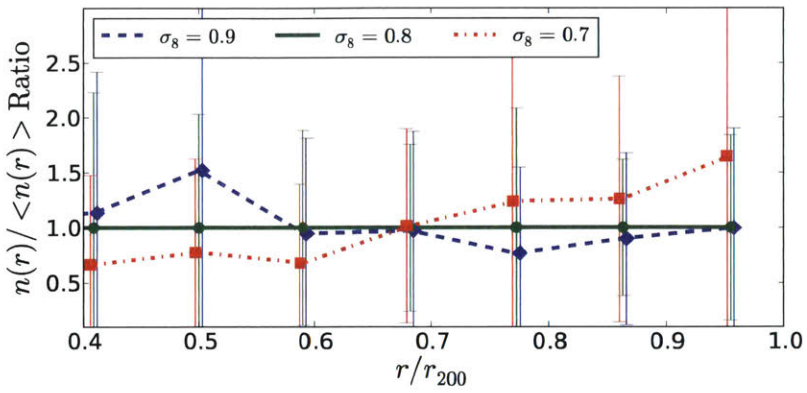

Figure 2-15 The averaged subhalo number density for matched subhalos as a function of r/r_{vir} normalized to the mean matched subhalo number density for cosmologies with varied σ_8 divided by the same function for the fiducial *WMAP7* simulation. This demonstrates the trend that matched subhalos in higher σ_8 cosmologies are closer to the halo center on average. Subhalos are 3.5% closer to the halo center in the $\sigma_8 = 0.9$ simulation than in the $\sigma_8 = 0.8$ simulation on average. Variation bars show 1 σ halo-to-halo variation for each data point.

2.3.4.3 Spatial Distribution

The spatial distribution of matched subhalos, characterized by number density as a function of r/r_{200} as in Figure 2-7, is shown in Figure 2-15. The local density is normalized by $<n>$, the average number density of matched subhalos within r_{200}. Relative to the $\sigma_8 = 0.8$ simulation, the matched subhalos within the $\sigma_8 = 0.9$ simulation are closer to the center of their host: the number density is higher below $r/r_{200} = 0.5$ and lower above $r/r_{200} = 0.8$. The $\sigma_8 = 0.9$ subhalos which accreted earlier on average have more time for dynamical friction effects to slow down their orbits and thus drag them closer to the halo center. The net average change of position is 3.5% closer. This result helps further confirm the conclusions made at the end of Section 2.3.4.1 that matched subhalos in the $\sigma_8 = 0.9$ simulation are closer to the end of their life and may be destroyed earlier, leaving their counterparts in the $\sigma_8 = 0.8$ simulation matchless.

2.4 Summary of Subhalo Property Changes

In this investigation, we found that the way in which σ_8, n_s and Ω_m affect subhalo properties can be more concisely summarized in terms of changes in M^*, which is positively correlated to all three parameters. The following list details how subhalo properties change with higher M^* cosmologies for all properties that vary with cosmological parameters:

- Subhalo V_{max} function as a function of host halo mass is greater.
- Subhalo concentration is greater.
- Subhalo formation time is earlier on average and on a matched subhalo to subhalo basis.
- Subhalo accretion time is earlier on a matched subhalo to subhalo basis.
- Subhalo peak mass is smaller on a matched subhalo to subhalo basis.
- Subhalo mass at $z = 0$ is smaller on a matched subhalo to subhalo basis.
- Subhalo radial distribution (number density as a function of r/r_{200}) is shifted toward the host's center on a matched subhalo to subhalo basis.

Individual subhalos in cosmologies with higher M^* are shown to have formed earlier, accreted into their eventual host earlier, both in agreement with Zentner and Bullock (2003); van den Bosch et al. (2005), and as a result of spending more time within their host, lost more mass due to tidal stripping and moved closer to the center of the host due to dynamical friction. This leads to subhalos in higher M^* cosmologies having, on average, less mass at $z = 0$. This is opposite of the trend for host halo masses, where large hosts have more mass in higher M^* cosmologies.

Including subhalo abundances, the averaged subhalo properties that remain unchanged under variations in cosmological parameters are:

- subhalo mass function as a function of host halo mass,

- subhalo abundance above $V_{max} = 30$ km s^{-1} as a function of host halo mass,
- subhalo mass fraction as a function of host halo mass,
- radial distribution of subhalos (number density as a function of r/r_{200}),
- subhalo accretion time probability distribution function, and
- distribution of maximal mass attained by subhalos before they enter their host's tidal field.

2.5 Discussion and Conclusion

Changes in the cosmological parameters σ_8, n_s, and Ω_m affect the abundance of dark matter halos in the universe. Larger values of n_s and Ω_m lead to a larger abundance of halos. Larger values of σ_8 lead to a higher abundance of halos above the current characteristic mass scale of collapse, M^*, and a lower abundance below that threshold. These differences, however, as tested by a suite of cosmological dark matter only simulations, are erased when considering local subhalo populations (all subhalos within a given host halo) as a function of their host halo mass. For any given host halo mass, the average subhalo mass function is independent of small variations in cosmological parameters on the order of the changes in the best estimate values of *WMAP* and *Planck* (Planck Collaboration et al., 2014). Thus, the expected abundance of dark matter subhalos by mass in the Milky Way, for example, should not change with the recent revision of parameters by *Planck*.

Subhalo concentrations, V_{max} functions, and formation times, on the other hand, retain the same trends in cosmological dependence as the host halos. In cosmologies where objects tend to collapse sooner (higher values of σ_8, n_s, and Ω_m), both host halos and subhalos have systematically more concentrated cores for all mass ranges considered and form earlier. More concentrated cores lead to subhalos with higher V_{max} relative to less concentrated halos of the same mass. Other cosmology induced differences arise on a subhalo to subhalo basis. Since all simulations used the same

initial random white noise field, it was possible to identify halos and subhalos which arose from the same density fluctuation. A comparison of these "matched" subhalos indicates that cosmological parameters do in fact have a significant effect on individual halos, even ones that become subhalos. In the higher σ_8 simulation, matched subhalos formed earlier, accreted earlier, are located closer to the host halo center, are more concentrated, and are smaller in mass.

Clearly, there is a disparity of results between averaging properties of all subhalos within a host and averaging only the subset of matched subhalos. This indicates that the mechanisms of subhalo accretion, mass loss, and ultimately destruction lead to a cosmology independent steady state distribution of subhalos within a host. While subhalos in the higher M^* simulations are accreted earlier on average, there is a steady rate of subhalo accretion that is the same in all cosmologies at low redshift for hosts of the same size. In addition, even though the overall mass function of halos changes, the mass function of subhalos at accretion for a given host halo does not change. Since subhalo mass and angular momentum loss is governed by cosmology independent physics, it follows that the resulting size, position, number, and accretion time of subhalos will be unaffected by cosmology. In contrast, the formation time and concentration, and thus V_{max}, of subhalos at accretion do vary with cosmology. These differences are not erased subhalo-host interactions.

The results of this study indicate that in simulations with pre-*Planck* cosmological parameters, characterizations of subhalo mass function, mass fraction, spatial distribution, accretion time, and peak mass at infall are all still correct. This indicates that expensive high resolution simulations remain valid when used to study subhalos as long as the properties are appropriately normalized to the mass of the host. There are a few exceptions for when subhalo concentration matters, such as in computing a dark matter annihilation rate estimate from substructure, or when subhalo formation time (defined by when halos first reach a threshold mass) is important, or to a lesser extent when the subhalo V_{max} function matters. In such cases a simple way to estimate how these properties change is to compute M^* for two cosmologies and know that in higher M^* cosmologies subhalos will be more concentrated and have formed

earlier. For example, between *WMAP7* and *Planck*, M^* increased by an amount similar to the effect of an increase of σ_8 of 0.07. The results also indicate that variations in cosmological parameters should have negligible effect on the magnitude of the missing satellite problem, and some affect on the too big to fail problem. Decreases in predicted subhalo V_{max} and concentration, which could come from lower values of n_s, σ_8, and Ω_m, both serve to alleviate the too big to fail problem. Polisensky and Ricotti (2013) confirms this for the cases of n_s, σ_8. Still, based on the study of subhalo properties over many host halos, the mass of the host halo and variations from halo to halo are much more important in controlling the distribution of subhalo properties within a galaxy than cosmological parameters.

Chapter 3

Enhanced Tidal Stripping of Satellites in the Galactic Halo from Dark Matter Self-Interactions

The content of this chapter was published in Monthly Notices of the Royal Astronomical Society *Volume 461, Issue 1 in May 2016. It has an arXiv location of* `https://arxiv.org/abs/1603.08919`. *The authors are: Gregory A. Dooley, Annika H.G. Peter, Mark Vogelsberger, Jesús Zavala, and Anna Frebel*

Disclaimer

This chapter was predominately a product of my own work. The simulations analyzed were run by Mark Vogelsberger and Jesús Zavala. All authors provided comments and feedback on the paper.

Acknowledgements

GAD acknowledges support from an NSF Graduate Research Fellowship under Grant No. 1122374 and support from the Dr. Pliny A. and Margaret H. Price Endowment Fund as a Price Visiting Graduate Student, administered through the Center for

Cosmology and AstroParticle Physics. The Dark Cosmology Centre is funded by the DNRF. JZ is supported by the EU under a Marie Curie International Incoming Fellowship, contract PIIF-GA-2013-62772.

Abstract

We investigate the effects of self-interacting dark matter (SIDM) on the tidal stripping and evaporation of satellite galaxies in a Milky Way-like host. We use a suite of five zoom-in, dark-matter-only simulations, two with velocity-independent SIDM cross-sections, two with velocity-dependent SIDM cross-sections, and one cold dark matter (CDM) simulation for comparison. After carefully assigning stellar mass to satellites at infall, we find that stars are stripped at a higher rate in SIDM than in CDM. In contrast, the total bound dark matter mass-loss rate is minimally affected, with subhalo evaporation having negligible effects on satellites for viable SIDM models. Centrally located stars in SIDM haloes disperse out to larger radii as cores grow. Consequently, the half-light radius of satellites increases, stars become more vulnerable to tidal stripping, and the stellar mass function is suppressed. We find that the ratio of core radius to tidal radius accurately predicts the relative strength of enhanced SIDM stellar stripping. Velocity-independent SIDM models show a modest increase in the stellar stripping effect with satellite mass, whereas velocity-dependent SIDM models show a large increase in this effect towards lower masses, making observations of ultrafaint dwarfs prime targets for distinguishing between and constraining SIDM models. Due to small cores in the largest satellites of velocity-dependent SIDM, no identifiable imprint is left on the all-sky properties of the stellar halo. While our results focus on SIDM, the main physical mechanism of enhanced tidal stripping of stars apply similarly to satellites with cores formed via other means.

3.1 Introduction

The cold dark matter (CDM) paradigm is part of a simple model that successfully describes the Universe on large scales. Assuming that gravity alone acts on dark matter particles, the theoretical CDM framework has been shown to adequately explain halo mass functions, correlation functions, and shapes of large galaxies, clusters, and filaments (for a review, see Frenk and White 2012). On smaller scales, however, tensions with observations still arise when using CDM to predict properties of substructure and dwarf galaxies. For instance, whereas simulations predict high central density

cusps in dark matter haloes (Navarro et al., 1997; Bullock et al., 2001; Wechsler et al., 2002), observations of low surface brightness galaxies (de Blok and McGaugh, 1997; de Blok et al., 2001; Kuzio de Naray and Spekkens, 2011), low-mass spiral galaxies, (Gentile et al., 2004; Simon et al., 2005; Oh et al., 2011; Castignani et al., 2012; Adams et al., 2014), and Milky Way dwarf spheroidal galaxies (Walker and Peñarrubia, 2011; Salucci et al., 2012; Breddels and Helmi, 2013) indicate that such galaxies have lower central densities and more shallow, or cored, inner density profiles. This problem of the central density of haloes is referred to as the cusp/core issue, initially identified by Flores and Primack (1994) and Moore (1994). Specifically, at low radii, $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln(\mathrm{r}) \approx -1$ for a cuspy density profile and $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln(\mathrm{r}) \approx 0$ for a cored profile.

These discrepancies have renewed interest in the idea that dark matter may not be collisionless, but could scatter with particles of the same species. The idea of self-interacting dark matter (SIDM) was first proposed by Carlson et al. (1992), Machacek et al. (1993), de Laix et al. (1995), and further explored by Spergel and Steinhardt (2000) and Firmani et al. (2000) in part to explain the cusp/core issue. In these models, the local particle scattering rate scales as

$$\Gamma(r) \propto \rho(r)\frac{\sigma}{m_x}v_{\mathrm{rms}}(r) \tag{3.1}$$

where ρ is the local density, $\frac{\sigma}{m_x}$ is the interaction cross-section per unit mass, and v_{rms} is the rms speed of dark matter particles. Increased scattering in high-density regions transfers energy from the outer halo to the inner halo, increasing velocity dispersion and decreasing density (Burkert, 2000; Yoshida et al., 2000a,b; Davé et al., 2001; Colín et al., 2002). Thus SIDM naturally leads to lower density cores that may better fit observations.

SIDM also provides a solution to the so-called too big to fail problem (Boylan-Kolchin et al., 2011, 2012), which refers to a population of high-density subhaloes in Milky Way-like simulations with no observed analogues. If such haloes did exist as simulated, they would be too big to fail to form stars, and as such should have

visible counterparts. More specifically, these haloes have a circular velocity profile which is too large to match the kinematics of known satellites. The problem exists for both the largest haloes at $z = 0$, initially pointed out in Read et al. (2006), and the largest haloes at infall and before reionization. For certain cross-sections, SIDM offers a solution to this issue by reducing the circular velocity profile of all haloes (Vogelsberger et al., 2012; Zavala et al., 2013; Vogelsberger et al., 2016).

Other solutions to the cusp/core and too big to fail issues have also been proposed that require no modifications to dark matter. Supernova driven outflows could erase cusps, as suggested and described by Navarro et al. (1996a); Read and Gilmore (2005); Governato et al. (2010); Oh et al. (2011); Governato et al. (2012); Pontzen and Governato (2012); Teyssier et al. (2013); Madau et al. (2014) and Pontzen et al. (2015). In order to simultaneously solve "too big to fail" though, it may require an impossibly large number of supernovae (Garrison-Kimmel et al., 2013). However, sufficiently early supernovae might create cores large enough to solve the problem (Amorisco et al., 2014). Zolotov et al. (2012) and Brooks and Zolotov (2014) also argue for early supernovae, adding in a further reduction of satellite central density due to tidal stripping to explain Milky Way observations. Additionally, late and bursty star formation could alleviate both cusp/core and too big too fail problems (Oñorbe et al., 2015), but not solve them entirely. Further possibilities include a combination of a reduced mass Milky Way (as proposed in González et al. 2014), baryon loss due to reionization and supernova outflows (Sawala et al., 2014), or other combinations of these effects (Pontzen and Governato, 2014; Brook and Di Cintio, 2015; Wetzel et al., 2016).

In spite of several plausible purely baryonic explanations, the too big to fail problem extends to the Local Group (Garrison-Kimmel et al., 2014a), and may extend to the nearby field (Klypin et al., 2015; Papastergis et al., 2015), although that claim is sensitive to the methodology used (Brook and Shankar, 2016). Ultimately, uncertainty still remains whether or not baryonic mechanisms alone can systematically erase all concerns. Read et al. (2016a) for instance argue that supernovae form cores in field galaxies the size of ultrafaint dwarfs, but rely on a stellar mass to halo mass relation

consistent with the abundance matching model of Behroozi et al. (2013a), which may greatly overpredict the relationship for low-mass galaxies (Garrison-Kimmel et al., 2014b). When using lower stellar masses, Peñarrubia et al. (2012) find no cores. Consequently, SIDM remains of interest as it could provide an attractive non-baryonic solution.

Several studies have set constraints on the strength of the SIDM interaction cross-section by searching for observable signatures. Since SIDM produces more spherical halo distributions, Miralda-Escudé (2002) used measurements of the ellipticity of galaxy clusters to enforce a stringent constraint of $\sigma/m_x < 0.02\,\mathrm{cm}^2/\mathrm{g}$. Randall et al. (2008) set another constraint by considering the offset between the galaxy centroid and peak position of total mass in the Bullet Cluster. Yoshida et al. (2000b) used simulations of core sizes in clusters, and Gnedin and Ostriker (2001) considered increased rates of subhalo evaporation due to SIDM to set constraints. A set of companion papers, Rocha et al. (2013) and Peter et al. (2013), reassessed these limits and added their own constraints through studying the expected core sizes of haloes and halo shapes, respectively. They relax previous constraints, concluding that a cross-section σ/m_x between 0.1 and $1\,\mathrm{cm}^2/\mathrm{g}$ is sufficient to produce large enough cores in dwarfs to match observations and still withstand constraints.

At higher cross-sections, tension arises with the measured ellipticity of galaxy clusters and with overproducing cores in large haloes. These tensions can be obviated with a velocity-dependent cross-section. Motivated from particle theory, such particles could have a Yukawa potential in which the cross-section decreases for higher relative particle velocities, as proposed by Feng et al. (2010); Buckley and Fox (2010); Loeb and Weiner (2011), simulated initially by Vogelsberger et al. (2012), and recently incorporated in an encompassing effective theory of structure formation by Vogelsberger et al. (2016) and Cyr-Racine et al. (2015). This velocity dependence results in cores on the small scales as needed, while mitigating effects on large scales where constraints are stronger (Rocha et al., 2013; Elbert et al., 2015; Kaplinghat et al., 2016).

Since velocity-dependent cross-sections are as of yet not well constrained, and

velocity-independent models are still viable, we seek to highlight a new approach where their effects could be observed. First, we investigate whether the presence of a core influences how stars in satellites of a Milky Way-like host are tidally stripped. The reduced central density of dark matter in satellites and increased central velocity dispersion both occur where stars reside. These two effects should have an impact on the stripping and distribution of stars based on the dynamics of tidal stripping and tidal steams (see Bovy (2014) for instance). Peñarrubia et al. (2010) have already demonstrated that cored satellites lose mass more quickly than cusped ones. They, however, assign cores manually and do not consider SIDM. They also emphasize tidal interactions of satellites with the disc of their host galaxy, whereas we predominantly study tidal forces arising from the host halo only.

Secondly, we investigate whether cores or subhalo evaporation change the total mass-loss rate of dark matter in satellites. Particles bound to subhaloes may be ejected after colliding with a dark matter particle in the background environment. The resulting mass-loss is called subhalo evaporation. Rocha et al. (2013) demonstrate hints of this, with the subhalo maximum circular velocity function suppressed in the inner half of a host's virial radius.

Finally, we search for signatures of SIDM imprinted on the stellar halo, and observable implications of any changes in stellar and total mass-loss rates. With these goals in mind, we run simulations of four different models of SIDM to search for their impact on the disruption of satellites. By running our simulations with DM only, we isolate the effects of SIDM core formation. A complete picture would require including baryons in addition to SIDM, but such simulations are computationally more expensive and sensitive to the strength of the baryonic feedback implemented (Vogelsberger et al., 2014b; Fry et al., 2015). As such, a suite of full hydrodynamic simulations with SIDM that resolve Milky Way-like substructure is beyond the scope of this paper.

This paper is organized as follows. Section 3.2 introduces the simulations, Section 3.3 discusses our method of tagging stellar mass to dark matter particles, Section 3.4 presents the differences in stellar tidal stripping between CDM and SIDM, Section

3.5 discusses subhalo evaporation, Section 3.6 provides a simple theory to explain and predict SIDM-driven enhancements in stellar stripping, and Section 3.7 investigates how these in turn affect the stellar halo, stellar mass function, and half-light radii. Finally, Section 3.8 summarizes our results.

3.2 Simulations

We use a suite of five simulations, each of which takes its initial conditions from the Aquarius Project Aq-A-3 halo of Springel et al. (2008). Two simulations are run with a velocity-independent cross-section (the **v-i** models), two with a velocity-dependent cross-section (the **v-d** models), and one is run with pure CDM as the baseline for comparison. Using the same initial conditions allows us to isolate the effects of each SIDM model. All of the simulations are the same as those initially used in and described in Vogelsberger et al. (2012). A summary of the simulations is found in Table 3.1. In the v-i cases, $\sigma_T^{\rm max}/m_\chi$ is the collisional cross-section per unit mass. In the v-d cases, it is the cross-section achieved when the relative particle velocity is $v_{{\rm max},\sigma}$, which is the velocity at which the quantity σv is maximized.

The transfer cross-section as a function of relative particle velocity for each of our four models is shown in Fig. 3-1, similar to fig. 1 of Zavala et al. (2013). Our velocity-dependent cross-sections scale approximately as $\sigma \propto v^{-4}$ at high velocity, and $\sigma \propto v^{-0.7}$ at low velocity for the range plotted. The full form of the transfer cross-section can be found in equation 1 of Vogelsberger et al. (2012). Our SIDM10 model, with a constant-cross section of $10\,{\rm cm}^2/{\rm g}$ is ruled out by constraints on the ellipticity of galaxy clusters, but we include it simply to more easily highlight the effects of SIDM. Each of the other three models are not ruled out, and are compatible with observed kinematics of the Milky Way dwarf spheroidal galaxies, including resolving too big to fail, as shown in Vogelsberger et al. (2012).

Each simulation has a particle mass of $m_{\rm p} = 4.911 \times 10^4\,{\rm M}_\odot$, and Plummer equivalent gravitational softening length of $\epsilon = 120.5\,{\rm pc}$. The cosmological parameters used are $\Omega_{\rm m} = 0.25$, $\Omega_\Lambda = 0.75$, $h = 0.73$, $\sigma_8 = 0.9$, and $n_{\rm s} = 1$.

Table 3.1 Summary of Simulations

Name	$\sigma_T^{\max}/m_\chi(\mathrm{cm}^2/\mathrm{g})$	$v_{\max,\sigma}(\mathrm{km/s})$
CDM	-	-
SIDM10	10	-
vdSIDMa	3.5	30
vdSIDMb	35	10
SIDM1	1	-

NOTE. — CDM and SIDM simulations listed with cross-section parameters used in this study. $v_{\max,\sigma}$, is the velocity at which the quantity σv is maximized in the velocity-dependent cases. Naming scheme adopted from Vogelsberger and Zavala (2013) and Zavala et al. (2013) for consistency.

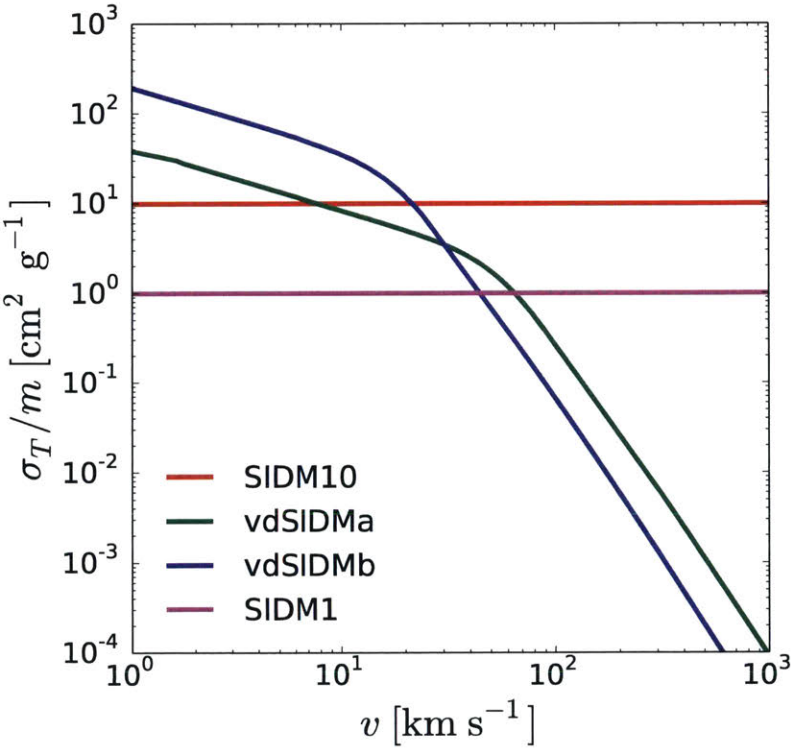

Figure 3-1 Cross-section as a function of relative particle velocity for each SIDM model considered (see Table 3.1). The colour scheme used here is repeated throughout the paper to more easily identify each model.

All self-bound haloes are found using a modified version of ROCKSTAR HALO FINDER (Behroozi et al., 2013b). The modifications include adding full iterative unbinding to improve halo finding accuracy, as described in Griffen et al. (2016b), and returning all bound particle IDs for each halo in order of boundedness. Merger trees were produced by ROCKSTAR CONSISTENT TREES (Behroozi et al., 2013c). We use the Bryan and Norman (1998) definition of the virial radius, $r_{\rm vir}$, which at $z = 0$ for our cosmological parameters is the radius such that the mean enclosed halo density is 94.2 times the critical density of the universe, $3H_0^2/8\pi$G. Unless otherwise clarified, any mention of mass refers to $m_{\rm vir}$, the gravitationally bound mass within $r_{\rm vir}$. The main halo at redshift $z = 0$ in each simulation has a mass of $\sim 2.2 \times 10^{12}\,{\rm M}_\odot$, and radius of 330 kpc.

3.3 Methods

3.3.1 Particle Tagging

To study the stellar component of satellites in dark-matter-only simulations, we tag particles with stellar mass. We choose to use particle tagging over hydrodynamical simulations due to the tremendous computational time saved. This allows us to better explore SIDM parameter space via running more simulations. Additionally, while particle tagging is an imperfect method, hydrodynamical simulations are not without their own uncertainties that depend on which baryonic physics models are implemented.

In developing a tagging technique, we take guidance from a long precedent of particle tagging in simulations. Dark matter particles in the most central part of subhaloes have been assigned stellar mass at the time of subhalo accretion (De Lucia and Helmi, 2008; Rashkov et al., 2012), subhalo peak mass (Bailin et al., 2014), and through live tagging where mass is added at each snapshot (Bullock and Johnston, 2005; Cooper et al., 2010; Le Bret et al., 2015). The fraction of stars tagged varies from the most bound 1% (Cooper et al., 2010; Rashkov et al., 2012; Bailin et al.,

2014), to 5% (Le Bret et al., 2015), to 10% (De Lucia and Helmi, 2008). Mass may be split evenly among particles per halo (De Lucia and Helmi, 2008; Rashkov et al., 2012; Bailin et al., 2014), or may be split differently among dark matter particles to mimic the light profile of galaxies (Bullock and Johnston, 2005; Cooper et al., 2010).

A consensus on what fraction of particles is best to tag does not exist. Bailin et al. (2014) find that tagging 1% versus 10% results in little qualitative difference in the stellar halo, Cooper et al. (2010) recommend $1 - 3$ per cent to avoid stellar profiles that are too concentrated or too diffuse, and Le Bret et al. (2015) suggest 5% is better than 1% to reproduce realistic stellar density profiles. Furthermore, there are cautions that particle tagging generically produces less concentrated, more structured, and more prolate galaxies (Bailin et al., 2014) and that live tagging schemes are superior at reproducing realistic stellar density and energy profiles (Le Bret et al., 2015).

While these studies demonstrate that particle tagging is an imperfect method for producing reliable quantitative results, they are still able to pick up on trends and qualitative results, which are of interest to the present study. The diversity of tagging prescriptions and conflicting recommendations on what fraction of most bound particles to tag suggests that any scheme needs to be tested within the context of the project goals.

We aim to tag enough particles for good resolution, and to probe a realistic range of where stars reside, particularly making sure to target where a core forms in SIDM. For field haloes at $z = 0$, we find that cores extend to $\sim 3\%$ of the most bound particles in $10^8\,\mathrm{M}_\odot$ haloes. In $10^{10}\,\mathrm{M}_\odot$ haloes, cores extend to 0.5% in vdSIDMb on the low end, and 1.5% in SIDM10 on the high end for our models (see Section 3.6 for our calculation of core sizes). Thus, based on previous studies of tagging, to probe satellite cores, and to ensure good resolution, we would ideally tag all stars up to the $\sim 2\%$ most bound particle in all satellites, satellites of satellites, etc. throughout cosmic history.

However, tagging particles in SIDM simulations introduces a new obstacle. Dark matter particles can elastically scatter, whereas stars should never scatter. We deal with this problem by only assigning mass to particles which do not scatter between

their time of tagging and the final $z = 0$ snapshot. Consequently, only a fraction of the central 2% most bound particles can be tagged. In order to avoid resolution problems, we resort to still tag 2% of the total bound particles, tagging all unscattered ones in order of boundedness until we reach our 2%. This, however, creates an additional problem. In haloes with high rates of interaction, stellar mass would be assigned at unrealistically large distances from the halo's centre. We solve this new issue by enforcing a cut-off at 5%. No particles that are greater than 5% in their boundedness ranking can be tagged. If fewer than 2% of the total number of halo particles can be tagged within the interval of the 5% most bound particles, then fewer than 2% are tagged. This represents a compromise between resolution and tagging the truly central regions of satellites.

The distribution of particles that remain unscattered within the 5% most bound interval varies between each SIDM simulation, and as a function of halo mass and infall time. For an unbiased comparison with CDM, it is imperative to tag particles in CDM in as similar a fashion as possible to the corresponding SIDM case. We divide all haloes into bins according to their infall time and infall mass. Each halo has a distribution recorded for it, cataloguing the fraction of particles that remain unscattered as a function of rank order particle boundedness. We consider particles, starting with the most bound, in intervals of 0.5% of the total number of bound particles, recording the fraction of unscattered particles in each interval. The distributions of all haloes within an infall scale and mass bin are then averaged. An example of four averaged distributions for the SIDM1 and vdSIDMa simulations are shown in the left-hand panel of Fig. 3-2. As expected, satellites which are more massive (higher central density) and fall in earlier experience more scattering and have fewer remaining particles to tag. This procedure results in a look-up table used to tag the CDM simulation in four ways, once for each of the four SIDM simulations. An infalling satellite in the CDM simulation is tagged to reproduce the right look-up table distribution, with specific particles within a particle boundedness interval being selected uniformly at random. In cases such as the $0.99 < z_{\rm infall} < 1.47$, $10^{9.7}\,{\rm M}_\odot < M < 10^{10.5}\,{\rm M}_\odot$ bin in the left-hand panel of Fig. 3-2, fewer than 2% of all bound particles can be tagged on

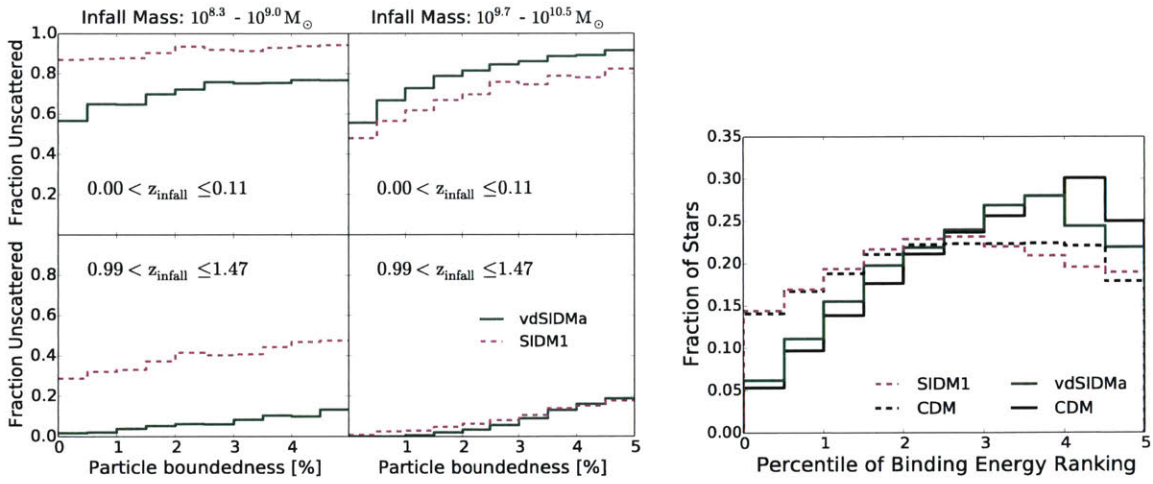

Figure 3-2 Left-hand panel: distribution of the fraction of particles that do not scatter between infall and $z = 0$ (and can thus be tagged) as a function of their rank order boundedness in satellites at infall. A sample of two infall time bins and two infall mass bins are shown for the SIDM1 simulation in magenta (dashed lines), and the vdSIDMa simulation in green (solid lines). See Table 3.1 for differences between the SIDM models. The unscattered distributions of all satellites within an infall mass and time bin are averaged to produce a look-up table used to tag the corresponding CDM case. Right-hand panel: the normalized distribution of the fraction of tagged particles per boundedness percentile at the time of tagging for all tagged particles in the simulation. We verify that our tagging of the CDM simulations mirror that of the target SIDM simulation. The dashed-line CDM distribution mirrors that of the corresponding SIDM1 distribution (dashed magenta line) and the solid-line CDM distribution mirrors that of the corresponding vdSIDMa distribution (solid green line).

average. Correspondingly, such haloes in the CDM simulation are tagged with fewer than 2% of all particles.

After completing this procedure, we verify that our tagging did not introduce significant biases. One particular test is shown in the right-hand panel of Fig. 3-2. This shows the normalized distribution of the boundedness percentile at which particles were tagged for all tagged particles in all subhaloes. We present the cases of vdSIDMa (solid green line), SIDM1 (dashed magenta line), and each of their corresponding CDM instances (solid black and dashed black lines respectively). Whereas the left-hand panels shows the unnormalized distribution of unscattered particles, this panel shows the normalized distribution of actually tagged particles. Importantly, the CDM simulation is tagged in different ways to mimic the features of the

different SIDM cases. The dashed CDM line mirrors that of SIDM1, and the solid CDM line mirrors that of vdSIDMa. Imperfections in the match are due mostly to slight differences in merger history.

Other tagging attempts which did not produce matches between CDM and SIDM pairs resulted in biases, such as stars tagged at systematically larger radii in SIDM than in CDM, resulting in them being more easily tidally stripped. For full confidence in our tagging method, we perform a variety of additional tests verifying that our primary results are not simply an artefact of our tagging technique. A summary of these robustness tests can be found in Appendix 3.9.

Upon selecting particles for tagging, we assign to them each a stellar mass according to the abundance matching prescription of stellar mass to halo mass in Moster et al. (2013). Since this defines a relationship for haloes at infall, we tag particles at infall. Infall is the snapshot before a halo first becomes a subhalo of the main host. A subhalo is defined as a halo that resides within the virial radius of its larger host. Given the virial mass of a halo at infall, we use the abundance matching relationship to specify total stellar mass, then split that evenly among particles tagged within the halo. This means particles within a halo have the same stellar mass, but particles in different haloes have different stellar masses.

At a halo mass of $10^8\,M_\odot$, the stellar mass is only $80\,M_\odot$. Due to this negligible contribution, we do not tag haloes with an infall mass below that value. In fact, the more recent stellar mass to halo mass investigation by Brook et al. (2014) suggests that $80\,M_\odot$ is a great overestimate. They find a much steeper slope going towards lower stellar masses below a halo mass of $10^{10.3}\,M_\odot$. Models by Behroozi et al. (2013a) and Sawala et al. (2015) on the other hand, predict much higher stellar mass to halo mass values. Sawala et al. (2015) for instance predicts a stellar mass of $8 \times 10^4\,M_\odot$ for $10^8\,M_\odot$ DM haloes at $z = 0$ that host galaxies, but argue that nearly all haloes that size and smaller remain dark since reionization suppresses star formation entirely. None the less, the uncertainty in what stellar mass to assign to low mass galaxies does not affect our primary results in Sections 3.4-3.6 since they are based on the ratio of mass in a satellite after infall to the mass at infall, which is independent of the stellar

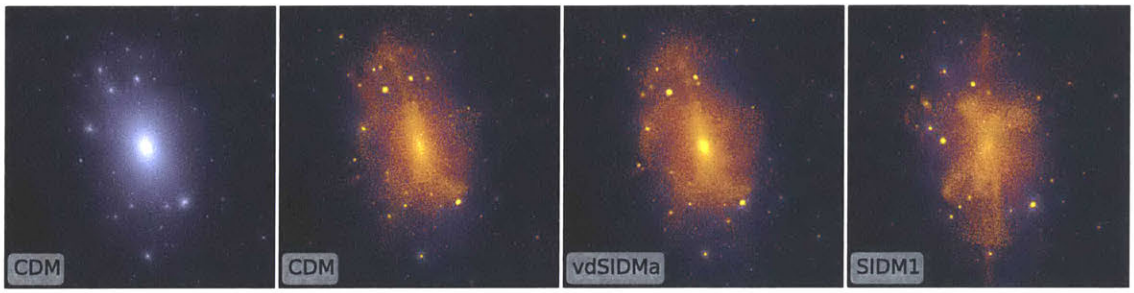

Figure 3-3 Logarithmic density projections of dark matter and tagged stellar mass in our simulations. From left to right: dark matter in the CDM simulation, dark matter with overlaid tagged stellar mass in the CDM simulation, dark matter with overlaid tagged stellar mass in the vdSIDMa simulation, dark matter with overlaid tagged stellar mass in the SIDM1 simulation. Due to core formation in the host of SIDM1, orbits are altered relative to the CDM simulation and the stellar halo deviates in appearance from that of CDM. In vdSIDMa, only a small core forms in the host, resulting in little difference in the stellar halo's appearance. All images have a box length of 500 kpc.

mass to halo mass ratio used. Uncertainty in this relationship does somewhat impact the implications of our findings in Section 3.7, and is accordingly discussed there.

A visual of a sample of our haloes with and without stellar tagging is shown in Fig. 3-3. Dark matter is shown from low to high density in blue to silver, and stars in magenta to yellow.

3.4 Stellar Stripping

Using the particle tagging technique described in Section 2.2, we quantify the stripping of stars in each of our SIDM models compared to the baseline CDM simulation. In Figs 3-4-3-6, we present the fraction of stellar mass remaining in satellites as a function of time since infall for the models vdSIDMa, vdSIDMb, and SIDM1, respectively. Since stripping will depend on the satellite orbit, infall time, and mass, we divide our samples into bins of different infall mass and mean pericentre achieved. The fraction of stars remaining versus time is then averaged over all subhaloes within each bin.

All satellites that fell into the host, whether extant at $z = 0$ or destroyed, are included in the sample. In total, there are around 800 satellites considered per simulation. Most satellites eventually merge with the host, i.e. they are no longer found

by the halo finder, or fall in more recently than 10 Gyr ago. Consequently, as time increases, there are fewer and fewer remaining satellites to average over for each data bin. We enforce a cut-off beyond 10 Gyr where the data become particularly limited.

The mean pericentre of satellite orbits is well correlated with the strength of disruptive tidal forces exerted on a satellite, even though it ignores the complexity of a full orbit. We therefore divide our sample into bins of $0 - 30\,\mathrm{kpc}$, $30 - 70\,\mathrm{kpc}$, and $> 70\,\mathrm{kpc}$ mean pericentre crossing. We choose 30 kpc as the upper limit of the first bin since it corresponds approximately to the radius at which a Milky Way-like disc causes divergent tidal stripping in cored versus cuspy subhaloes as studied in Peñarrubia et al. (2010). Both cored and cuspy satellites experience more disruption with a disc, but cored satellites with low pericentres experience a much greater increase in disruption than cuspy ones. Since our simulations do not include a disc in the host, the effects of stripping in both CDM and SIDM cases in the $d_{\mathrm{peri}} < 30\,\mathrm{kpc}$ bin will be underestimated relative to the other bins.

We choose our infall mass bins as $10^{8.0}\,\mathrm{M}_\odot < M_{\mathrm{infall}} < 10^{8.5}\,\mathrm{M}_\odot$, $10^{8.5}\,\mathrm{M}_\odot < M_{\mathrm{infall}} < 10^{9.1}\,\mathrm{M}_\odot$, and $10^{9.1}\,\mathrm{M}_\odot < M_{\mathrm{infall}} < 10^{10}\,\mathrm{M}_\odot$ to highlight interesting ranges of the halo mass dependence for each SIDM model, and to have enough satellites per bin to see trends of the average effect. Specifically, the mass $M_{\mathrm{infall}} < 10^{9.1}\,\mathrm{M}_\odot$ is chosen to highlight a transition in tidal effects on the v-d models from significant to insignificant. On halo mass scales above $10^{10}\,\mathrm{M}_\odot$, there are only ~ 10 subhaloes that ever enter the host. In these cases, the specific orbits of the satellites drive the differences in evolution more than SIDM, and cannot be averaged out. Additionally, dynamical friction drags massive haloes more rapidly to low pericentre orbits, leaving few examples with large orbits.

Due to the difficulty of particle tagging in a simulation with scattering particles, there are a few considerations that must be taken into account.

First, in the model SIDM10, there are too few non-scattered particles to tag as stars to extract meaningful results. Our goal of tagging 2% of all bound particles within the interval of the 5% most bound particles can only be achieved for satellites falling into the host within ~ 0.9 Gyr of $z = 0$. Tagging 1% of particles is no longer

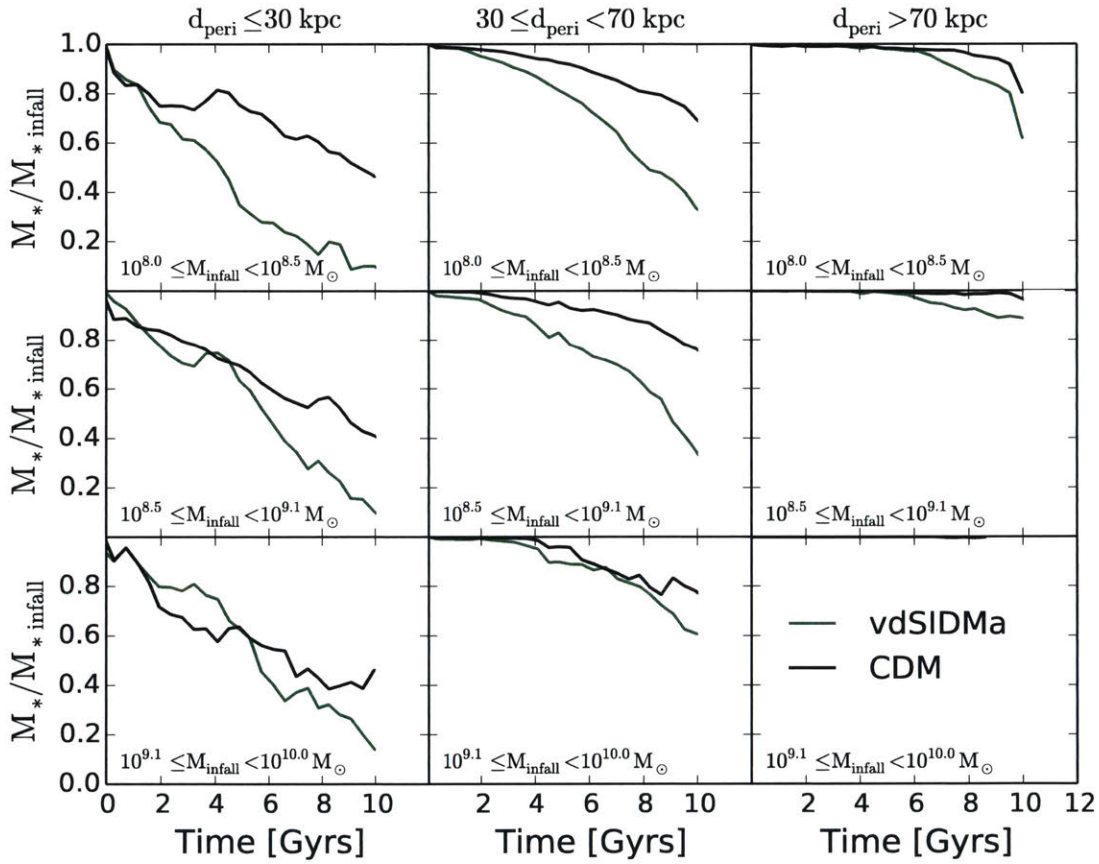

Figure 3-4 Fraction of stellar mass remaining in satellites as a function of time since infall, averaged over all satellites for vdSIDMa compared to CDM. Satellites are divided into different bins according to their mean pericentre distance, $d_{\rm peri}$ (columns), and infall mass, $M_{\rm infall}$ (rows). Satellites with $d_{\rm peri} < 30\,{\rm kpc}$ would pass through the host's disc in a real Milky Way analogue, but do not in our dark-matter-only simulations. Stripping in both CDM and SIDM should therefore be increased in this interval. Stellar mass stripping is enhanced in the vdSIDMa model relative to CDM for all orbit bins, particularly for lower-mass satellites.

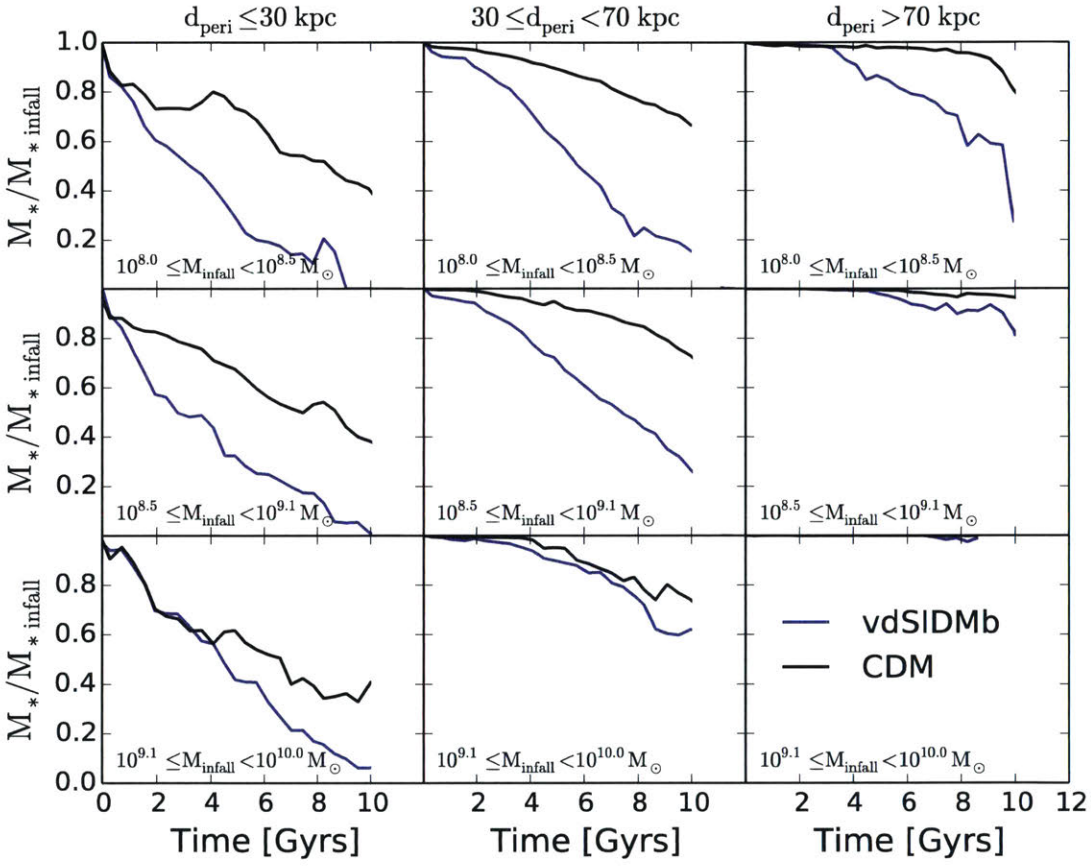

Figure 3-5 Same as Fig. 3-4, except for model vdSIDMb. Stellar mass stripping is enhanced for all orbit bins in vdSIDMb relative to CDM, with a strong dependence on satellite mass. The effect is the largest of all our models for low-mass haloes, and the smallest for high-mass haloes.

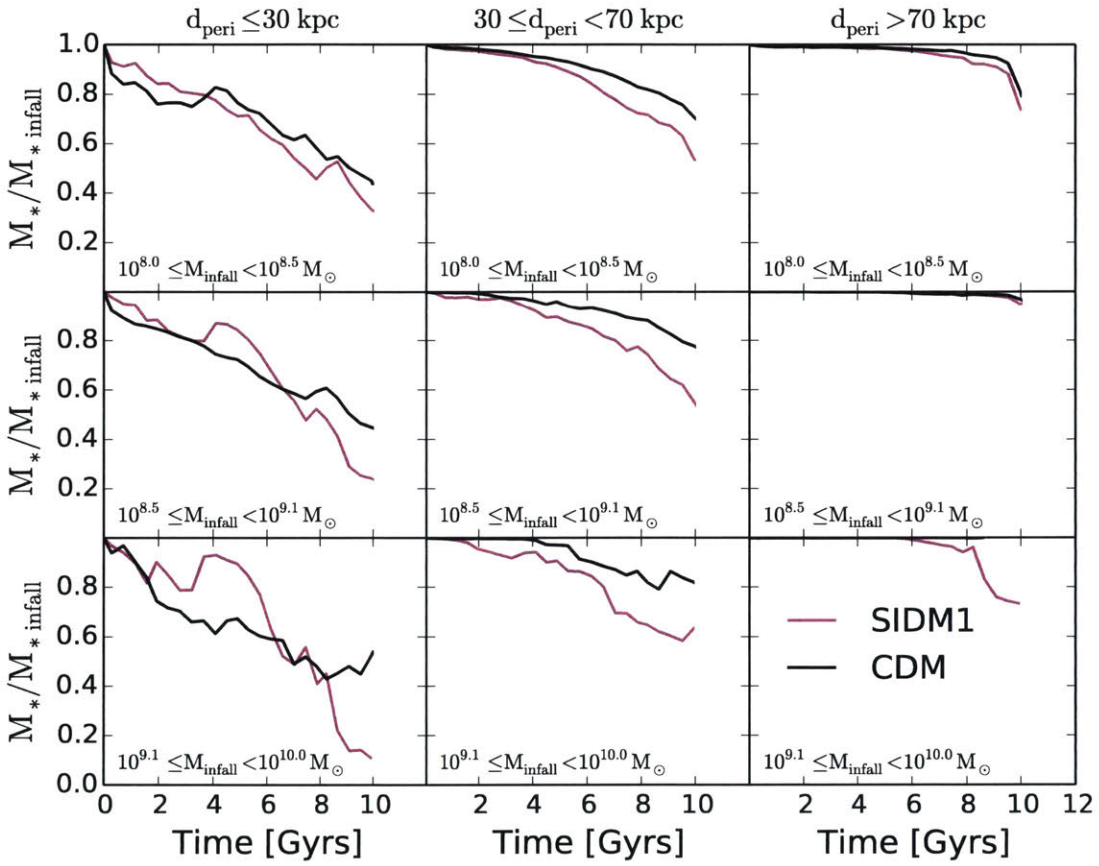

Figure 3-6 Same as Fig. 3-4, except for model SIDM1. Stellar mass stripping is enhanced in SIDM1 relative to CDM for orbit bins with $d_{\rm peri} > 30\,$kpc, with a weak dependence on satellite mass. More massive satellites experience a greater enhancement. Below 30 kpc, a reduced tidal field due to the host halo's core in SIDM counteracts the effects of a core in the subhaloes. The combined effect makes stripping looks similar to CDM.

possible for infalls earlier than ~ 2.5 Gyr before $z = 0$. At earlier times, the fraction of particles that can be tagged quickly approaches 0%. We therefore do not include SIDM10 in our results here, but expect that its effects will be a more extreme version of SIDM1.

Secondly, due to the different distributions of particles tagged in each SIDM model, it is important to only make comparisons with the CDM data tagged in accordance with the corresponding SIDM model. In spite of this caution, the stripping of particles in CDM is not substantially altered between different tagging instances, as seen by the black lines in Figs 3-4-3-6. Any differences that do exist are small compared to the difference in stripping due to the SIDM effects. This indicates that primarily dark matter physics drives the signals we study, not our tagging technique.

Thirdly, we combine all satellites regardless of infall time. Satellites with an earlier infall time have systematically higher rates of stripping in both SIDM and CDM. This is not physical, but instead due to particles being tagged further from the subhalo centre as compared to recently accreted satellites. Fewer unscattered particles remain in satellites that fall in early since there is more time for scattering to occur between infall and $z = 0$. Consequently, we must go to larger distances on average to find unscattered particles. In spite of this effect, the relative difference between SIDM and CDM stripping remains the same independent of infall time. It does, however, mean that the fraction of stars stripped in our simulation data is likely an overestimate of the fraction that would be stripped in a real Milky Way satellite.

With concerns of tagging artefacts assuaged, we look at the physical effects of SIDM. In the v-d models, we see a large effect of SIDM enhancing tidal stripping relative to CDM for low mass haloes, and the effect diminishing for higher mass haloes. The high cross-section in vdSIDMb at low relative velocities typical for haloes with $M_{\rm infall} \leq 10^{8.5}\,{\rm M}_\odot$ (for $v_{\rm max} \leq 13.5\,{\rm km/s}$, $\sigma/m_x \geq 25.6\,{\rm cm}^2/{\rm g}$), results in a very large effect, even for haloes with a mean pericentre $> 70\,{\rm kpc}$, where tidal forces are small. The strength of stripping for halo masses below $10^{9.1}\,{\rm M}_\odot$ in vdSIDMa is weaker in comparison, particularly for mean pericentres above $30\,{\rm kpc}$.

In SIDM1, for our intermediate range of mean pericentres, there is a clear enhance-

ment of tidal stripping relative to CDM. This effect gradually increases with larger mass haloes. Furthermore, this enhancement overtakes the effect seen in vdSIDMa and vdSIDMb for haloes between $10^{9.1}\,M_\odot$ and $10^{10}\,M_\odot$.

For haloes in SIDM1 with mean pericentre above 70 kpc, there is neither appreciable stripping in SIDM nor in CDM. For haloes with mean pericentre below 30 kpc, we see an interesting effect where self-interactions do not cause any additional enhancement in stripping. This is dramatically different than the velocity-dependent models where a strong enhancement is apparent. A large core forms in the host in SIDM1, reducing the local tidal field that causes stripping, whereas only a small core forms in the host of the v-d models.

Extending the trend of diminishing enhancement of stripping towards higher halo masses in the v-d models, there should be little difference between SIDM and CDM for haloes above $10^{10}\,M_\odot$. Indeed, by matching individual haloes to each other via identifying common particle IDs between simulations, we find no significant difference in stripping in the few haloes that follow similar orbits. For SIDM1, the opposite effect of enhanced stripping increasing towards higher masses exists. We therefore expect some difference to exist for haloes above $10^{10}\,M_\odot$, but the SIDM effect is sub-dominant to the more critical orbital parameters and infall time. By matching large subhaloes to each other in SIDM1 and CDM, we could not isolate a clear SIDM related effect. The implications of stellar stripping in these largest systems on the stellar halo are discussed in Section 3.7.3.

3.5 Subhalo Evaporation and DM stripping

Historical interest in SIDM stems from subhalo evaporation. As a satellite orbits within the background dark matter in the host, high velocity collisions with background particles can eject subhalo particles. Consequently, subhalo evaporation could suppress the subhalo mass function and alleviate the missing satellite problem, as originally proposed in Spergel and Steinhardt (2000).

We therefore investigate the strength of subhalo evaporation in our models by

studying the mass-loss rate of the total bound mass of subhaloes. In turn, we assess whether it plays an important role in causing the increased rate of stellar mass-loss in SIDM already reported in Section 3.4. Significant subhalo evaporation can reduce the total mass of a satellite, lowering its binding energy, and increasing the likelihood of stars being stripped. However, tidal forces alone could also increase the total mass-loss rate of cored SIDM satellites relative to cuspy CDM ones, and we must disentangle the two effects. Fortunately, as will become clear, we see no significant effect of tidal forces causing divergent total mass-loss rates for the subhaloes we study, and can therefore isolate the effects of subhalo evaporation.

Previously, Vogelsberger et al. (2012) studied the effect of subhalo evaporation on the subhalo mass function in the models SIDM10, vdSIDMa, and vdSIDMb and found no change in the two v-d models. They did, however, find mass function suppression in SIDM10, especially for subhaloes closer to the host's centre. Rocha et al. (2013) also investigated subhalo evaporation, but in the context of the cumulative maximum circular velocity function. They find a small suppression in SIDM1 for subhaloes in $10^{12}\,\mathrm{M}_{\odot}$ hosts, again magnified for subhaloes in the inner regions.

We expand upon these studies, looking at total mass-loss over time as a function of mean orbital pericentre, d_{peri}, and mass, M_{infall}, as in Section 3.4, instead of a static $z = 0$ subhalo count. Since we track all mass, not just stellar mass, particle tagging does not affect the results. Whereas previously we had to eliminate haloes in which no unscattered particles remained for tagging, now we can include all haloes in our sample. Consequently, we can study the SIDM10 model, shown in Fig. 3-7. Here, we see significant effects of subhalo evaporation. In all cases where $d_{\mathrm{peri}} > 30$ kpc, there is a clear and consistent break from CDM 1 Gyr after infall. In the lowest mass bins, the separation continues with SIDM haloes having an additional 10% of their original mass lost. By 10 Gyr, this translates into having only half of the mass of their CDM counterparts. In the highest mass bins, there are hints of increased mass-loss, but with only ~ 20 subhaloes per bin, the statistics are weak and subject to orbit dependent variations. In contrast, our lowest mass bins have ~ 200 haloes per bin. For satellites with $d_{\mathrm{peri}} \leq 30\,\mathrm{kpc}$, no SIDM satellites last longer than 6 Gyr, whereas

some CDM satellites remain throughout the duration of accretion activity for nearly 11 Gyr.

Fig. 3-8 shows results for vdSIDMa. Aside from halo finding noise at low pericentres, the SIDM results are strikingly similar to those of CDM. Similarly, nothing different happens in the case of vdSIDMb. This indicates that subhalo evaporation is not significant for the v-d models, and also that tidal stripping does not affect the DM loss rate differently in cored versus cuspy subhaloes in the regime presented. Only for subhaloes with low pericentre will a core cause increased tidal stripping in the total DM mass, as discussed in Fig. 3-12. Without the added tidal forces of a disc in the host (see Section 3.7.5) and with few subhaloes ever achieving very low pericentre, $r_{\rm peri} < 10\,{\rm kpc}$, the average mass-loss over time is not substantially altered by differential tidal stripping in cored and cuspy subhaloes. Lastly, in Fig. 3-9, SIDM1 shows hints of a weak effect in some instances, with stripping increased relative to CDM by just a few percent after 10 Gyr in orbit. However, as with SIDM10, the highest mass bins have a limited sample of haloes and orbits are perturbed relative to CDM more in the v-i models than in the v-d models. In particular, the extra mass-loss seen in the highest mass and largest pericentre bin of SIDM1 arises mostly due to different satellite orbits, not because of subhalo evaporation.

These results all fit expectations when considering the typical time-scale for a subhalo particle to collide with the background under different orbits. Using the scattering rate of equation (3.1), cross-sections shown in Fig. 3-1, and assuming constant orbital velocity, $v_{\rm orbit}$, the mean time between collisions for a given particle, τ, is given by

$$\tau = \left(\rho(r)\frac{\sigma(v_{\rm orbit})}{m_\chi}v_{\rm orbit}\right)^{-1} \tag{3.2}$$

This time-scale is shown in Fig. 3-10 for each of our SIDM models as a function of radius for circular orbits. The value of $\rho(r)$ is taken from each simulation's respective host at $z = 0$, averaging out spherical asymmetries, and used in turn to compute $v_{\rm orbit}$. Since we do not use an evolving host potential, yet host haloes grow in time, τ is underpredicted for a true dynamical system. We draw a line at $t_{\rm Hubble} = 13.7\,{\rm Gyr}$

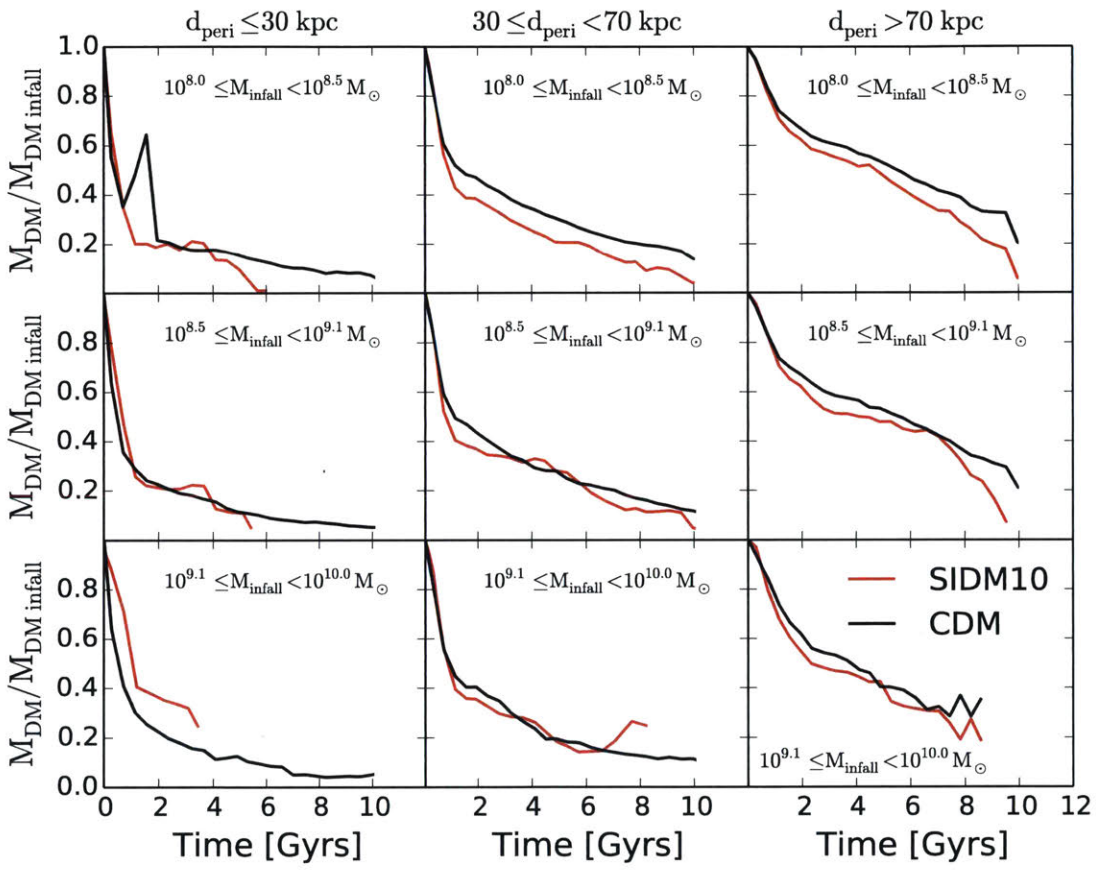

Figure 3-7 Total mass remaining in satellites over time binned by mean pericentre, $d_{\rm peri}$, and infall mass, $M_{\rm infall}$. Scattering is high enough for subhalo evaporation to reduce the mass of satellites in SIDM10 faster than in the CDM case.

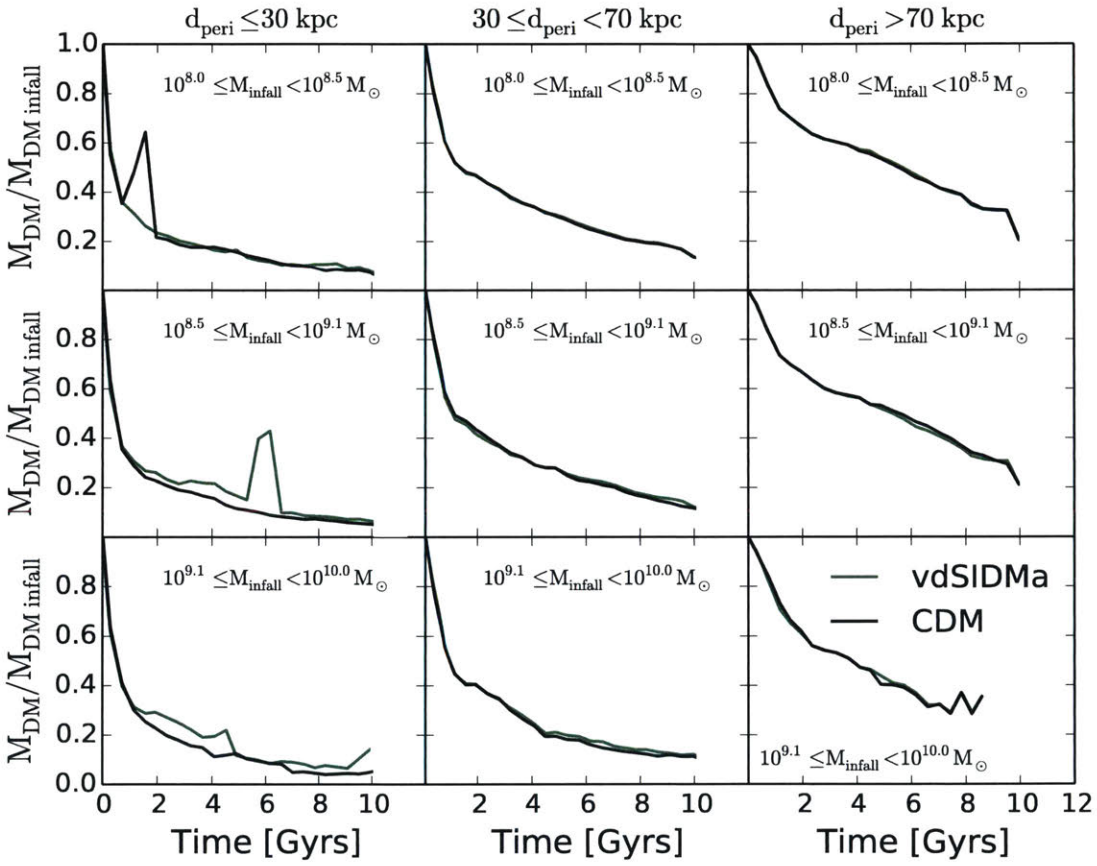

Figure 3-8 Total mass remaining in satellites over time binned by mean pericentre, $d_{\rm peri}$, and infall mass, $M_{\rm infall}$. Subhalo evaporation is small enough to render no observable difference between vdSIDMa and CDM. Simulation vdSIDMb (not shown) shows identical results with no difference with CDM.

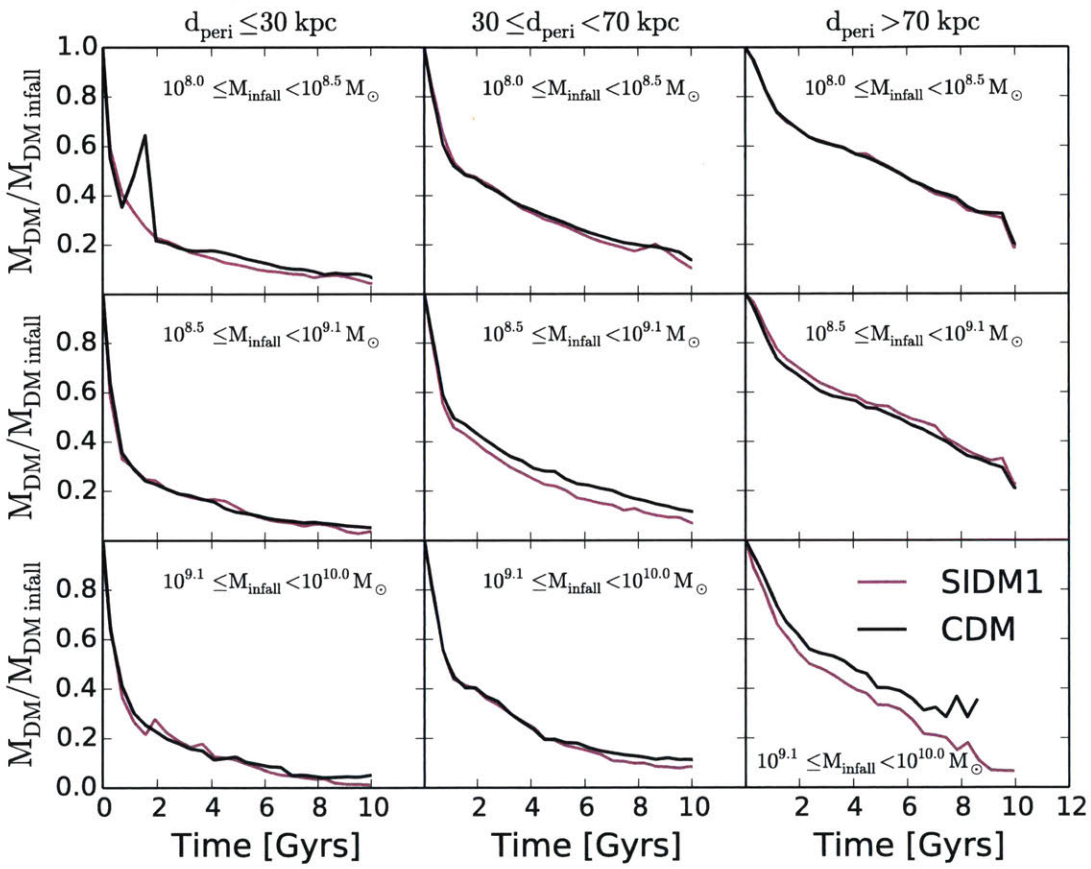

Figure 3-9 Total mass remaining in satellites over time binned by mean pericentre, $d_{\rm peri}$, and infall mass, $M_{\rm infall}$. Subhalo evaporation has a very weak effect in increasing the rate of mass-loss in SIDM1 relative to CDM. The extra mass-loss seen in the highest mass and largest pericentre bin arises mostly due to different satellite orbits of a limited sample, and not due to subhalo evaporation.

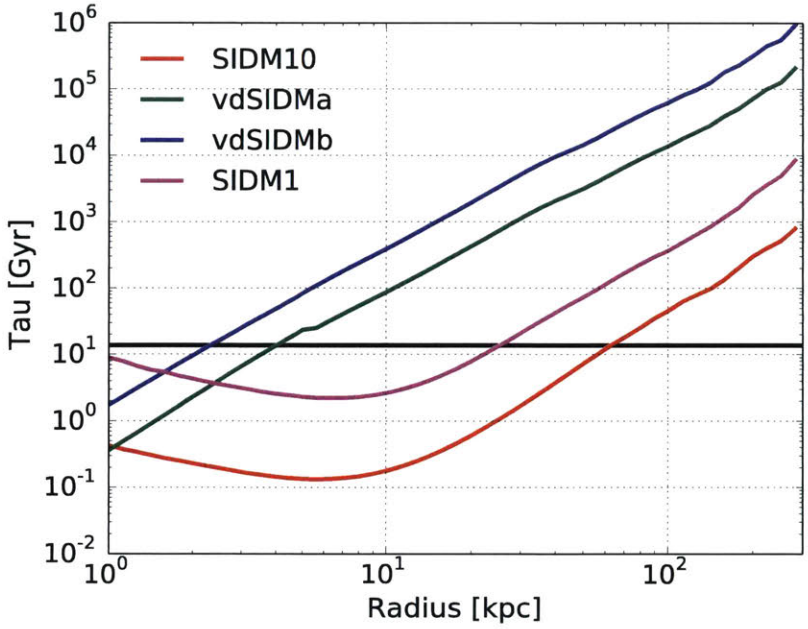

Figure 3-10 Mean time between collisions between background host halo particles and particles within a subhalo on a circular orbit as a function of orbital radius. The black horizontal line is the age of the universe.

to highlight the maximum length of time a satellite can spend in a Milky Way-like system and experience evaporation. Subhalo evaporation becomes important for all radii where $\tau < t_{\rm Hubble}$.

For SIDM10, $\tau < t_{\rm Hubble}$ for all orbits within 60 kpc of the host. Many subhaloes spend a large proportion of their time within this distance, meaning evaporation should noticeably increase the mass-loss rate. SIDM1 has the next strongest level of interaction, a full order of magnitude weaker. Here $\tau = t_{\rm Hubble}$ at a radius of 25 kpc. Haloes that spend a Hubble time within this distance will be fully disrupted due to tidal forces and thus not exist today. Yet the radius is still large enough for satellites to pass within during pericentre and survive with a few percent increase in mass-loss. The v-d models have such long time-scales for scattering that this simple time-scales argument suggests that subhalo evaporation should not have an observable effect on total mass-loss. All three of these predictions agree with the results of Figs 3-7-3-9. Interestingly, due to the host halo density flattening out in the core of our v-i simulations, τ plateaus and even rises as R decreases below 10 kpc.

Since subhalo evaporation does not substantially affect satellites in our viable SIDM models, we rule out the possibility of evaporation contributing to the enhanced loss of stars. Additionally, the lack of change in total mass-loss rate in the v-d models also indicates that tidal stripping does not affect cored SIDM subhaloes more than CDM subhaloes with regard to total DM mass-loss. Instead, another mechanism involving an outward migration of stars must exist. Since tidal stripping preferentially removes mass from the outer halo, but stars originate in the inner halo, any outward transport of stars can lead to increased stellar mass-loss without a simultaneous increase in total mass-loss. We discuss this mechanism in the next section.

3.6 Analytic Explanation for Enhanced Stellar Stripping

To explain the results of Section 3.4 without subhalo evaporation, we turn to the fundamental physics of tidal stripping. External to subhaloes, the strength of the gravitational tidal field controls how much mass is stripped from satellites. This plays a role in determining the tidal radius, $r_{\rm t}$, which defines an approximate distance beyond which mass is no longer bound to a subhalo. Naturally, a smaller tidal radius results in more stripping. Internal to subhaloes, which stellar particles are stripped is controlled by dynamics at the centre of the halo. To compare CDM against SIDM, the size of the halo core, $r_{\rm core}$, is critical since it correlates with the strength of dynamical differences due to self-interactions. SIDM reduces halo central density and increases halo central velocity dispersion, reducing the binding energy of particles within the core. Consequently, an increase in stellar mass-loss manifests in two ways.

First, stars initially formed in the centre of the subhalo are more likely to disperse out to larger radii, and thus more likely to pass beyond the tidal radius. In a hydrodynamical simulation of a field dwarf galaxy with SIDM, Vogelsberger et al. (2014b) demonstrated that the distribution of stars within the core is tied to the dark matter distribution, and that the dark matter halo core size grows with time. As a result,

stars tend to spread out over time. We show this for our simulations in Fig. 3-11, plotting the ratio of the half-mass radius of the stellar component, $r_{1/2}$, as a function of time to the half-mass radius of the stellar component at infall. Just as in Figs 3-4-3-6, we take the average of this radius ratio for all haloes belonging to a certain infall mass and mean pericentre bin. In the vdSIDMa case, stars quickly puff out to larger radii whereas in the corresponding CDM case, $r_{1/2}$ grows at a much slower rate over time or remains constant. The relative strength with which SIDM stars spread out in each infall mass and mean pericentre bin is well correlated with the strength of the enhancement of stellar stripping for all SIDM models. As such, we only include the example of vdSIDMa.

Secondly, for a fixed subhalo mass and tidal field, the tidal radius is diminished in cored haloes for orbits where $r_{\rm t} \leq r_{\rm core}$. Since the tidal radius lies within the core only for strong tidal fields at low galactocentric radii, the second effect is applicable only for a small fraction of satellites. We highlight this effect in Fig. 3-12, showing the tidal radius of a satellite at 30, 15, and 5 kpc from the host halo centre for a cored SIDM and a cuspy CDM subhalo.

With $r_{\rm t}$ and $r_{\rm core}$ as controlling factors, we postulate that the relative strength of stellar mass stripping will follow the ratio $r_{\rm core}/r_{\rm t}$, an indicator of how close stripping is to where stars reside. Using simple models, we compute this ratio as a function of subhalo mass and dark matter model.

As discussed in Rocha et al. (2013), the radius of the cores in SIDM satellites can be estimated by the radius at which particles are expected to scatter once per Hubble time, or approximately once per 10 Gyr halo lifetime since haloes first form after a few Gyr. They find this predicts the magnitude of core sizes well for the equivalent of our SIDM1 simulation. However, Rocha et al. (2013) fit haloes with a Burkert profile (Burkert, 1995) and use the Burkert scale radius, $r_{\rm b}$, to define the core size for SIDM1, commenting that it is likely a lucky coincidence that this works for SIDM1, and would not work as well for other models. We instead choose to define core sizes as the radius at which the density profile of an SIDM satellite diverges from that of a CDM satellite with otherwise equivalent conditions. This definition then accurately

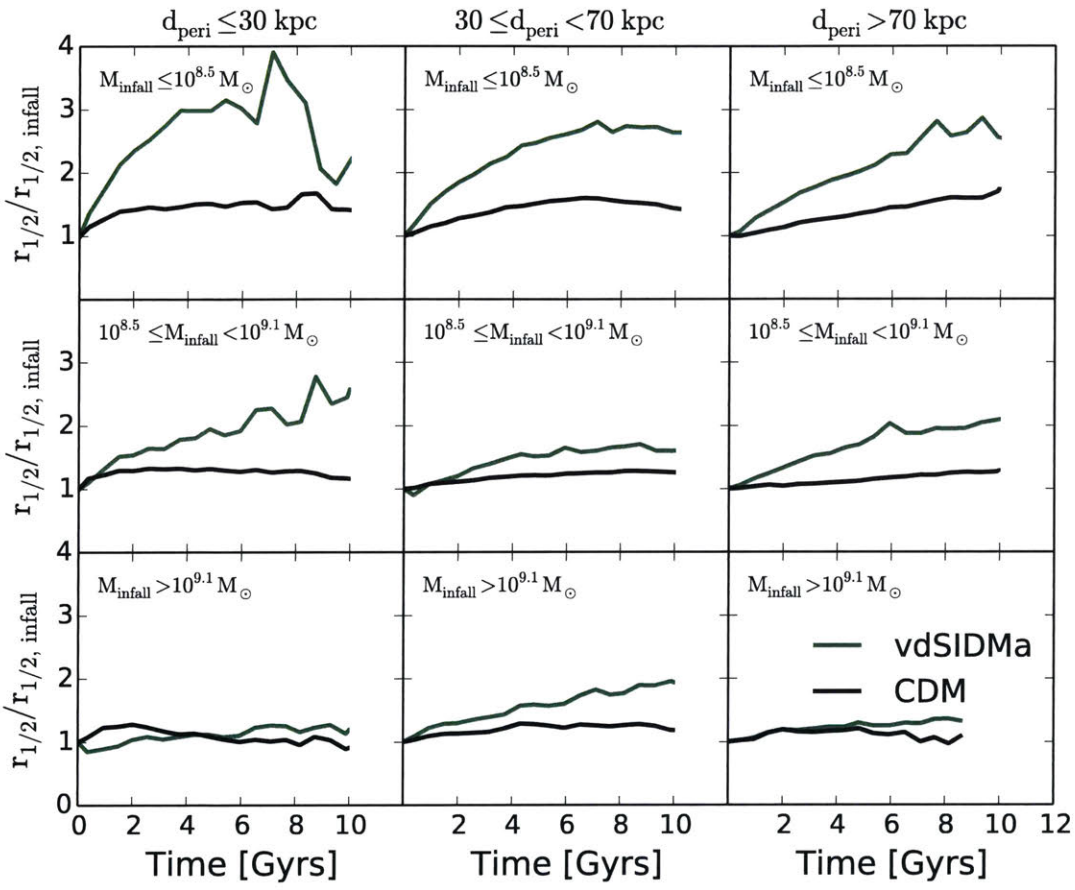

Figure 3-11 The ratio of the half-mass radius of stars that remain bound to a satellite as a function of time to the half-mass radius at infall. Satellites are binned into groups of different mean orbital pericentre and infall mass, and their distributions averaged per bin. Stars in SIDM disperse out to larger radii than in CDM, starting immediately after tagging. This is a primary driver of the increase in stellar mass stripping.

probes where binding energies will differ due to SIDM.

By matching isolated field haloes from the same initial overdensity of SIDM with those in CDM, we find the SIDM density profile typically breaks away from that of CDM when the slope of the log-log profile is

$$\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln\mathrm{r} = -1. \tag{3.3}$$

Our simulation's core sizes measured in this manner best align with the purely analytic estimate when we choose the core size as the radius at which particles scatter once per Gyr instead of once per 10 Gyr. Each of these two methods agree roughly in magnitude, and agree with the scaling of $r_{\rm core}$ with halo mass. Thus we define $r_{\rm core}$ as the distance where

$$\frac{\rho(r_{\rm core})}{m_\chi}\langle\sigma v\rangle(r_{\rm core}) = 1\,\mathrm{Gyr}^{-1}. \tag{3.4}$$

The value $\langle\sigma v\rangle(r)$ is the local thermal average of the transfer cross-section times the relative velocity, estimated as in Vogelsberger et al. (2012) by an average over a Maxwell Boltzman distribution function:

$$\langle\sigma v\rangle(r) = \frac{1}{2\sigma^3_{\rm vel}(r)\sqrt{\pi}}\int(\sigma v)v^2\mathrm{e}^{-\mathrm{v}^2/4\sigma^2_{\rm vel}(\mathrm{r})}\,\mathrm{dv}, \tag{3.5}$$

where $\sigma_{\rm vel}(r)$ is the local velocity dispersion.

Following Vogelsberger et al. (2012), we model our haloes as Hernquist profiles (Hernquist, 1990) with density varying as

$$\rho(r) = \frac{Ma}{2\pi\mathrm{r}}\frac{1}{(r+a)^3} \tag{3.6}$$

and velocity dispersion $\sigma^2_{\rm vel}(r,a,M)$ taken directly from Hernquist (1990). M is an indication of the halo mass, and a is the scale radius. Using our CDM simulation and definition of virial radius, we find that the halo concentration $c \equiv r_{\rm vir}/a$ varies as $c = 80.1 \times \left(\frac{M}{\mathrm{M}_\odot}\right)^{-1/11}$ at $z = 0$. The value of M in terms of our virial mass, $m_{\rm vir}$

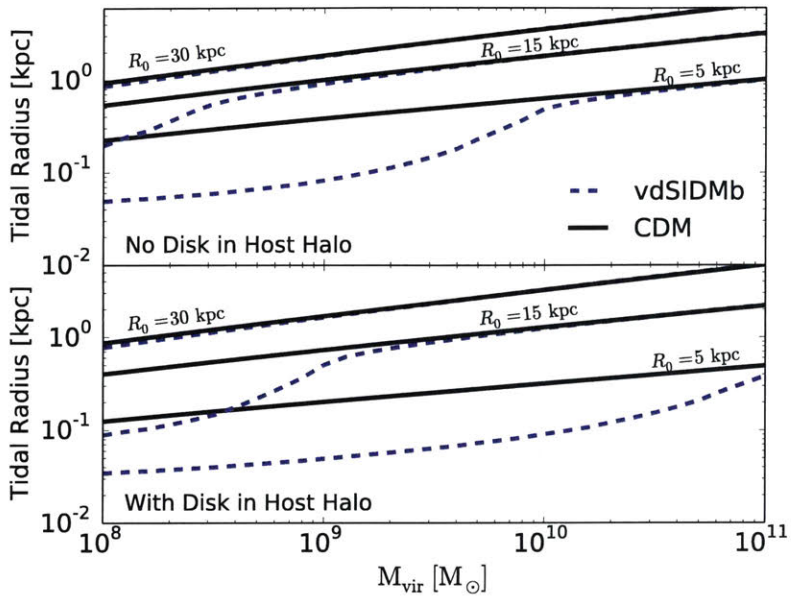

Figure 3-12 Approximate tidal radius of subhaloes in the gravitational field of a $2.2\times10^{12}\,M_\odot$ Hernquist profile CDM-like host at distances of 30, 15, and 5 kpc. Shown in blue is the tidal radius for subhaloes with cores of the size predicted for vdSIDMb. At low orbital distances and low halo masses, the tidal radius begins to lie within the low density core, resulting in a reduction in r_t compared to cuspy haloes (solid black lines). If the host halo includes a disc potential (bottom panel), tidal forces are increased at low radii and r_t is reduced even further compared to cuspy haloes. We assume a disc of mass $1\times10^{11}\,M_\odot$ of the same form as used in Peñarrubia et al. (2010). For orbital distances larger than 30 kpc in all SIDM models, r_t is negligibly affected by cores in subhaloes.

is computed as

$$M = m_{\rm vir}\frac{(1+c)^2}{c^2} \tag{3.7}$$

and the scale radius is simply $a = r_{\rm vir}/c$.

Now for any $m_{\rm vir}$ we use equation (3.4) to compute the core sizes of each of our SIDM models, as shown in Fig 3-13. In the two velocity-independent models, the core size scales with halo mass. As central density increases, scattering increases. In the velocity-dependent models, there is a competing effect of higher densities and higher velocity dispersions (and thus lower cross-section, see Fig. 3-1) making the core size vary only slightly with halo mass. Consequently, core sizes of vdSIDMa and vdSIDMb are lower than that of SIDM1 for high mass haloes, then overtake SIDM1 between 10^9 and $10^{10}\,M_\odot$ going towards lower halo mass. Due to its cross-section's

steep dependence on velocity, vdSIDMb even overtakes SIDM10 for haloes smaller than $10^8\,\mathrm{M}_\odot$ in spite of having the smallest cores above $10^{10}\,\mathrm{M}_\odot$.

To continue our estimate of $r_{\mathrm{core}}/r_{\mathrm{t}}$, we need to compute the tidal radius. In a three body system of host, satellite, and star, the last closed effective potential surface indicates the region in which stars will still be bound to the subhalo. The distance to this surface at the L_3 Lagrange point is a good approximation of r_t. For a circular orbit and spherically symmetric host and subhalo, r_{t} is given in Binney and Tremaine (2008) and rewritten here as

$$r_{\mathrm{t}} = R_0 \left(\frac{m(r_{\mathrm{t}})}{M(R_0)(3 - \frac{d\ln M}{d\ln R}|_{R=R_0})} \right)^{1/3}, \tag{3.8}$$

where R_0 is the radius of the satellite's orbit, $m(r_{\mathrm{t}})$ is the mass enclosed by the satellite at a distance r_{t} from its centre, and $M(R)$ is the mass enclosed by the host at distance R from the host's centre.

Isolating how the tidal radius varies with satellite mass, we compute and plot in Fig. 3-12 the tidal radius for subhaloes within a $2.2 \times 10^{12}\,\mathrm{M}_\odot$ CDM-like host at various fixed distances to the host halo's centre. The tidal radius increases with both satellite mass, m_{sat}, and galactocentric distance, R_0, closely following the approximate equation

$$r_{\mathrm{t}} \approx 0.7 \left(\frac{m_{\mathrm{sat}}}{M_{host}} \right)^{1/3} R_0 \tag{3.9}$$

for Hernquist haloes. Since the tidal radius depends on the mass enclosed within a satellite at a given radius, and the mass enclosed is reduced within a core, the tidal radius is suppressed for SIDM haloes when $r_{\mathrm{t}} \leq r_{\mathrm{core}}$. Just beyond the core radius, there is a buildup of excess mass in SIDM haloes relative to their equivalent CDM haloes since mass displaced from the halo centre is not removed entirely. As a result, the same total mass is enclosed at all larger radii, and there is no difference in r_{t} for SIDM and CDM haloes. We verify this effect by modifying the Hernquist profile for our satellites to enforce $\mathrm{d}\ln\rho(\mathrm{r})/\mathrm{d}\ln\mathrm{r} = -0.3$ within r_{core} while preserving the same total virial mass. While a true core has a logarithmic density profile slope of 0.0, the

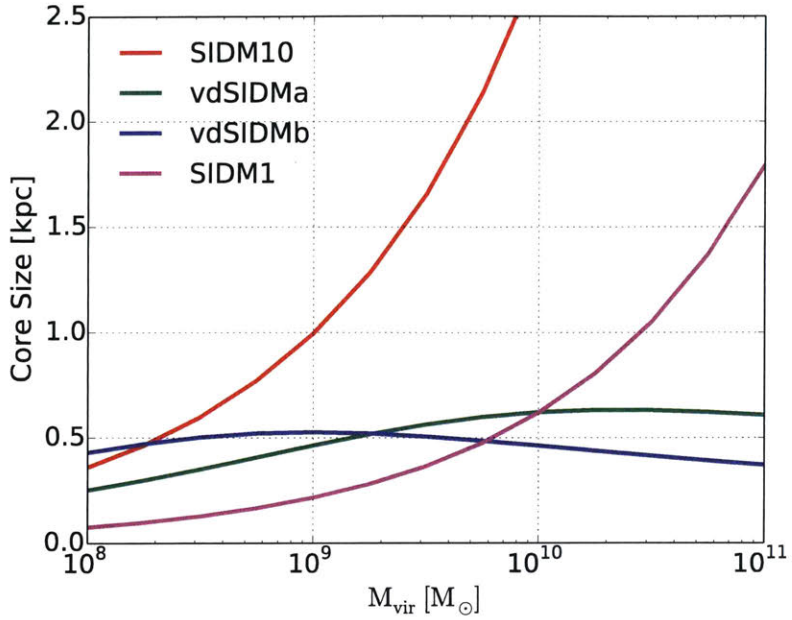

Figure 3-13 Core sizes as a function of virial mass predicted for haloes in each SIDM model. Core size is estimated as the radius where particles scatter on average once per Gyr for an isolated halo.

profile of satellites in our simulations is rarely completely flat and instead closer to -0.3.

As seen in the blue dashed lines of Fig. 3-12, the tidal radius of cored satellites is reduced relative to that of a cuspy halo (black solid lines) for orbits where $R_0 \lesssim 30\,\mathrm{kpc}$, and the reduction becomes increasingly important for lower galactocentric radii. The effect is also magnified as the satellite core size increases relative to the tidal radius of the same mass cuspy satellite. For the vdSIDMb model illustrated, this happens towards lower satellite masses. The bottom panel of Fig. 3-12 illustrates how the presence of a disc in the host halo would enhance these effects. Here, we assume a disc of mass $1 \times 10^{11}\,\mathrm{M}_\odot$ of the same form as used in Peñarrubia et al. (2010), and reduce the host's dark matter mass so the sum total of halo plus disc mass is still $2.2 \times 10^{12}\,\mathrm{M}_\odot$. Beyond 30 kpc for all SIDM models, the change in r_t is negligible even with the presence of a disc.

In Fig. 3-14, we now compute the ratio $r_\mathrm{core}/r_\mathrm{t}$ with r_t found for a satellite at a distance of 45 kpc from the host. The ratio shows the intricacies of the effects of each

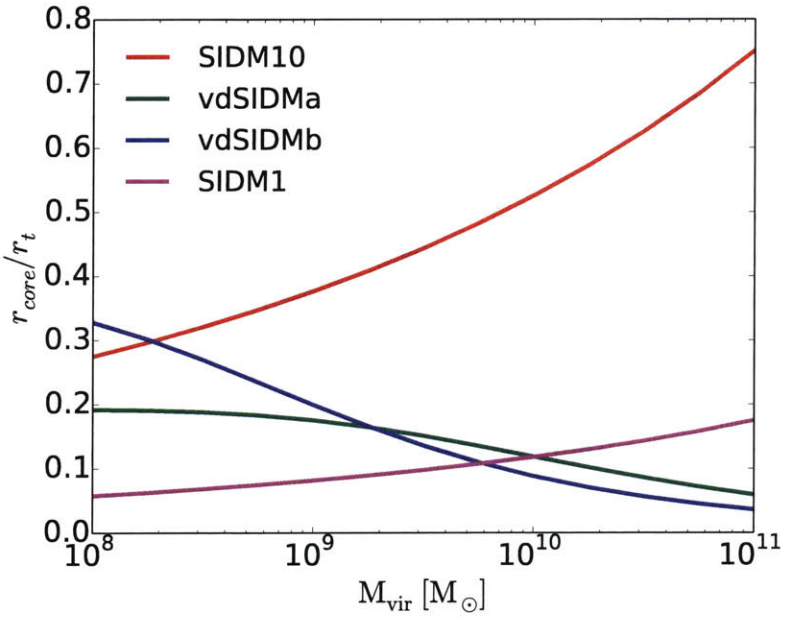

Figure 3-14 Ratio of the core size of subhaloes to tidal radius in a $2.2 \times 10^{12}\,M_\odot$ host with subhaloes placed at a distance of 45 kpc from the host's centre. Higher values of this ratio indicate a higher propensity for stellar mass stripping. The inverse dependence of cross-section on relative velocity in vdSIDMa and vdSIDMb leads to more enhanced stripping in low-mass satellites, while the opposite is true in constant cross-section models.

SIDM model and the mass dependence of the tidal radius. In the v-i models, due to $r_{\rm core}/r_{\rm t}$ increasing as a function of halo mass, we expect the disparity in stripping relative to CDM to increase with halo mass. Indeed, as shown in Section 3.4, SIDM1 does have a higher rate of stripping relative to CDM for larger mass haloes when $d_{\rm peri} > 30\,{\rm kpc}$. This is driven by core sizes that increase in size faster than the tidal radius for more massive subhaloes. The opposite is true of the v-d models. Here, core sizes vary only weakly with mass, allowing for larger haloes to more effectively shield their inner region via larger tidal radii. The vdSIDMb model particularly demonstrates this where haloes above $10^{10}\,{\rm M}_{\odot}$ should exhibit little difference with stripping in CDM, but haloes with a mass of $10^{8}\,{\rm M}_{\odot}$ magnify the effects of SIDM enhanced stripping the most of any model, including the high cross section case of SIDM10. This is precisely what the simulation data in Figures 3-4 and 3-5 show. The SIDM enhancement of tidal stripping is large for low mass haloes, and quickly drops to become a small effect for more massive haloes.

By itself, Figure 3-14 does not explain the different stellar stripping behaviour in SIDM1 from the v-d models for the interval of $d_{\rm peri} \leq 30\,{\rm kpc}$. The behaviour arises from differences in the host halo potential. In the v-i models, the host halo develops a substantial core, as can be imagined by extrapolating the core sizes of Fig. 3-13 to a 2.2×10^{12} ${\rm M}_{\odot}$ halo. In SIDM1, this core causes a reduction in the mass enclosed versus radius profile relative to CDM out to 17 kpc. In the v-d models, the collisional cross-section drops enough in the high velocity dispersion region of the host's centre to affect only $\sim 2\,{\rm kpc}$ in vdSIDMa and $\sim 1\,{\rm kpc}$ in vdSIDMb. Since the strength of the tidal force, as approximated in equation (3.8), depends on the mass enclosed in the host for a fixed radius, haloes in the SIDM1 experience a reduction in $r_{\rm t}$ relative to the v-d models when they pass near or within 17 kpc. Since haloes have extended distributions, a subhalo need only pass near 17 kpc for part of its extremities to pass within the critical radius where tidal forces are reduced relative to CDM. Therefore, any subhalo in SIDM1 orbiting in this range will face competing effects of the subhalo core increasing the likelihood for stars to escape relative to CDM, but also the host halo core reducing tidal forces and decreasing this likelihood. The result seen in the

first two rows of column 1 in Fig. 3-6 suggests that these effects cancel each other out enough to produce a CDM level of stripping. In the case of the highest mass bin at low pericentre, low number statistics lead to more chaotic behaviour.

3.7 Implications

In this section, we highlight the consequences of SIDM on the $z = 0$ ensemble of satellites, and the stellar halo.

3.7.1 Mass Functions

As demonstrated in Section 3.4, SIDM causes an increase in the mass-loss rate of stellar material but does not change significantly the mass-loss rate of the total dark matter mass. In turn, the subhalo mass function in terms of dark matter is not affected, but the stellar mass function should be suppressed. Due to limitations of not tagging the same total number of haloes in SIDM as in CDM, however, we cannot directly show the suppression.

Instead, in Fig. 3-15 we show the fraction of stars stripped since infall for $z = 0$ satellites in SIDM, $f_{\rm SIDM}$, divided by the fraction of stars stripped since infall for $z = 0$ satellites in CDM, $f_{\rm CDM}$, versus the infall mass of satellites. For reference, we include the approximate stellar mass per satellite on the upper axis based on Moster et al. (2013) abundance matching. The signals seen for all satellites over the entire accretion history of the host persist at the single snapshot of the present day. In all three SIDM models, there is a clear increase in stellar mass stripped relative to CDM. In the v-d models, the velocity dependence of the cross-section shows its effects clearly. Lower mass satellites with larger cross-sections lead to increased core sizes relative to their subhalo radius and a stronger enhancement of stripping. The effect of SIDM1 is more constant with mass. While these trends do appear, there is also tremendous scatter from halo to halo arising from different infall times and orbits.

In spite of nearly all stripping occurring while satellites are within 50 kpc of the host, there is not a strong correlation of $f_{\rm SIDM}/f_{\rm CDM}$ with the $z = 0$ galactocentric

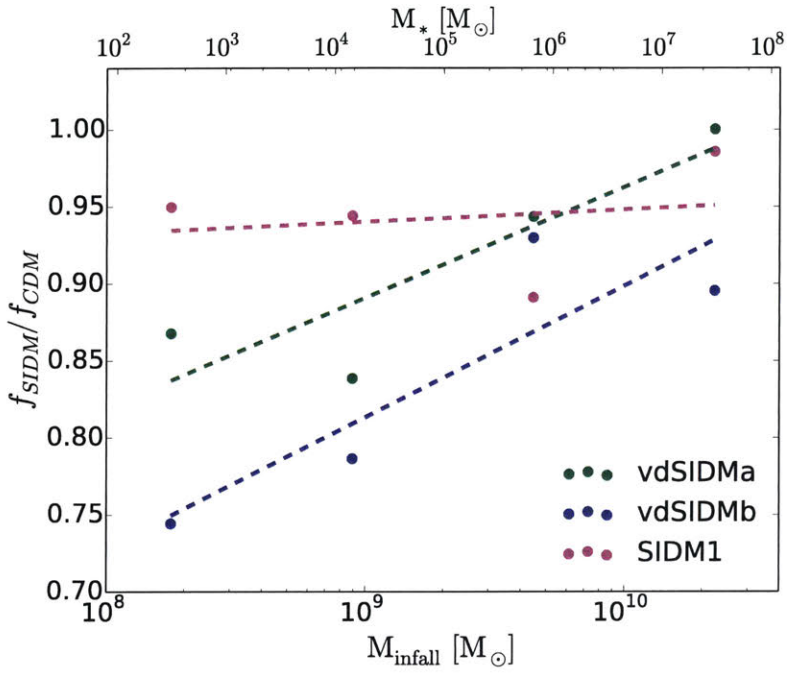

Figure 3-15 Ratio of the fraction of stars remaining in SIDM haloes to that in CDM haloes at $z = 0$ as a function of halo infall mass. SIDM haloes lose more stars, which in turn causes a suppression in the stellar mass function. The strength of this effect correlates well with the size of the core radius relative to the tidal radius as a function of mass (see Section 3.6). The upper axis shows the approximate stellar mass per satellite based on Moster et al. (2013) abundance matching.

distance of satellites all the way out to 300 kpc. Beyond that, satellites in both SIDM and CDM are stripped by just a few percent, if at all. Orbits are elliptical enough to disperse affected satellites out to large radii. We caution that due to the nature of our particle tagging, we cannot definitely declare the magnitude of the signal, only the trends. A full hydrodynamical simulation, or careful injection of high resolution, non-scattering star particles would be needed to establish the magnitude properly.

3.7.2 Stellar half-mass radius

In Section, 3.6 we demonstrated that stars migrate to larger distances on average in SIDM than in CDM. This results in an increased half-light radius, $r_{1/2}$, an effect also seen for cored haloes in Vogelsberger et al. (2014b) and Errani et al. (2015). We therefore measure the half-mass radius, a proxy for half-light radius, of all satellites

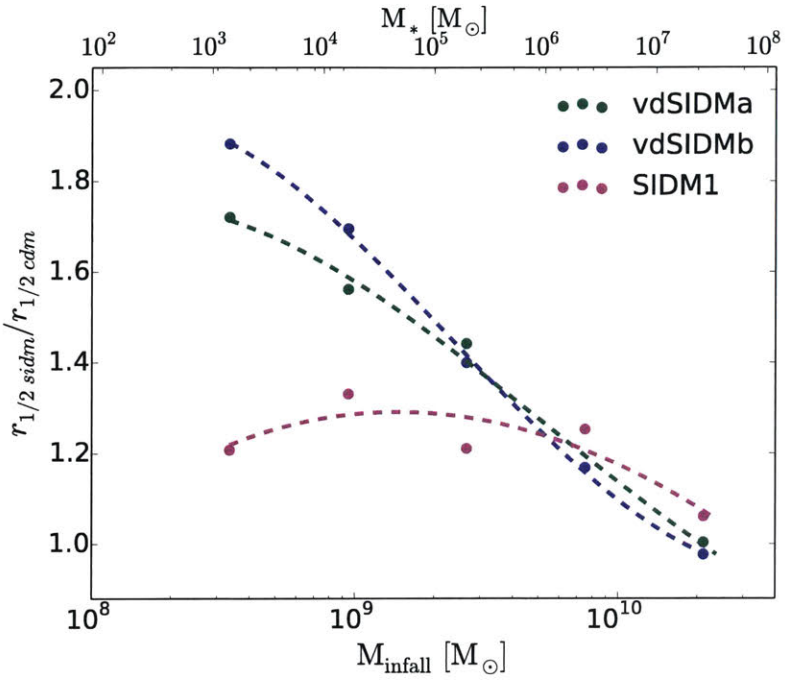

Figure 3-16 SIDM to CDM ratio of the half-mass radius of stars in all satellites at $z = 0$ as a function of their infall halo mass. The half-mass radius increases relative to that of CDM in all SIDM models due to stars drifting outwards as satellite cores grow. The upper axis shows the approximate stellar mass per satellite based on Moster et al. (2013) abundance matching.

in our simulations at $z = 0$. We then take the ratio of the half-mass radius for SIDM to that of CDM and plot it as a function of the subhalo infall mass in Fig. 3-16. Once again we see the same ordering of effects for each of our SIDM models. In all cases, SIDM leads to a larger $r_{1/2}$. For masses above $10^{10}\,M_\odot$, SIDM1 has the largest effect, then vdSIDMa, then vdSIDMb. Towards lower halo masses, the effect increases rapidly for the v-d models. We caution that the magnitude of the increase relative to CDM is not independent of our tagging technique. The trends and relative effect of each model however, reinforced visually by the dashed lines, are much more robust. As in Fig. 3-15, we include the approximate stellar mass per satellite on the upper axis based on Moster et al. (2013) abundance matching.

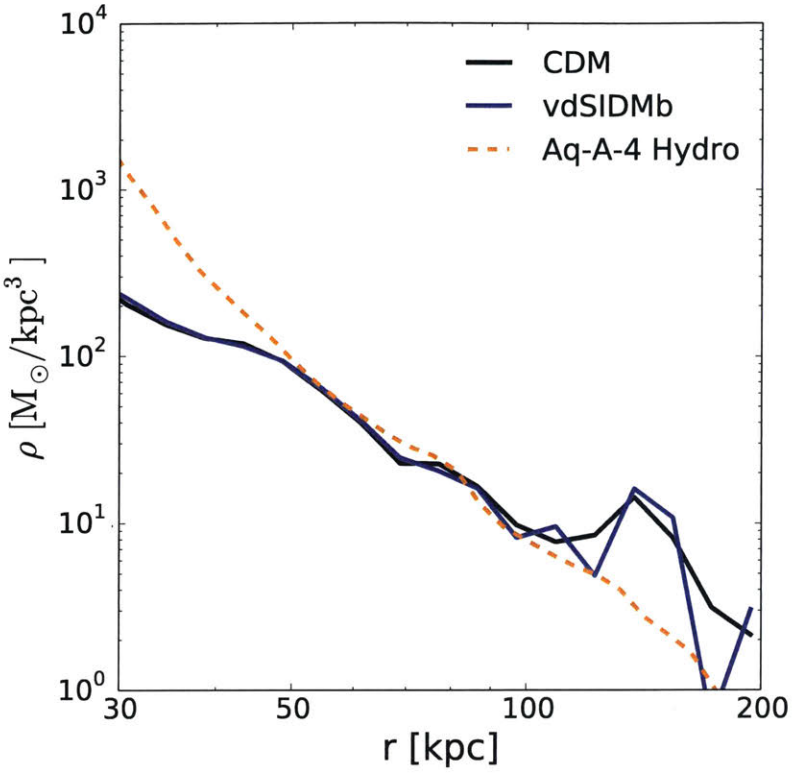

Figure 3-17 Stellar density profile of tagged particles in our host halo due to satellite accretion for our vdSIDMb and corresponding CDM simulations in the outer halo. For comparison we show the stellar density profile of the hydro Aquarius run (Marinacci et al., 2014) when scaled down in magnitude by a factor of 11 to match the Moster et al. (2013) stellar mass-halo mass relationship. Our tagging technique reproduces the slope very well between 50 and 100 kpc where effects of the disc are unimportant. We do not see differences in the stellar density profile between CDM and SIDM in of our simulations.

3.7.3 Stellar Halo

The stellar halo consists of all stars beyond the influence of the host halo's disc. Since our simulations only trace stars formed in accreting satellites, we focus our predictions on the outer halo where accreted stars dominate over stars formed within the host halo. Pillepich et al. (2015) find that beyond $\sim 30\,\mathrm{kpc}$, more than 95% of stellar mass comes from accreted stars in a Milky Way-like host. We therefore consider the stellar halo as the region between 30 kpc and r_{vir} of our hosts.

One major result is the lack of an effect the v-d models have on the stellar halo. In Fig. 3-17, we show the stellar density profile of vdSIDMb compared to the corre-

sponding CDM case. The two profiles are remarkably similar, with no apparent effect on the density profile due to SIDM. We find the same result for the radial velocity dispersion profile of stars (not shown). To verify that our particle tagging properly reflects features of the stellar halo, we compare the $z = 0$ stellar density profile of our host halo to that of the Marinacci et al. (2014) simulation, which is a hydrodynamic resimulation of the same initial conditions we used but at a lower DM particle resolution, Aq-A-4. From looking at their $z = 0$ field haloes, we find that the slope of their stellar mass-halo mass function is consistent with that of Moster et al. (2013), but is a factor of 11 higher in magnitude. After uniformly scaling their stellar masses down by that factor of 11 to better mimic the Moster et al. (2013) relation, our stellar density profile matches reasonably well the hydro run (orange line) between 50 and 100 kpc.

The v-d models make such little influence due to their small interaction cross-sections in haloes of mass scales above $10^{10}\,\mathrm{M}_{\odot}$. The host halo forms only a very small core, leading to minimal perturbations of the orbits of satellites and the spherically averaged DM dynamics of the host halo in the outer regions. As discussed in Sections 3.5 and 3.6, satellites with $M_{\mathrm{infall}} > 10^{10}\,\mathrm{M}_{\odot}$ have a low $r_{\mathrm{core}}/r_{\mathrm{t}}$ ratio, experiencing minimal enhancement in stellar stripping, and have no subhalo evaporation. Yet 90% of the stellar mass accreted on to the host comes from such satellites. Thus the stellar halo is dominated by of the order of 10 large satellites which are not substantially affected by SIDM, in a host halo potential which is minimally changed, resulting in no global changes to the stellar halo.

In order to observe any effects of the v-d models, smaller satellites must be isolated. We highlight one such instance of extracting a signal for vdSIDMb in Fig. 3-18. Here, we show a measure of the smoothness of tidally stripped stars across the sky inspired by methods in Bell et al. (2008). We pixelate the sky into 12288 pixels and find the total amount of stellar mass per pixel that is not bound to an identifiable subhalo. We then compute the ratio of the stellar mass density, ρ, to the mean stellar mass density per pixel, $\bar{\rho}$, and take the logarithm to yield $\log_{10}(\rho/\bar{\rho})$. We plot a histogram of these values for stars that came from satellites with infall mass between

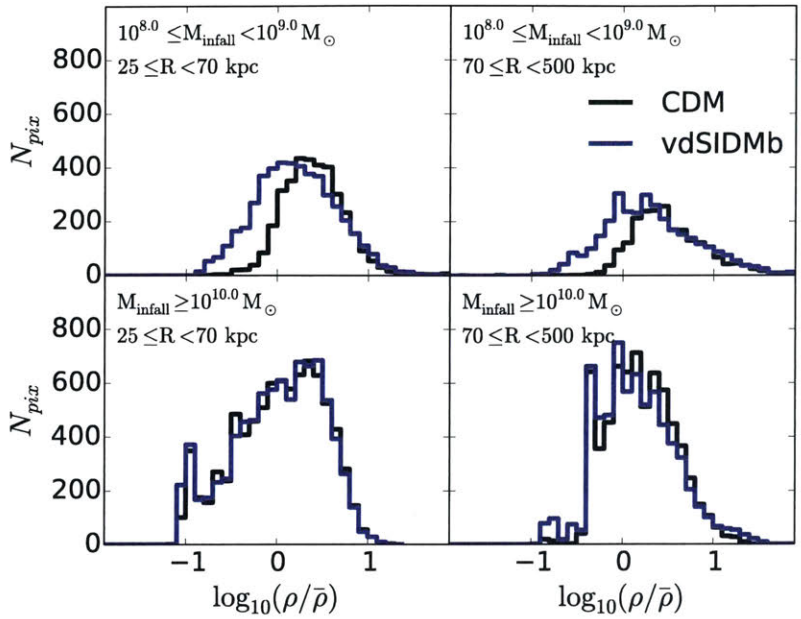

Figure 3-18 Histogram of the density of all stripped stellar mass beyond 30 kpc relative to the mean stellar density, as observed in 3.4 deg^2 pixels. Stars are divided into four samples according to the infall mass of the satellite where they came from. We remove stars bound, thus not yet stripped, in subhaloes above $10^8 M_\odot$ in this sample. Density is given as $\log_{10}(\rho/\bar{\rho})$. Shown is the case of vdSIDMb.

$10^8 < M_{\rm infall} < 10^9\,M_\odot$ (top row), and stars from satellites with infall mass between $10^{10} < M_{\rm infall} < 10^{13}\,M_\odot$ (bottom row). Consistent with expectations, no difference exists in the distribution for stars coming from the larger satellites which dominate the total signal. For the lower mass satellites, we do find enhanced stellar stripping. With more stars dispersed across the sky, there is a greater tail of sky patches with a lower than mean density of stars, and fewer sky patches completely devoid of tidally stripped stars (60 per cent for vdSIDMb versus 70 per cent for CDM).

Both of these distributions hold independent of the radial range of the stellar halo studied. In spite of nearly all stars being stripped from satellites while they are within 50 kpc of the host, elliptical orbits carry the effect of enhanced stellar stripping to much larger radii. This is seen in a comparison of the left-hand column which isolates the stellar halo between $25 \leq R < 70$ kpc, and the right-hand column which isolates the stellar halo between $70 \leq R < 500$ kpc. We caution that this figure serves primarily to illustrate a point and not to find a regime of maximal signal.

Different scale sizes of patches on the sky should produce different signals. With limited particle resolution, particularly from stars originating in low-mass satellites, any further investigation is difficult.

In the case of SIDM1, some differences do arise in the stellar density profile and velocity dispersion profile, but primarily result from perturbations of satellite orbits due to a core in the host halo. These perturbations are visually apparent in Fig. 3-3. The largest satellites in SIDM1 do have significant cores and consequently have enhanced stripping relative to CDM for a fixed orbit. They also produce tidal streams with a higher velocity dispersion perpendicular to the orbital plane, as demonstrated in Errani et al. (2015) for satellites with cores. However, the global level characteristics of the stellar halo considered are dominated by the particular orbits of the handful of most massive satellites. We are unable to isolate effects directly indicative of SIDM without detailed knowledge of the orbits. On smaller mass scales, SIDM1 does not produce large enough cores as in the v-d models to create significant features like those shown in Fig. 3-18 for the vdSIDMb model.

3.7.4 Alternate Abundance Matching Models

Due to limited observations, and stochasticity in the stellar mass halo mass relation, abundance matching for haloes with total dark matter mass below $10^{10}\,M_\odot$ is very uncertain. For the sake of assigning approximate stellar masses, we used the Moster et al. (2013) abundance matching relationship which influences only Figs 3-17 and 3-18, and the upper x-axis of Figs 3-15 and 3-16. A steeper stellar mass halo mass relationship as suggested in Brook et al. (2014) would mean the largest satellites dominate the stellar halo even more, washing out any signals from v-d models more than already presented. A more shallow relationship, as suggested in Behroozi et al. (2013a), would do the opposite. What stellar mass is associated to haloes also affects the luminosity of galaxies in the Milky Way which probe the regimes of influence by each SIDM model. For instance, where v-d models produce the greatest signal, in haloes with infall mass between 10^8 and $10^9\,M_\odot$, Moster et al. (2013) suggests a stellar mass of $10^2 < M_* < 2 \times 10^4\,M_\odot$, coinciding with ultrafaint dwarfs. The

relationship of Eq. 2 in Brook et al. (2014) extrapolated to lower masses suggests a stellar mass of $2 < M_* < 2 \times 10^3\,\mathrm{M}_\odot$, perhaps too low to be observed. In contrast, the relationship in fig. 4 of Sawala et al. (2015) suggests such haloes could host galaxies of approximately $M_* = 10^5\,\mathrm{M}_\odot$, making observations much more tractable.

3.7.5 Effect of a Disc

A significant effect not modelled in our simulations is the presence of a baryonic disc in the host halo. Peñarrubia et al. (2010) show that a disc increases the tidal forces on all satellites, and widens the gap in tidal stripping between cuspy and cored satellites. As demonstrated in Fig. 3-12, when the tidal radius starts to probe within a satellite core, it drops to lower values with cores than with cusps. Strong tidal forces during low pericentric passages are needed to reach this regime. Since a disc + DM halo increases tidal forces relative to just a DM halo at low radii, it serves to enhance the separation of stellar stripping in cored versus cusped systems.

We therefore expect all satellites with pericentre $r_{\mathrm{peri}} \lesssim 30\,\mathrm{kpc}$ to have more stellar stripping, with SIDM satellites affected more than CDM satellites. Approximately 1/3 of all existing $z = 0$ subhaloes with $m_{\mathrm{sub}} > 10^8\,\mathrm{M}_\odot$ within the virial radius of 330 kpc have passed within 30 kpc of the host. To fully quantify the effect, a set of simulations with a disc is needed.

3.8 Conclusion

SIDM may offer a resolution to the too big to fail problem and cusp/core issues of satellite and field galaxies in the Local Group (Vogelsberger et al., 2012; Zavala et al., 2013). Scattering between particles drives mass away from dark matter halo central density cusps, creating lower density cores. Currently, constraints on velocity-independent SIDM places the interesting range of the interaction cross-section at $0.1 \leq \sigma/m_x \leq 1\,\mathrm{cm}^2/\mathrm{g}$ (Rocha et al., 2013). However, the class of velocity-dependent SIDM models, in which the cross-section scales approximately as $\sigma \propto 1/v^\alpha$ for some α, remain largely unconstrained. In this paper, we search for new signatures of

SIDM on satellite galaxies and their interactions with a host halo to ultimately aid in identifying or constraining SIDM.

We employ a suite of five cosmological DM simulations of Milky Way sized galaxies from Vogelsberger et al. (2012), two with velocity-dependent SIDM models, two with velocity-independent SIDM models, and one with pure CDM. We then carefully tag stellar mass to central particles in satellites at infall, and investigate how SIDM changes their dynamics. Our main results are as follows.

1. **SIDM causes an enhanced level of stellar mass-loss in satellites.** Normalized by time since infall and mean orbital pericentre, satellites lose stellar mass at a faster rate in all SIDM models considered relative to CDM. The discrepancy between CDM and SIDM grows towards lower masses in velocity-dependent models, with large differences in satellites with $M_{\rm infall} < 10^{9.1}\,{\rm M}_{\odot}$, and little difference in satellites with $M_{\rm infall} > 10^{9.1}\,{\rm M}_{\odot}$. In velocity-independent cases, the difference increases weakly with higher mass subhaloes.

2. **Subhalo evaporation does not affect MW satellites for viable SIDM models.** Based on the mean time for collisions between a satellite particle and a host halo background particle, subhalo evaporation has virtually no effect on the mass-loss rate of satellites in v-d SIDM models. In the v-i SIDM1 model, a very small effect may exist, backed by simulations hinting at an increased mass-loss of a few percent in satellites 10 Gyrs after accretion. The effect, however, is substantially less important than the precise orbit traversed by satellites. Only for the ruled out cross-section of $10\,{\rm cm}^2/{\rm g}$ do effects of subhalo evaporation become important.

3. **The level of enhanced stellar stripping can be predicted by the ratio $r_{\rm core}/r_{\rm t}$.** The relative strength of the difference in stellar mass loss between SIDM and CDM is proportional to the SIDM core radius, $r_{\rm core}$, divided by the tidal radius, $r_{\rm t}$. In v-d models, $r_{\rm core}$ grows much more slowly with subhalo mass than $r_{\rm t}$, leading to more CDM-like behaviour for larger haloes. In v-i models, $r_{\rm core}$ grows slightly faster than $r_{\rm t}$.

4. **In velocity-independent SIDM models, tidal stripping is much less efficient for low pericentre orbits than in velocity-dependent models.** A large core forms in the host halo in v-i models, whereas only a small core forms in the v-d models. Consequently, tidal forces in v-i models are reduced for orbits with $r_{\rm peri} \lesssim 20\,{\rm kpc}$ compared to CDM and v-d SIDM. The reduced tidal forces counteract the reduced central binding energy of stars in satellites resulting in little stellar mass-loss relative to CDM. For very low pericentre orbits, $r_{\rm peri} \lesssim 15\,{\rm kpc}$, and large satellite cores in the v-d models, $r_{\rm t}$ can lie within a halo's core, reducing $r_{\rm t}$ relative to cuspy haloes and further accentuating the increase in stellar stripping.

5. **SIDM suppresses the stellar mass function of satellites.** As a consequence of increased stellar mass-loss, the stellar mass function is reduced at all satellite mass scales in accordance with the ratio $r_{\rm core}/r_{\rm t}$ for some characteristic orbit.

6. **SIDM increases the half-light radius of satellites.** Due to growing core sizes, stars tend to disperse out to larger radii in SIDM than in CDM. In addition to increasing their likelihood of being tidally stripped, the dispersal leads to larger half-light radii, which scales in accordance with the size of $r_{\rm core}$.

7. **SIDM does not produce easily identifiable global signatures on the stellar halo.** The stellar halo is dominated by $\sim$ 10 large accretions with $M_{\rm infall} > 10^{10}\,{\rm M}_{\odot}$. In our v-d models, these large satellites form small enough cores to leave no imprint on the stellar density profile, velocity dispersion profile, nor spatial distribution of stars on an all sky projection. In SIDM1, there are likely small imprints if all other variables remain unchanged, but the specific orbits of the largest accretions dominate any SIDM signal.

8. **Ultrafaint dwarfs may be key to observationally distinguishing between v-i SIDM, v-d SIDM, and purely baryonic core formation.** Since the cusp/core issue and too big to fail problem mostly concern galaxies on the

scale of the Milky Way classical dwarfs, leading theories of core formation all have mechanisms to resolve the central density issues on mass scales of approximately $10^{9.5} < M_{\rm infall} < 10^{10.5}\,{\rm M}_\odot$. Extending the theories to lower mass satellites, in particular those with $M_{\rm infall} < 10^{9.1}\,{\rm M}_\odot$, leads to divergent predictions among the possibilities. Finding evidence of large cores in ultrafaint dwarfs would give support to v-d SIDM. Finding small cores would be consistent with v-i SIDM, and finding cuspy density profiles would support pure CDM, and in turn strengthen the case for purely baryonic means of core formation in galaxies with sufficient star formation. By comparing half-light radii and enclosed dynamical masses of ultrafaint dwarfs to those of cored and cuspy satellites in simulations, Peñarrubia et al. (2010) find stronger evidence for cuspy haloes. Cored satellites lose too much mass to stripping at low galactocentric distances. Further studies incorporating more recently discovered ultrafaint dwarfs will be critical to solidify an answer.

While our simulations and theory are sufficient to justify each of these conclusions, full hydrodynamic simulations are needed to better determine the precise magnitude of each of effect, subject to uncertainty in the strength of baryonic core formation mechanisms on top of SIDM. Additionally, even though we discuss all consequences of cores in terms of SIDM, similar concepts will apply to cores formed via other means. In particular, the migration of stars to larger radii, enhanced stellar stripping, suppression of the stellar mass function, and increase in the half-light radius would all apply generically to satellites with cores that grow gradually over time. The ratio $r_{\rm core}/r_{\rm t}$ would still be an indicator of the strength of each of these effects.

3.9 Appendix: Particle Tagging Verifications

In Section 2.2, we described our particle tagging technique and the challenges arising from scattering particles in SIDM simulations. In Fig. 3-2, we presented one test of our methodology, showing how particles were selected in similar fashion according to their boundedness percentile in each SIDM-CDM pair. In Section 3.7.3, we verified

that the slope of the stellar density profile in the outer halo matches that of a hydro simulation. Here, we list the additional tests taken to ensure our results do not arise from tagging artefacts.

1. We checked the distance to the median tagged particle and the furthest tagged particle versus infall mass for each CDM and SIDM simulation pair. Since the central density of haloes in SIDM is reduced relative to haloes in CDM, there are fewer particles within a fixed radius. This raises the concern that while particles are tagged in the same distribution in terms of rank order boundedness, particles in SIDM haloes are tagged at systematically further radii. Thus, enhanced stripping could be due to star particles starting at larger radii and nothing else. Per infall mass and infall time bin, we find SIDM particles are either tagged at no further distance than CDM particles or are tagged at fractionally larger radii. These are small differences in all cases, which are negligible compared to the growth in $r_{1/2}$ in SIDM relative to CDM seen in Fig. 3-16.

2. We tested if haloes with large amounts of internal scattering, and thus very few particles to tag, produce biased results. We isolated all infalling haloes where we could tag $< 100\%$, $< 50\%$, and $< 20\%$ of our target of 2% of the most bound particles. Hereafter, we call the fraction of the target number particles to tag that can actually be tagged the tagging yield. When removing each of these subsets, all trends of enhanced stellar stripping as seen in Figs 3-4-3-6 remain the same.

3. We further tested if haloes with very few particles to tag, either due to their small size or large fraction of scattered particles produce biased results. We therefore removed all haloes with $< N$ particles tagged from our data, with N ranging from 5 to 10, and again find the trends in enhanced stellar stripping unchanged.

4. We tested for bias due to ranges of infall time and infall mass where many SIDM haloes have zero particles to tag. In our tagging process, we divide

haloes into bins by infall time and infall mass, and use the mean distribution of what particles can be tagged in each bin in the SIDM case to apply to the corresponding CDM case. As a result, the group of CDM haloes in each bin are all tagged with the same distribution, whereas the SIDM haloes are tagged with some variation about that distribution. In bins where the typical tagging yield is below 20%, a fraction of the SIDM haloes have no particles to tag, and are thus not included in further analysis, whereas the CDM haloes are all tagged with the same low tagging yield. Concerned about this mismatch in the number of haloes included in our two samples, we investigated the trends in enhanced stellar stripping after excluding bins with substantial mismatches. Once again, this did not change the trends.

5. We searched for biases in haloes with early and late infall times. For a fixed infall halo mass, haloes with early infall times have more internal scattering between infall and $z = 0$ than those with late infall times. The early infall haloes have particles typically tagged at larger distances, and in turn the fraction of particles stripped is consistently higher. However, the difference in the strength of stripping between SIDM and CDM remains the same for both early and late infall time samples.

Chapter 4

An observer's guide to the (Local Group) dwarf galaxies: predictions for their own dwarf satellite populations

The content of this chapter was submitted to the Monthly Notices of the Royal Astronomical Society *on September 27th, 2016. It has an arXiv location of* `https://arxiv.org/abs/1610.00708`. *This version has most, but not all, edits made in response to a referee report. Check the arXiv location for a more up to date version, and a link to the eventual published version. The authors are Gregory A. Dooley, Annika H.G. Peter, Tianyi Yang, Beth Willman, Brendan F. Griffen, and Anna Frebel.*

Disclaimer

This chapter was predominately a product of my own work. The idea that initiated this project came from the senior thesis written by Tianyi Yang and advised by Beth Willman and Annika Peter. The names and properties of isolated dwarf galaxies were mostly taken from that thesis. All authors provided comments and feedback on the

paper.

Acknowledgements

G.A.D. acknowledges support from an NSF Graduate Research Fellowship under Grant No. 1122374. Support for A.H.G.P. was provided in part by NASA through the HST theory grant HST-AR-13896.005-A from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. BW acknowledges support by NSF Faculty Early Career Development (CAREER) award AST-1151462. We thank Shea Garrison-Kimmel for private correspondence regarding abundance matching implementations, and Chris Barber for providing data we used for our default reionization model.

Abstract

A recent surge in the discovery of new ultrafaint dwarf satellites of the Milky Way has inspired the idea of searching for faint satellites, $10^3\,\mathrm{M}_\odot < M_* < 10^6\,\mathrm{M}_\odot$, around less massive field galaxies in the Local Group. Such satellites would be subject to weaker environmental influences than Milky Way satellites, and could lead to new insights on low mass galaxy formation. In this paper, we predict the number of luminous satellites expected around field dwarf galaxies by applying several abundance matching models and a reionization model to the dark-matter only *Caterpillar* simulation suite. For three of the four abundance matching models used, we find a $> 99\%$ chance that at least one satellite with stellar mass $M_* > 10^5\,\mathrm{M}_\odot$ exists around the combined five Local Group field dwarf galaxies with the largest stellar mass. When considering satellites with $M_* > 10^4\,\mathrm{M}_\odot$, we predict a combined $5 - 25$ satellites for the five largest field dwarfs, and $10 - 50$ for the whole Local Group field dwarf population. Because of the relatively small number of predicted dwarfs, and their extended spatial distribution, a large fraction each Local Group dwarf's virial volume will need to be surveyed to guarantee discoveries. We compute the predicted number of satellites in a given field of view of specific Local Group galaxies, as a function of minimum satellite luminosity, and explicitly obtain such values for the Solitary Local dwarfs survey. Uncertainties in abundance matching and reionization models are large, implying that comprehensive searches could lead to refinements of both models.

4.1 Introduction

Hierarchical structure formation in the Lambda Cold Dark Matter Universe predicts that galaxies like the Milky Way (MW) and M31 are orbited by satellite galaxies (Frenk and White, 2012). Observations have long supported this hierarchical accretion model, starting with identifying that the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC) are within close proximity of the Milky Way (Shapley, 1922, 1924). An additional nine MW satellites with luminosity $L_* > 10^5\,\mathrm{L}_\odot$, the classical dwarfs, were discovered next, followed by a class of satellite galaxies with luminosity $10^3 < L < 10^5\,\mathrm{L}_\odot$, the ultrafaint dwarfs (UFDs), initially discovered in the Sloan Digital Sky Survey (Willman et al., 2005b; Zucker et al., 2006b; Belokurov et al., 2006, 2007, 2008, 2010; Irwin et al., 2007; Walsh et al., 2007). Even smaller "hyperfaint" galaxies ($L_* < 10^3\,\mathrm{L}_\odot$) have also been discovered, galaxies so tiny they can only be found very near the Sun (Willman et al., 2005a; Zucker et al., 2006a; Belokurov et al., 2009). The window into UFD and hyperfaint satellites of the MW is opening up dramatically, as the Dark Energy Survey, PanSTARRS, ATLAS, and MagLiteS surveys have found ~20 new UFD satellite candidates in the past two

years (Bechtol et al. 2015; Drlica-Wagner et al. 2015; Kim et al. 2015; Kim and Jerjen 2015; Koposov et al. 2015a; Laevens et al. 2015; Martin et al. 2015; Luque et al. 2016; Torrealba et al. 2016b; Drlica-Wagner et al. 2016), and will likely continue to find more.

These recent discoveries, along with follow up observations, have opened the door to better understand low mass galaxy formation. Whereas classical dwarfs have recent star formation, many UFDs have been confirmed to be "fossil" galaxies, meaning that $> 70\%$ of their stars formed before reionization (Brown et al., 2012, 2014a,b). Consequently, they contain very old stellar populations (Kirby et al., 2008; Norris et al., 2010; Frebel and Bromm, 2012) and are ideal targets to learn about early universe galaxy formation. Both classical dwarfs and UFDs serve as probes on the interplay of ionizing radiation, supernova feedback, star formation, and halo size. Low-mass galaxies are susceptible to losing gas from reionization and supernovae, which can turn them into fossil galaxies or even leave them entirely dark. However, due to strong environmental effects on dwarf galaxy evolution, and halo-to-halo variation, a large sample size of dwarf galaxies in different environments is necessary to probe the halo size scale where these effects begin.

To constrain star formation models in dwarf galaxies, there is significant value in even just the number counts of UFDs and classical dwarfs. Completeness in discovery around the MW will further refine the so-called "missing satellite problem" (Moore et al., 1999; Klypin et al., 1999b) and its many proposed solutions, which come in both baryonic and dark-matter flavors. Baryonic solutions include the effects described above, to suppress the formation of stars in small dark-matter halos. With a large sample of UFDs, we can test abundance-matching-derived $M_* - M_{\rm halo}$ relationships to much smaller mass scales than those for which the relations were observationally inferred. A change in the slope or scatter of abundance-matching relations would have significant implications for the drivers of star-formation efficiency in small galaxies.

Discovering a large sample of UFDs is also critical in revealing a diversity of chemical enrichment pathways, such as r-process enhancement or lack thereof (Ji et al., 2016). By finding more UFDs in particular, we should discover older and more

metal poor stellar populations. We also create more opportunities to measure internal halo structure, which has important implications on the cusp/core debate (de Blok, 2010) and the nature of dark matter (Elbert et al., 2015; Dooley et al., 2016a).

Given the importance of finding more low-mass galaxies, it is natural to consider searching for them beyond the MW. Already, many satellites have been discovered around Andromeda (Zucker et al., 2004, 2007; McConnachie et al., 2008, 2009; Majewski et al., 2007; Irwin et al., 2008; Martin et al., 2009; Bell et al., 2011; Slater et al., 2011; Richardson et al., 2011), and around a handful of nearby galaxies and clusters (Jang and Lee, 2014; Sand et al., 2014; Crnojević et al., 2016). Several of these recent discoveries have been of dwarf galaxy satellites of dwarf galaxies themselves (Sand et al., 2015; Carlin et al., 2016, e.g.). This opens the question if there could exist satellites around isolated dwarf galaxies within the Local Group itself. Due to a lower mass host, they would experience weaker environmental influences than those in the MW or M31. Tidal and ram pressure stripping are reduced, so satellites would retain more of their original stars, gas, and dark matter. Reionization would also proceed differently, as the nature of the closest source of ionizing photons would change. These differences would provide an opportunity to better isolate the internal drivers of low mass galaxy formation. While isolated galaxies would have even weaker environmental effects, hierarchical galaxy formation dictates that the density of low-mass galaxies is greater around a larger galaxy than in areas of complete isolation.

We therefore set out to characterize the abundance of satellites around Local Group field dwarfs, or "dwarf-of-dwarf", systems as a guide to current and future surveys. We predict the number of satellites of dwarf galaxies given simple, physically motivated prescriptions for how dwarfs populate dark-matter halos in the canonical cold-dark-matter model. We outline observational strategies for finding dwarf-of-dwarf satellites and discuss how to interpret observations in light of models for star formation in small halos. We focus specifically on the satellite systems of Local Group field dwarf galaxies, because the proximity of these galaxies enables the discovery of very low luminosity satellites as overdensities of resolved stars. We include specific predictions for the fields of view of the Solitary Local dwarfs survey (Solo), a recent

survey of all isolated dwarfs within 3 Mpc of the Milky Way (Higgs et al., 2016). Though the main goals of the Solo survey do not include finding satellites, it likely already has at least one lurking in its data. Furthermore, our results can be used to estimate the number of dwarf-of-dwarf satellites which Sagittarius, Fornax, and the SMC brought into the Milky Way at infall. Values for the LMC, which is larger than the mass range of hosts we consider in this paper, will be presented in future work.

Sales et al. (2013) and Wheeler et al. (2015) have made similar calculations to ours, predicting the probability of an UFD satellite around a dwarf galaxy within a $10^{10}\,\mathrm{M}_{\odot}$ dark matter halo, as $40-50\%$ and 35%, respectively. We perform a more in-depth study over a larger parameter space, finding the likelihood of satellites existing around dwarf galaxy hosts of a range of host masses, the mean number of satellites around hosts as a function of satellite stellar mass, and the full probability distribution of the number of satellites around known field dwarfs. Due to uncertainty in the $M_{*}-M_{\mathrm{halo}}$ relationship for low luminosity systems, we use a variety of abundance matching models rather than just one model for star formation. We additionally determine the sensitivity of predictions on input parameters, including reionization, a study not previously conducted.

Our paper is organized as follows: In Section 4.2, we outline our methods for modeling satellite populations in isolated dwarf galaxies. In Section 4.3, we validate our methods with predictions for the Milky Way, predict how many luminous satellites should exist around dwarf galaxies, and compute the probability of finding one or more satellites per host. In Section 4.4, we provide a model for the number of satellites within a line of sight as a function of field of view, and comment on observational strategies. In Section 4.5, we show how sensitive our predictions are to uncertainties. Finally, we summarize our key findings and present a plan for future directions in Section 4.6.

4.2 Methods

To predict the number of luminous satellites of Local Group dwarf galaxies, we apply a suite of abundance matching models and a parameterized reionization recipe to dark-matter-only simulations of Milky Way-like halos (and their surrounding environments). This simple scheme is fast to implement, unlike fully hydrodynamic simulations, and allows us to quickly explore different models for how dwarf galaxies populate halos. With this scheme, we can generate many realizations of dwarf satellite systems, so we can define a probability distribution for the satellite populations for each model.

We use dark-matter only simulations to predict the subhalo mass functions (SHMFs) of satellites around isolated field dwarfs in the vicinity of a Milky Way-mass galaxy, and use these SHFMs to generate Poisson samples of subhalos around each dwarf galaxy host. Next, we model the effects of reionization by assigning each subhalo a probability that it hosts stars or remains dark. We then apply abundance matching prescriptions from the literature to assign stellar mass to the luminous subhalos.

In the following subsections we elaborate on the simulations, abundance matching models, reionization methodology, and mass functions used.

4.2.1 *Caterpillar* Simulation Suite

We use a sample of 33 high particle resolution ($m_{\rm p} = 3 \times 10^4\,{\rm M}_\odot$) and high temporal resolution (320 snapshots) zoom-in simulations of Milky Way-sized galaxies from the *Caterpillar* simulation suite (Griffen et al., 2016b). The simulations are used to determine the typical SHMF, radial dependence of the SHMF, subhalo infall distribution times, and dark fraction of halos due to reionization. We perform these calculations on both the Milky Way sized host halo and smaller nearby field halos. We consider field galaxies as halos with virial mass between 10^{10} and $10^{11.5}\,{\rm M}_\odot$ at $z = 0$ that are outside of the virial radius of the MW sized host, and within the uncontaminated volume of each simulation. We choose the mass range to reflect that of real Local Group field galaxies. In total there are 148 field halos across the 33 simulations.

All self-bound haloes are found using a modified version of the ROCKSTAR HALO FINDER (Behroozi et al., 2013b) which includes full iterative unbinding to improve halo finding accuracy, as described in Griffen et al. (2016b). Merger trees were produced by ROCKSTAR CONSISTENT TREES (Behroozi et al., 2013c). Any mention of *virial* refers to the Bryan and Norman (1998) definition of the virial radius, $R_{\rm vir}$, which at $z = 0$ for our cosmological parameters is the radius such that the mean enclosed halo density is 104 times the critical density of the universe, $\rho_c = 3H_0^2/8\pi$G. $M_{\rm vir}$ refers to the gravitationally bound mass within $R_{\rm vir}$, and any mention of R_Δ or M_Δ refers to the radius and mass of a halo where the mean enclosed density is Δ times the critical density.

4.2.2 Abundance Matching Models

Abundance matching (AM) is a technique employed to determine an approximate stellar mass to halo mass ($M_* - M_{\rm halo}$) relationship for galaxies. Given a set of observed galaxies within a volume down to some luminosity completeness limit, galaxies are matched in a one-to-one fashion with dark matter halos from a simulation of the same volume. They traditionally assume a monotonic relationship of stellar mass and dark matter halo mass to create a function $M_*(M_{\rm halo})$ that satisfies the condition

$$\int_{M_*(m_1)}^{M_*(m_2)} \frac{\mathrm{dN_*}}{\mathrm{dM}}(M_*)\mathrm{dM_*} = \int_{\mathrm{m_1}}^{\mathrm{m_2}} \frac{\mathrm{dN}}{\mathrm{dM}}(\mathrm{M_{halo}})\mathrm{dM_{halo}} \tag{4.1}$$

where $\frac{\mathrm{dN_*}}{\mathrm{dM}}(M_*)$ is the differential stellar mass function, and $\frac{\mathrm{dN}}{\mathrm{dM}}$ is the differential halo mass function (Yang et al., 2003; Vale and Ostriker, 2004; Kravtsov et al., 2004; Tasitsiomi et al., 2004; Vale and Ostriker, 2006; Guo et al., 2010; Moster et al., 2010; Kravtsov, 2010; Wang and Jing, 2010; Yang et al., 2012; Moster et al., 2013; Behroozi et al., 2013a; Brook et al., 2014; Garrison-Kimmel et al., 2014b). For galaxies with $M_* > 10^8\,\mathrm{M}_\odot$, abundance matching relationships produce relatively consistent results with each other. However, at smaller masses, incomplete surveys of low luminosity galaxies and a more stochastic process of star formation in halos leads to larger uncertainty in the $M_* - M_{\rm halo}$ relationship. We highlight this in Fig. 4-2, showing

the relationship for several recently proposed models. Different extrapolations of the function down to low masses, how stochastic star formation is, and what simulations and observations were compared lead to very different predictions. We briefly describe each of the models and details on their implementation in the following paragraphs. The names in bold indicate how we refer to the models in the rest of the paper.

Moster: Moster et al. (2013) match observed stellar mass functions at different redshifts from Sloan Digital Sky Survey (SDSS), Spitzer Space Telescope, Hubble Space Telescope and Very Large Telescope to dark matter halos in the Millennium (Springel et al., 2005) and Millennium-II (Boylan-Kolchin et al., 2009) simulations to produce a redshift dependent AM model. For subhalos, they define $M_{\rm halo}$ as $M_{200}^{\rm infall}$, the mass of a halo at first infall enclosed by a volume that is $200 \times \rho_c$. To account for the redshift dependence in their model, we find the infall time distribution of all $z = 0$ satellites in *Caterpillar* and use it to assign random infall times to our subhalos in subsequent analysis. We investigate whether the infall time distribution changes when considering different ranges of subhalo masses, but find at most a weak dependence on subhalo mass that results in a $< 1\%$ influence on our final estimates of luminous satellites. We therefore use an infall time distribution independent of subahlo mass. We find the distributions for satellites in MW-sized halos and field halos are consistent with each other, as seen in Fig. 4-1, and match well in form to fig. 3 of Barber et al. (2014). Barber et al. (2014) finds that when selecting only subhalos which form stars as opposed to all subhalos, the mean infall time is shifted ~ 1 Gyr earlier. This adjustment makes little difference to the predictions made by the Moster model, as discussed in Section 4.5, so we use the distribution for all subhalos.

Behroozi: Behroozi et al. (2013a) deduce stellar mass functions from $z = 0$ to $z = 8$ with results from SDSS, *GALEX*, and PRIMUS surveys. They match these to halo mass functions from the *Bolshoi* (Klypin et al., 2011), *MultiDark* (Riebe et al., 2011), and *Consuelo* simulations to produce a redshift dependent $M_* - M_{\rm halo}$ model which defines $M_{\rm halo}$ as $M_{\rm vir}^{\rm peak}$, the maximal virial mass achieved by the subhalo over its history. We show the function in Fig. 4-2, but do not include the model in

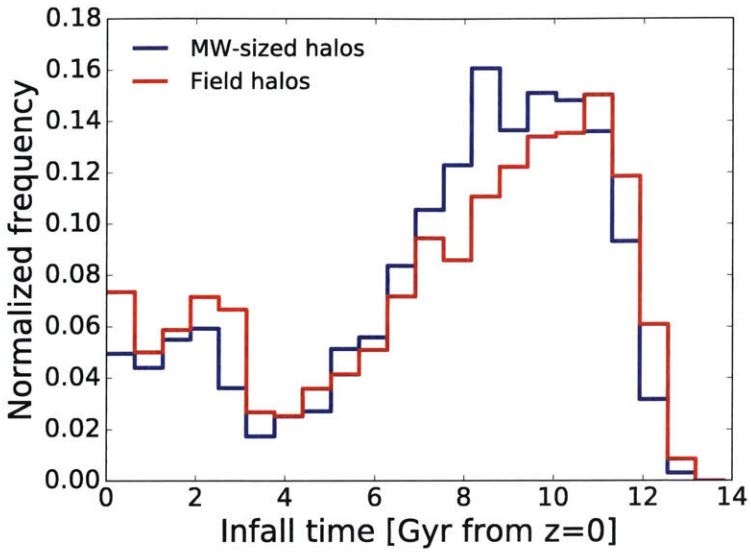

Figure 4-1 Infall time distribution for all $z = 0$ subhalos averaged over all MW-sized hosts in 33 *Caterpillar* simulations, and over all field halos. Time is given by the duration between infall and $z = 0$. The distribution does not vary significantly with host halo mass, nor range of subhalo mass.

our results since it overpredicts the abundance of low-mass galaxies and is otherwise incorporated in the next two models.

GK14: Garrison-Kimmel et al. (2014b) match galaxies from the SDSS to dark matter halos in their ELVIS suite to create an AM model of the same functional form as that in Behroozi et al. (2013a), but with a steeper logarithmic slope on the low mass end. They identify that Behroozi et al. (2013a) overestimates the number of galaxies with $M_* < 10^{8.5}\,\mathrm{M}_\odot$ at $z = 0$ due to using a now outdated stellar mass function, and correct for it. Below the completeness limit of $M_* = 10^8\,\mathrm{M}_\odot$, GK14 extrapolates their relationship with a constant slope. Like Behroozi, they define $M_{\rm halo}$ as $M_{\rm vir}^{\rm peak}$. For their cosmology, $R_{\rm vir}$ encloses a volume that has density $97\times R_{\rm vir}$, making it marginally larger than $R_{\rm vir} = R_{104}$ in our cosmology. We find the discrepancy small enough to not take into account in detail.

GK16: Many hydrodynamic simulations have demonstrated that there can be significant scatter about a mean $M_* - M_{\rm halo}$ relationship (Munshi et al., 2013; Sawala et al., 2015; Wheeler et al., 2015; O'Shea et al., 2015). The scatter increases to-

wards lower masses, making the default abundance matching assumption of a monotonic relationship problematic particularly for dwarfs (Power et al., 2014; Ural et al., 2015). Garrison-Kimmel et al. (2016) explicitly model the scatter, proposing a range of stochastic abundance matching relationships. They build off the GK14 model, changing the slope of the best fit relationship as a function of the 1-sigma level of lognormal scatter, $\sigma_{\rm scat}$. With higher levels of scatter, more galaxies are upscattered above the observed completeness luminosity than are scattered below it due to the increasing abundance of DM subhalos at lower masses. Consequently, their mean $M_* - M_{\rm halo}$ relationship must be steeper and thereby lower for low mass halos. They match galaxies in the Local Group to the ELVIS suite down to a completeness limit of $M_* > 4.5 \times 10^5\,{\rm M}_\odot$, and define $M_{\rm halo}$ as $M_{\rm vir}^{\rm peak}$ as in GK14.

Our default implementation of this model is to use the "growing scatter" model, in which $\sigma_{\rm scat}$ grows for decreasing halo masses. We make this choice because simulations such as those of O'Shea et al. (2015) and Sawala et al. (2015) support a growing scatter more than a constant scatter. The level of growth is dictated by a parameter, γ, as in Eq. (3) of Garrison-Kimmel et al. (2016). We choose a default value of $\gamma = -0.2$. In Section 4.5, we discuss how results change when varying γ. We implement the scatter by sampling a lognormal offset from the mean $M_* - M_{\rm halo}$ relationship randomly from a Gaussian of width $\sigma_{\rm scat}(M_{\rm halo})$ for each subhalo considered.

Brook14: Brook et al. (2014) proposes an even steeper slope than Garrison-Kimmel et al. (2014b), which, when extrapolated to $M_* < 10^7\,{\rm M}_\odot$, estimates lower stellar masses for a fixed DM mass. They match observed galaxies in the Local Group to the CLUES simulation suite (Gottloeber et al., 2010). Instead of $M_{\rm vir}^{\rm peak}$, they define $M_{\rm halo}$ as $M_{350}^{\rm peak}$, the peak mass achieved by a subhalo measured within a volume that has density $350 \times \rho_c$. We implement their model which has an $M_* - M_{\rm halo}$ log-log slope of 3.1, and normalization factor $M_0 = 79.6$. For stellar masses $M_* > 10^8\,{\rm M}_\odot$, the Brook model is unspecified, so we linearly interpolate values in log-log space between $M_* = 10^8\,{\rm M}_\odot$ in the Brook model and $M_* = 3 \times 10^9\,{\rm M}_\odot$ in the Moster model, then switch to values from Moster. No satellites considered in this paper have $M_* > 10^8\,{\rm M}_\odot$, but a few host galaxies do, and a function to estimate their dark

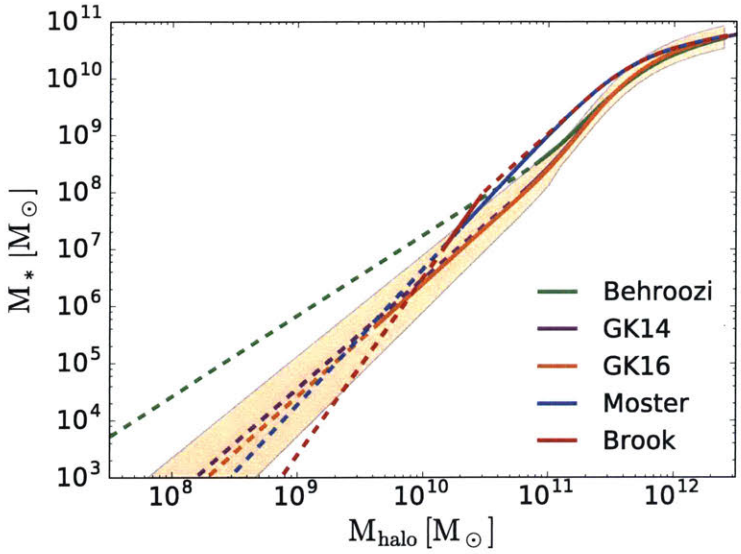

Figure 4-2 Abundance matching stellar mass-halo mass relationships for several recent models. Solid lines indicate ranges of each model where they were matched to observations, dashed lines indicate ranges of extrapolation. The large variation in predictions, particularly for halos with $M_{\rm halo} < 10^{10}\,{\rm M}_\odot$, results in very different predictions for the number of low mass satellite galaxies that could be discovered in the Local Group. While similar, the definition of $M_{\rm halo}$ is different for each of the models, making purely visual comparisons between functions not entirely accurate. The mass definitions are listed in Table 4.1. The shaded area around GK16 indicates the $\pm 1\sigma$ lognormal scatter we implement.

matter halo mass from stellar mass is needed.

4.2.3 Reionization

UV photons emitted by the first stars during reionization are able to ionize hydrogen atoms and prevent sufficient cooling and gas accretion needed for star formation (Efstathiou, 1992; Thoul and Weinberg, 1996; Gnedin, 2000; Wiersma et al., 2009; Pawlik and Schaye, 2009). In low mass halos, they can also heat gas enough to gravitationally escape, sometimes before any star formation begins (Barkana and Loeb, 1999; Shapiro et al., 2004; Okamoto et al., 2008). The combination of effects renders many halos entirely dark, an effect recently simulated and emphasized in Sawala et al. (2013, 2015, 2016). Simply assuming that all dark matter subhalos host luminous galaxies would therefore wildly overestimate the number of visible satellites

(Bullock et al., 2000; Somerville, 2002; Benson et al., 2002).

We model the effects of reionization by randomly assigning halos to host stars or remain dark with probabilities that depend on the halo's mass. Using data obtained from Barber et al. (2014), we produce a smoothed curve indicating the fraction of halos that are luminous at $z = 0$ as a function of $M_{200}^{\rm infall}$, as plotted in Fig. 4-3. The function follows from a semi-analytic model applied to the level-2 halos of the Aquarius simulation suite (Springel et al., 2008). The model has reionization proceeding from $z = 15$ to $z = 11.5$. Below a redshift dependent filtering mass, it models photoevaporation by removing baryons from halos. Full details of the semi-analytic model are given in Starkenburg et al. (2013).

Since the abundance matching models use different definitions for $M_{\rm halo}$, we produce a different luminous fraction function for each definition. We do this by randomly assigning halos to be dark or luminous in our simulation according to their $M_{200}^{\rm infall}$, then collecting the values of $M_{\rm vir}^{\rm peak}$ and $M_{350}^{\rm peak}$ for those same halos in the merger tree. Repeating the random assignments for many instances generates a list of dark and luminous halos paired with each mass definition, which is then turned into the desired function. In all cases, reionization suppresses the number of satellites with $M_* < 10^5\,{\rm M}_\odot$, but has little effect suppressing larger systems.

Many details of reionization, including the redshift of occurrence, environmental effects, H_2 shielding, and the efficiency of photoevaporation, remain uncertain (Oñorbe et al., 2016), which adds variability to the number of luminous subhalos produced. We therefore investigate alterations to the reionization model and subsequent effects on our results in Section 4.5.

For that investigation, and for use in determining the radial distribution of satellites which survive reionization, we employ a simple model inspired by Lunnan et al. (2012) and Peter and Benson (2010). While exploring parameter space with detailed hydrodynamic or semi-analytic reionization models would be more rigorous, it is beyond the scope of this paper. For a halo to form stars by $z = 0$, it must either reach a critical size for H_2 cooling and atomic line cooling before reionization, or become massive enough after reionization to reaccrete and cool gas. We define the size

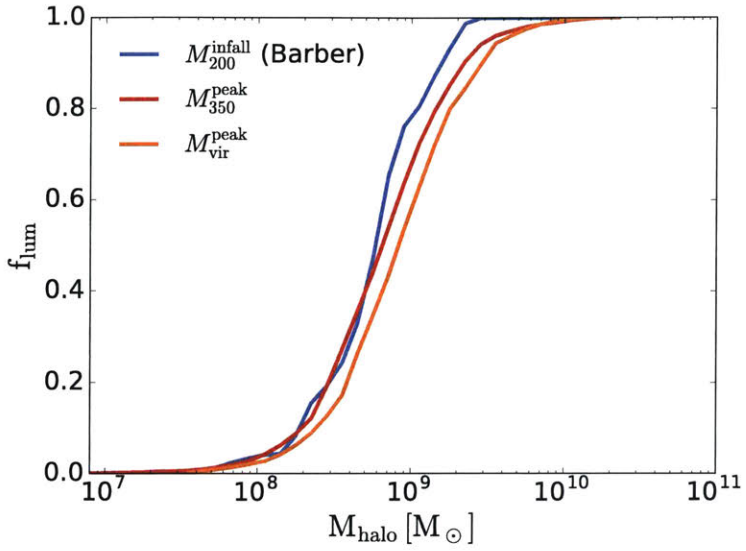

Figure 4-3 Fraction of dark matter halos that host luminous galaxies by $z = 0$ as a function of different mass definitions. The blue line is smoothed data from Barber et al. (2014), which uses a semi-analytic model applied to dark matter only simulations. The green and red lines are functions for alternate mass definitions, as inferred from the baseline model and *Caterpillar* merger trees.

thresholds in terms of $v_{\rm max}$, calling them $v_{\rm max}^{\rm pre}$ and $v_{\rm max}^{\rm filt}$ respectively. Fixing reionization to happen instantaneously at $z = 13.3$, approximately the mean redshift in the model used in Barber et al. (2014), we conduct a parameter search to minimize the difference between the Barber et al. (2014) luminous fraction function of Fig. 4-3 and the one produced by applying the $v_{\rm max}$ cuts to the merger history in all *Caterpillar* simulations. We achieve a close fit ($< 4\%$ difference at any point) with $v_{\rm max}^{\rm pre} = 9.5$ and $v_{\rm max}^{\rm filt} = 23.5\,{\rm km/s}$.

Our value for $v_{\rm max}^{\rm pre}$ is consistent with expectations from the literature. While the threshold for H_2 cooling is more accurately weakly redshift dependent, it occurs around $M_{200} = 10^6\,M_\odot$ or $T_{\rm vir} = 2000 - 3000\,{\rm K}$ (Tegmark et al., 1997; Madau et al., 2008; Power et al., 2014), corresponding to $4 - 7\,{\rm km/s}$ at $z = 13.3$ in our simulations. Atomic line cooling occurs for larger mass halos, near $M_{200} = 10^8\,M_\odot$ or $T_{\rm vir} \approx 4000\,{\rm K}$ (Bromm and Yoshida, 2011), corresponding to $16 - 26\,{\rm km/s}$ at $z = 13.3$ in our simulations. Halos which reach the atomic line cooling limit before reionization nearly universally form and retain stars by $z = 0$, whereas H_2 cooling minihalos may or

may not retain stars due to reionization and supernova (Power et al., 2014). We thus expect $v_{\rm max}^{\rm pre}$ to lie between the minimum H_2 cooling threshold and the atomic cooling threshold, which it does. Moreover, it agrees closely with Okamoto and Frenk (2009) who inferred a value of $v_{\rm max}^{\rm pre} \approx 12\,{\rm km/s}$ from hydrodynamic simulations where reionization occurs at $z = 8$. If we shift reionization to $z = 8$, our best fit value becomes 12.6 km/s.

Our value for $v_{\rm max}^{\rm filt}$ is consistent with the low end of expectations from the literature. This threshold for star formation to proceed after reionization has been termed the "filtering mass", with initial values placed at $20 < v_{\rm max} < 30\,{\rm km/s}$ (Bovill and Ricotti, 2009; Okamoto and Frenk, 2009; Bovill and Ricotti, 2011a,b). More recent publications have used higher values of $30 < v_{\rm max} < 50\,{\rm km/s}$ (Peter and Benson, 2010; Lunnan et al., 2012; Garrison-Kimmel et al., 2014b; Griffen et al., 2016a), highlighting uncertainty in how to model reionization.

4.2.4 Mass Functions and Monte Carlo Sampling

Using *Caterpillar*, we identify the mean SHMF for all isolated field galaxies and MW analogs. Since the AM models use different mass definitions for $M_{\rm halo}$, we correspondingly find different SHMFs. In each case, the differential number of halos in a given mass interval, $\frac{\rm dN}{\rm dM_{sub}}$, follows the form

$$\frac{\rm dN}{\rm dM_{sub}} = K_0 \left(\frac{M_{\rm sub}}{\rm M_\odot}\right)^{-\alpha} \frac{M_{\rm host}}{\rm M_\odot} \tag{4.2}$$

as has been identified in several previous studies (Gao et al., 2004; van den Bosch et al., 2005; Dooley et al., 2014). The best fit values of α and K_0 do depend weakly on the host halo mass range, but change negligibly within a one dex host mass interval. Since Milky Way-like hosts are more than one dex larger than field halos, we separately compute best fit values of α and K_0 for satellites of field halos and satellites of Milky Way-like hosts. These values are computed for each mass definition and shown in Table 4.1. The mass functions count all self-bound subhalos (excluding subhalos of subhalos) within $R_{\rm vir}$ at $z = 0$ regardless of the mass definition. $M_{\rm host}$, however, uses

Table 4.1 Subhalo mass functions for MW size halos and dwarf field halos

Mass Definition	α (field)	K_0 (field)	α (MW analog)	K_0 (MW analog)	AM Model(s)
$M_{200}^{\rm infall}$	1.81	0.000635	1.84	0.000854	Moster
$M_{\rm vir}^{\rm peak}$	1.82	0.000892	1.87	0.00188	GK14, GK16, Behroozi
$M_{350}^{\rm peak}$	1.81	0.000765	1.87	0.00200	Brook

NOTE. — Values of α and K_0 in equation (4.2) for the mean subhalo mass function for various definitions of subhalo mass. Columns two and three designate values for isolated dwarf field galaxies, and four and five for Milky Way analogs. The parameters of the mass function are approximately independent of host halo mass over a one dex interval. Milky Way analogs and field dwarf halos need to be separated, but within each category one set of parameters is sufficient. The abundance matching models employing each definition are indicated.

the same mass definition as $M_{\rm sub}$. We also considered a SHMF form where $\frac{\rm dN}{\rm dM_{sub}}$ is a power law function of $M_{\rm sub}/M_{\rm host}$ rather than being directly proportional to the host mass, but find the best fit parameters are more sensitive to the host halo mass range considered.

The mean number of dark matter subhalos, $\bar{N}$, around a host of mass $M_{\rm host}$ is found by integrating equation (4.2) from $M_{\rm min}$ to $M_{\rm host}$ to yield:

$$\bar{N} = \frac{K_0 M_{\rm host}}{\alpha - 1} \left(M_{\rm min}^{1-\alpha} - M_{\rm host}^{1-\alpha} \right) \tag{4.3}$$

where $M_{\rm min}$ is a halo mass at which no star formation occurs in any model due to reionization. We choose a conservative value of $10^{7.4}\,{\rm M}_\odot$. Since subhalo abundances are approximately Poisson distributed around the mean, we generate random realizations of the number of subhalos between $M_{\rm min}$ and $M_{\rm host}$ according to a Poisson distribution with a mean $\lambda = \bar{N}$, and then randomly assign halo masses to them according to the SHMF. More accurately, subhalo abundances follow a negative binomial distribution with the variance of $\bar{N}$ increasing relative to that of a Poisson distribution as $M_{\rm sub}/M_{\rm host}$ decreases (Boylan-Kolchin et al., 2010; Mao et al., 2015; Lu et al., 2016). However, Boylan-Kolchin et al. (2010) find that the super Poissonian spread only becomes important for $M_{\rm sub}/M_{\rm host} \leq 10^{-3}$. Otherwise Poisson statistics remain a good approximation, as verified in the *Caterpillar* suite. The largest field halo we consider is $\sim 6 \times 10^{10}\,{\rm M}_\odot$, and only subhalos with $M_{\rm vir} \gtrsim 10^8\,{\rm M}_\odot$ con-

tribute to satellites with $M_* > 10^3\,\mathrm{M}_\odot$, so the minimum ratio relevant to this study is $M_{\rm sub}/M_{\rm host} = 1.7 \times 10^{-3}$. In this instance, $\sigma_{\bar{N}}/\sigma_{\rm Poisson} \leq 1.2$. For all other field halos and for abundances of more luminous satellites, our Poisson approximation is even more accurate.

Once halo masses are assigned, they are chosen to be luminous or dark with probabilities following the luminous fraction as a function of halo mass as shown in Fig. 4-3. The luminous halos are then assigned a stellar mass according to the $M_* - M_{\rm halo}$ relationship and scatter (if any) of the model under consideration.

4.2.5 Inferring $M_{\rm halo}$ from M_*

4.3 Results

In the following subsections we compute the mean expected number of luminous satellite galaxies, $\bar{N}_{\rm lum}$, above a given stellar mass threshold, $M_*^{\rm thresh}$, and the probability of at least one satellite existing above $M_*^{\rm thresh}$ as a function the host galaxy's stellar mass. For the rest of the paper, we consider galaxy stellar masses above $M_* > 10^3\,\mathrm{M}_\odot$, since smaller "hyperfaint" galaxies have so few stars that they will be difficult to detect above the background. All values are found using the methodology presented in Section 4.2, generating 30000 random realizations of satellite populations per host. Due to uncertainty in mass to light ratios, we strictly report on stellar mass, not stellar luminosity.

4.3.1 Validation of our models with the Milky Way satellite system

To verify our model implementations, we predict the number of satellites around a Milky Way-sized galaxy. We fix the host's dark matter mass to $M_{\rm vir} = 1.4 \times 10^{12}\,\mathrm{M}_\odot$ and plot the mean number of satellites as a function of $M_*^{\rm thresh}$ in the upper panel of Fig. 4-4. We include predictions of each abundance-matching model with and without reionization to indicate how much reionization suppresses the formation

of low mass galaxies. The abundances are all scaled to a radius of 300 kpc using Eq. (4.4) (presented later in Section 4.4) since that distance corresponds to published completeness limits for satellite surveys of the Milky Way (Walsh et al., 2009; Drlica-Wagner et al., 2015). The models are all consistent with the high-mass end of the MW satellite stellar mass function. There are 11 MW satellites with $M_* > 4.5 \times 10^5\,\mathrm{M}_\odot$ within 300 kpc of the Galactic Center (Garrison-Kimmel et al., 2016); all models are consistent with 11 such satellites within 1σ Poisson errors. Since no model was calibrated exclusively to the MW satellites, and there is uncertainty in the mass of the Milky Way from $\approx 0.5 - 2.5 \times 10^{12}\,\mathrm{M}_\odot$ (Wang et al., 2015), an exact agreement with the MW luminosity function is not expected. For instance, while the Brook model is consistent with Milky Way classical dwarf satellite counts for the fiducial MW halo mass, the model would better match the MW system with a higher assumed mass.

On the faint end, the models predict a median of 66, 71, 67, and 37 satellites with $M_* > 10^3\,\mathrm{M}_\odot$ for the Moster, GK14, GK16, and Brook models respectively. The first three models are in strong agreement with a prediction by Hargis et al. (2014), who estimated 69 satellites above $10^3\,\mathrm{L}_\odot$ in the MW out to 300 kpc from observations and completeness limits. Drlica-Wagner et al. (2015) also make a consistent prediction of 70 UFDs in the MW when excluding the sub-substructure of UFDs in the Large and Small Magellanic Clouds.

If ≥ 70 UFDs, or more than 80 satellites with $M_* > 10^3\,\mathrm{M}_\odot$ in total are discovered in the MW in the future, it would strongly disfavor the Brook model as implemented. It is difficult to reconcile the Brook model even if the MW halo is on the massive side. At a MW mass of $2 \times 10^{12}\,\mathrm{M}_\odot$, the Brook model predicts 53 satellites, more than 2σ less than 80 where σ is taken from the negative binomial distribution in (Boylan-Kolchin et al., 2010). If reionization was shut off entirely, the predicted number of satellites in the Brook model could increase by 20%. A more likely maximal increase of 10% would still mean the model predicts the MW has an unusually high number of UFDs for its size. The rarity of LMC and SMC sized systems in a MW-like halo (Boylan-Kolchin et al., 2010; Busha et al., 2011b), though, could be evidence in favor of that argument.

In stark contrast, reionization has a large impact on the Moster, GK14, and GK16 models. In these models it is absolutely necessary to include reionization or else an additional 110 to 205 luminous satellites with $M_* > 10^3\,\mathrm{M}_\odot$ would be predicted. Based on the MW comparison, we expect numbers predicted by the Brook model to be on the low end of possibilities for satellites of dwarf galaxies, and values from the other models to be closer to median expectations.

4.3.2 How many satellites of dwarfs are there?

In the lower panel of Fig. 4-4 we again show the satellite abundance as a function of minimum M_*^{sat}, but now fix the host halo stellar mass to $M_* = 2.7 \times 10^8\,\mathrm{M}_\odot$ to reflect the largest field dwarf galaxy, IC 5152. Without a direct measurement of total baryonic plus dark matter mass available, we convert from stellar mass to halo mass for each AM model as described in Section 4.2.5. The radius out to which subhalos are counted is the R_{vir} associated with the halo's total mass, unlike in the top panel. Whereas the Moster, GK14, and GK16 models make similar predictions when the host's total mass is fixed, as in the upper panel, their predictions diverge when instead the host's stellar mass is fixed, as in the lower panel. This is due to the GK models having a lower $M_* - M_{\mathrm{halo}}$ relationship in the range of field halos, which leads to predicting more massive halos for fixed stellar mass. Furthermore, the GK16 and GK14 models are separated due to the larger scatter in the GK16 model, which leads to the GK16 model preferring smaller host halo masses than does the GK14 model. The models predict a mean of $\sim 2-6$ satellites with M_* above $10^3\,\mathrm{M}_\odot$, $\sim 1-4$ above $10^4\,\mathrm{M}_\odot$, and $\sim 1-2$ above $10^5\,\mathrm{M}_\odot$ for a galaxy like IC 5152.

In Fig. 4-5, we show the dependence of satellite abundances on the host galaxy's size in terms of its stellar mass. The top panel plots the median number of satellites with $M_*^{\mathrm{sat}} > 10^3\,\mathrm{M}_\odot$, the middle panel with $M_*^{\mathrm{sat}} > 10^4\,\mathrm{M}_\odot$, and the bottom panel with $M_*^{\mathrm{sat}} > 10^5\,\mathrm{M}_\odot$. The dashed vertical lines correspond to the stellar masses of the five largest known field galaxies in the Local Group, IC 5152, IC 4662, IC1613, NGC 6822, and NGC 3109. For IC 5152 for instance, the models predict a mean of 1.3 satellites with $M_* > 10^4\,\mathrm{M}_\odot$ on the low end in the Brook model, to 4.3 satellites

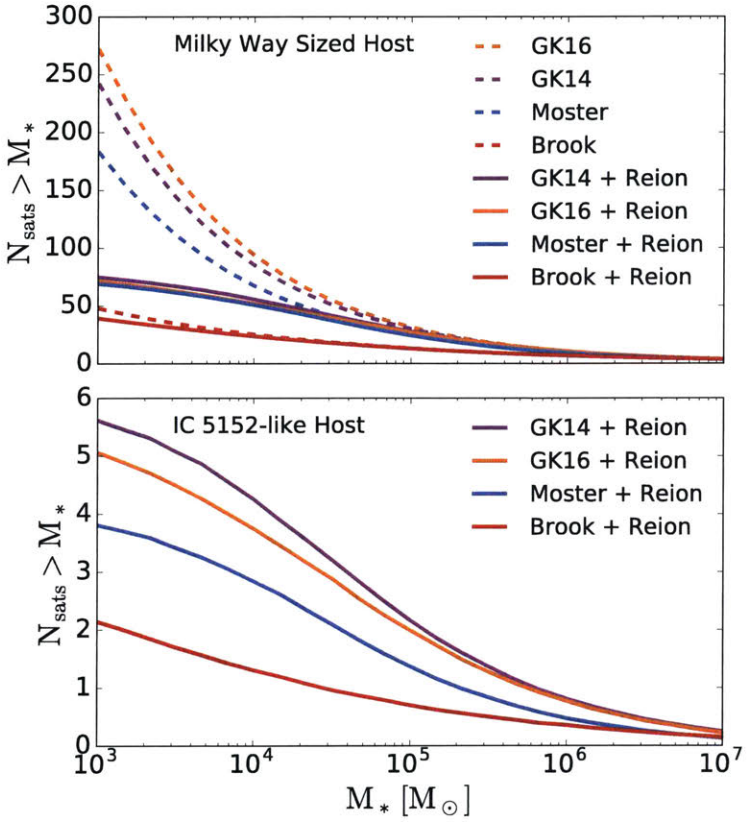

Figure 4-4 Upper Panel: Mean number of satellites around a MW-sized host with total mass $1.4 \times 10^{12}\,M_\odot$ as a function of satellite stellar mass. While the Moster, GK14, and GK16 models predict similar numbers of satellites after accounting for reionization, the Brook model predicts significantly fewer. Due to its steep $M_* - M_{\rm halo}$ relationship, the Brook model is also minimally affected by reionization whereas the other models are reduced by more than 60% on the low mass end. Lower Panel: Mean number of satellites around a host of $M_* = 2.7 \times 10^8\,M_\odot$, approximately that of IC 5152, as a function of satellite stellar mass. While the Moster, GK14, and GK16 models make similar predictions if the host's total mass is fixed, they predict different abundances when stellar mass is fixed.

in the GK14 model.

This figure shows that the number of expected UFDs is a strong function of stellar mass of the host. Below a host stellar mass of $10^7\,M_\odot$, there is a mean of less than one UFD sized satellite according to all models, indicating many hosts will have no satellites above $10^3\,M_\odot$. For a host of $10^8\,M_\odot$ and larger, at least one satellite per host is expected.

A list of known isolated field galaxies and the mean number satellites within their virial volume is listed in Table 4.2. Galaxies and stellar masses were compiled from McConnachie (2012), and supplemented with values from Karachentsev et al. (2014), McQuinn et al. (2015), and Karachentsev et al. (2015) for KK 258, Leo P, and KKs 3 respectively. We give an indication of the probabilistic distribution of satellites by including the $20^{\rm th}$ and $80^{\rm th}$ percentile of abundance, and the probability that at least one satellite exists, $P(\geq 1)$.

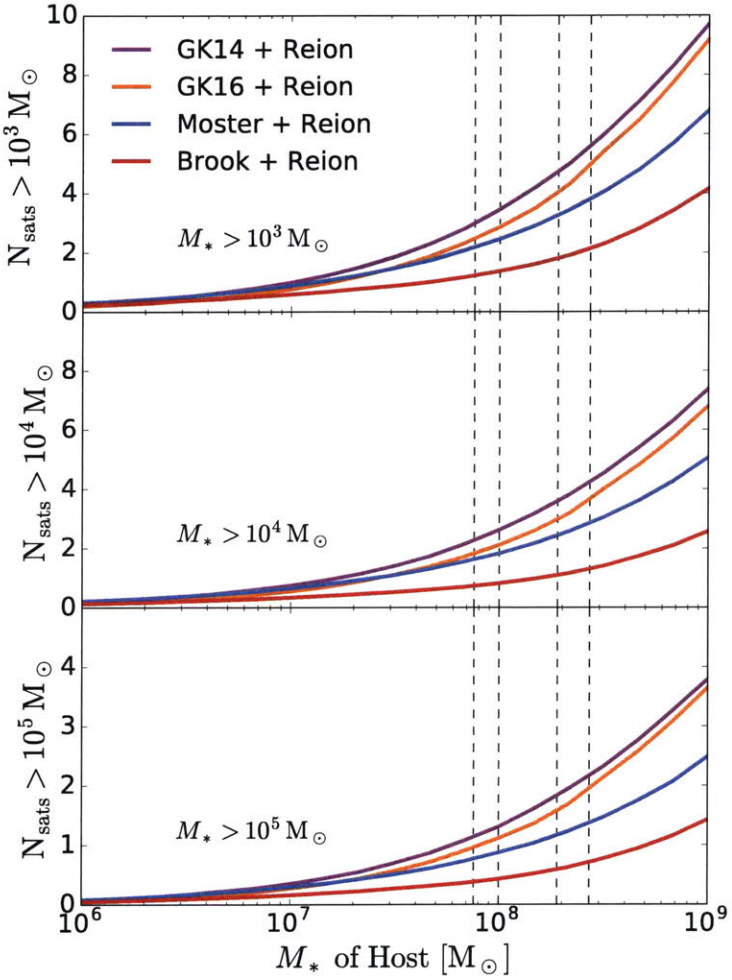

Figure 4-5 Mean number of satellites with stellar mass above $10^3\,\mathrm{M}_\odot$ (upper panel), $10^4\,\mathrm{M}_\odot$ (middle panel), and $10^5\,\mathrm{M}_\odot$ (lower panel) as a function of a host halo's total stellar mass, M_*. Each abundance matching model predicts different values, but all agree that satellites should exist around hosts with $M_* \geq 10^8\,\mathrm{M}_\odot$. Vertical lines correspond to the stellar masses of the five field galaxies listed in Table 4.2 with the largest stellar masses

.

Table 4.2 Mean number of satellites with $M_* > 10^4\,M_\odot$ around Local Group dwarf field galaxies

Name	M_*	**Moster**			**GK14**			**GK16**			**Brook**		
	$[10^6\,M_\odot]$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$
Leo T	0.14	0.08	0/0	0.07	0.05	0/0	0.05	0.03	0/0	0.04	0.05	0/0	0.05
And XXVIII	0.21	0.10	0/0	0.09	0.07	0/0	0.07	0.05	0/0	0.05	0.06	0/0	0.06
KKR 3	0.54	0.16	0/0	0.15	0.13	0/0	0.12	0.09	0/0	0.09	0.10	0/0	0.10
Tucana	0.56	0.17	0/0	0.15	0.13	0/0	0.12	0.09	0/0	0.09	0.10	0/0	0.09
Leo P	0.56	0.17	0/0	0.15	0.13	0/0	0.12	0.09	0/0	0.09	0.10	0/0	0.10
And XVIII	0.63	0.17	0/0	0.16	0.15	0/0	0.13	0.10	0/0	0.09	0.11	0/0	0.10
Phoenix	0.77	0.19	0/0	0.18	0.17	0/0	0.15	0.11	0/0	0.11	0.12	0/0	0.11
KKH 86	0.82	0.19	0/0	0.18	0.17	0/0	0.15	0.12	0/0	0.11	0.12	0/0	0.11
Antlia	1.3	0.25	0/1	0.22	0.23	0/1	0.20	0.16	0/0	0.15	0.15	0/0	0.14
KKR 25	1.4	0.26	0/1	0.23	0.24	0/1	0.21	0.16	0/0	0.15	0.15	0/0	0.14
Aquarius	1.6	0.28	0/1	0.24	0.26	0/1	0.22	0.18	0/0	0.16	0.16	0/0	0.15
DDO 113	2.1	0.32	0/1	0.27	0.30	0/1	0.26	0.21	0/0	0.19	0.18	0/0	0.17
Cetus	2.6	0.35	0/1	0.29	0.35	0/1	0.29	0.24	0/1	0.21	0.20	0/0	0.18
ESO 294-G010	2.7	0.36	0/1	0.30	0.35	0/1	0.29	0.25	0/1	0.22	0.21	0/0	0.18
Sagittarius dIrr	3.5	0.41	0/1	0.33	0.41	0/1	0.34	0.30	0/1	0.25	0.23	0/1	0.20
ESO 410-G005	3.5	0.41	0/1	0.33	0.41	0/1	0.34	0.29	0/1	0.25	0.23	0/1	0.21
KKH 98	4.5	0.46	0/1	0.37	0.48	0/1	0.38	0.35	0/1	0.29	0.25	0/1	0.22
Leo A	6	0.52	0/1	0.41	0.56	0/1	0.43	0.41	0/1	0.33	0.28	0/1	0.25
GR 8	6.4	0.54	0/1	0.41	0.58	0/1	0.44	0.43	0/1	0.34	0.29	0/1	0.25
Pegasus dIrr	6.6	0.54	0/1	0.42	0.59	0/1	0.45	0.43	0/1	0.35	0.29	0/1	0.25
UGC 9128	7.8	0.59	0/1	0.45	0.65	0/1	0.47	0.48	0/1	0.38	0.31	0/1	0.27
UGC 4879	8.3	0.60	0/1	0.46	0.67	0/1	0.49	0.49	0/1	0.39	0.32	0/1	0.27
KK 258	14	0.77	0/1	0.53	0.90	0/2	0.59	0.67	0/1	0.49	0.39	0/1	0.32
UGCA 86	16	0.82	0/1	0.56	0.97	0/2	0.62	0.74	0/1	0.53	0.41	0/1	0.33

Table 4.3 *continued* Mean number of satellites with $M_* > 10^4\,M_\odot$ around Local Group dwarf field galaxies

Name	M_*	**Moster**			**GK14**			**GK16**			**Brook**		
	$[10^6\,M_\odot]$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$	$\bar{N}_{\rm lum}$	20/80%	$P(\geq 1)$
DDO 99	16	0.82	0/1	0.56	0.98	0/2	0.62	0.73	0/1	0.52	0.41	0/1	0.34
UKS 2323-326	17	0.84	0/2	0.57	1.01	0/2	0.63	0.76	0/1	0.54	0.41	0/1	0.34
UGC 8508	19	0.88	0/2	0.58	1.07	0/2	0.66	0.82	0/1	0.56	0.43	0/1	0.35
KKs 3	23	0.96	0/2	0.62	1.19	0/2	0.69	0.92	0/2	0.60	0.47	0/1	0.37
NGC 4163	37	1.19	0/2	0.70	1.53	0/3	0.78	1.22	0/2	0.70	0.55	0/1	0.42
WLM	43	1.28	0/2	0.72	1.67	1/3	0.81	1.34	0/2	0.74	0.59	0/1	0.44
Sextans A	44	1.28	0/2	0.72	1.70	1/3	0.82	1.34	0/2	0.74	0.60	0/1	0.45
DDO 125	47	1.32	0/2	0.74	1.75	1/3	0.83	1.40	0/2	0.76	0.61	0/1	0.46
DDO 190	51	1.37	0/2	0.74	1.84	1/3	0.84	1.47	0/2	0.77	0.62	0/1	0.47
Sextans B	52	1.37	0/2	0.75	1.85	1/3	0.84	1.48	0/2	0.78	0.64	0/1	0.47
IC 3104	62	1.50	0/2	0.77	2.04	1/3	0.87	1.64	1/3	0.81	0.68	0/1	0.49
NGC 3109	76	1.63	1/3	0.80	2.26	1/3	0.90	1.81	1/3	0.84	0.73	0/1	0.52
NGC 6822	100	1.84	1/3	0.84	2.61	1/4	0.93	2.12	1/3	0.88	0.82	0/2	0.56
IC 1613	100	1.84	1/3	0.84	2.62	1/4	0.93	2.13	1/3	0.88	0.83	0/2	0.56
IC 4662	190	2.44	1/4	0.91	3.61	2/5	0.97	2.99	1/4	0.95	1.09	0/2	0.67
IC 5152	270	2.83	1/4	0.94	4.25	2/6	0.98	3.73	2/5	0.98	1.30	0/2	0.73

NOTE. — Mean number of satellites with $M^* > 10^4\,M_\odot$ expected to exist within the virial volume of known Local Group dwarf irregular and dwarf spheroidal galaxies as predicted with various AM models. The $20^{\rm th}$ and $80^{\rm th}$ percentile of the satellite abundance distributions are included in the second column. Also shown is the probability of finding at least one satellite around each galaxy, $P(\geq 1)$.

4.3.3 Likelihood of finding at least one satellite

Another important metric in determining the merit of searching for satellites of field dwarfs galaxies is the probability that at least one satellite exists around a host. In Fig. 4-6 we show the probability that at least one satellite with $M_*^{\rm sat} > 10^3, 10^4$ and $10^5\,{\rm M}_\odot$ exists around a host as a function of M_* of the host. Dotted vertical lines again show the stellar masses of the five largest known field galaxies in the Local Group. For the largest field galaxy, IC 5152, the probability of a satellite with $M_*^{\rm sat} > 10^4\,{\rm M}_\odot$ is $> 90\%$ according to the Moster, GK14, and GK16 models. It drops to 73% for the Brook model. For the 5^{th} largest field galaxy, the probability remains above 80% in the first three models, and is 51% for the Brook model.

The high likelihoods indicate that a comprehensive search of Local Group dwarf galaxies is likely to yield at least one new satellite discovery, even if only the five most massive field dwarfs are surveyed. For the Moster, GK14, and GK16 models, the probability of one satellite with $M_* > 10^5\,{\rm M}_\odot$ is $> 99\%$ if all five of the largest field galaxies are surveyed to their full virial volume. We choose to highlight the five largest because of this fact. In the Brook model, the probability is 92%, but goes to $> 99\%$ for $M_* > 10^4\,{\rm M}_\odot$.

In Fig. 4-7, we show a probability distribution function for the total number of satellites expected around these five largest field galaxies. The three panels show values for abundances above stellar mass thresholds of 10^3, 10^4, and $10^5\,{\rm M}_\odot$. When comparing to Fig. 4-8, which shows the same distribution except for all 38 field dwarfs listed in Table 4.2, we demonstrate that surveying just the five largest galaxies would reveal $\sim 1/3$ of the total population of satellites of field dwarfs.

The shape of the distributions is driven mostly by Poisson statistics in the number of dark matter subhalos. The randomness of reionization suppressing star formation contributes a smaller additional component. In the case of the Moster model, there is an another perturbation due to differing halo infall times, and in the case of the GK16 model, from scatter in $M_* - M_{\rm halo}$. The combined uncertainty arising from counting statistics and abundance matching models results in a broad distribution of

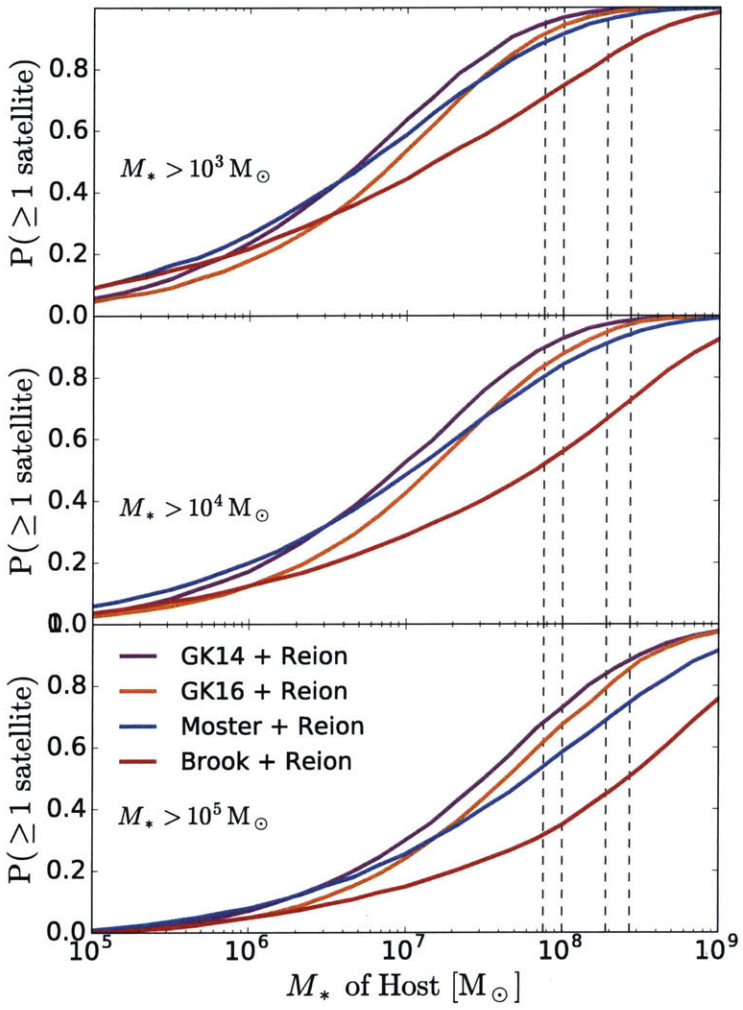

Figure 4-6 Probability that at least one satellite with stellar mass above $10^3\,\mathrm{M}_\odot$ (upper panel), $10^4\,\mathrm{M}_\odot$ (middle pannel), or $10^5\,\mathrm{M}_\odot$ (lower panel) exists around a host with stellar mass M_*. The Moster, GK14, and GK16 models all predict a high likelihood of at least one satellite with $M_* > 10^4\,\mathrm{M}_\odot$ existing around each of the five largest field galaxies, whose stellar masses are indicated by dotted lines (and listed in Table 4.2). The Brook model predicts lower probabilities, but still $> 50\%$ for each of these galaxies.

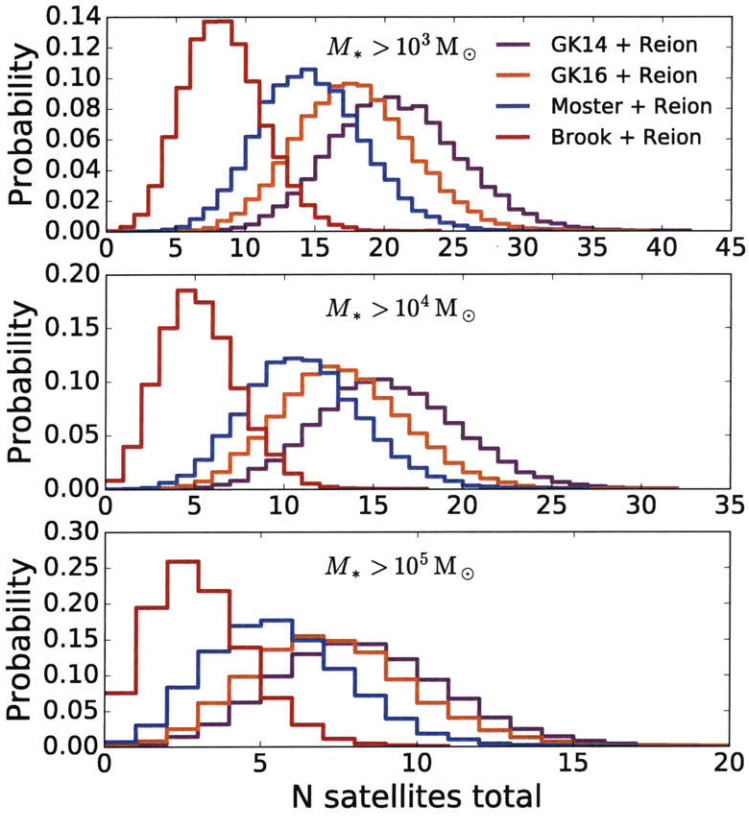

Figure 4-7 Probability distribution for the total number of satellites expected to be found with $M_* > 10^3, 10^4$ and $10^5\,\mathrm{M}_\odot$ around the combined largest five Local Group field galaxies. The GK14 and GK16 models predict over twice the number of satellites as the Brook model, whose predictions are on the lowest end possible to be consistent with MW satellites, and thus are likely a lower limit to the number of satellites of dwarf galaxies. According to the Moster, GK14 and GK16 models, there is a $> 99\%$ that at least one satellite with $M_* > 10^5\,\mathrm{M}_\odot$ exists. These five largest galaxies contain $\sim 1/3$ of the total number of satellites of field dwarfs.

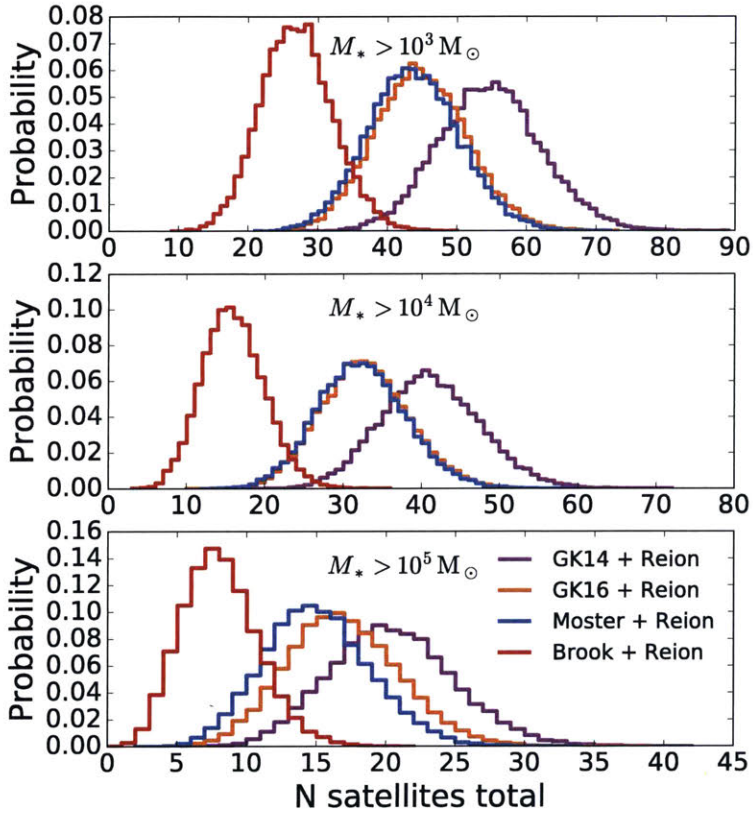

Figure 4-8 Probability distribution for the total number of satellites expected to be found with $M_* > 10^3, 10^4$ and $10^5\,\mathrm{M}_\odot$ around all Local Group field galaxies listed in Table 4.2. There is an even greater separation of the Brook and GK14 models than in Fig. 4-7. Also in comparison, the GK16 model makes predictions more similar to the Moster model due to a greater scatter in M_* for the lower mass halos causing the median predicted M_{halo} to be less for a given halo stellar mass.

possible satellite tallies. The distinct curves from each abundance matching model, in particular comparing the Brook and GK14 models, indicates how observing satellites could provide important insight into ruling out or improving AM models. This is particularly true for satellites above $10^5\,M_\odot$ since reionization has little effect on the number of satellites above that threshold, and those satellites are easier to discover. Probing down to satellites near $10^3\,M_\odot$ or $10^4\,M_\odot$ would help constrain the effects reionization when compared with the number of satellites found above $10^5\,M_\odot$, but would be observationally more challenging.

4.4 Survey Strategy: Dependence on Field of View

Due to the geometry of a line of sight, the values we report in Section 4.3 for the number of satellites within a virial volume do not directly translate into expectations for an observed field of view. For instance, the number of dwarfs-of-dwarfs that may exist in the Solo survey depends on their field of view, the distance to target galaxies, the radius of target galaxies, and the radial distribution of satellites within a galaxy. Future campaigns must account for these details when designing a strategy for limited amounts of telescope time. Both targeting larger mass host galaxies and focusing on the innermost region will increase the projected density of satellites. If the observing goal is to discover as many as possible, would it be better to observe the inner region of a lower mass host, or the outer regions of a higher mass host? We address this issue in detail by making predictions specifically for the Solo survey, and discussing a survey strategy.

We begin by computing a scaling factor to convert the predicted values of the mean number of luminous satellites within a host's virial radius to values expected from surveying circular apertures centered on a host. As found in Han et al. (2016) and confirmed in our simulation suite, satellites are distributed approximately spherically symmetrically around a host and the normalized distribution does not vary with host halo size. This allows us to express the normalized cumulative abundance of satellites generically around a host as $K(r)$. We compute $K(r)$ from the subhalos which are

deemed luminous in our reionization model across all 33 *Caterpillar* simulations at $z = 0$. Selecting only luminous satellites is crucial, because they are more centrally concentrated than the full sample of dark matter subhalos, as previously found and discussed in Gao et al. (2004), Starkenburg et al. (2013), Barber et al. (2014), and Sawala et al. (2016).

The distribution of all dark matter subhalos and the subset of luminous satellites is plotted in Fig. 4-9. For comparison, we also plot the cumulative distribution of known MW satellites with stellar mass above the observed completeness limit of $M_* = 2 \times 10^5\,\mathrm{M}_\odot$ that are within 300 kpc of the Galactic Center. We exclude the LMC and SMC since they occur rarely in MW-sized galaxies (Boylan-Kolchin et al., 2010; Busha et al., 2011b) and they are spatially correlated. The positions and stellar masses of satellites were taken from McConnachie (2012),[1] Our predicted radial distribution of satellite galaxies fits remarkably well to the MW satellites. In contrast, the distribution inferred from all dark matter subhalos does not, demonstrating the importance of subhalo selection effects due to reionization.

For ease of use, we find a very tight match to the data with a piecewise analytic function. It takes the form

$$K(R) = \begin{cases} k_1 R + k_2 R^2 + k_3 R^3 & R < 0.2 \\ k_4 \arctan(\frac{R}{k_5} - k_6) & 0.2 \leq R \leq 1.5 \end{cases} \tag{4.4}$$

with best fit values of the constants $k_1 = -0.2615$, $k_2 = 6.888$, $k_3 = -7.035$, $k_4 = 0.9667$, $k_5 = 0.5298$, and $k_6 = 0.2055$. We find that luminous satellites with a lower peak mass are slightly more centrally concentrated than more massive luminous satellites since they are subject to greater selection effects from reionization. However, the shift in $K(R)$ is $\lesssim 10\%$ which is small compared to Poisson noise and much smaller than the difference between all subhalos and only luminous subhalos. Since galaxies are self-similar, and the radial distribution varies only weakly with satellite mass range, it is possible to simply multiply $K(R)$ by the expected number of satellites in

[1]Available online at `http://www.astro.uvic.ca/~alan/Nearby_Dwarf_Database.html`.

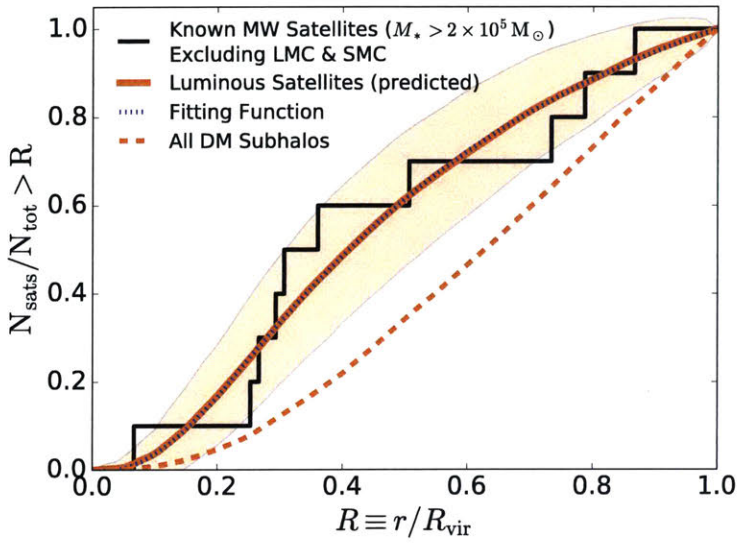

Figure 4-9 Normalized radial distribution of satellites scaled to a host with $R_{\rm vir}$ = 300 kpc. The radial distribution of satellites which survive reionization in our simulations (and are thus luminous) agree well with the radial distribution of known MW satellites, whereas the radial distribution of all dark matter subhalos does not. The sample of known MW satellites includes all satellites with stellar mass above the completeness limit of $M_* = 2 \times 10^5\,{\rm M}_\odot$, except for the LMC and SMC which are known to be spatially correlated and very rare. One sigma variation about the prediction for luminous satellites is shown with a shaded band. It is important to take reionization into account when predicting the radial distribution of satellites, and to not assume it follows the distribution of all dark matter subhalos. A fitting function for the predicted radial distribution is given by equation (4.4) and plotted with a dotted line.

any mass interval within the virial radius of a host (taken from Fig. 4-5 for instance) to yield the number of satellites within a radius $r/R_{\rm vir}$.

By integrating the density function $\frac{1}{4\pi r^2}\frac{{\rm dK}}{{\rm dr}}$ over a cylinder of radius $R \equiv r/R_{\rm vir}$ and half depth $Z \equiv z/R_{\rm vir}$ where r and z are cylindrical coordinates centered on the host galaxy, one gets the number of satellites expected in a line of sight relative to the number within $R_{\rm vir}$. We call this quantity $K_{\rm los}(R)$ We numerically integrate the function and show the result in Fig. 4-10, fixing Z to a value of 1, approximately where halos are likely to be bound to the host halo, and 1.5. $Z = 1.5$ encompasses the "splashback" radius for slowly accreting halos, defined as the distance to first apocenter of orbiting bound satellites (More et al., 2015). It also represents a distance

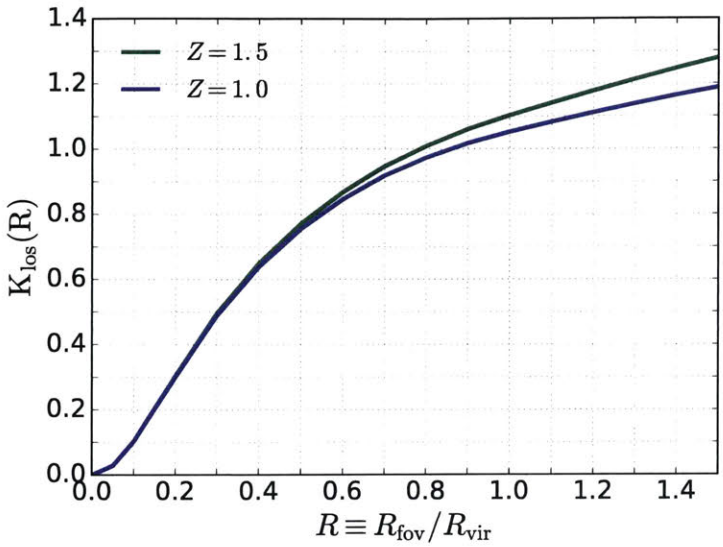

Figure 4-10 Multiplicative value used to scale the expected number of satellites within a halo's virial radius to the expected number of satellites within an observed column whose field of view has the radius $R_{\rm fov}$ on the sky at the distance of the target host galaxy. The column is centered on the galactic center and has a depth of $2 \times Z \equiv 2 \times z/R_{\rm vir}$. A value of $Z = 1$ means the center of the column extends from the near edge of the virial radius to the far edge. For instance, observing a field of view that has $R = 1/2, Z = 1$ corresponds to a multiplicative factor of 0.76.

beyond which the density of additional satellites diminishes rapidly towards zero. Since galaxies at all values of z are in the line of sight, $Z = 1.5$ is a more accurate reflection of what satellites can be observed.

There are several important results from this exercise. First, if one were to search out to the projected virial radius of the host, one would find $\sim 10\%$ more dwarfs than one would expect based on the number of dwarfs strictly within the spherical virial volume. This comes from dwarfs inhabiting a region outside the virial radius that can be imagined by circumscribing a sphere by a cylinder, and then extending the depth (height) of the cylinder to the splashback radius. Second, the number of expected dwarfs rises sublinearly with aperture radius for $R > 0.5$, regardless of the depth of the line-of-sight, Z, under consideration. This is in contrast to the survey area, which grows as the square of the aperture radius. Therefore, in the absence of increased backgrounds like an extended stellar halo, which is unlikely to be significant for dwarfs

(Pillepich et al., 2014), pointings centered on the host galaxy will have higher satellite yields than individual pointings in the outskirts of the host halos.

In light of this, the most efficient strategy for finding dwarf-of-dwarf satellites is to first target the innermost region of the most massive dwarf hosts. Additional observations should map out the whole virial volume of the biggest dwarfs before targeting down the field dwarf stellar mass function, unless environment is a strong motivating factor in the dwarf-of-dwarf search.

What is the transition between a "big" and a "small" dwarf host? The optimal transition point from searching the innermost region of smaller galaxies to the whole volume of larger galaxies depends on which galaxies are targeted and the telescope's field of view. It can be calculated from Fig. 4-10 and Table 4.4. Similar to $K(R)$, $K_{\rm los}(R, Z)$ can be multiplied by the number of satellite galaxies within the virial radius of any host, for any satellite mass interval, to yield the number of satellites within a specified line of sight. Table 4.4 lists the distances to all isolated field galaxies, repeats the mean number of galaxies within the virial radius, $\bar{N}_{\rm lum}$, for convenience, and lists the virial radius of each galaxy inferred from the AM models. $R_{\rm vir}$ refers to the Bryan and Norman (1998) definition, consistent with the radius used to determine all SHMFs. We infer $M_{\rm halo}$ from the AM model and host halo's stellar mass, then convert $M_{\rm halo}$ to $M_{\rm vir}$ assuming an NFW halo density profile, and finally find $R_{\rm vir}$ from $M_{\rm vir}$.

As an example of computing the number of galaxies in a field of view, IC 5152 is located at a distance of 1.7 Mpc (Zijlstra and Minniti, 1999). DECam has a 2.2° field of view, resulting in a 33 kpc observed radius. At a Moster-model-estimated total halo mass of $5.2 \times 10^{10}\,{\rm M}_{\odot}$, IC 5152 has a virial radius of 101 kpc. Using Fig. 4-10 with $Z = 1.5$, the number of satellites reported for the virial volume is multiplied by 0.54 to yield 1.5 expected satellites in the line of sight. An equivalent calculation for the GK16 model produces 1.8 expected satellites, 2.0 for the GK14 model, and 0.7 for the Brook model. If we were to survey the whole virial volume of IC 5152, we would require an additional 8 distinct pointings, and would expect to find only $1 - 2$ more satellites. This example also demonstrates that part of the disparity between

the Moster, GK14, and GK16 models is generated from predicting different virial volumes for a host with fixed stellar mass. When normalizing to more equivalent volumes, as exemplified here, their predictions start converging.

We apply the same calculations to all of the galaxies in Table 4.4 using an approximate field of view from the Solo dwarf survey. Solo has a $1° \times 1°$ camera (Higgs et al., 2016), which we approximate as a circular aperture with radius 0.56°. We present all of the values expected for Solo in the $\bar{N}_{\rm fov}$ column. Our table contains all but two of the galaxies they target, Perseus and HIZSS3A(B), for which we could not obtain stellar masses. Due to a small field of view, nearly all galaxies have a mean expected number of satellites of fewer than 1. However, enough galaxies were surveyed that in aggregate multiple satellites with $M_* > 10^4\,{\rm M}_\odot$ could be in their observations.

Table 4.4 Mean number of satellites with $M_* > 10^4\,M_\odot$ within a 0.56° radius field of view around Local Group dwarf field galaxies

Name	$D_\odot$ **[kpc]**	**Moster**			**GK14**			**GK16**			**Brook**		
		$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]
Leo T	417	0.08	0.01	34	0.04	0.01	31	0.03	0.00	30	0.04	0.00	41
And XXVIII	661	0.10	0.02	36	0.06	0.02	33	0.05	0.01	32	0.06	0.01	43
KKR 3	2188	0.16	0.13	41	0.13	0.10	39	0.09	0.07	37	0.10	0.07	48
Tucana	887	0.18	0.05	42	0.13	0.04	40	0.09	0.03	37	0.10	0.02	48
Leo P	1620	0.17	0.10	42	0.13	0.08	40	0.09	0.06	37	0.10	0.05	48
And XVIII	1355	0.18	0.09	42	0.14	0.08	40	0.10	0.05	38	0.11	0.05	49
Phoenix	415	0.20	0.02	43	0.16	0.01	42	0.11	0.01	39	0.11	0.01	50
KKH 86	2582	0.20	0.16	44	0.17	0.14	42	0.12	0.11	39	0.12	0.09	50
Antlia	1349	0.25	0.11	47	0.23	0.10	46	0.15	0.07	43	0.15	0.06	53
KKR 25	1905	0.26	0.16	47	0.24	0.15	46	0.17	0.11	43	0.15	0.08	53
Aquarius	1072	0.28	0.09	48	0.26	0.08	47	0.18	0.06	44	0.16	0.05	54
DDO 113	2951	0.32	0.26	50	0.31	0.26	50	0.21	0.18	46	0.18	0.14	56
Cetus	755	0.36	0.06	52	0.35	0.06	52	0.24	0.05	48	0.20	0.03	57
ESO 294-G010	2032	0.36	0.22	52	0.35	0.21	52	0.24	0.16	48	0.21	0.12	57
Sagittarius dIrr	1067	0.39	0.11	54	0.42	0.11	54	0.29	0.09	51	0.23	0.05	59
ESO 410-G005	1923	0.41	0.23	54	0.41	0.23	54	0.29	0.17	51	0.23	0.12	59
KKH 98	2523	0.46	0.32	56	0.46	0.31	57	0.34	0.25	53	0.25	0.16	60
Leo A	798	0.52	0.08	58	0.55	0.08	60	0.41	0.07	56	0.28	0.04	62
GR 8	2178	0.55	0.32	59	0.58	0.32	61	0.42	0.25	56	0.29	0.16	63
Pegasus dIrr	920	0.55	0.10	59	0.58	0.11	61	0.44	0.09	57	0.29	0.05	63
UGC 9128	2291	0.59	0.34	60	0.65	0.37	63	0.47	0.29	58	0.32	0.18	64
UGC 4879	1361	0.60	0.19	61	0.67	0.21	63	0.50	0.17	59	0.32	0.10	65
KK 258	2230	0.77	0.41	66	0.89	0.45	69	0.68	0.37	65	0.39	0.20	68
UGCA 86	2965	0.81	0.55	67	0.97	0.63	71	0.74	0.50	67	0.41	0.27	69

Table 4.5 *continued:* Mean number of satellites with $M_* > 10^4\,M_\odot$ within a 0.56° radius field of view around Local Group dwarf field galaxies

Name	$D_\odot$ **[kpc]**	**Moster**			**GK14**			**GK16**			**Brook**		
		$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]	$\bar{N}_{\rm lum}$	$\bar{N}_{\rm fov}$	$R_{\rm vir}$ [kpc]
DDO 99	2594	0.82	0.49	67	0.98	0.56	71	0.75	0.45	67	0.40	0.23	69
UKS 2323-326	2208	0.84	0.43	67	0.99	0.48	72	0.77	0.39	68	0.41	0.20	70
UGC 8508	2582	0.90	0.53	69	1.06	0.58	73	0.82	0.48	69	0.44	0.25	71
KKs 3	2120	0.97	0.45	70	1.22	0.52	76	0.92	0.42	72	0.46	0.21	72
NGC 4163	2858	1.17	0.69	75	1.56	0.84	82	1.22	0.70	78	0.56	0.33	76
WLM	933	1.26	0.16	77	1.69	0.18	85	1.32	0.16	80	0.58	0.07	77
Sextans A	1432	1.28	0.32	77	1.70	0.37	85	1.33	0.31	81	0.60	0.15	78
DDO 125	2582	1.31	0.68	78	1.73	0.81	86	1.41	0.70	82	0.61	0.31	78
DDO 190	2793	1.37	0.76	79	1.87	0.94	87	1.47	0.77	83	0.63	0.35	79
Sextans B	1426	1.40	0.33	79	1.86	0.38	87	1.48	0.33	83	0.64	0.15	79
IC 3104	2270	1.49	0.64	81	2.03	0.77	90	1.63	0.65	86	0.67	0.29	81
NGC 3109	1300	1.65	0.31	84	2.29	0.37	93	1.85	0.32	89	0.75	0.14	83
NGC 6822	459	1.85	0.05	87	2.61	0.07	98	2.09	0.05	94	0.82	0.02	86
IC 1613	755	1.84	0.14	87	2.60	0.16	98	2.16	0.14	94	0.81	0.06	86
IC 4662	2443	2.46	0.95	95	3.61	1.18	109	2.96	1.02	105	1.10	0.43	94
IC 5152	1950	2.81	0.75	100	4.28	0.94	115	3.74	0.84	113	1.29	0.35	99

NOTE. — $\bar{N}_{\rm fov}$ indicates the mean number of luminous satellites with $M_* > 10^4\,M_\odot$ within a field of view of radius 0.56° (corresponding to a footprint of equal area as the Solo Dwarfs Project) centered on target Local Group dwarf galaxies, as predicted with various AM models. Galaxies with a larger heliocentric distance and smaller AM model inferred virial radius, $R_{\rm vir}$, will have a larger fraction of their volume surveyed in the field of view. The total mean number of satellites within each galaxy's virial volume, $\bar{N}_{\rm lum}$, is listed for comparison.

4.5 Systematic Uncertainties

Here we determine what the largest sources of uncertainty are for our luminous dwarf-of-dwarf number count predictions. We show some of them in Fig. 4-7: differences between AM models, counting statistics, and halo-to-halo variations. However, there are other sources of uncertainty that we quantify in this section. We introduce $\bar{N}^5_{\rm lum}$ which is the mean total number of satellites with $M_* > 10^4\,{\rm M}_\odot$ that exist around the five largest field galaxies according to the Moster model. Using $\bar{N}^5_{\rm lum}$ as a baseline, the GK14 model predicts 45% more satellites than the Moster model, and the Brook model predicts 55% less. Statistical fluctuations, driven mostly by Poisson noise, contribute a standard deviation of $\pm 30\%$. Not accounted for are systematic uncertainties in the stellar mass of the host, the total mass of the host, the SHMF, the infall time distribution for the Moster model, the magnitude of scatter in $M_* - M_{\rm halo}$ for the GK16 model, and the reionization model used. Each of these uncertainties must be estimated before a robust prediction can be made regarding whether or not satellites of dwarfs could be discovered, and how many. We therefore estimate typical uncertainties in each input variable and the resultant effect of uncertainty on $\bar{N}^5_{\rm lum}$. Finally, we compare which uncertainties contribute most.

Table 4.6 lists each input variable and variation along with the corresponding % change in $\bar{N}^5_{\rm lum}$. The first value reports the % change in the number of satellites with $M_* > 10^4\,{\rm M}_\odot$, and the second value in parenthesis is for $M_* > 10^3\,{\rm M}_\odot$. When considering scatter in $M_* - M_{\rm halo}$, $\bar{N}^5_{\rm lum}$ refers to the GK16 model, otherwise all instances refer to the Moster model.

Uncertainty in the total halo mass, $M^{\rm host}_{\rm halo}$, is computed from the likelihood distribution for $M_{\rm halo}$ given M_* as described in Section 4.2.5. For the Moster model, one σ uncertainty in the largest field dwarf is $+22\%$ and -18%. Uncertainty in the GK14 and Brook models is similar, but in the GK16 model which has higher scatter in $M_* - M_{\rm halo}$, one σ is $+28\%$ and -35%. Consistent with $\bar{N}$ being directly proportional to $M^{\rm host}_{\rm halo}$ to first order, as written in equation (4.3), a 22% change in halo mass results in a 23% change in $\bar{N}^5_{\rm lum}$.

Table 4.6 Systematic Errors

Parameter Change	% change in $\bar{N}^5_{\rm lum}$	
	$M_* > 10^4\,{\rm M}_\odot$	$M_* > 10^3\,{\rm M}_\odot$
Reionization at $z = 14.4$	-5	-12
Reionization at $z = 11.3$	8	26
Reionization at $z = 9.3$	20	65
$v^{\rm filt}_{\rm max}$, $v^{\rm pre}_{\rm max} \uparrow 25\%$	-38	-47
$v^{\rm filt}_{\rm max}$, $v^{\rm pre}_{\rm max} \downarrow 25\%$	27	70
$M^{\rm host}_{\rm halo} \uparrow 22\%$	23	26
$M^{\rm host}_{\rm halo} \downarrow 18\%$	-18	26
$M^{\rm host}_{*} \uparrow 25\%$	10	10
$M^{\rm host}_{*} \downarrow 25\%$	-12	-12
$M^{\rm infall}_{200}$ SHMF $\alpha = 1.88$	-12	-9
$M^{\rm infall}_{200}$ SHMF $\alpha = 1.75$	10	8
GK16, $\gamma = 0.0$	24	5
GK16, $\gamma = -0.5$	-32	-24
$t_{\rm infall} \uparrow 0.23$ Gyr	< 1	< 1
$t_{\rm infall} \downarrow 0.23$ Gyr	< 1	< 1

Systematic uncertainties in model input variables and their effect on the predicted number of luminous satellites. Per cent change is reported on $\bar{N}^5_{\rm lum}$, the mean number of satellites with $M_* > 10^4\,{\rm M}_\odot$ found around the largest five galaxies for the Moster model. The values in the final column refer to satellites with $M_* > 10^3\,{\rm M}_\odot$. Adjusting scatter, γ, in the $M_* - M_{\rm halo}$ relationship applies only to the GK16 model. More negative values of γ indicate more scatter. Our baseline reionization occurs at $z = 13.3$, our baseline α in the SHMF is 1.81, and our baseline γ is -0.2. $v^{\rm filt}_{\rm max}$ is the maximum circular velocity of halos above which all halos are assumed to have formed stars, regardless of when this $v_{\rm max}$ was reached. $v^{\rm pre}_{\rm max}$ is the maximum circular velocity of halos which, if reached by the redshift of reionization, indicates star formation will proceed.

For stellar mass, we choose an uncertainty of 25%, representative of the uncertainty ranges presented in Roediger and Courteau (2015). Due to the steep dependence of M_* on $M_{\rm halo}$ in the AM models, a change in M_* results in a smaller per cent change in the inferred $M_{\rm halo}$. As a result, a 25% change in M_* yields just a $10-12\%$ change in $\bar{N}^5_{\rm lum}$, whereas the same change in $M^{\rm host}_{\rm halo}$ yields a 26% change.

To estimate uncertainty in the infall times of satellites and the SHMF of field halos, we use a jackknife method with one simulation removed at a time. We find the mean infall time of satellites counting back from $z = 0$ to be $t_{\rm infall} = 7.45 \pm 0.23$ Gyr for the field halos. Shifting the entire infall time distribution by 0.23 Gyr results in a negligible $< 1\%$ effect on $\bar{N}^5_{\rm lum}$. For the SHMF, $\alpha = 1.81 \pm .065$. We therefore adjust α to 1.88 and 1.75 and find the best fit value of K_0 for each slope. A steeper slope reduces M_* at the low $M_{\rm halo}$ end, causing fewer expected luminous satellites. A shallower slope does the opposite, with a total uncertainty of $\sim \pm 11\%$.

Next, we change the level of scatter in the GK16 model. A value of $\gamma = 0$ produces a constant lognormal scatter about $M_* - M_{\rm halo}$ of $\sigma_{\rm scat} = 0.2$ dex. Increasing γ to -0.5 increases the level of scatter relative to our baseline of $\gamma = -0.2$. Three different effects influence the overall outcome of modifying scatter. First, increasing scatter leads to smaller inferred halo mass for a fixed stellar mass, and thus fewer satellites. Second, in GK16's model, increasing scatter requires a steeper $M_* - M_{\rm halo}$ slope, which reduces M_* for a fixed $M_{\rm halo}$, which subsequently reduces our predicted $\bar{N}_{\rm lum}$. Third, increasing scatter causes more of the more numerous lower mass halos to upscatter above a detection threshold than higher mass halos to downscatter, which increases $\bar{N}_{\rm lum}$. However, this third effect diminishes on the mass scale where reionization suppresses low mass galaxies. The interplay of all three effects is complex and model dependent, but ultimately results in less change to $\bar{N}^5_{\rm lum}$ than exists between abundance matching models.

Lastly, we consider modifications to reionization. Using the model presented in Section 4.2.3, we are able to adjust the redshift of reionization to $z = 14.3$, 11.3, and 9.3 and catalog how $\bar{N}^5_{\rm lum}$ responds in Table 4.6. We additionally adjust both $v_{\rm max}$ thresholds ($v^{\rm pre}_{\rm max}$ and $v^{\rm filt}_{\rm max}$) up by 25%, both of which create more dark halos, and

down by 25%, both of which create more luminous halos.

Modifications to reionization can have enormous implications for the abundance of UFD satellites, but little effect on larger satellites. For instance, shifting reionization to later times, $z = 9.3$, increases $\bar{N}^5_{\rm lum}$ by 65% for $M_* > 10^3\,{\rm M}_\odot$. For $M_* > 10^4\,{\rm M}_\odot$, it increases it by 20%, and for $M_* > 10^5\,{\rm M}_\odot$ there is only a 1% level effect. The same trends are true for adjusting $v^{\rm pre}_{\rm max}$ and $v^{\rm filt}_{\rm max}$. For $M_* > 10^3\,{\rm M}_\odot$ there can be as high as a 70% increase in satellites, while for $M_* > 10^5\,{\rm M}_\odot$ the increase is just 13%.

For an individual halo, the approximate uncertainty from each input for $M_* > 10^4\,{\rm M}_\odot$ is as follows: reionization - 33%, total halo mass - 20%, stellar mass of host - 11%, and SHMF - 11%. We consider the scatter in $M_* - M_{\rm halo}$ as part of the spectrum of AM models. When combined in quadrature, the uncertainty reaches 42%, commensurate with the differences between abundance matching models, but less than the Poisson noise of systems with $\bar{N}_{\rm lum} \leq 4$. Since uncertainty in total halo mass and stellar mass are not fully correlated from one halo to the next, their contribution can be mitigated by observing a larger sample of galaxies.

Consequently, the dominant contributors to uncertainty are the abundance matching model, reionization, and Poisson noise. While Poisson noise is uncontrollable, reionization and AM models will improve with future observations and better models. For satellites with $10^3 < M_* < 10^4\,{\rm M}_\odot$, uncertainty in reionization is the single most important model dependent factor. However, for satellites with $M_* > 10^5\,{\rm M}_\odot$, reionization has little influence, and differences between abundance matching models dominate.

Although uncertainties are large, even the most conservative estimates for the existence of satellites suggest that at least one satellite with $M_* > 10^4$ exists around the largest field dwarf galaxies. For $M_* > 10^5\,{\rm M}_\odot$, the lowest estimate comes from the Brook model with a combined uncertainty from reionization, halo mass, stellar mass, and SHMF of 15% for the number of satellites around the five largest field dwarfs. Here, reionization only makes a 2% contribution on its own. Reducing the mean expected predictions by 15% and including Poisson noise, this lower limit still predicts an 88% chance that at least one satellite with $M_* > 10^5\,{\rm M}_\odot$ exists around

one of the targets.

4.6 Conclusions

We have made predictions for the number of luminous satellites around galaxies using SHMFs derived from the *Caterpillar* simulation suite, a model for reionization, and four different AM methods. We find the number of satellites as a function of their minimum stellar mass as well as their host galaxy's total stellar mass. We predict a combined ~ 70 ultrafaint and classical satellites ($M_*^{\rm sat} > 10^3\,{\rm M}_\odot$) around a MW-sized galaxy, consistent with observational expectations. This result is devoid of any missing satellite problem.

For the more massive isolated Local Group dwarf field galaxies, our predictions overwhelmingly indicate that at least one satellite with $M_* > 10^5\,{\rm M}_\odot$ exists, and that many exist with $M_* > 10^4\,{\rm M}_\odot$. Specifically, when observing the virial volumes of the five largest field galaxies combined, there is a $> 99\%$ chance of discovering at least one satellite with $M_* > 10^5\,{\rm M}_\odot$, when employing the AM models of Moster, GK14, and GK16, each paired with a reionization model. The existence of at least one such satellite is even supported by our most conservative case, i.e. using the Brook model that underpredicts the number of MW satellites by $\approx 2\sigma$ compared to other AM models. The Brook model still predicts a $92 \pm 4\%$ chance of one satellite with $M_* > 10^5\,{\rm M}_\odot$ within the combined virial volumes of the five largest field galaxies, IC 5152, IC 4662, IC 1613, NGC 6822 and NGC 3109. If probing down to $M_*^{\rm sat} > 10^4\,{\rm M}_\odot$, $5 - 25$ satellites may even exist.

We therefore conclude that deep, wide-field searches for faint ($10^3 < M_* < 10^6\,{\rm M}_\odot$) satellites around the known, isolated Local Group field dwarf galaxies should result in the discovery of satellites. This bears the opportunity to study satellites that would likely have environments different from any of the known MW or M31 satellites. A smaller mass host implies they would experience reduced ram pressure stripping, reduced tidal forces, and a different local reionization field. An additional implication is that the MW satellites Fornax ($M_* = 2.5 \times 10^7\,{\rm M}_\odot$) and Sagittarius ($M_* \sim 10^8\,{\rm M}_\odot$

at infall; Niederste-Ostholt et al., 2010) more likely than not had one own small satellite galaxy before getting accreted by the MW. Along the same vein, predictions for the numbers of satellites expected around the SMC and LMC will be reported on in a separate paper.

We estimate how many satellites are in an observable line of sight of specific Local Group galaxies as a function of the radius of the field of view. Making specific predictions will help ensure that the most promising targets are observed to avoid spending significant amounts of telescope time on many (potentially unnecessary) pointings to cover the full virial volume of a given field galaxy. We explicitly calculate expected numbers for galaxies included in the Solo dwarf survey Higgs et al. (2016), which has recently observed all isolated dwarf field galaxies within 3 Mpc of the MW. For instance, a single pointing at IC 4662 with their $1^\circ \times 1^\circ$ camera would cover a volume that encompasses a mean of $\sim 0.9-1.2$ satellites with $M_* > 10^4\,\mathrm{M}_\odot$ according to the Moster, GK16, and GK14 models. No satellites of dwarfs have been reported yet in the Solo survey, but given our predictions, we expect that some will be found in their fields of view.

In order to maximize chances of discovering as many satellites as possible, we find the best observing strategy to be to search within small radii of the largest galaxies. After all, the density of satellites is much higher towards their galactic centers than the outskirts. In addition, multiple galaxies should be surveyed this way since observing just one galaxy could yield a non-detection of satellites if it were to have an unusually low halo mass for its stellar mass. If results from a complete survey of all field dwarfs would be available, especially if observations would also cover the full virial volumes of many galaxies including the MW and M31, constraints on the epoch of reionization and the various AM models could be derived.

AM models contribute a large $\sim 50\%$ uncertainty relative to the Moster model to the predictions of satellite abundances. Reionization, which preferentially suppresses star formation in low mass galaxies, can contribute up to 70% uncertainty for $M_*^{\mathrm{sat}} > 10^3\,\mathrm{M}_\odot$, but only $\sim 10\%$ for $M_*^{\mathrm{sat}} > 10^5\,\mathrm{M}_\odot$. Measurements to a completeness limit of $10^5\,\mathrm{M}_\odot$ would primarily constrain the $M_* - M_{\mathrm{halo}}$ relationship for galaxies and thus

the accuracy of the abundance matching models. If observations could ever yield a census of satellites down to $10^3\,M_\odot$, reionization could also be constrained by placing limits on the ability of reionization to suppress (or not) star formation in these small halos.

Finally, we emphasize that discovering no satellites at all is highly unlikely. Nevertheless, a non-detection would imply at least two of the following: a very strong suppression of star formation by reionization, a low $M_* - M_{\rm halo}$ relationship for low mass galaxies, a high $M_* - M_{\rm halo}$ relationship for galaxies with masses typical of field dwarfs, and a MW which has an abnormally large number of luminous satellites with $M_* < 10^6\,M_\odot$. We therefore conclude that there are almost certainly many small satellites of dwarf galaxies waiting to be discovered and that their discovery will help refine not just AM models and reionization but our understanding of low mass galaxy formation.

Chapter 5

A missing satellite problem around the LMC

The content of this chapter will be submitted to the Monthly Notices of the Royal Astronomical Society *with at most minor edits. The authors are Gregory A. Dooley, Annika H.G. Peter, Jeffrey L. Carlin, Anna Frebel, Keith Bechtol, and Beth Willman*

Disclaimer

This chapter was predominately a product of my own work. Jeff Carlin compiled the names and properties of isolated LMC-sized galaxies. He also compiled the names and properties of all currently known Milky Way satellites, and provided the text and references introducing them. All authors provided comments and feedback on the paper.

Acknowledgements

AHGP acknowledges support from National Science Foundation (NSF) Grant No. AST-1615838. BW acknowledges support by NSF Faculty Early Career Development (CAREER) award AST-1151462. AF acknowledges support from NSF-CAREER grant AST-1255160.

Abstract

We predict satellite abundances around Magellanic Cloud-mass host galaxies, specifically modeling the volume within 50 kpc of the Large Magellanic Cloud (LMC). Using simple abundance matching and reionization models applied to the *Caterpillar* simulations, we predict a mean of 4-8 satellites with stellar mass $M_* > 10^4\,\mathrm{M}_\odot$, and a mean of 3-4 satellites with $80 < M_* \leq 3000\,\mathrm{M}_\odot$ in this volume around the LMC. Surprisingly, all 12 currently known satellite candidates have stellar masses of $80 < M_* \leq 3000\,\mathrm{M}_\odot$. The dearth of any larger satellites may indicate a steep $M_* - M_{\mathrm{halo}}$ relationship, or strong tidal stripping. The large abundance of known very faint galaxies, which is expected to increase with ongoing surveys, may rule out an early reionization redshift for the Local Group ($z_{\mathrm{reion}} \lesssim 9$ preferred), or indicate an abrupt flattening of the $M_* - M_{\mathrm{halo}}$ relationship at $10^3\,\mathrm{M}_\odot$. Despite large uncertainty in the overall number of satellites predicted, we more robustly predict that within 50 kpc of the LMC today, $\sim 33\%$ of galaxies accreted with the LMC, $\sim 20\%$ with the Small Magellanic Cloud, and $\sim 47\%$ were only ever Milky Way satellites. Observing satellites in isolated LMC-sized field galaxies can help to decipher how unusual the satellite population near the LMC is, and better constrain the $M_* - M_{\mathrm{halo}}$ relationship. Modeling known LMC-sized galaxies within 8 Mpc, we predict 1-6 (2-12) satellites with $M_* > 10^5\,\mathrm{M}_\odot$ ($M_* > 10^4\,\mathrm{M}_\odot$) within the virial volume of each, and 1-3 (1-7) within a single 1.5° diameter field of view, making their discovery likely.

5.1 Introduction

The hierarchical structure formation predicted in Lambda Cold Dark Matter theories implies that dark matter halos of all scales contain substructure from past accretion events (Springel et al., 2008). The Milky Way and M31 each fit this paradigm, both orbited by many known satellites. Moving down roughly an order of magnitude in total halo mass, the Large Magellanic Cloud should similarly contain its own substructure. Lynden-Bell (1976) made an initial proposition of the Large and Small Magellanic Clouds (LMC and SMC) as a dynamically linked group, followed by speculation that additional MW satellites and globular clusters were part of a greater Magellanic stream (Lynden-Bell, 1982; Lynden-Bell and Lynden-Bell, 1995). In the years since, there has been a vigorous debate in the literature as to how many, if any, of the MW satellites were originally satellites of the LMC (D'Onghia and Lake, 2008; Nichols et al., 2011; Sales et al., 2011). One way to make progress is to probe the volume of space near the Clouds, especially the volume that remains tidally bound

to them, for new satellites, as suggested by Sales et al. (2011).

Recently, surveys including Dark Energy Survey (DES), SMASH, Pan-STARRS, ATLAS, and MagLiteS have in fact revealed $\sim$ 20 new satellite candidates, many currently in the vicinity of the LMC (Bechtol et al., 2015; Drlica-Wagner et al., 2015; Kim et al., 2015; Kim and Jerjen, 2015; Koposov et al., 2015a; Laevens et al., 2015; Martin et al., 2015; Luque et al., 2016; Torrealba et al., 2016b; Drlica-Wagner et al., 2016). There is an ongoing debate as to which of these new satellites is physically associated with the Clouds. Yozin and Bekki (2015), Koposov et al. (2015a) and Drlica-Wagner et al. (2015) each identify that the clustering of satellites near the LMC is highly improbable unless many satellites are dynamically associated with the Magellanic system. Deason et al. (2015) considered satellites on an individual basis, comparing their positions and velocities to N-body simulations of LMC-sized galaxies and their substructure in a MW-sized system, and found that $2-4$ out of the 9 then known DES satellites were likely part of the LMC. Jethwa et al. (2016) followed with slightly different methodology, injecting LMC-sized galaxies into a MW potential with the known kinematics of the LMC, and found that $7-12$ of 14 satellites were consistent with the LMC. Sales et al. (2016) conducted a follow-up to the Sales et al. (2011) study, and this time found 8 of the 20 new satellites possibly consistent with the LMC, with one, Horologium I, definitely consistent. They agree with Koposov et al. (2015b), who had previously also concluded that Horologium I is a satellite of the LMC.

There remains much uncertainty over the fundamental theoretical question of how many satellites the LMC should have brought with it at infall. Jethwa et al. (2016) predicts that a total of 70^{+30}_{-40} were accreted with the LMC and SMC. This means as many as 1/3 of all MW satellites actually entered as part of the Magellanic Cloud system. In stark contrast, Sales et al. (2016) predicts that only 5% should have entered with the LMC, and Deason et al. (2015) estimates a value of 7%, but allows for a range from a low of 1% to a high of 25%.

It is hard to reconcile how many satellites should be observed around the LMC and SMC with how many are actually observed without a consistent and transparent

prediction of satellite populations. As more satellites are discovered and/or confirmed, a prediction with clearly defined inputs can be tested to constrain the underlying galaxy formation physics in the model. The goal of this paper, a companion paper to Dooley et al. (2016b), is to show what simple, well-motivated theories predict for satellite populations in the LMC and SMC, as well as isolated galaxies of similar size to the LMC and SMC. Hence, we predict how many satellites were likely brought in with the LMC and SMC at infall, and how many satellites should exist within the local vicinity of the LMC today where many of the recent dwarfs were discovered. Any mention of the "LMC" by itself refers to the actual LMC. Any mention of "LMC-size", "LMC-scale", or "LMC analog" galaxies refers to galaxies which have a similar stellar mass to the LMC, but are not the LMC. The same applies to the SMC.

It is important to compare the LMC satellite population with those of isolated galaxies of comparable stellar mass, which provides a great opportunity to study satellite populations without the complicating dynamical issues of tidal stripping, ram pressure, and complex orbits of those which are in close proximity to the LMC, SMC, and MW. Additionally, the ambiguity of associating satellites with their original host does not exist for isolated galaxies. In that regard, isolated hosts are a cleaner sample through which observations can provide better empirical constraints on the estimates for the satellite population of our own LMC. These systems are also test grounds to study the effect of environment on dwarf galaxy star formation (Wetzel et al., 2015), where gravitational effects are reduced and reionization may proceed differently than near the MW and M31. On the other hand, isolated galaxies are too distant to detect the extremely low luminosity satellites that are being found near the Magellanic Clouds. Thus to fully understand the satellite populations of LMC-sized hosts as well as to test our model for populating hosts with satellites in general, it is important to observe and analyze both LMC-size isolated galaxies and the LMC itself.

The prospects for finding dwarf galaxies around nearby LMC analogs in and beyond the Local Group is good, as new surveys prove that even ultrafaint dwarf galaxies can be identified at distances of $\sim$ a few Mpc. Recent surveys of resolved stellar pop-

ulations in ~Milky Way-mass halos are revealing faint dwarf satellites and their remnants (e.g., Cen A: Crnojević et al. 2014, 2016; M 81: Chiboucas et al. 2013; NGC 253: Sand et al. 2014; Romanowsky et al. 2016; NGC 891: Mouhcine et al. 2010). There are even some isolated examples of dwarfs around dwarfs. For instance, the dwarf galaxy Antlia B ($M_V \sim -9.7$) is located near the Local Group galaxy NGC 3109, which has a stellar mass similar to the SMC (Sand et al., 2015). The tidally disrupting dwarf galaxy NGC 4449B and its associated tidal stream (Martínez-Delgado et al. 2012; Rich et al. 2012; Toloba et al. 2016) has been discovered near NGC 4449 ($D = 4.3$ Mpc) which has a stellar mass similar to the LMC. Finally, the $M_V = -7.7$ ($M_* \sim 10^5\,M_\odot$) dwarf MADCASH J074238+652501-dw has been found around the $M_* \sim 2 \times M_*^{LMC}$ host galaxy NGC 2403 ($D = 3.2$ Mpc; Carlin et al. 2016). To date, systematic searches for satellite companions of lower-mass hosts are lacking. However, the available capabilities of wide-field imagers on large aperture telescopes to resolve stellar populations to nearly the outer reaches of the Local Volume ($D \lesssim 8$ Mpc) are beginning to be exploited (e.g., Carlin et al. 2016) to study the halos of LMC analogs beyond the Local Group.

Here, we provide the theoretical context for both the LMC satellite population and Magellanic Cloud analogs. In Section 5.2, we describe our theoretical approach. In Section 5.3, we present our main results. First, we highlight some surprising results when we compare simple theoretical models of the satellite population around the LMC to the observed population. Namely, we show that the vicinity of the LMC has a statistically significant dearth of satellites with $10^4\,M_\odot < M_* < 10^7\,M_\odot$, especially compared with the number of satellites with $M_* \lesssim 10^3\,M_\odot$. This primary result is seen in Fig. 5-2. We present different hypotheses to explain this discrepancy. Finally, we make predictions for satellite abundances around specific target galaxies over a range of host galaxy stellar masses near that of the LMC that are located between 2 and 8 Mpc from the MW. Observations of these systems can help unravel the "missing satellite" puzzle of why so few (currently zero) satellites with $10^4\,M_\odot < M_* < 10^7\,M_\odot$ are found within 50 kpc of the LMC. To guide observers, we further provide estimates on the line of sight radial dependence of satellite abundances to motivate and compare

to observational searches. In Section 5.4, we summarize our key findings and make both theory- and observation-oriented recommendations to solve this puzzle of the dwarf satellite population near the LMC.

5.2 Methods

We apply the same techniques used in Dooley et al. (2016b) to make predictions for the distribution of possible satellites around host galaxies. Given the stellar mass of a host galaxy, we generate a random realization of satellites around it according to the following procedure:

1. Convert the host's stellar mass to a total halo mass (see section 2.5 of Dooley et al. (2016b) for details).

2. Determine the typical subhalo mass function (SHMF) for host galaxies.

3. Sample the SHMF to generate a random realization of dark matter subhalos for the host.

4. Model reionization by assigning subhalos to be dark or luminous according to a probability function that depends on a halo's peak or infall mass.

5. Assign a stellar mass to each luminous subhalo according to an abundance matching (AM) model.

6. *(Optionally)* Assign distances to each satellite according to a radial distribution profile of luminous satellites.

We repeat this procedure $30,000$ times for each calculation of interest in order to obtain a convergence of the mean and to sample the variance.

The SHMF, fraction of galaxies that survive reionization, radial distribution of satellites, and infall distribution of satellites are found by analyzing simulated galaxies from the *Caterpillar* simulation suite (Griffen et al., 2016b). This consists of 33 high

particle resolution ($m_{\rm p} = 3 \times 10^4\,{\rm M}_\odot$) and high temporal resolution (320 snapshots) zoom-in simulations of Milky Way-sized galaxies.

As in Dooley et al. (2016b), we model reionization's ability to leave dark matter halos entirely dark as follows: a halo's maximum circular velocity must reach a critical value, $v_{\rm max}^{\rm pre}$, before the redshift of reionization, $z_{\rm reion}$, or reach a larger critical value after reionization, $v_{\rm max}^{\rm filt}$, in order to form stars. Applying this model to the *Caterpillar* suite leads to a function indicating the fraction of halos that are luminous as a function of halo mass, as seen in fig. 3 of Dooley et al. (2016b). We choose values to replicate the model of Barber et al. (2014) with $z_{\rm reion} = 13.3$, finding $v_{\rm max}^{\rm pre} = 9.5\,{\rm km/s}$ and $v_{\rm max}^{\rm filt} = 23.5\,{\rm km/s}$. Unless otherwise stated, we use this as our "baseline" reionization model.

A reionization redshift of $z = 13.3$ is relatively early, especially compared to the value of $7.8 \leq z_{\rm reion} \leq 8.8$ estimated by Planck Collaboration et al. (2016a). However, in order to remain consistent with the model used in Dooley et al. (2016b), which leads to a prediction for the number of MW satellites with $M_* > 10^3\,{\rm M}_\odot$ consistent with that of completeness corrected observations (Hargis et al., 2014), we stick with $z_{\rm reion} = 13.3$ in our baseline reionization model. We then also explore models where reionization occurs later for comparison. Furthermore, the physical meaning of $z_{\rm reion}$ in our model is imprecise since decreasing it is degenerate with decreasing $v_{\rm max}^{\rm pre}$. Decreasing either parameter increases the fraction of halos which form stars.

Due to large differences in AM models, particularly at low stellar masses, we implement a total of five AM models. These models are the Brook model (Brook et al., 2014), the Moster model (Moster et al., 2013), the GK14 model (Garrison-Kimmel et al., 2014b), the GK16 model (Garrison-Kimmel et al., 2016), and the Behroozi model (Behroozi et al., 2013a).

We define the halo virial radius, $R_{\rm vir}$, using the Bryan and Norman (1998) fitting function for the radius at which a halo is virialized in the spherical tophat model. At $z = 0$ for our cosmological parameters, $R_{\rm vir}$ is the radius such that the mean enclosed halo density is 104 times the critical density of the universe, $\rho_c = 3H_0^2/8\pi{\rm G}$. $M_{\rm vir}$

refers to the gravitationally bound mass within $R_{\rm vir}$, and any mention of R_Δ or M_Δ refers to the radius and mass of a halo where the mean enclosed density is Δ times the critical density.

All of our steps are identical to those in Dooley et al. (2016b) except for two adjustments. First, we use a slightly different SHMF. We find a single function in which abundances are directly proportional to the host halo mass is sufficient to describe the SHMF over a host mass interval of 1 dex, but begins to become less accurate outside the range it was calibrated to. Since the total halo mass of the SMC and LMC is around 1 dex smaller than the MW mass, we calibrate the mass function specifically to halos in a mass range that encompasses the estimated masses of the LMC and SMC. We select a subset of isolated field halos in the simulations (outside the virial radius of the MW-like host and within the contamination radius) which have masses in the range $5 \times 10^{10} < M_{\rm vir} < 5 \times 10^{11}\,{\rm M}_\odot$. The differential abundance of subhalos follows the form

$$\frac{{\rm dN}}{{\rm dM_{sub}}} = K_0 \left(\frac{M_{\rm sub}}{{\rm M}_\odot}\right)^{-\alpha_{\rm mf}} \frac{M_{\rm host}}{{\rm M}_\odot}, \tag{5.1}$$

where the mass definition of $M_{\rm host}$ and $M_{\rm sub}$ vary according to AM model. The total number of subhalos counted, however, always refers to the number within the Bryan and Norman (1998) virial radius. The function is controlled by two fit parameters: the logarithmic slope $\alpha_{\rm mf}$ and normalization factor K_0. The values we obtain and use are $\alpha_{\rm mf} = 1.93$, $K_0 = 0.00588$ for $M_{\rm sub} = M_{\rm vir}^{\rm peak}$ and $M_{\rm host} = M_{\rm vir}$ which is needed for the GK14 and GK16 models, $\alpha_{\rm mf} = 1.88$, $K_0 = 0.00219$ for $M_{\rm sub} = M_{200}^{\rm infall}$ and $M_{\rm host} = M_{200}$ which is needed for the Moster model, and $\alpha_{\rm mf} = 1.88$, $K_0 = 0.00282$ for $M_{\rm sub} = M_{350}^{\rm peak}$ and $M_{\rm host} = M_{350}$ which is needed for the Brook model. These values lead to a $\sim 20\%$ reduction in abundance predictions compared to the values used for the smaller dwarf field halos in Dooley et al. (2016b). For predictions of the MW, we use the same SHMF parameters as listed in Dooley et al. (2016b).

Second, the LMC is massive enough such that we can no longer assume the subhalo abundances are Poisson distributed. As pointed out in Boylan-Kolchin et al. (2010),

and confirmed in Mao et al. (2015) and Lu et al. (2016), subhalo abundances more accurately follow a negative binomial distribution where the variance in the mean number of satellites increases relative to that of a Poisson distribution as $M_{\rm sub}/M_{\rm host}$ decreases. More quantitatively, the variance of $N(> M_{\rm sub}/M_{\rm host})$ is $\sigma^2 = \sigma_P^2 + \sigma_I^2$, where $\sigma_P^2 = \langle N \rangle$ is the Poisson variance, and $\sigma_I^2 = s_I^2 \langle N \rangle^2$ is an intrinsic scatter. The value of the fractional intrinsic scatter, s_I, was found to be 0.18 in Boylan-Kolchin et al. (2010). Studying the variance in *Caterpillar* halos with our own mass functions, we find $s_I = 0.14$ to be a better fit. We implement this by generating random samples of dark matter halos in logarithmic bins of $M_{\rm sub}/M_{\rm host}$. The number of halos per bin is chosen according to a negative binomial distribution whose variance increases for lower mass ratios. For our scale of galaxies, the variance in the number of satellites with $M_* > 10^3\,{\rm M}_\odot$ is larger than that of a Poisson distribution, but is roughly the same for satellites with $M_* > 10^5\,{\rm M}_\odot$.

5.3 Results

In the next two subsections, we consider two applications of our methods. First, we predict the satellite galaxy distribution within the vicinity of the LMC, and compare it to currently known dwarf galaxies in that region. To do this, we approximate the LMC and SMC as once-isolated galaxies that accreted on to the MW. Second, we make predictions for isolated LMC and SMC mass galaxies. This allows us to estimate and tabulate the number of satellites that may be found in distant galaxies of similar size. The two investigations are mutually-reinforcing in that improved understanding of one can lead to improved modelling and predictions for the other.

Both subsections require an estimate of the total halo mass of the LMC and SMC if they were isolated at $z = 0$, as well as the virial radius of each. We assume stellar mass values of $2.6 \times 10^9\,{\rm M}_\odot$ for the LMC (calculated, as are all stellar masses of Local Volume galaxies in this work, based on the luminosity derived using the $K_{\rm S}$-band magnitude and distance from Karachentsev et al. 2013[1], assuming $M_{{\rm Ks},\odot} = 3.28$,

[1]Updated version available at `http://www.sao.ru/lv/lvgdb/`.

and $M/L = 1$; see also van der Marel et al. 2002; Harris and Zaritsky 2009), and $7.1 \times 10^8\,\mathrm{M}_\odot$ for the SMC. From these stellar masses, we infer the total halo mass as in Dooley et al. (2016b), using the AM models with an assumed 0.2 dex 1σ log-normal scatter in the $M_* - M_{\rm halo}$ relationship. The GK16 model is an exception, in which its scatter is larger and increases towards lower mass halos. For the LMC, we find a value of $M_{\rm vir} = 2.3 \times 10^{11}\,\mathrm{M}_\odot$ for the GK models, $2.1 \times 10^{11}\,\mathrm{M}_\odot$ for the Behroozi model, $1.7 \times 10^{11}\,\mathrm{M}_\odot$ for the Moster model, and $1.8 \times 10^{11}\,\mathrm{M}_\odot$ for the Brook model. This corresponds to a virial radius of 156, 153, 141 and 146 kpc. For the SMC, we find $M_{\rm vir} = 1.3 \times 10^{11}\,\mathrm{M}_\odot$ for the GK models, $1.1 \times 10^{11}\,\mathrm{M}_\odot$ for the Behroozi model, $9.2 \times 10^{10}\,\mathrm{M}_\odot$ for the Moster model, and $9.5 \times 10^{10}\,\mathrm{M}_\odot$ for the Brook model. The corresponding radii are 132, 123, 116, and 117 kpc.

The masses we infer for the LMC are slightly lower than the value of $M_{200} = 2 - 2.5 \times 10^{11}\,\mathrm{M}_\odot$ (which converts approximately to $M_{\rm vir} = 2.3 - 3.9 \times 10^{11}\,\mathrm{M}_\odot$) estimated by Nichols et al. (2011), Peñarrubia et al. (2016), and Jethwa et al. (2016), but larger than the LMC analog mass of $3.6 \times 10^{10}\,\mathrm{M}_\odot$ used in Sales et al. (2016).

5.3.1 Number of LMC and SMC satellites

In this section, we consider the actual LMC and SMC, quantifying the number of their satellites. We want to estimate the number of satellites they each had upon accretion onto the MW, which corresponds to the number of present-day MW satellites once associated with them. To do so, we must estimate their halo mass at infall, which we infer from their stellar mass. However, estimating their stellar masses at infall accurately is difficult due to uncertainty over their infall time. Depending on the bulk velocity of the LMC, the mass of the MW and the Clouds, the evolving MW gravitational potential, models for the Magellanic Stream, and LMC-SMC interactions, the accretion time of the LMC could be as recent as 1 Gyr ago (Busha et al., 2011a), longer than 4 Gyr ago (Bekki, 2011), or anywhere from $1 - 12$ Gyr ago (Shattow and Loeb, 2009; Besla et al., 2007; Kallivayalil et al., 2013).

Fortunately, this very large uncertainty does not significantly impact the number of satellites we predict. If the LMC accreted onto the MW 5 Gyr ago for instance,

its stellar mass at infall would be $\sim 30\%$ less than today according to the LMC star formation history computed in Harris and Zaritsky (2009). A 30% decrease in the stellar mass of an LMC-sized host only decreases the predicted abundance of satellites by 13% (as seen in Figure 5-6). When limiting predictions to fixed volumes rather than the virial radius, the difference drops yet again. Within 50 kpc, the change is 8%. Moreover, uncertainty in the present day stellar mass of the LMC and SMC is large, at $-30\%/+70\%$ according to Harris and Zaritsky (2009). With no obvious best solution for identifying the stellar mass at infall, we simply use the present day estimate of stellar mass for the LMC and SMC, acknowledging that this uncertainty leads to a $\sim \pm 15\%$ uncertainty in satellite abundances in our model.

In Fig. 5-1, we show our prediction for the number of satellites within an isolated halo's virial radius as a function of the minimum satellite stellar mass. We do this for hosts with stellar masses of the LMC and SMC. Particularly for low stellar masses of the satellites, the number of satellites predicted varies greatly due to different abundance matching models, super Poissonian noise, and uncertainty in reionization. Regarding AM models, most of the discrepancy between the GK models and the Moster model arises from their different predictions of total halo mass and virial radius, with the Moster model predicting smaller values. The Brook model assigns lower stellar masses to halos than the rest, so much so that its predictions for UFD MW satellites are below the completeness limit estimates of Hargis et al. (2014) and Drlica-Wagner et al. (2015) as shown in Dooley et al. (2016b). In that regard it can be considered a lower limit abundance matching model for the number of satellites with $M_* > 10^3\,\mathrm{M}_\odot$. The Behroozi model, on the other hand, is known to overpredict the number of MW satellites with $M_* > 10^5\,\mathrm{M}_\odot$ (Garrison-Kimmel et al., 2014b; Dooley et al., 2016b), and be seen as an upper limit in that mass range.

We highlight uncertainty due to counting statistics and reionization on the GK16 model since it is calibrated down to a lower M_* than the other models, is the only model to explicitly model scatter in the $M_* - M_{\mathrm{halo}}$ relationship, and is the most recent model. The shaded orange band indicates the $\pm 1\sigma$ range due to counting statistics about the mean. With the GK16 model and our baseline reionization, we predict

that $7-14$ and $4-9$ satellites with $M_* > 10^3\,\mathrm{M}_\odot$ are accreted into the Milky Way with the LMC and SMC respectively. For satellites with $M_* > 10^5\,\mathrm{M}_\odot$, we predict $2-6$ and $1-4$.

Reionization in our baseline model occurs relatively early, at $z = 13.3$. We demonstrate how delaying reionization affects the abundance of faint satellites by including predictions for the GK16 model with reionization occurring at $z = 11.3$ and $z = 9.3$. When reionization begins later, it suppresses fewer low mass halos leading to a large increase in galaxies with $M_* < 10^5\,\mathrm{M}_\odot$. The effect is diminished for larger galaxies, with almost no change occurring for satellites with $M_* > 10^5\,\mathrm{M}_\odot$. More observations of low mass satellites are needed to provide better joint AM and reionization model constraints.

5.3.1.1 Stellar mass function of LMC vicinity satellites

Fig. 5-1 predicts the number of satellites around LMC and SMC-sized galaxies if they were isolated. However, the actual LMC and SMC are located only 50 and 60 kpc from the MW Galactic Center. Using a MW mass of $1.4\times10^{12}\,\mathrm{M}_\odot$, and approximating the dark matter halo density distributions as Hernquist profiles (Hernquist, 1990), we estimate their tidal radii as the distance to the L_3 Lagrange point and get $15-17$ kpc for each. Thus nearly all of the former Magellanic satellites would now no longer be gravitationally bound to either the LMC or the SMC. Even if not bound though, satellites that accreted close to the LMC and SMC are likely still spatially correlated with them. This is especially true if the Magellanic Clouds are on their first pericentric passage, as favored by Besla et al. (2007); Busha et al. (2011a); Kallivayalil et al. (2013) and Sales et al. (2016), although they could be on second or third passages (Shattow and Loeb, 2009; Bekki, 2011).

Since most of the recently discovered satellites near the LMC are within 50 kpc of it, and in particular all 12 satellites that Jethwa et al. (2016) found to have a probability of $> 50\%$ to be associated with the LMC are within 50 kpc, we focus our predictions on that subvolume. We approximate that the number of Magellanic satellites within 50 kpc of the LMC and SMC at infall is still the same number within

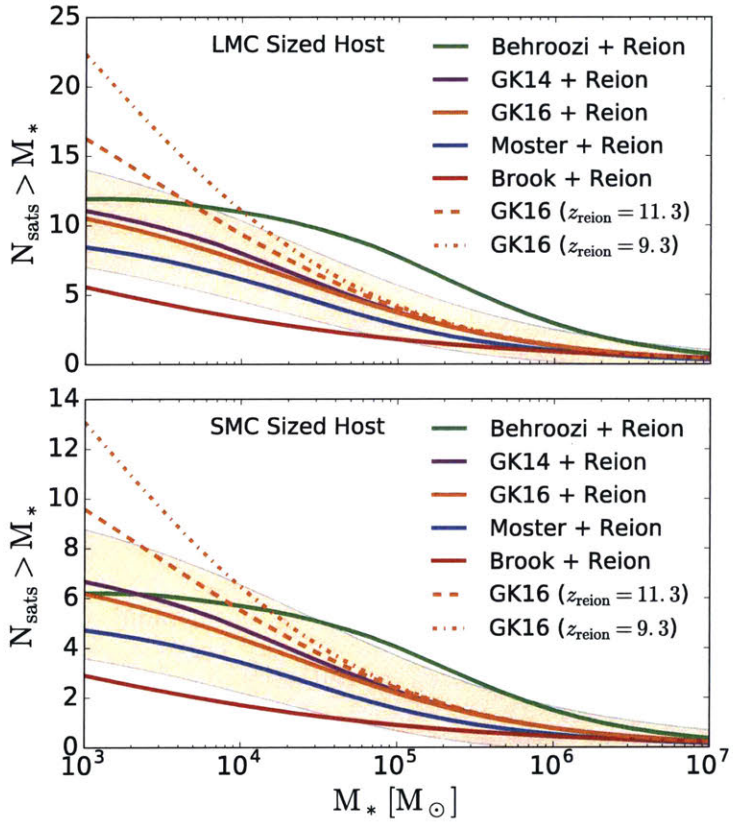

Figure 5-1 *Upper Panel:* Mean number of satellites around an isolated LMC-sized host galaxy as a function of the minimum satellite stellar mass. *Lower Panel:* Mean number of satellites around an isolated SMC-sized host as a function of the minimum satellite stellar mass. The shaded band in each panel shows the one σ variation in satellite abundance for the GK16 model. The effect of delaying reionization by shifting its starting redshift in our model from $z = 13.3$ (solid lines) to $z = 11.3$ (dashed orange line) and $z = 9.3$ (dot-dashed orange line) is shown for the GK16 model. Reionization controls the fraction of low-mass galaxies which can form stars.

those distances presently. In this regard, our predictions are an upper limit, and are more accurate the more recently the LMC and SMC accreted. We comment on the effect of satellites migrating away from the LMC and SMC later in our analysis.

The expected number of satellites in our target subvolume come from three sources: the MW, the LMC, and the SMC. Assuming a reionization model, a MW mass of $1.4 \times 10^{12}\,\mathrm{M}_{\odot}$, and an isotropic satellite distribution, we integrate the radial distribution of satellites from each source over the LMC subvolume, placing the MW at 50 kpc from the LMC, and the SMC at 24 kpc from the LMC. The radial distribution of satellites is determined from the *Caterpillar* simulations, selecting only subhalos which we identify as luminous in our baseline reionization model. This is a very critical step, since the radial distribution of luminous satellites is much more concentrated than that of all subhalos, one of the biases discussed in Sawala et al. (2016). We would under predict the combined satellite abundance near the LMC by a factor of ~ 2.5 if we used the radial distribution of all $z = 0$ subhalos.

We find that the normalized radial distribution of satellites within the host halo virial radius does not depend on the mass of the host. The distribution does depend weakly on subhalo peak mass, with satellites having $M_{\mathrm{vir}}^{\mathrm{peak}} > 10^{9}\,\mathrm{M}_{\odot}$ being less centrally concentrated than those with $M_{\mathrm{vir}}^{\mathrm{peak}} < 10^{9}\,\mathrm{M}_{\odot}$, which are subject to more selection effects from reionization due to being smaller. However, the difference between these distributions leads to only a 10% change in our predictions for the LMC, which is within the 1σ uncertainty of the mean of our radial distribution. We therefore use a single radial distribution as a function of r/R_{vir} calibrated to all luminous satellites. We tested for this on all isolated halos within the *Caterpillar* suite with a mass of $10^{10} < M_{\mathrm{vir}} < 3 \times 10^{12}\,\mathrm{M}_{\odot}$. The functional form of our radial distribution, and a plot comparing the distribution of luminous satellites against all subhalos, can be found in Dooley et al. (2016b).

In Fig. 5-2, we show the predicted cumulative satellite stellar mass function for galaxies in the LMC subvolume. We show this for the Behroozi, GK16, and Brook models. Due to restricting our predictions to a small fixed volume, the Moster, GK14, and GK16 models all predict nearly identical values, so we show just the GK16 model

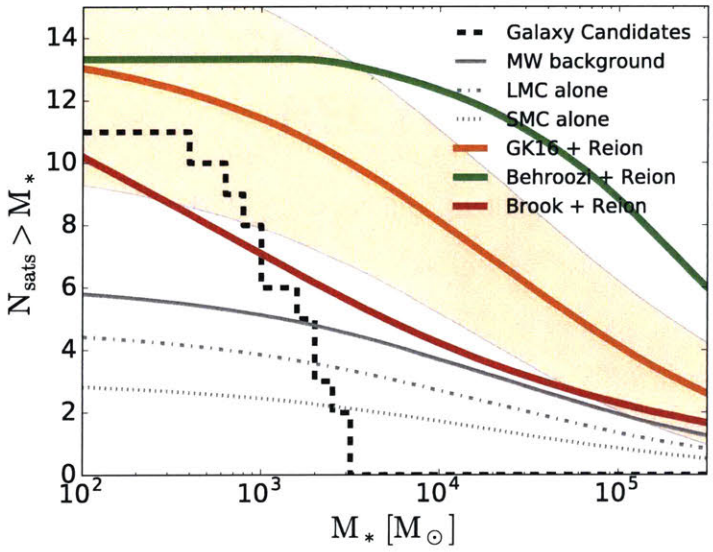

Figure 5-2 Predicted and actual number of satellites within 50 kpc of the LMC as function of minimum stellar mass. Dashed black line shows the stellar mass function of all currently known candidate galaxies. At an estimated stellar mass of just $82\,M_\odot$, Cetus II falls off the mass range plotted. For galaxies with $M_* < 10^4\,M_\odot$, this can be regarded as a lower limit since more low stellar mass galaxies are expected to be found in ongoing surveys. Also plotted are predictions for the stellar mass function according to the Behroozi, GK16, and Brook AM models. For the GK16 model in orange, we include shaded bands indicating 1σ variance due to counting statistics and uncertainty in the radial distribution. Predictions from the GK14 and Moster models are nearly identical to the GK16 model and therefore not shown. The GK16 prediction is broken down into contributions of satellites from the MW background, the LMC, and the SMC.

for simplicity. As the most recent and versatile model, we split the predictions for the GK16 model into contributing components from the MW, LMC, and SMC, as shown in grey. The sum of all three components is shown by the solid orange line, with 1σ variation due to counting statistics shown by the shaded orange band.

We compare these predictions to the dashed black line, which shows the cumulative stellar mass function for all currently known satellite candidates within 50 kpc of the LMC. A major caveat is that this function is likely a lower limit for satellites with $M_* < 10^4\,M_\odot$, since the full extent of the volume we are considering has not yet been surveyed and analyzed for such faint galaxies (Drlica-Wagner et al., 2016). Not included in the observed sample are the LMC and SMC themselves, since we

purposefully choose our small subvolume to contain these two rare galaxies. Whether we include Magellanic Cloud sized systems in our predictions or not is unimportant, since so few are predicted in all models.

The catalog of satellites, their positions and distances (from which we calculate the distance to the LMC) were taken from the compilation of McConnachie (2012),[2] which we supplemented with the following systems discovered more recently than the last update of the McConnachie catalog in Sept 2015: Virgo I (Homma et al., 2016), Pictor II (Drlica-Wagner et al., 2016), Crater 2 (Torrealba et al., 2016a), Aquarius 2 (Torrealba et al., 2016b), and DES J0225+0304 (Luque et al., 2016). Stellar masses were calculated based on the derived V-band absolute magnitudes, doubling the luminosity to account for M_V being the luminosity within the half-light radius, and assuming a stellar mass to light ratio of one. Systematic uncertainty in the luminosity of satellites up to a factor of two makes little difference in the interpretation and qualitative nature of Fig. 5-2. A factor of 10 increase in luminosity would be needed to significantly alter the interpretation.

After selecting only those within 50 kpc of the LMC, our sample (in order of smallest to largest distance from the LMC) consists of Pic II, Ret II, Tuc IV, Tuc V, Tuc III, Tuc II, Hor I, Hor II, Ret III, Cet II, Gru II, and Eri III. Of these, only three have been spectroscopically confirmed as dwarf galaxies: Reticulum II (Koposov et al., 2015b; Simon et al., 2015; Walker et al., 2015), Horologium I (Koposov et al., 2015b) and Tucana II (Walker et al., 2016). This means some candidates in our sample may end up not being true dwarf galaxies. For instance, spectroscopic follow-up of Tucana III suggests it is a tidally stripped dwarf galaxy, but its status as a galaxy is not definitive (Simon et al., 2016). Ongoing spectroscopic analysis, however, suggests that most, if not all candidates, will be confirmed as galaxies (Fermi-LAT et al., 2016). Uncertainty in satellite positions could mean gaining or $0-2$ satellites with $M_* < 10^3\,\mathrm{M}_\odot$, which would not be enough to change the interpretation nor qualitative nature of Fig. 5-2.

More galaxies are also expected, since this population comes from an incomplete

[2]Available online at `http://www.astro.uvic.ca/~alan/Nearby_Dwarf_Database.html`.

survey area around the LMC. The Magellanic Satellites Survey (MagLiteS) is likely to uncover more ultrafaint satellites near the LMC in a footprint not already surveyed by DES (Drlica-Wagner et al., 2016). It is also possible that within the DES footprint, faint dwarf galaxies (those with $M_* < 10^4\,\mathrm{M}_\odot$) will continue to be discovered. Between the likelihood that candidate galaxies will be confirmed as real galaxies, and the expectation of discovering of more galaxies, the observed stellar mass function in Fig. 5-2 should be considered a lower limit.

Even with these caveats, there is a dramatic disagreement between the predicted and observed stellar mass functions. All AM models greatly overestimate the number of satellites with $M_* > 3 \times 10^3\,\mathrm{M}_\odot$, or more importantly, with $M_* > 10^4\,\mathrm{M}_\odot$, a mass range where few, if any, new satellites are likely to be discovered. Following assumptions in Bechtol et al. (2015), a galaxy with stellar mass of $M_* = 10^4\,\mathrm{M}_\odot$ corresponds to $M_\mathrm{V} \sim -5.1$, which in the year 1 DES survey would be expected to be found with high efficiency out to 300 kpc. To remain undetected, a satellite of this size would need a half-light radius larger than ~ 300 kpc, making the surface brightness comparable to or lower than that of the lowest surface brightness MW satellites (Torrealba et al., 2016a). Since the DES survey covered approximately half of the volume within 50 kpc of the LMC and found zero $M_* > 10^4\,\mathrm{M}_\odot$ satellites (Drlica-Wagner et al., 2016), it is unlikely the the remaining volume contains more than two such satellites.

In contrast to zero observed $M_* > 10^4\,\mathrm{M}_\odot$ satellites, the Brook model predicts ~ 4, the GK16 model ~ 8, and the Behroozi model ~ 12. Additionally, all models predict far too few satellites in the interval $10^2 < M_* < 3 \times 10^3\,\mathrm{M}_\odot$. Once again, the Behroozi model is in particularly strong disagreement, a fact not surprising in light of its shortcomings in predicting satellite galaxies already discussed in Garrison-Kimmel et al. (2014b) and Dooley et al. (2016b). Quantifying the disagreement in terms of random chance, the GK16 model predicts 7.9 ± 3.0 satellites with $M_* > 10^4$, and a 0.04% chance of zero galaxies. For the Brook model, odds are improved, but only to 1.5%.

There is some evidence that this subvolume happens to have fewer satellites with

$M_* > 10^4\,M_\odot$ by chance. A total of 8 such satellites (excluding the LMC and SMC) are known within the MW out to a galactic distance of 100 kpc. If they were isotropically distributed and follow the radial distribution from our model, 1.4 satellites would be expected in the 50 kpc radius volume under consideration. An additional easing of tensions could be made if the Canis Major overdensity ($M_* \approx 4.5 \times 10^7\,M_\odot$) is in fact a dwarf galaxy since it is within 50 kpc of the LMC. However, too much contention exists regarding whether it is (e.g., Martin et al., 2004; Bellazzini et al., 2004; Martínez-Delgado et al., 2005; Bellazzini et al., 2006; de Jong et al., 2007), or is not (e.g., López-Corredoira 2006; Moitinho et al. 2006; Momany et al. 2006; Rocha-Pinto et al. 2006; for a recent summary of this debate, see Yanny and Newberg 2016) a galaxy to include it in our sample.

5.3.1.2 Reconciling theory and observations?

Since we are analyzing a range satellites far fainter than those used to calibrate any AM model, it is perhaps not surprising that our predictions do not agree with observations. We therefore explore a range of possible alterations to our model which improve alignment with the data, and assess the plausibility of each.

To make discussions easier, we define satellites with $M_* > 10^4$ as "large UFDs," and satellites with $10^2 < M_* < 3 \times 10^3$ as "small UFDs." In these terms, the problem with our predicted satellite abundance is a matter of too many large UFDs, and too few small UFDs.

There are several ways in which the number of predicted large UFDs can be reduced which we identify as important to consider:

1. Lower mass MW, LMC and/or SMC.
2. Original LMC and SMC satellites have migrated to larger distances.
3. Less centrally concentrated radial distribution of satellites.
4. Tidal stripping.
5. Steeper $M_* - M_{\rm halo}$ relationship.

Each of these options simultaneously exacerbate the problem of predicting too few small UFDs. The predicted number of small UFDs can be increased without substantially increasing the predicted number of large UFDs in the following three ways:

1. Reionization occurs later.

2. The halo size threshold needed for star formation before reionization is reduced.

3. The $M_* - M_{\rm halo}$ relationship deviates from a power law. It is "bent" near a stellar mass of $10^3\,{\rm M}_\odot$ to have a more flat slope.

Reduced MW/LMC/SMC Mass

Any reduction in the mass used to model the MW, LMC, or SMC would decrease the predicted number of satellites at all stellar mass scales. However, this can not by itself be an explanation for the over-prediction of large UFDs. In the extreme event that the MW, LMC, and SMC are all one-half the total halo mass we use, the number of large UFDs predicted for the LMC subvolume in the GK16 model decreases 30% from 7.9 to 5.6, still a large statistical discrepancy from zero.

LMC/SMC satellite migration

We have so far assumed that the number of satellites within 50 kpc of the LMC and SMC at infall is equal to the number within 50 kpc today. Especially if the LMC/SMC pair is on its second or third pericentric passage around the MW, the positions of the original Magellanic satellites would be changed by the MW gravitational potential, becoming less concentrated in the immediate vicinity of the LMC and SMC. This would reduce the predicted number of LMC and SMC satellites, but would once again not be sufficient by itself to explain the over-prediction of large UFDs. If the number of LMC and SMC satellites within 50 kpc of the LMC are reduced by one half due to satellite migration, the baseline GK16 model would still predict 5.7 large UFDs.

Radial Distribution

A radial distribution that is less centrally concentrated would reduce the predicted contribution of satellites from all three of the MW, LMC, and SMC. The uncertainty in the mean of the normalized radial distribution of luminous satellites that we determine from the *Caterpillar* halos leads to a 1σ uncertainty of 28% in the number of satellites within 50 kpc of the LMC, regardless of satellite mass range. A reduction by this amount changes the GK16 prediction of large UFDs from 7.9 to 5.7, still too high to be statistically consistent with zero.

Tidal Stripping

Tidal stripping can reduce the stellar mass of predicted large UFDs, helping to alleviate tension with observations by shifting the predicted stellar mass function curve to the left. The fact that most UFDs fall on the same mass-metallicity relationship as larger galaxies (Kirby et al., 2013b) suggests that most UFDs are not tidally stripped. However, there is evidence of possibly significant tidal stripping in the UFDs Segue II (Kirby et al., 2013a), Hercules (Roderick et al., 2015; Küpper et al., 2016), Leo V (Collins et al., 2016), and Tucana III (Drlica-Wagner et al., 2015; Simon et al., 2016). Of those, only Tucana III is in the sample of satellites within 50 kpc of the LMC. However, most of the satellites in the sample have not yet been studied in follow-up campaigns to determine what levels of tidal stripping they may have undergone. It is therefore plausible that more satellites will show evidence of tidal stripping.

In the top panel of Fig. 5-3 we show the original GK16 stellar mass function prediction and the prediction if all satellites are stripped. Only a massive amount of tidal stripping, $> 95\%$ of stellar mass stripped, could explain the excess of large UFDs by itself. This would reduce the number of large UFDs from ~ 8 to ~ 2. Using particle tagging in the *Caterpillar* simulations, we conduct a quick estimate of the fraction of stars that could be stripped. We select satellites which form stars before reionization and survive to $z = 0$ with the stellar mass of UFDs. We then tag their 2% most bound particles at the time of peak mass, and assume no star formation

proceeds thereafter. This is roughly consistent with their classification as reionization fossils, in which $> 70\%$ of their stars are estimated to have formed before reionization (Brown et al., 2012, 2014a,b). The satellites which are near 50 kpc from the MW at $z = 0$, consistent with the mean distance of the observed LMC vicinity satellites to the MW, have on average 30% of their stars stripped. We include a model of 30% stellar mass stripping in the top panel of Fig. 5-3 and find it has little quantitative effect on the predicted number of satellites at any mass scale.

Our tagging scheme ignores the complexities of satellite orbits and additional tides due to the LMC/SMC system, and the gravitational potential of the MW disk, both of which would increase the fraction of stellar mass stripped. Additionally, the subhalos in our simulations have cuspy density profiles, whereas if UFD satellites are cored (for instance due to self-interacting dark matter (Dooley et al., 2016a)) they would be much more susceptible to tidal stripping (Peñarrubia et al., 2010; Zolotov et al., 2012; Brooks et al., 2013). Even so, we find it highly implausible that $\geq 95\%$ of stellar mass is stripped on average in the LMC vicinity satellite galaxies without evidence of more tidal tails or strong deviation from the mass-metallicity relation.

Steeper $M_* - M_{\rm halo}$ relationship

Lastly, we consider a steeper $M_* - M_{\rm halo}$ relationship as part of the solution to over-predicting large UFDs. For stellar masses with $M_* < 10^8\,{\rm M}_\odot$, $M_* \propto M_{\rm halo}^{\alpha}$ in all AM models considered. Already from comparing the Brook AM model ($\alpha = 3.1$) to the GK16 model ($\alpha = 1.97$) in Fig. 5-2, it is apparent that a steeper $M_* - M_{\rm halo}$ relationship reduces the predicted number of low stellar mass satellites. We can easily study this using the GK16 model because it allows for a range of possible $M_* - M_{\rm halo}$ slopes that all are consistent with the classical dwarf scale MW and M31 satellites. In the model, a steeper slope is paired with higher scatter in the $M_* - M_{\rm halo}$ relationship so that the reduction of classical dwarf-sized galaxies due to the steeper slope is compensated by an increase in the number of low mass dark matter halos that upscatter to host high stellar mass galaxies.

In the upper panel of Fig. 5-4, we show our original choice of $\alpha = 1.97$ along with

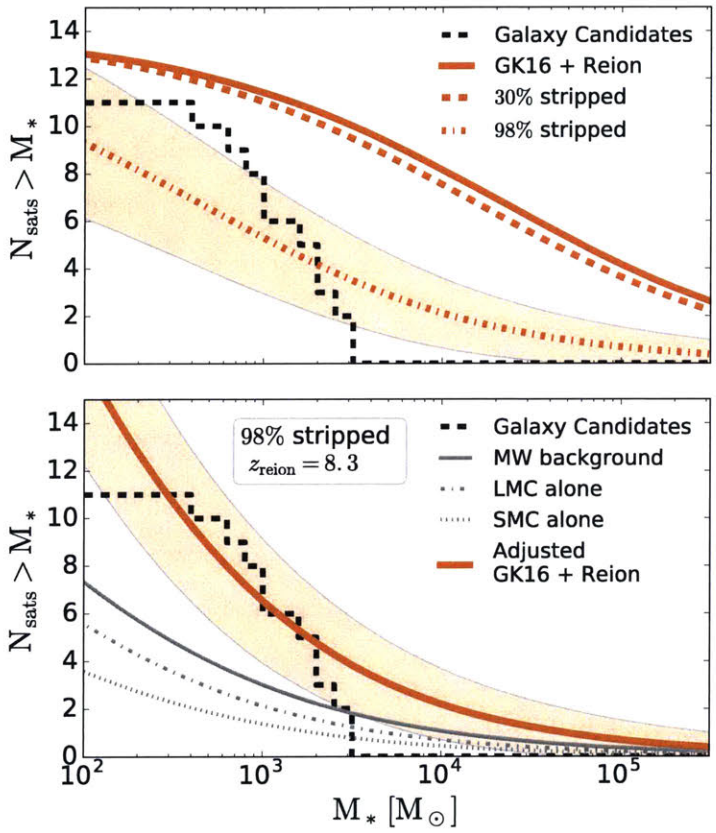

Figure 5-3 *Upper panel:* Predictions for the stellar mass function of satellites within 50 kpc of the LMC according to the GK16 model with 0%, 30%, and 98% of stellar mass stripped in satellites. Only with extreme amounts of stripping are the number of predicted satellites with $M_* > 10^4\,M_\odot$ diminished enough to be within 2σ of the number of observed satellites, zero. One sigma variation is shown around the 98% stripped model with a shaded band. *Lower panel:* Same prediction of the GK16 model with 98% of stellar mass stripped except with reionization shifted from $z = 13.3$ to $z = 8.3$. A similar result could be achieved by lowering the halo size threshold needed for the first galaxies to form.

values of $\alpha = 2.43$ and 3.31 in the GK16 model. The steepest slope greatly diminishes the number of large UFDs, down to 2.6. As with other possible solutions, however, it also exacerbates the problem of predicting too few small UFDs.

We elaborate on possible solutions to predicting too few small UFDs in the next three subsections.

Reionization occurs later

Delaying the start of reionization allows for more halos to grow large enough to form stars before reionization suppresses star formation. We illustrate the effect of delaying reionization in Fig. 5-3. In the upper panel, the GK16 model with 98% of stellar mass stripped predicts 9 satellites with $M_* > 10^2\,\mathrm{M}_\odot$, fewer than the 11 observed and the 11+ expected after completion of the MagLiteS and DES surveys. In the lower panel, shifting reionization from $z = 13.3$ to $z = 8.3$ increases the number of predicted small UFDs, bringing the prediction into better alignment with observations. Similarly, in the upper panel of Fig. 5-4, the GK16 model predicts too few small UFDs, especially when the $M_* - M_{\mathrm{halo}}$ relationship slope is 3.31. In the lower panel, we show again how delaying reionization to $z = 8.3$ increases the predicted number of small UFDs. It does, however, also slightly increase the number of large UFDs, making that discrepancy worse.

Recent analysis from the Planck collaboration estimates that reionization occured between $z = 7.8$ and $z = 8.8$, and that less than 10% of hydrogen in the Universe was ionized before $z = 10$ (Planck Collaboration et al., 2016a). This suggests that reionization occuring between $z = 15$ and 11.5 as in Barber et al. (2014) is too early, and that shifting reionization to a lower redshift may be part of the solution to predicting more small UFDs.

Star formation begins in smaller halos

Similar to delaying reionization, lowering the size threshold needed for halos to form stars increases the fraction of halos that host luminous galaxies. In our model, this

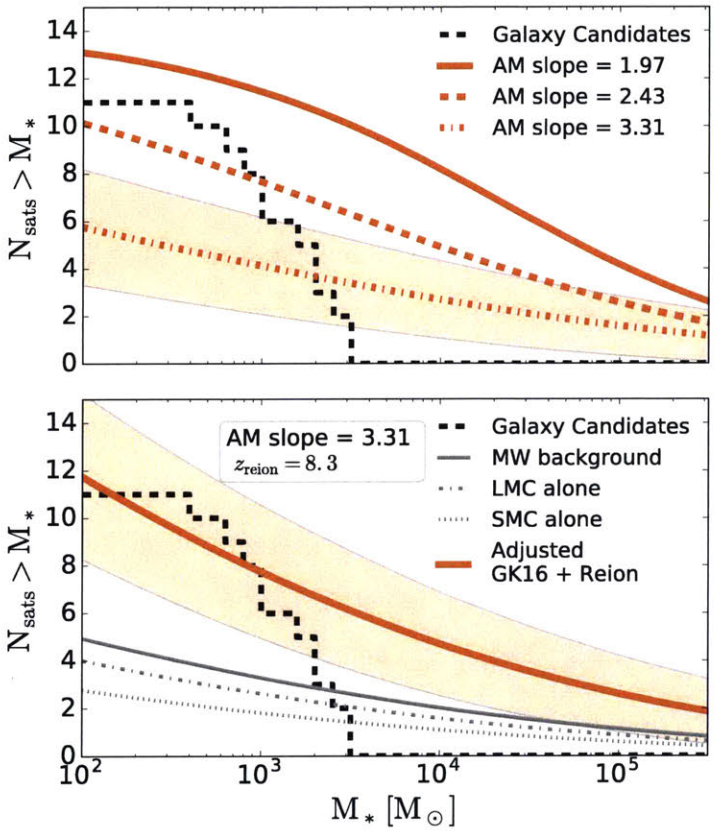

Figure 5-4 *Upper panel:* Predictions for the stellar mass function of satellites within 50 kpc of the LMC according to the GK16 model with an $M_* - M_{\rm halo}$ logarithmic slope of 1.97, 2.43, and 3.31. For reference, the Book model has a slope of 3.1, but less scatter in the $M_* - M_{\rm halo}$ relationship. One sigma variation is shown around the 3.31 slope model with a shaded band. *Lower panel:* Same prediction of the GK16 model with a logarithmic slope of 3.31 except with reionization shifted from $z = 13.3$ to $z = 8.3$. A similar result could be achieved by lowering the halo size threshold needed for the first galaxies to form.

means reducing $v_{\rm max}^{\rm pre}$. The resultant effect of increasing the number of predicted low mass satellites with $M_* < 10^5$ is nearly identical and therefore degenerate with that of lowering $z_{\rm reion}$. Having to lower $v_{\rm max}^{\rm pre}$ would be natural if UFDs formed in H_2 cooling minihalos before reionization, as hypothesized by Salvadori and Ferrara (2009) and Kirby et al. (2013a), rather than forming in larger atomic line cooling halos (Bromm and Yoshida, 2011; Power et al., 2014).

Bent $M_* - M_{\rm halo}$ relationship

Increasing the number of small UFDs could also be achieved by introducing a "bend" in the $M_* - M_{\rm halo}$ relationship, as shown in Fig. 5-5. This would be similar in form to that proposed in Sawala et al. (2015) (which uses a mass definition of $M_{\rm halo} = M_{200}^{\rm peak}$), except instead of the $M_* - M_{\rm halo}$ relationship flattening out near $M_* = 10^5\,{\rm M}_\odot$, it would have to flatten our near $M_* = 10^3\,{\rm M}_\odot$. In fact, the AM model suggested by Sawala et al. (2015) would only exacerbate the problem of predicting too many large UFDs. A bend can only increase the number of predicted small UFDs without greatly increasing the number of predicted large UFDs if the scatter in the $M_* - M_{\rm halo}$ relationship is sufficiently small. In Fig. 5-5, the scatter is tuned to a constant ± 0.4 dex to best predict the observed stellar mass function (with some room for more discoveries) as shown in the lower panel. If the scatter is as large as that predicted in the baseline GK16 model, ± 0.9 dex at $M_{\rm vir}^{\rm peak} = 10^8\,{\rm M}_\odot$, it would once again predict too many large UFDs at a value of 4.1. A bend at $M_* = 10^3\,{\rm M}_\odot$ with little scatter would need to be physically justified. It could, for instance, result from a minimum threshold of stellar mass being created in any single pre-reionization star forming event.

The multitude of ways in which our prediction for the stellar mass function of LMC vicinity satellites can be adjusted leads to innumerable possible solutions when combinations of each effect are considered. Furthermore, observations are not even complete, and the LMC vicinity represents just a single measurement of a satellite stellar

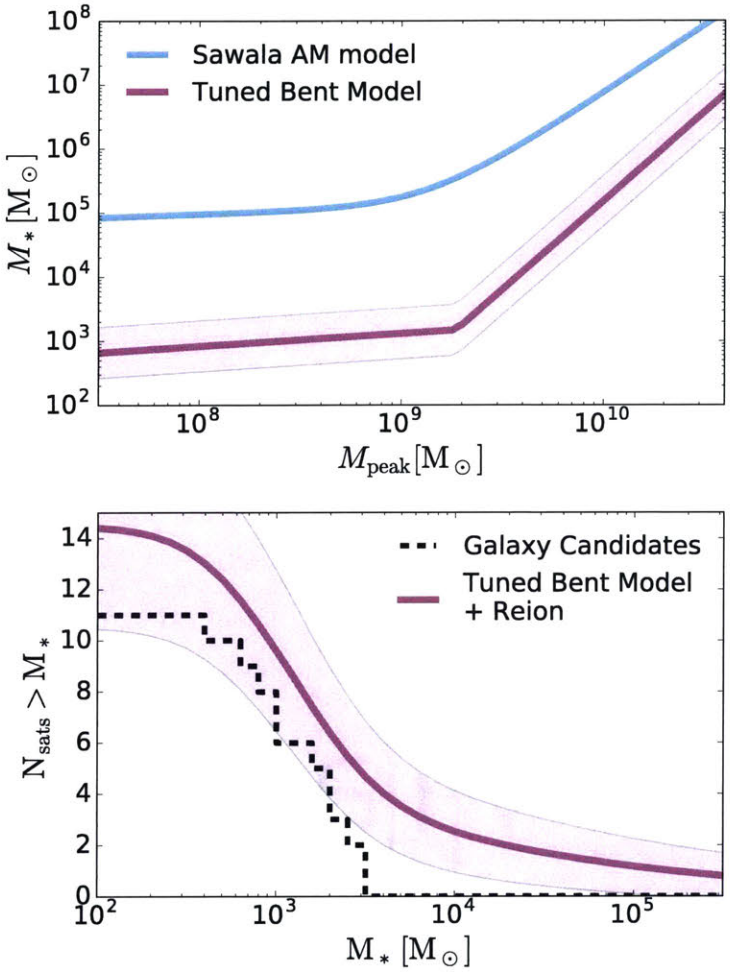

Figure 5-5 Upper Panel: $M_* - M_{\mathrm{halo}}$ relationships with a bend introduced. The Sawala AM model (Sawala et al., 2015) derived from hydrodynamic simulations includes a flattening of the relationship for $M_* \lesssim 10^5\,\mathrm{M}_\odot$. A similar flattening added to the GK16 model with a steep logarithmic slope of 2.77, except at $M_* \approx 10^3\,\mathrm{M}_\odot$, is one of multiple ways in which to predict few satellites within 50 kpc of the LMC that have $M_* > 10^4\,\mathrm{M}_\odot$, but many with $M_* < 3 \times 10^3\,\mathrm{M}_\odot$, as is required to best match observed satellites. However, it is also necessary that the scatter in the $M_* - M_{\mathrm{halo}}$ relationship be small in the bend (shown here as ± 0.4 dex), which is inconsistent with that of the GK16 model. Lower Panel: Prediction of the stellar mass function of satellites within 50 kpc of the LMC according to the bent $M_* - M_{\mathrm{halo}}$ relationship from the upper panel, and our baseline reionization model. A bend with little scatter is one way to create a steep rise in the mass function around $M_* = 10^3\,\mathrm{M}_\odot$, as is found with observed satellite galaxies.

mass function, subject to large variance from counting statistics. We therefore do not propose any single favored model. Instead, we hope identifying the discrepancy and ways to alter predictions will engender future efforts to better explain the issue. One additional valuable constraint that will also improve with future observations is the stellar mass function of UFDs in the full MW volume. Just as adjustments to our models change predictions for the LMC vicinity, they also change predictions for the MW.

5.3.1.3 Ratio of MW:LMC:SMC satellites

In spite of shortcomings in predicting the LMC vicinity stellar mass function, the *ratio* of satellites between the MW, LMC and SMC is more certain. In the GK16 model, the ratio of satellites within 50 kpc of the LMC is 46% from the MW, 33% from the LMC, and 21% from the SMC. The ratio does not change with adjustments to tidal stripping, reionization, nor the slope of the $M_* - M_{\rm halo}$ relationship. It also does not change with the range of stellar mass considered. In the Brook and Moster models, the ratios are nearly the same at 49% from the MW, 32% from the LMC, and 19% from the SMC. More concretely, of the 12 UFDs near the LMC, the most consistent ratio of integers would be 6 or 5 from the MW, 4 from the LMC, and 2 or 3 from the SMC. This is broadly consistent with Jethwa et al. (2016) finding 6/9 of the UFDs considered within 50 kpc of the LMC having a $> 70\%$ chance of association with the Magellanic Clouds, and Sales et al. (2016) finding 6/10 of them possibly consistent with the Clouds. The fraction of LMC and SMC satellites would decline if their original satellites at infall spread out to larger distances. Changes in the halo mass used for the three galaxies would also adjust the ratio predicted.

Considering the entire virial volume of the MW, we expect that $15 - 25\%$ of all MW satellites, regardless of the mass range considered (so long as it is less than that of the SMC), originated within the virial radius of the LMC or SMC before their infall. The Moster and Brook models predict closer to 15%, and the GK models closer to 25%. From the LMC alone, we expect $10 - 15\%$. This range is higher than the 5% LMC contribution predicted by Sales et al. (2016), who used a lower mass

LMC, and within the $1-25\%$ range suggested by Deason et al. (2015). The combined Magellanic contribution we predict is less than the 33% predicted by Jethwa et al. (2016).

5.3.2 Satellites in isolated LMC analogs

Isolated LMC-sized galaxies offer a much cleaner way to probe satellite populations. Instead of three overlapping hosts of satellite galaxies, there is just a single host. Consequently, satellite-host membership is unambiguous, there is no concern about the infall time of the LMC and SMC, and sub-subhalo orbits do not need to be considered. Since isolated LMC-sized galaxies are much less massive than the MW, and satellites can be considered out to the full virial radius beyond 50 kpc, the strength of tidal stripping is greatly diminished. Also, uncertainty in the radial distribution of galaxies is decreased when considering volumes near the full virial radius. Lastly, measuring many host-satellite population pairs can be used to help control for halo-to-halo variance, and statistical uncertainty in the mass of the hosts. Of the ways suggested to modify the predicted LMC-vicinity satellite stellar mass function presented in Section 5.3.1.1, only the $M_* - M_{\rm halo}$ relationship and reionization remain significant. Therefore, searching for satellites in isolated LMC-sized galaxies can provide valuable information to help constrain models, and to determine how typical the satellite population near our own LMC is.

The majority of known isolated LMC-sized field galaxies are more than 2 Mpc away, as cataloged in Table 5.1. At that distance, current sensitivities are likely to only find galaxies with $M_* \gtrsim 10^5\,{\rm M}_\odot$. For example, the $M_{\rm V} = -7.7$ ($M_* \sim 10^5\,{\rm M}_\odot$) dwarf at $D \sim 3.2$ Mpc discovered by Carlin et al. (2016) had only $\sim 25-30$ stars resolved to ~ 1.5 mags below the RGB tip, which is near the minimum number of resolved stars needed to algorithmically detect galaxies (Walsh et al., 2009). Reionization has little impact on galaxies with $M_* > 10^5\,{\rm M}_\odot$, leaving the AM model and Poisson noise as the dominant sources of uncertainty. As such, searching for satellites at current detection thresholds is most helpful in constraining AM models and the slope of the $M_* - M_{\rm halo}$ relationship. We focus on AM model constraints with $10^5\,{\rm M}_\odot$ satellites

here, and discuss the science return on probing smaller satellites later.

Since isolated field galaxies span a range of masses, we show in Fig. 5-6 the dependence of satellite abundance on host halo stellar mass. In the upper panel we include all satellites with $M_* > 10^5\,\mathrm{M}_\odot$ within the virial volume, while in the lower panel we limit the volume to a sphere of radius 100 kpc. As seen in Fig. 5-1, within 1σ about the mean of the GK16 AM model we predict $0.7 - 3.7$ and $1.7 - 5.7$ such satellites around SMC and LMC-sized galaxies respectively. For the Brook model, it drops to $0 - 1.9$ and $0.4 - 3.1$. If the lack of any satellites with $M_* > 3 \times 10^3\,\mathrm{M}_\odot$ in the LMC vicinity is reflective of typical satellite abundances, the predictions of the Brook model may be favored.

The SHMF is approximately directly proportional to the host halo mass, and $M_{\mathrm{halo}} \propto M_*^{0.42}$ in the vicinity of both the LMC's and SMC's stellar masses using the GK16 model. Thus doubling the stellar mass of a host results in a halo with 1.34 times the number of satellites in any satellite mass interval. Much of this increase arises from a larger virial volume being considered. When fixing the volume to 100 kpc, doubling the halo mass results in a 1.24 factor of increase in satellites. For the Moster and Brook models, the dependence is slightly steeper with doubling the host stellar mass resulting in 1.4 and 1.3 times the satellite abundance within the virial volume and 100 kpc. Although we only show functions for satellites with $M_* > 10^5\,\mathrm{M}_\odot$, the shapes of the curves are identical for different stellar mass ranges. Thus the plot shown can be multiplied by a ratio of the abundance of satellites with $M_* > 10^5\,\mathrm{M}_\odot$ to any other range, as can be deduced from Fig. 5-1.

When observing distant galaxies, the geometry is dictated by a line of sight, not spherical volumes as considered so far. We give a sense of the cumulative radial distribution of satellites as a function of observed radius perpendicular to the line of sight in physical units for an LMC-sized isolated galaxy in Fig. 5-7. Using results from Dooley et al. (2016b), we predict how many satellites are within a line of sight with a circular aperture, counting satellites in the axis of the line of sight out to the splashback radius (More et al., 2015) of $1.5\times R_{\mathrm{vir}}$. The abundance of satellites is concentrated towards the center, with an observed area out to $0.5\times R_{\mathrm{vir}}$ containing

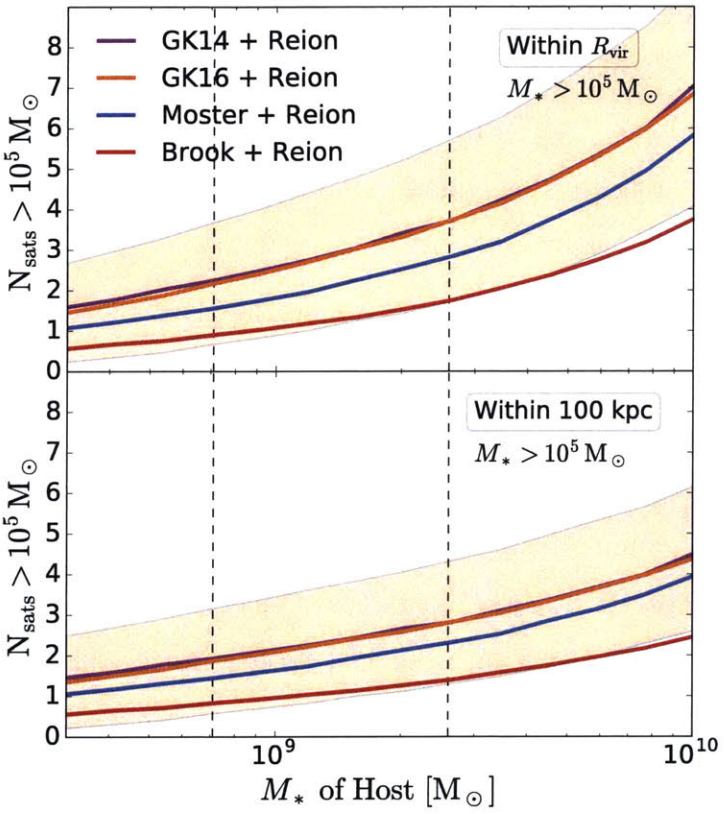

Figure 5-6 Mean number of satellites with stellar mass above $10^5\,M_\odot$ as a function of a host halo's total stellar mass, M_*, for several AM models. The upper panel shows the number of satellites within the halo's full virial radius. The lower panel shows the number of satellites within 100 kpc. Within a fixed volume, the number of satellites is less sensitive to the host's total mass. The shape of the curves are the same up to a multiplicative factor for any other stellar mass range considered. 1σ halo-to-halo variation is shown for the GK16 model with a shaded band. Vertical lines correspond to the stellar masses of the SMC and the LMC.

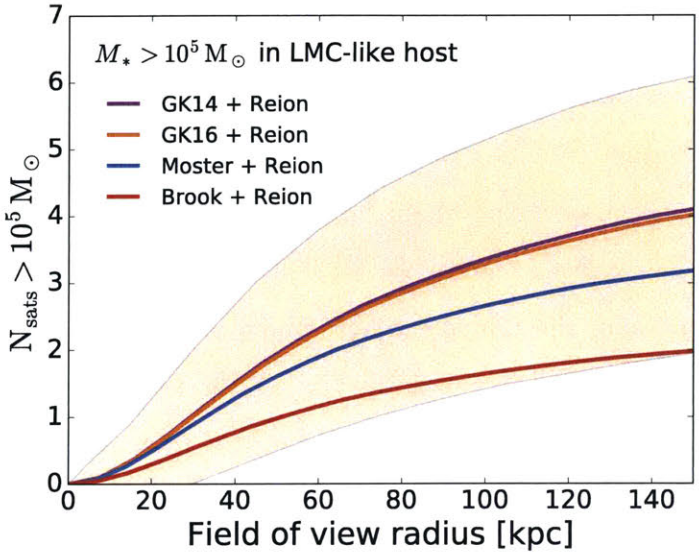

Figure 5-7 Mean number of satellites with $M_* > 10^5\,\mathrm{M}_\odot$ as a function of observed radius perpendicular to the line of sight assuming a circular field of view for an LMC-sized galaxy. We count all halos within $1.5\times R_{\rm vir}$ of the host in the line of sight direction. 1σ halo-to-halo variation is shown for the GK16 model with a shaded band.

70% of the total number of satellites within a field of view encompassing the full $R_{\rm vir}$. In our LMC-sized galaxy example, observing out to just 50 kpc is enough to expect 1 satellite with $M_* > 10^5\,\mathrm{M}_\odot$ in all abundance matching models. Complete information on computing the radial dependence for any host galaxy can be found in Dooley et al. (2016b).

In Table 5.1, we catalog known galaxies with stellar masses on the order of $10^9 - 10^{10}\,\mathrm{M}_\odot$ that are 2−8 Mpc away, have Galactic latitudes $|b| > 25°$, have low extinction ($E(B-V) < 0.15$), and are categorized as relatively isolated in the Karachentsev et al. catalog. For each galaxy, we indicate the mean number of expected satellites with $M_* > 10^5\,\mathrm{M}_\odot$ within the virial radius. We also list the 20^{th} and 80^{th} percentile number of satellites, the inferred virial radius of the host, and the mean number of satellites within one pointing of a 1.5° and 2.2° diameter field of view centered on the host. These angular fields of view correspond to the Hyper SuprimeCam (Subaru 8.2m) and DECam (CTIO Blanco 4m) imagers, the largest current imagers capable of resolving the stellar halos of the galaxies in Table 5.1. We focus on the Brook and GK16

models since the GK14 model is nearly identical to the GK16 model, and the Moster model predictions fall in between those of the Brook and GK16 models. The galaxies are sufficiently distant, and have a large enough predicted satellite population that a single pointing of either camera at each galaxy covers an area expected to contain a mean of at least one observable satellite in most cases. Some of the more distant targets are expected to have as many as four satellites within a single pointing of a 2.2° diameter camera in the GK16 model. While a survey of many host galaxies complete to $M_* = 10^5\,\mathrm{M}_\odot$ satellites is needed to test of abundance matching models, satellites may be abundant enough that targeting just a few galaxies will reveal new satellites to study in a different environment than the MW.

We note that if the satellite abundance of LMC analogs mimics the true LMC, we expect fewer $10^5\,\mathrm{M}_\odot$ galaxies orbiting these systems than predicted by the Brook model. Discovering very few $10^5\,\mathrm{M}_\odot$ satellites in these systems would favor a steep $M_* - M_{\mathrm{halo}}$ relationship.

If continued improvement in imagers enables the detection of satellites down to $M_* > 10^4\,\mathrm{M}_\odot$, we could gain valuable information on how reionization effects galaxy formation, and in turn help unravel the "small UFD" piece of the LMC-vicinity puzzle. This is especially true if the $M_* - M_{\mathrm{halo}}$ relationship for $M_* > 10^5\,\mathrm{M}_\odot$ is better constrained, leaving reionization or a bend in the $M_* - M_{\mathrm{halo}}$ relationship as the most significant poorly understood controls on the population of lower mass satellites in isolated LMC-sized hosts. For instance, the ratio of $M_* > 10^4\,\mathrm{M}_\odot$ to $M_* > 10^5\,\mathrm{M}_\odot$ satellites in the GK16 model is 2.0 using our baseline reionization model, or 2.6 using $z_{\mathrm{reion}} = 9.3$. Measuring the ratio of $M_* > 10^5\,\mathrm{M}_\odot$ satellites to lower thresholds is a valuable tool for studying the effects of reionization.

Table 5.1 Mean number of observable satellites with $M_* > 10^5\,M_\odot$ around isolated field dwarf galaxies

Name	M_*	$D_\odot$	**Brook**					**GK16**				
	$[10^8\,M_\odot]$	[Mpc]	$\bar{N}_{\rm lum}$	20/80%	$\bar{N}_{\rm fov}^{1.5}$	$\bar{N}_{\rm fov}^{2.2}$	$R_{\rm vir}$	$\bar{N}_{\rm lum}$	20/80%	$\bar{N}_{\rm fov}^{1.5}$	$\bar{N}_{\rm fov}^{2.2}$	$R_{\rm vir}$
SMC	7.1	···	0.89	0/2	···	···	117	2.16	1/3	···	···	131
NGC5023	10.2	6.05	1.09	0/2	0.97	1.18	124	2.53	1/4	2.11	2.61	138
NGC5585	10.7	5.7	1.13	0/2	0.96	1.18	125	2.60	1/4	2.07	2.58	139
NGC1800	11	8	1.13	0/2	1.16	1.34	126	2.62	1/4	2.54	2.99	139
ESO383-087	11.2	3.19	1.14	0/2	0.60	0.85	126	2.64	1/4	1.26	1.81	140
NGC5253	12	3.44	1.19	0/2	0.67	0.92	127	2.71	1/4	1.39	1.96	141
NGC4204	13.2	8	1.24	0/2	1.25	1.46	129	2.81	1/4	2.69	3.18	143
HolmII	15.8	3.47	1.37	0/2	0.75	1.04	134	3.03	1/4	1.51	2.15	146
NGC4144	17.8	6.89	1.44	0/2	1.30	1.57	136	3.19	2/5	2.75	3.37	148
NGC0404	18.2	2.98	1.46	0/2	0.66	0.96	137	3.20	2/5	1.31	1.96	149
IC5052	18.6	5.5	1.49	0/2	1.16	1.46	137	3.23	2/5	2.39	3.04	149
IC2574	21.4	3.93	1.60	0/3	0.94	1.27	140	3.42	2/5	1.85	2.58	152
NGC0045	21.4	6.64	1.59	0/3	1.38	1.69	140	3.42	2/5	2.83	3.50	152
NGC4136	24.6	7.9	1.71	1/3	1.62	1.92	144	3.64	2/5	3.31	3.98	155
NGC0300	25.7	2.09	1.76	1/3	0.47	0.77	145	3.69	2/5	0.88	1.48	156
LMC	26	···	1.76	1/3	···	···	145	3.70	2/5	···	···	156
NGC7713	26.9	7.8	1.79	1/3	1.67	1.99	146	3.75	2/5	3.37	4.06	157

Mean number of satellites with $M^* > 10^5\,M_\odot$ expected to exist within the virial volume of known isolated field galaxies as predicted with the Brook and GK16 models. The $20^{\rm th}$ and $80^{\rm th}$ percentile of the satellite abundance distributions are listed, indicating a > 80% chance of at least one satellite in even the smallest galaxies considered. $\bar{N}_{\rm fov}^{1.5}$ and $\bar{N}_{\rm fov}^{2.2}$ indicate the mean number of satellites within a field of view of diameter 1.5° and 2.2° respectively. These are the apertures of Hyper SuprimeCam and DECam. Also shown is the inferred virial radius of each galaxy.

Table 5.2 *continued:* Mean number of observable satellites with $M_* > 10^5\,M_\odot$ around isolated field dwarf galaxies

Name	M_*	$D_\odot$	**Brook**					**GK16**				
	$[10^8\,M_\odot]$	[Mpc]	$\bar{N}_{\rm lum}$	20/80%	$\bar{N}^{1.5}_{\rm fov}$	$\bar{N}^{2.2}_{\rm fov}$	$R_{\rm vir}$	$\bar{N}_{\rm lum}$	20/80%	$\bar{N}^{1.5}_{\rm fov}$	$\bar{N}^{2.2}_{\rm fov}$	$R_{\rm vir}$
NGC4242	29.5	7.9	1.88	1/3	1.75	2.09	148	3.91	2/6	3.51	4.24	159
NGC4395	29.5	4.76	1.89	1/3	1.25	1.64	148	3.89	2/6	2.43	3.24	159
NGC0024	30.2	7.31	1.91	1/3	1.70	2.05	149	3.93	2/6	3.36	4.12	159
NGC0055	30.2	2.11	1.90	1/3	0.49	0.81	149	3.92	2/6	0.92	1.55	159
NGC0247	31.6	3.72	1.96	1/3	1.02	1.43	150	4.02	2/6	1.96	2.80	160
NGC3239	33.1	7.9	2.00	1/3	1.84	2.21	151	4.08	2/6	3.62	4.39	161
NGC4244	33.1	4.31	2.02	1/3	1.20	1.62	151	4.10	2/6	2.30	3.17	161
NGC1313	37.1	4.31	2.13	1/3	1.24	1.69	154	4.30	2/6	2.38	3.28	164
NGC4656	38.9	5.4	2.19	1/3	1.55	1.99	156	4.35	2/6	2.93	3.82	165
NGC4236	40.7	4.41	2.25	1/3	1.32	1.80	157	4.43	3/6	2.47	3.41	166
IC5332	41.7	7.8	2.25	1/3	2.01	2.43	158	4.48	3/6	3.88	4.74	166
NGC4449	47.9	4.27	2.47	1/4	1.37	1.89	162	4.77	3/7	2.52	3.53	170
NGC4605	50.1	5.55	2.50	1/4	1.74	2.25	163	4.87	3/7	3.26	4.26	171
NGC5102	50.1	3.66	2.51	1/4	1.18	1.70	163	4.88	3/7	2.17	3.17	171
NGC7793	50.1	3.63	2.52	1/4	1.17	1.69	163	4.88	3/7	2.14	3.15	171
NGC5068	60.3	5.45	2.79	1/4	1.86	2.43	169	5.30	3/7	3.41	4.50	176
NGC2403	72.4	3.19	3.08	2/5	1.13	1.73	174	5.82	4/8	2.03	3.16	181

Mean number of satellites with $M^* > 10^5\,M_\odot$ expected to exist within the virial volume of known isolated field galaxies as predicted with the Brook and GK16 models. The $20^{\rm th}$ and $80^{\rm th}$ percentile of the satellite abundance distributions are listed, indicating a $> 80\%$ chance of at least one satellite in even the smallest galaxies considered. $\bar{N}^{1.5}_{\rm fov}$ and $\bar{N}^{2.2}_{\rm fov}$ indicate the mean number of satellites within a field of view of diameter 1.5° and 2.2° respectively. These are the apertures of Hyper SuprimeCam and DECam. Also shown is the inferred virial radius of each galaxy.

5.4 Conclusions

Using abundance matching (AM) models, a model for reionization, and simulations from the *Caterpillar* suite, we predict the abundance of satellites around galaxies with stellar masses comparable to the SMC and LMC. We additionally predict the abundance of satellites within 50 kpc of the actual LMC, modelling the total population as a superposition of satellites originally belonging to the MW, LMC, and SMC. Independent of adjustments to our reionization model, the $M_* - M_{\rm halo}$ relationship for satellite galaxies, and the stellar mass range of satellites considered, we find the ratio of satellites in the 50 kpc LMC vicinity to be approximately 47% from the MW, 33% from the LMC, and 20% from the SMC. Moreover, within 300 kpc of the MW, we estimate $10 - 15\%$ of all satellites smaller than the SMC were accreted with the LMC, and $5 - 10\%$ with the SMC. These values are somewhat sensitive to the mass of the host galaxies used, which were $M_{\rm vir} = 1.4 \times 10^{12}\,{\rm M}_\odot$ for the MW, $\sim 2 \times 10^{11}$ for the LMC at infall, and $\sim 1 \times 10^{11}$ for the SMC at infall.

Curiously, we find that all twelve currently known satellite candidates within 50 kpc of the LMC have a stellar mass of $M_* \lesssim 3 \times 10^3\,{\rm M}_\odot$. This creates two significant discrepancies with our predicted satellite stellar mass functions:

1. We predict too many $M_* > 10^4\,{\rm M}_\odot$ satellites (large UFDs).

2. We predict too few $10^2 < M_* < 3 \times 10^3\,{\rm M}_\odot$ satellites (small UFDs).

For instance, combining our reionization model with the AM model from Garrison-Kimmel et al. (2016) (GK16), we predict ~ 8 large UFDs in the LMC vicinity which has only a 0.04% statistical chance of being consistent with the zero currently known. The same model predicts just ~ 3 small UFDs, much less than the 11 currently known. Furthermore, the problem is only expected to get worse since the full volume under consideration has not yet been surveyed. The ongoing MagLiteS survey is expected to increase the number of known small UFDs near the LMC, but not increase the number of known large UFDs by more than one or two.

We explore a variety of model adjustments that would produce a prediction more consistent with observations. However, all options have limitations, and combined

with incomplete observations, the parameter space of possible adjustments is too large to select any preferred solution. The lack of any known large UFDs near the LMC could for instance be indicative of some combination of a steeper $M_* - M_{\rm halo}$ relationship, lower mass MW, LMC, and SMC, extreme amount of tidal stripping, a less centrally concentrated radial satellite distribution than we predict, and/or migration of original LMC/SMC satellites to farther distances since infall. The large abundance of small UFDs could be indicative of some combination of reionization occurring later (e.g. $z_{\rm reion} \approx 8$ instead of $z_{\rm reion} \approx 13$), a reduced $v_{\rm max}$ threshold for dark matter halos to first form galaxies (such as UFDs beginning as minihalos instead of atomic cooling halos), or a "bend" in the $M_* - M_{\rm halo}$ relationship at $M_* \approx 10^3\,{\rm M}_\odot$ (which could arise from a minimum amount of stellar mass being created in any luminous galaxy).

We also make predictions for the abundance of satellites in known isolated LMC-sized galaxies which fall between 2 and 8 Mpc away. We find that searching these galaxies for satellites is worthwhile not only because of the likelihood of discovering new satellites in different environments, but because it can provide valuable information to help constrain AM models. The stellar mass function of satellites near the LMC cannot be fully understood without a better grasp of the statistical fluctuations in satellite abundances around hosts. We find the target galaxies selected likely contain $1-6$ satellites with stellar mass $> 10^5\,{\rm M}_\odot$ within their virial volumes. Perhaps more importantly, using a single pointing of a 1.5° diameter field of view camera is sufficient to expect a mean of $1-3$ such satellites around most targets. If the number of satellites discovered is consistent with or less than predictions from the Brook AM model Brook et al. (2014), as suggested by the stellar mass function of satellites within 50 kpc of the LMC, it would favor a steep $M_* - M_{\rm halo}$ relationship. Reionization could not be part of the explanation since it makes little impact on the abundance of satellites with $M_* > 10^5\,{\rm M}_\odot$. Alternatively, if predictions from the GK16 model are more accurate, a steep $M_* - M_{\rm halo}$ relationship would no longer be a viable part of the explanation for the dearth of large UFDs near the LMC. If surveys are eventually able to discover galaxies down to $M_* = 10^4\,{\rm M}_\odot$ or lower, a better understanding of reionization and its ability to suppress star formation can

be achieved. If reionization occurs later, or the size threshold needed for galaxies to first form stars is lower, there would be an increase in the ratio of galaxies with $10^4 < M_* < 10^5\,\mathrm{M}_\odot$ to galaxies with $M_* > 10^5\,\mathrm{M}_\odot$.

The path forward towards understanding the LMC vicinity stellar mass function will require completion of the MagLiteS survey, further analysis of DES data for more galaxies, improved luminosity completeness limits on searches for satellites within the entire MW volume, and conducting searches for satellites around as many other hosts beyond the MW as possible. It will also require improving constraints on reionization and understanding how it influences low mass galaxy formation. Better constraints on the MW mass and orbital histories of the LMC and SMC will additionally help. In spite of many uncertainties, we find that the discrepancies between the predicted and observed LMC-vicinity stellar mass function are surprisingly large and may require significant effort to arrive at a satisfactory solution. We encourage others to join the effort.

Chapter 6

The *Caterpillar* Project: A Large Suite of Milky Way Sized Halos

The content of this chapter was published in The Astrophysical Journal *Volume 818, Issue 1 on February 2nd, 2016. It has an arXiv location of* `https://arxiv.org/abs/1509.01255v1`. *The authors are Brendan F. Griffen, Alexander P. Ji, Gregory A. Dooley, Facundo A. Gómez, Brian W. O'Shea, and Anna Frebel.*

Disclaimer

The majority of the writing and analysis in this chapter was conducted by Dr. Brendan Griffen. Alex Ji and I worked alongside Brendan from 2012 to 2016 to help make key decisions, obtain and manage computing resources, run the simulations, test for accuracy, and produce a diverse set of analysis tools. This crucially included running the ROCKSTAR halo finder and ROCKSTAR CONSISTENT-TREES codes, identifying and fixing several bugs in both of those codes, writing software to load all output information in a way to ease analysis, and writing a large assortment of generalized functions to analyze data from the raw simulation data, halo finders, and merger trees. I additionally contributed reviews and edits of the text in the paper.

Acknowledgements

BG would like to thank Paul Hsi for assistance with the compute cluster at MKI. He would also like to thank Phillip Zukin and Paul Torrey for helpful discussions. The authors thank Oliver Hahn for making the initial conditions code, MUSIC, publicly available. The authors also thank Volker Springel for making GADGET-2 publicly available and for providing a version of GADGET-3/GADGET-4 for our use. The authors thank Peter Behroozi for making ROCKSTAR and CONSISTENT-TREES publicly available and additionally thank him for technical support in modifying ROCKSTAR.

Support for this work was provided by XSEDE through the grants (TG-AST120022, TG-AST110038). BG and AF acknowledges support of the compute cluster of the Astrophysics Division which was built with support from the Kavli Investment Fund administered by the MIT Kavli Institute for Astrophysics and Space Research. GD acknowledges support by NSF Grant 1122374. BWO and FG were supported through the NSF Office of Cyberinfrastructure by grant PHY-0941373 and by the Michigan State University Institute for Cyber-Enabled Research (ICER). BWO was supported in part by by NSF grant PHY 08-22648 (Physics Frontiers Center/Joint Institute for Nuclear Astrophysics) and NSF Grant PHY-1430152 (JINA Center for the Evolution of the Elements). AF acknowledges support from the Silverman (1968) Family Career Development professorship.

Abstract

We present the largest number of Milky Way sized dark matter halos simulated at very high mass ($\sim 10^4$ $M_\odot$/particle) and temporal resolution ($\sim$5 Myrs/snapshot) done to date, quadrupling what is currently available in the literature. This initial suite consists of the first 24 halos of the *Caterpillar Project*[1] whose project goal of 60 – 70 halos will be made public when complete. We resolve $\sim$20,000 gravitationally bound subhalos within the virial radius of each host halo. Over the ranges set by our spatial resolution our convergence is excellent and improvements were made upon current state-of-the-art halo finders to better identify substructure at such high resolutions (e.g., on average we recover $\sim$4 subhalos in each host halo above 10^8 $M_\odot$ which would have otherwise not been found using conventional methods). For our relaxed halos, the inner profiles are reasonably fit by Einasto profiles ($\alpha = 0.169 \pm 0.023$) though this

[1]Project Website: `http://www.caterpillarproject.org`

depends on the relaxed nature and assembly history of a given halo. Averaging over all halos, the substructure mass fraction is $f_{m,subs} = 0.121 \pm 0.041$, and mass function slope is $\mathrm{d}N/\mathrm{d}M \propto M^{-1.88\pm0.10}$ though we find scatter in the normalizations for fixed halo mass due to more concentrated hosts having less subhalos at fixed subhalo mass. There are no biases stemming from Lagrangian volume selection as all Lagrangian volume types are included in our sample. Our detailed contamination study of 264 low resolution halos has resulted in obtaining very large and unprecedented, high-resolution regions around our host halos for our target resolution (sphere of radius $\sim 1.4 \pm 0.4$ Mpc) allowing for accurate studies of low mass dwarf galaxies at large galactocentric radii and the very first stellar systems at high redshift ($z \geq 10$).

6.1 Introduction

Under the current paradigm of structure formation (White and Rees, 1978) stellar halos of large galaxies such as the Milky Way are believed to be primarily formed as a result of the accumulation of tidal debris associated with ancient as well as recent and ongoing accretion events (Helmi 2008, Pillepich et al. 2015). In principle, the entire merger and star formation history of our Galaxy and its satellites can be probed with their stellar contents (i.e., the "fossil record"; Freeman and Bland-Hawthorn 2002) because this information is not only encoded in the dynamical distribution of the different Galactic components, but also in the stellar chemical abundance patterns (e.g., Font et al. 2006; Gómez et al. 2010).

To further map out the structure and composition of the various components of the Milky Way, large scale observational efforts are now underway. Several surveys such as RAVE (Steinmetz et al., 2006), SEGUE (Yanny et al., 2009), APOGEE (Majewski et al., 2010), LAMOST (Deng et al., 2012) and GALAH (Freeman, 2012) have collected medium-resolution spectroscopic data on some four million stars primarily in the Galactic disk and stellar halo. There are also ongoing large-scale photometric surveys such as Pan-STARRS (Kaiser et al., 2010) and SkyMapper Southern Sky Survey (Keller et al., 2013) mapping nearly the entire sky. Soon, the GAIA satellite (Perryman et al., 2001) will provide precise photometry and astrometry for another one billion stars.

Studies of individual metal-poor halo stars have long been used to establish prop-

erties of the Galactic halo, such as the metallicity distribution function, to learn about its history and evolution. More recently, the discoveries of the ultra-faint dwarf galaxies (with $L_{\rm tot} \leq 10^5\,L_{\odot}$) in the northern Sloan Digital Sky Survey (SDSS) and the southern Dark Energy Survey (DES) have shown them to be extremely metal-deficient systems which lack metal-rich stars with [Fe/H] $\gtrsim -1.0$. To some extent they can be considered counterparts to the most metal-poor halo stars. They extend the metallicity-luminosity relationship of the classical dwarf spheroidal galaxies down to $L_{\rm tot} \sim 10^3\,L_{\odot}$ (Kirby et al., 2008), and due to their relatively simple nature, they retain signatures of the earliest stages of chemical enrichment in their stellar population(s). Indeed, the chemical abundances of individual stars in the faintest galaxies suggest a close connection to metal-poor halo stars in the Galaxy (Frebel and Norris, 2015).

This comes at a time when there is still uncertainty over what role dwarf galaxies play in the assembly of old stellar halos because the true nature of the building blocks of large galaxies (e.g., Helmi and de Zeeuw 2000, Johnston et al. 2008, Gómez et al. 2010) are not yet fully understood. Nevertheless, observations of the, e.g., the Segue 1 ultra-faint dwarf suggest that these faintest satellites could be some of the the universe's first galaxies (presumably the building blocks) that survived until today (Frebel and Bromm, 2012; Frebel et al., 2014). They would thus be responsible for the Milky Way's oldest and most metal-poor stars.

This wealth of observational results offers unique opportunities to study galaxy assembly and evolution and will thus strongly inform our understanding of the formation of the Milky Way. Along with it, the current dark energy plus cold dark matter paradigm (ΛCDM) can be tested at the scales of the Milky Way and within the Local Group. But to fully unravel the Galaxy's past and properties, theoretical and statistical tools need to be in place to make efficient use of data.

For over three decades now, numerical simulations of structure formation have consistently increased in precision and physical realism (see Somerville and Davé 2015 for a review). Originally, they began as a way to study the evolution of simple N-body systems (e.g., merging galaxies; Aarseth 1963, Toomre and Toomre 1972, White

1978 and globular clusters; Hénon 1961) but with the advent of better processing power and more sophisticated codes (e.g., Springel 2010, Hopkins 2015, Bryan et al. 2014), N-body solvers are now fully coupled to hydrodynamic solvers allowing for a comprehensive treatment of the evolution of the visible Universe (e.g., Vogelsberger et al. 2014a, Schaye et al. 2015).

The most efficient method of studying volumes comparable to the Local Group whilst maintaining accurate large scale, low-frequency cosmological modes is via the *zoom-in* technique (Katz et al. 1994, Navarro and White 1994). This technique allows one to efficiently model a limited volume of the Universe at an extremely high resolution. Owing to the extreme dynamic range offered by such simulations, both the inside of extremely low mass, gravitationally bound satellite systems can be studied along side the hierarchical assembly of their host galaxy (e.g., Stadel et al. 2009). Gravity solvers which use hybrid tree-particle-mesh techniques (e.g., GADGET-2, Springel 2005) are ideally suited to carrying out such calculations on these scales. In addition to tailored codes for studying Milky Way sized halos, halo finders used for identifying substructure contained within them have also drastically improved over the past 30 years. Simple friends-of-friends (FoF) algorithms (e.g., Davis et al. 1985) have now evolved into parallel, fully hierarchical FoFs algorithms adopting six phase-space dimensions and one time dimension allowing shape-independent, and noise-reduced identification of substructure (Behroozi et al. 2013b). These tools are very robust methods for accurately identifying bound substructures (e.g., Onions et al. 2012), though Behroozi et al. (2015) has recently highlighted the difficulty in connecting halos during merger events. These efforts demonstrate that only algorithms that combine phase-space *and* temporal information should be used.

Two primary groups have performed zoom-in N-body simulations of the growth of Milky Way sized halos in extremely high resolution – the *Aquarius* project of Springel et al. (2008) and the *Via Lactea* simulations of Diemand et al. (2008). Whilst these works have been thoroughly successful and made it possible to quantify the formation of the stellar halo, for example, both the *Aquarius* and *Via Lactea* projects are limited in a number of respects.

The first of these is that they adopted the now observationally disfavored *Wilkinson Microwave Anisotropy Probe's* first set of cosmological parameters (*WMAP-1*, Spergel et al. 2003). The advent of the *Planck satellite* (Planck Planck Collaboration et al. 2014) with three times higher resolution and better treatment of the astrophysical foreground (owing in large part to using nine frequency bands instead of five with *WMAP*) has allowed even more precise estimates of key cosmological parameters. In particular, the most crucial of these for accurate cosmological simulations are the baryon density (Ω_b), the matter density (Ω_c), the dark energy density (Ω_Λ), the density fluctuations at 8 h^{-1} Mpc (σ_8) and the scalar spectral index (n_s). Dooley et al. (2014) showed through a systematic studies of structure formation using different cosmologies that the maximum circular velocities, formation and accretion times of a given host's substructure are noticeably different between cosmologies. σ_8 in *WMAP-1* for example is much higher ($\sigma_{8,WMAP1} = 0.9$ vs. $\sigma_{8,Planck} = 0.83$) which shifts the peak in cosmic star formation rate to lower redshift, resulting in slightly bluer galaxies at $z = 0$ (Jarosik et al. 2011, Guo et al. 2013, Larson et al. 2015).

The second major drawback and perhaps more significant is that the *Aquarius* and *Via Lactea* simulations were simply limited in number. The *Aquarius* project consists of six well-resolved Milky Way mass halos, while the *Via Lactea* study focused on only one such halo.

There exists significant halo-to-halo scatter in, e.g., the substructure shape and abundance owing to variations in accretion history and environment, (Springel et al. 2008, Cooper et al. 2010, Boylan-Kolchin et al. 2010), with the dispersion appearing significant (a factor > 3, Lunnan et al. 2012). But based on such a small sample, the extent cannot be well-quantified, although determining the distributions of substructure properties of galaxy halos is critical for interpreting the various observations of dwarf galaxy populations of all large galaxies, including the Milky Way and Andromeda.

More recently, Garrison-Kimmel et al. (2014b) have produced a suite of 36 Milky Way halos (24 isolated analogues, 12 Local Group analogues; *ELVIS* suite) at a resolution of $\sim 10^5$ $M_\odot$ per particle ($\sim$Aquarius level-3). Studies using this suite

have again highlighted the case for the *too big to fail problem* (Boylan-Kolchin et al. 2011) by showing that the so called "massive failures" (i.e. halos with $V_{\rm max} \geq 25$ km s^{-1} that became massive enough to have formed stars in the presence of an ionizing background, $V_{\rm peak} \geq 30$ km s^{-1}) do not disappear when larger numbers of halos across a range of host masses are simulated (Garrison-Kimmel et al., 2014b). Despite the *ELVIS* suite's utility, it unfortunately lacks the extra mass resolution required to study the formation of minihalos and very small dwarf galaxies ($\sim 10^6$ M$_\odot$), both at the present day and their evolution since the epoch of reionization. Also, *ELVIS* is not suitable[2] for using the *particle tagging* technique whereby a few per cent of the central dark matter particles of accreting systems are assigned stellar properties to study the assembly of the stellar halo (e.g., Cooper et al. 2015). If we are to understand the origin of the first stellar systems (including their chemical constituents) and to locate their descendants at the present day, higher resolution as well as particle tagging is of critical importance.

Whilst previous simulations all have their own merits and drawbacks, one issue prevalent across nearly all previous studies is that they introduced bias in selecting their halos. Usually halo candidates studied using the zoom-in technique meet three criteria: isolation, merger history and Lagrangian volume. From a computational standpoint, if one can obtain a compact Lagrangian region, a quiet merger history and keep the halo relatively isolated, the savings in CPU-hours can be immense. Ultimately, however this three-pronged approach introduces a selection bias. Whilst constructing a simulation with these three key criteria in place will generate an approximate Milky Way analogue, one will not gain an understanding of how the results from studying this halo will compare to halos more generally selected from a pool in the desired mass range (e.g., $1 - 2 \times 10^{12}$ M$_\odot$).

The first requirement is that the halos have a quiescent merger history, which is usually defined by the host having no major merger since a given redshift, e.g.,

[2]Particle tagging usually requires 1–5% of the most bound particles of a satellite to be tagged. For a simulation which resolves 10^8 M$_\odot$ hosts with ~100 particles (i.e., *ELVIS*), this means one can only use a single particle to contain all the baryonic information which is insufficient for modelling multiple stellar populations.

$z = 1$ (e.g., Springel et al. 2008). Constraining the merger history of a simulation suite severely limits the capabilities of reconstructing the formation history of the Milky Way. Indeed, by statistically contrasting observational data sets to mock data extracted from a set of Milky Way-like dark matter halos, coupled to a semi-analytical model of Galaxy formation (Tumlinson, 2010), Gómez et al. (2012) showed the best-fitting input parameter selection strongly depends on the underlying merger history of the Milky Way-like galaxy. For example, even though for every dark matter halo it is always possible to find a best-fitting model that tightly reproduces the Milky Way satellite luminosity function, these best-fitting models generally fail to reproduce a second and independent set of observables (see Gómez et al. 2014). It is thus critical to sample a wide range of evolutionary histories. The second requirement that the Lagrangian volume of the halo's particles be compact also in part biases the merger history of the halo. For a fixed $z = 0$ virial mass, the smaller the Lagrangian volume of a halo, the less likely that halo will have a late major-merger event. This bias further compounds the aforementioned issues of selecting halos with quiet merger histories. Lastly, the isolation criteria preferentially selects halos in low density environments, resulting in decreased substructure (Ragone-Figueroa and Plionis, 2007) and higher angular momentum (Avila-Reese et al. 2005, Lee 2006).

In light of all of these issues, we are motivated to create a comprehensive dataset consisting of 60–70 dark matter halos of approximately Milky Way mass in extremely high spatial and temporal resolution with a more relaxed selection criteria to not just understand the origin and evolution of the Milky Way, but additionally how it differs to other galaxies of similar mass *in general*. Moreover, this new simulation set (unlike the *Aquarius* and *Via Lactea* which were very specific in nature) lends itself well to studying the substructure and stellar halos of $\sim 10^{12}$ $M_\odot$ galaxies such as those being studied in the recently completed GHOSTS survey (de Jong et al., 2007; Monachesi et al., 2013, 2015).

We call this simulation suite *The Caterpillar Project* owing to the similarity between each of the individual halos and how they work together towards a common purpose. Due to the extreme computational requirement for a project of this size

(~14M CPU hours and ~700 TB of storage), we are staggering our release. For this first paper, we focus on the general $z = 0$ properties of the first 24 halos of the *Caterpillar* suite in order to clearly demonstrate data integrity and utility. In Section 2, we outline the simulation suite parameters, numerical techniques, and halo properties. In Section 3 we present a variety of initial results drawn from the suite. In Section 4, we present our primary conclusions from our initial subset of halos. Lastly, we present an Appendix with details of our convergence study and parameters used in the construction of our initial conditions.

6.2 The CaterPillar Suite

6.2.1 Simulation & Numerical Techniques

The *Caterpillar* suite was run using P-GADGET3 and GADGET4, tree-based N-body codes based on GADGET2 (Springel 2005). For the underlying cosmological model we adopt the ΛCDM parameter set characterised by a *Planck* cosmology given by, $\Omega_m = 0.32$, $\Omega_\Lambda = 0.68$, $\Omega_b = 0.05$, $n_s = 0.96$, $\sigma_8 = 0.83$ and Hubble constant, H = 100 h km s^{-1} Mpc^{-1} = 67.11 km s^{-1} Mpc^{-1} (Planck Collaboration et al. 2014). All initial conditions were constructed using MUSIC (Hahn and Abel 2011). We identify dark matter halos via ROCKSTAR (Behroozi et al. 2013b) and construct merger trees using CONSISTENT-TREES (Behroozi et al. 2012). ROCKSTAR assigns virial masses to halos, $M_{\rm vir}$, using the evolution of the virial relation from Bryan and Norman (1998) for our particular cosmology. At $z = 0$, this definition corresponds to an over-density of 97 × the critical density of the Universe. We have modified ROCKSTAR to output *all* particles belonging to each halo so we can reconstruct any halo property in post-processing if required. We have also improved the code to include iterative unbinding (see Section 6.2.5). In this work, we restrict our definition of virial mass to include only those particles which are bound to the halo.

6.2.2 Parent Simulation, Zoom-ins & Contamination

Initially a parent simulation box (see Fig. 6-1) of width 100 h^{-1} Mpc was run at $N_p = 1024^3$ ($m_p = 8.72 \times 10^7\ h^{-1}\ \mathrm{M}_\odot$) effective resolution (see MUSIC/P-GADGET3 parameter files on project website) to select viable candidate halos for re-simulation (i.e., ~10,000 particles per host). The candidates for re-simulation were selected via the following mass and isolation criteria:

- Halos were selected between $0.7 \times 10^{12}\ \mathrm{M}_\odot \leq M_{\mathrm{vir}} \leq 3 \times 10^{12}\ \mathrm{M}_\odot$ (Smith et al. 2007, Xue et al. 2008, Tollerud et al. 2012, Boylan-Kolchin et al. 2013, Sohn et al. 2013, Piffl et al. 2014, Guo et al. 2015, Peñarrubia et al. 2016, see Wang et al. 2015 for review).

- No halos with $M_{\mathrm{vir}} \geq 7 \times 10^{13}\ \mathrm{M}_\odot$ within 7 Mpc (Li and White 2008. van der Marel et al. 2012). We avoid halos near large clusters which would greatly enhance our Lagrangian volumes, making our ability to run simulations at our desired resolution impossible.

- No halos with $M_{\mathrm{vir}} \geq 0.5 \times M_{host}$ within 2.8 Mpc (Karachentsev et al. 2004, Tikhonov and Klypin 2009). We currently avoid pairs in our sample owing to the difficulty of running them at our desired resolution at the present time. We have nevertheless selected equivalent pairs of our current isolated sample but will be examining those in future work.

This results in 2122 candidates being found (from an original sample of 6564 within the specified mass range). We use an extremely weak selection over merger history such that we require no halo to have had a major merger (1:3 mass ratio) since $z = 0.05$ (<5%). Our overall aim is to construct a representative sample of $10^{12}\ \mathrm{M}_\odot$ halos and not specifically require Milky Way analogues *a priori* as has been done in previous studies (Diemand et al. 2007, Stadel et al. 2009, Springel et al. 2008, Garrison-Kimmel et al. 2014b, Sawala et al. 2014). This also allows us to apply statistical tools to constrain semi-analytic models in future work (e.g., Gómez et al. 2014). We place our halos into three mass bins with the largest number of halos

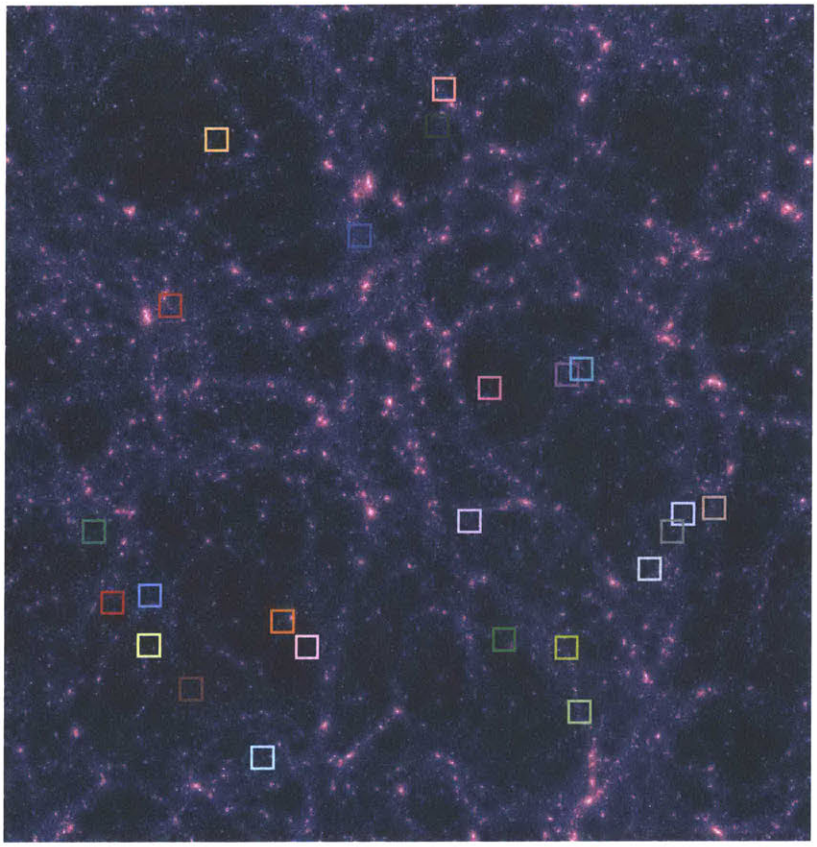

Figure 6-1 Projected dark matter density at $z = 0$ of the parent simulation, from which the 70 Caterpillar halos were selected. The box width is 100 h^{-1} Mpc and the color represents the logarithm of the dark matter density. The colored circles correspond to the location of the first 24 *Caterpillar* halos. The color for a given halo is kept identical for all figures throughout this work.

centered on the most likely mass for the Milky Way ($M_{200} = 1.6^{+0.5}_{-0.4} \times 10^{12}$ M$_\odot$, Piffl et al. 2014);

$$M_i = \begin{cases} 0.7 - 1.0 \times 10^{12} \text{ M}_\odot : 7 \text{ halos} \\ 1.0 - 2.0 \times 10^{12} \text{ M}_\odot : 46 \text{ halos} \\ 2.0 - 3.0 \times 10^{12} \text{ M}_\odot : 7 \text{ halos} \end{cases}$$

For this paper we are only considering a subset of the total sample in preparation, specifically 21 halos within the $1 - 2 \times 10^{12}$ M$_\odot$ mass range and 3 halos within the $0.7 - 1.0 \times 10^{12}$ M$_\odot$ mass range.

6.2.3 Contamination Study

As has been highlighted by Onorbe et al. (2014), a great deal of care has to be taken when carrying out re-simulations of this kind so as to avoid contamination of the main halo of interest by low-resolution particles at $z = 0$. If mass from low resolution particles contributes more than $\sim$ 2% of the total host mass there can be offsets to estimates of the halo profile, shape, spin, and especially gas properties in hydrodynamic runs. To avoid contamination in our sample we custom built a Python GUI (using TraitsUI), named *Caterpillar Made Easy* (CME: `http://github.com/bgriffen/cme`) for running and analyzing cosmological simulations (for both single and multi-mass simulations). This tool allowed us to carry out an extensive contamination study (i.e., using $\sim$264 low resolution test halos with a particle mass of $\sim 10^7$ $M_\odot$) specifically for the halos to be re-simulated. We have *automated* the monotony of constructing hundreds of qualitatively similar *but* quantitatively distinct cosmological simulations with the added benefit of being able to interactively select over initial condition parameters, cosmologies, halo finders and merger trees. This procedure was carried out self-consistently across all runs allowing for a systematic study of which simulation parameters produce the most computationally inexpensive to run, uncontaminated halos.

Using CME we tested eleven Lagrangian geometries (e.g., convex hull, ellipsoid, expanded ellipsoids, cuboids and expanded cuboids) so as to ensure a sphere of radius $\sim$1 h^{-1} Mpc exists of purely uncontaminated (high-resolution) particles centered on the host halo at $z = 0$. Our need to run eleven different Lagrangian geometries for each halo is motivated by the fact that the geometries vary substantially from halo to halo (due partially to their varied merger histories) and we wished to minimize the computation cost whilst achieving our contamination goals. It must also be highlighted that unlike many other studies, we did not select one Lagrangian geometry for all halos but used a specific geometry for a given halo depending on the needs of its simulation.

In Table 6.1 we show the various geometries we used for constructing our initial

conditions. We modified MUSIC to be able to produce expanded Lagrangian volumes rather than the bounded volumes with which it was originally published. In Figure 6-2 we show four examples of Lagrangian geometries for halos selected for re-simulation. In some cases the geometries are reasonably compact allowing for the traditional minimum cuboid enclosing to be used. Some larger regions however are extremely non-spherical (e.g., bottom right of Fig. 6-2) and so a minimum ellipsoid was used. For each geometry we take the enclosed volume at $z = 0$ denoted by either $4 \times R_{\rm vir}(z = 0)$ and $5 \times R_{\rm vir}(z = 0)$. We run each of these halos to $z = 0$, run our modified ROCKSTAR and determine at what distance the closest low resolution or contamination particle (type = 2) resides in each case. With the knowledge that the high-resolution volume distance decreases at higher levels of refinement (Onorbe et al., 2014), we ensure a minimum contamination distance of $\sim$1 h^{-1} Mpc at $z = 0$ at our lowest resolution re-simulation with the desire to have uncontaminated spheres of radius, $\sim$1 Mpc at our highest resolution re-simulation. In cases where four times the virial radius enclosure created contaminated halos but five times the virial radius created too large a simulation to run, we opted for an expanded ellipsoid of the minimum enclosing ellipsoid. In some cases there were a handful of offending particles far away from the primary Lagrangian volume (e.g., Figure 6-2 top right and bottom left) making no standard geometry enclosure feasible, expanded or otherwise. Here we trimmed the Lagrangian volume by hand and simulated the new geometry to $z = 0$ to ensure it had no contamination. Traditionally these types of halos are avoided but since we do not want to bias our sample, we dealt with complicated geometries in this specialized manner and have included them in our sample. Using this tailored approach, our highest resolution runs obtain very large, high-resolution regions with spheres of radius $\sim 1.4 \pm 0.4$ Mpc of solely high-resolution particles.

In Figure 6-3 we show box plots of the median contamination distance and respective quartiles for all 264 of our test halos using each of our selected geometries. Typically, the best performing geometry (i.e., the largest uncontaminated volume with the cheapest computational expense) was the expanded ellipsoid which enclosed all particles within 4 or 5 times the virial radius of the host in the parent simulation

Table 6.1 The contamination suite used on the first refinement level (i.e., LEVELMAX = 11) for every halo in the *Caterpillar* suite.

Name	$nR_{vir}(z=0)^a$	Geometry	Factor[b]
CA4	4	Convex Hull	–
CA5	5	Convex Hull	–
EA4	4	Bounded Ellipsoid	–
EA5	5	Bounded Ellipsoid	–
EX4	4	Expanded Ellipsoid	1.05
EX5	5	Expanded Ellipsoid	1.05
EB4	4	Expanded Ellipsoid	1.1
EC4	4	Expanded Ellipsoid	1.2
BA4	4	Minimum Cuboid	–
BA5	5	Minimum Cuboid	–
BB4	4	Expanded Cuboid	1.1

[a]The multiple of the $z = 0$ virial radius that we used to construct the Lagrangian volume.

[b]The factor we increased the original volume (e.g., 1.2 means the ellipsoid was expanded by 20% in size. A dash represents the minimum ellipsoid/cuboid/hull exactly). These values were arbitrarily chosen with the only requirement being that the initial condition files were not overly large in size (i.e., a few hundred megabytes at LX11).

at $z = 0$.

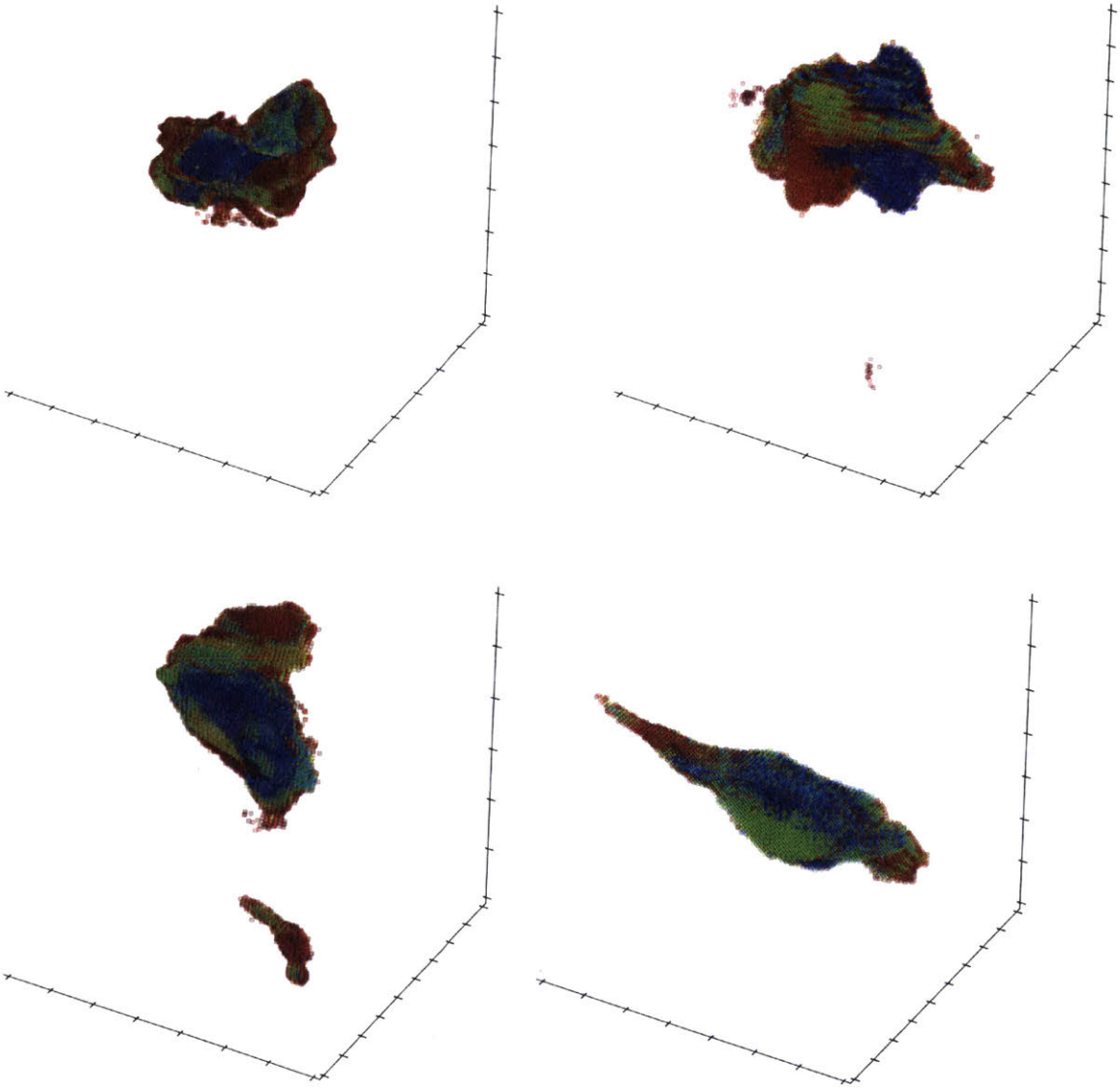

Figure 6-2 Sample Lagrangian volumes (at $z = 127$) of halos from the parent simulation. Some geometries are easily bounded by a minimum cuboid (e.g., top left) but others require an ellipsoid (e.g., bottom right). In some cases, there reside particles well away from the primary Lagrangian volume (e.g., top right and bottom left). For these difficult situations, we carefully trimmed the Lagrangian geometry to ensure no contamination of low resolution particles within 1 h^{-1} Mpc of the host at $z = 0$. Particles are color coded by distance from the host at $z = 0$ where red particles are $5\times R_{\rm vir}$ and blue particles are within the virial radius.

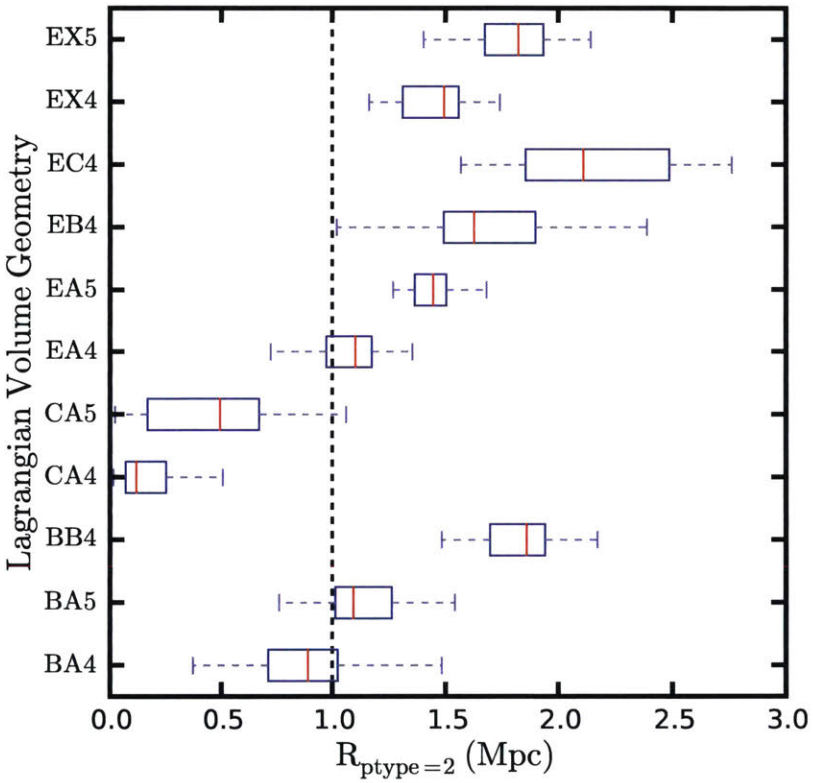

Figure 6-3 Box plot of the distances to the first low resolution particle (particle type 2) for each of the prospective initial condition geometries (i.e., all run at LEVELMAX = 11 in MUSIC, one level above our parent simulation). Table 6.1 contains the key for each geometry. The red lines indicate the median, edges of the box represent 25% and 75% quartiles and the outer tick marks represent the maximum and minimum distance. The dashed line represents the threshold we used to determine if a geometry was viable for a higher level re-simulation, though this was balanced against computational cost. We select the geometry which used the fewest CPU hours but maintained the largest uncontaminated volume.

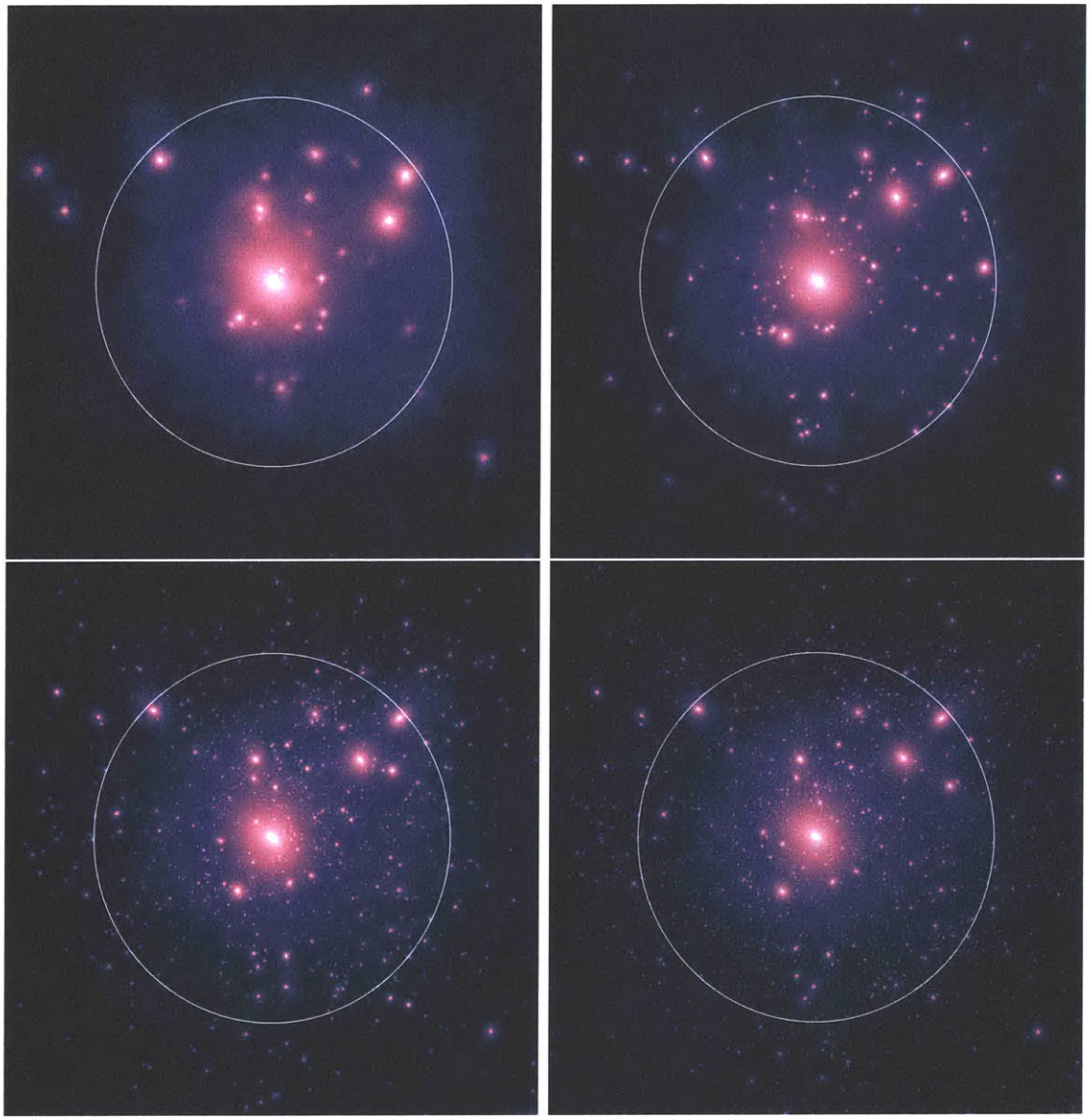

Figure 6-4 Projected dark matter density at $z = 0$ of Cat-1 at successively higher resolutions (increasing by a factor of 8 in mass resolution each time) from left to right. The upper left panel is LX = 11 ($m_p = 1.53 \times 10^8 M_\odot$) and the lower right panel is LX = 14 ($m_p = 2.99 \times 10^4 M_\odot$). The white circle represents the virial radius, $R_{\rm vir}$. The image brightness is proportional to the logarithm of the dark matter density squared (i.e., $\log(\rho^2)$).

Table 6.2 The resolution levels of the *Caterpillar* suite.

LX	N_p	m_p ($M_\odot$)	ϵ (h^{-1} pc)
15	32768^3	3.7317×10^3	38
14	16384^3	2.9854×10^4	76
13	8096^3	2.3883×10^5	152
12	4096^3	1.9106×10^6	228
11	2048^3	1.5285×10^7	452
10	1024^3	1.2228×10^8	904

LX represents the effective resolution ($N_p = (2^X)^3$) of the high resolution region given by parameter, LEVELMAX in MUSIC. m_p is the particle mass and ϵ is the Plummer equivalent gravitational softening length. The parent simulation parameters are also shown in the last row. Only a select sample of halos are run at resolution level LX15. These runs will be presented in future works.

6.2.4 Zoom-in Simulations

Starting from our parent simulation resolution, we re-simulated each halo at iteratively higher resolutions (a factor of 8x increase in particle number for each level) to ensure we did not obtain contaminated particles within the host halo (the uncontaminated volume shrinks with an increase in the ratio of the zoom-in resolution to that of the parent simulation resolution). In Figure 6-4, we show the dark matter distributions for iteratively higher resolutions of the same halo. One can clearly identify the same subhalos across all resolutions indicating the qualitative success of our numerical techniques. Regarding computational resources, each halo at our highest resolution took between 150 – 300K hours on TACC/Stampede and occupy ~5 – 10 TB of storage for both the raw HDF5 snapshots and halo catalogue. Table 6.2 shows the mass and spatial resolution for each of our refinement levels. Our softening length is $\epsilon \sim 76\ h^{-1}$ pc for our fiducial resolution.

We space our snapshots (320 per simulation) in the logarithm of the expansion factor until $z = 6$ (~5 Myrs/snapshot) and then linear in expansion factor down to $z = 0$ (~50 Myrs/snapshot). The motivation for this piecewise stitching of the two temporal schemes is two-fold. At $z > 6$, we enter the era of mini-halo formation and

the reionization epoch. If we wish to model the transport of Lyman-Werner (LW) radiation semi-analytically from the first mini-halos (e.g., Agarwal et al. 2012), we require a temporal resolution on par with the mean free path of LW photons in the intergalactic medium and the lifetime of a massive Population III star ($\leq$ 10 Myrs). Secondly, we also wish to resolve the disruption of low mass dwarf galaxies at low redshift, which requires a temporal resolution of order $\sim$ 50 Myrs (e.g., Segue I has a disruption time scale of $\sim$ 50 Myrs). These time scales are also required if one is attempting to determine subhalo orbital pericenters which can be input into semi-analytic models of tidal disruption (e.g., Baumgardt and Makino 2003). While we intend to use the capabilities offered by finely sampled snapshots in future work, this initial paper primarily focuses on the $z = 0$ halo properties.

6.2.5 Iterative Unbinding In ROCKSTAR

ROCKSTAR is able to find any overdensity in 6D phase space including both halos and streams. To distinguish gravitationally bound halos from other phase space structures, ROCKSTAR performs a single-pass energy calculation to determine which particles are gravitationally bound to the halo. Over-densities where at least 50% of the mass is gravitationally bound are considered halos, with the exact fraction a tuneable parameter (`unbound_threshold`) of the algorithm (Behroozi et al., 2013b).

This definition is generally very effective at identifying halos and subhalos –but it fails in two important situations. First, if a subhalo is experiencing significant tidal stripping, the 50% cutoff can remove a subhalo from the catalog that should actually exist. We have found that changing the cutoff can recover the missing subhalos, but the best value of the cutoff is not easily determined. Second, ROCKSTAR is occasionally *too* effective at finding substructure in our high resolution simulations. In particular, it often finds velocity substructures in the cores of our halos that are clearly spurious based on their mass accretion histories and density profiles. Importantly, these two issues do not just affect low mass subhalos, but they can also add or remove halos with $V_{\rm max} > 25$ km s^{-1}.

Both of these problems can be alleviated by applying an iterative unbinding proce-

dure. We have implemented such an iterative unbinding procedure within ROCKSTAR. At each iteration, we remove particles whose kinetic energy exceeds the potential energy from other particles in that iteration. The potential is computed with the ROCKSTAR Barnes-Hut method (see Appendix B of Behroozi et al. 2013b). We iterate the unbinding until we obtain a self-bound set of particles. Halos are only considered resolved if they contain at least 20 self-bound particles. All halo properties are then computed as usual, but with the self-bound particles instead of the one-pass bound particles. The iterative unbinding recovers the missing subhalos and removes most but not all of the spurious subhalos. Across 13 of our *Caterpillar* halos, we recover 52 halos with subhalo masses above 10^8 $M_\odot$ which would have otherwise been lost using the conventional ROCKSTAR. Figure 6-5 demonstrates how these large haloes can be recovered when our iterative unbinding procedure is used.

To remove the remaining spurious subhalos, we also remove halos if $r_{\rm max}$ (i.e., the radius at which the velocity profile reaches its maximum) of the subhalo is larger than the distance between the subhalo and host halo centers. The downside to adding iterative unbinding is that it increases the run time for ROCKSTAR by $\sim$50%. In the rest of this paper, we only consider subhalos with at least 20 self-bound particles passing the $R_{\rm max}$ cut.

6.3 Results

6.3.1 Host Halo Properties

In Table 6.3, we provide the basic properties of our first 24 re-simulated halos. This includes the simulation name, the halo virial mass, the halo virial radius, concentration, maximum circular velocity, the radius at which the maximum circular velocity occurs, the formation time (defined as when the halo reaches half its present day mass), the redshift of the last major merger (1:3 mass ratio), the fraction of the host mass contained within subhalos, the axis ratios defining the halo shape, and the distance to the closest contamination particle from the host. We adopt a simple naming

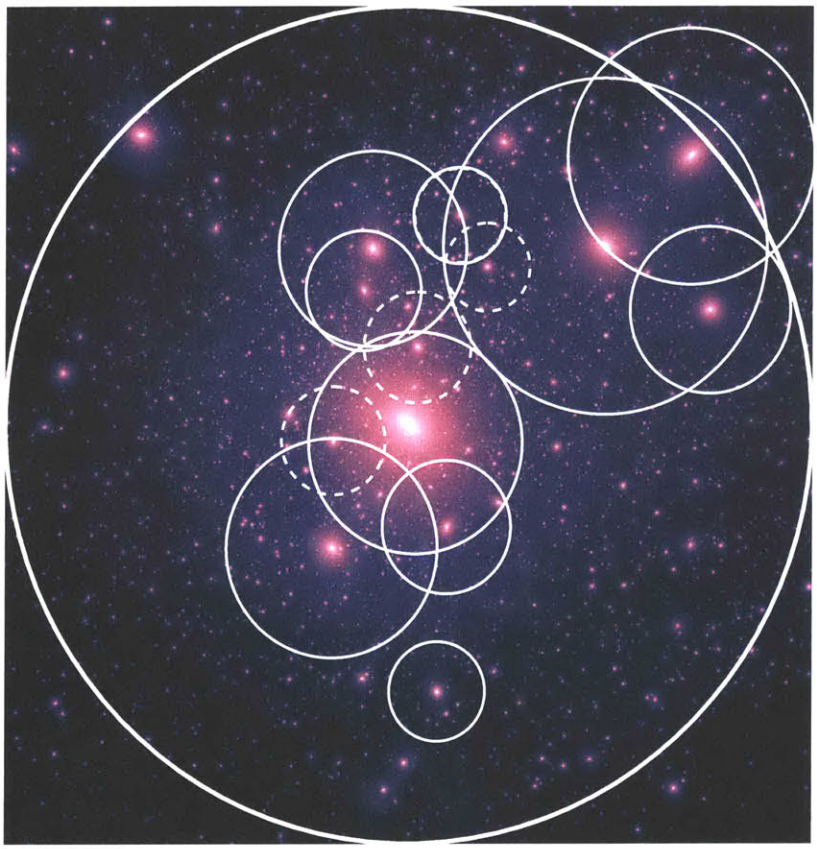

Figure 6-5 Density projection of the Cat-1 halo with subhalos with $V_{\rm max} > 30$ km s^{-1} highlighted by white circles. Their size corresponds to their virial radii. The full circles indicate subhalos found by the default parameters of the halo finding algorithm ROCKSTAR. The dashed circles indicate where halos are recovered when iterative unbinding is used.

convention based on when the halos were post processed (1 – N_{halos}). Where required, we use a shorthand reference to the resolution of the simulation. These refer to the parameter LEVELMAX inside the IC generation code MUSIC (e.g., LEVELMAX = 14 is simply LX14). This means LX14 represents an effective resolution of $N_p = (2^{14})^3$ ($\sim$ *Aquarius* level-2), LX13 $\rightarrow N_p = (2^{13})^3$ ($\sim$ *Aquarius* level-3 or $\sim$ *ELVIS* resolution), LX12 $\rightarrow N_p = (2^{12})^3$ and LX11 $\rightarrow N_p = (2^{11})^3$. Unless otherwise stated, all halos in the analysis of this paper are the LX14 halos (i.e., our flagship resolution). All halos have similar $z = 0$ properties except Cat-7 whose properties can be tied to the fact that it has recently undergone a massive merger (1:3 mass ratio at $z = 0.03$). We obtain extremely large uncontaminated volumes ($\sim$1.4 Mpc) in all but one of our halos (Cat-18 is $\sim$ 3% contaminated by mass). The fraction of mass held in subhalos

Table 6.3 The halo properties of the first 24 *Caterpillar* halos.

Halo Name	$M_{\rm vir}$ ($\times 10^{12}$ M$_\odot$)	$R_{\rm vir}$ (kpc)	c [a]	$V_{\rm max}$ (km s^{-1})	$R_{\rm max}$ [b] (kpc)	z_{form} [c]	z_{lmm} [d]	$f_{\rm m,subs}$ [e]	c/a	b/a	$R_{\rm hires}$ [f] (Mpc)
Cat-1	1.559	306.378	7.492	169.756	34.083	0.894	2.157	0.207	0.841	0.869	0.998
Cat-2	1.791	320.907	8.374	178.851	55.268	0.742	0.731	0.148	0.636	0.719	1.463
Cat-3	1.354	292.300	10.170	172.440	31.701	0.802	0.802	0.136	0.865	0.927	1.894
Cat-4	1.424	297.295	8.573	164.344	53.466	0.936	0.922	0.175	0.671	0.739	1.531
Cat-5	1.309	289.079	12.108	176.399	32.103	0.564	0.510	0.069	0.552	0.815	1.608
Cat-6	1.363	292.946	10.196	171.647	33.632	1.161	1.295	0.153	0.508	0.528	1.295
Cat-7	1.092	272.099	1.757	134.148	157.438	0.070	0.032	0.735	0.151	0.207	1.477
Cat-8	1.702	315.466	13.507	198.564	40.819	1.516	2.235	0.078	0.605	0.787	1.540
Cat-9	1.322	289.987	12.401	177.414	30.336	1.255	1.236	0.094	0.513	0.762	2.080
Cat-10	1.323	290.119	11.714	174.989	39.721	1.644	2.010	0.103	0.559	0.703	1.775
Cat-11	1.179	279.187	12.522	172.723	53.187	1.059	4.368	0.215	0.597	0.867	1.135
Cat-12	1.763	319.209	11.402	191.259	52.717	1.336	9.616	0.073	0.584	0.645	1.162
Cat-13	1.164	277.938	12.850	171.222	33.757	1.161	11.092	0.090	0.578	0.645	1.566
Cat-14	0.750	240.119	9.135	137.437	26.660	1.144	4.258	0.113	0.705	0.859	2.178
Cat-15	1.505	302.787	8.983	174.124	37.043	1.144	3.165	0.126	0.849	0.877	1.119
Cat-16	0.982	262.608	11.737	155.362	28.768	1.315	3.165	0.106	0.618	0.792	0.671
Cat-17	1.319	289.800	12.765	179.056	38.329	1.846	1.976	0.093	0.664	0.881	1.299
Cat-18	1.407	296.099	7.887	163.920	57.217	0.493	0.435	0.159	0.676	0.816	0.397
Cat-19	1.174	278.770	10.468	164.726	29.112	1.541	2.118	0.169	0.664	0.937	1.712
Cat-20	0.763	241.484	13.324	149.672	30.417	1.492	5.427	0.099	0.601	0.733	1.311
Cat-21	1.881	326.206	10.618	190.683	50.954	1.126	1.198	0.118	0.482	0.611	1.453
Cat-22	1.495	302.116	10.666	180.647	35.860	0.841	29.488	0.080	0.512	0.694	1.744
Cat-23	1.607	309.524	12.489	190.705	32.421	1.161	9.616	0.094	0.607	0.784	1.207
Cat-24	1.334	290.866	11.378	176.911	36.800	1.144	3.608	0.090	0.689	0.734	1.102
Mean*	1.368	291.791	10.903	173.167	38.886	1.144	4.410	0.121	0.634	0.771	1.402
$\pm 1\sigma$	0.285	21.610	1.761	13.441	9.530	0.329	6.112	0.041	0.103	0.102	0.409

[a] Concentration defined by ratio of the virial radius and the scale radius; $R_{\rm vir}/R_{\rm s}$.

[b] The radius at which the $V_{\rm max}$ occurs.

[c] Redshift of host formation defined as when the host main branch progenitor mass equals $0.5M_{\rm vir}(z=0)$.

[d] Redshift of last major merger defined as a halo with 1/3 mass merging into the main branch of the host.

[e] Fraction of the host mass in subhalos.

[f] Distance to the closest contamination particle from the host.

*Means and deviations were calculated over all halos *except* Cat-7 as it has undergone a very recent major merger.

across our sample is $f_{m,subs} = 0.121 \pm 0.041$ (1σ), though this excludes Cat-7.

In Figure 6-6 we plot the concentration-mass (c-M) relation of the parent simulation for similarly sized halos ($11.5 < \log_{10} M_{\odot} < 12.5$, grey band indicating the 1σ dispersion) and overlay the concentration (R_{vir}/R_s) and host mass of the high resolution halos. This shows that for nearly all of the halos, we are sampling within 68% of the average c-M relation at a fix halo mass. Again, Cat-7 is an outlier with an extremely low concentration because it recently underwent a major merger and has an extremely large substructure mass fraction so its concentration is not meaningful. For this reason we do not include it in the quantitative analysis in terms of determining average halo profile shapes or the mass function slopes. Its properties are still shown and plotted in the various tables and figures, however. Recently, Buck et al. (2015) found that the thickness of planes of satellites depends on the concentration of the host halo. Specifically, they found the thinnest planes are only found in the most concentrated, and hence earliest formed halos. The fact that we sample relatively average concentrations for halos of this mass range means that it is less likely that these hosts will contain planes of satellites, or if they do, their thicknesses will be quite large (Ji et al., in prep.). As the *Caterpillar* sample grows, we will eventually sample overly concentrated halos, enabling us to see in better detail how this concentration-plane relation holds.

6.3.2 Visualizing The Halos & Their Assembly Histories

In Figures 6-7 and 6-8 we show images of the dark matter distribution in each of our 24 high-resolution halos at redshift $z = 0$. The brightness of each pixel is proportional to the logarithm of the dark matter density squared (i.e. $\log(\rho^2)$projected along the line of sight. To enhance the density contrast, each panel has a different maximum density. We note that similarly colored pixel in one panel does not necessarily mean the density is the same for another panel. The panel width is 1 Mpc and the local dark matter density of the particles in each pixel is estimated with an SPH kernel interpolation scheme based on the 64 nearest neighbor particles. Upon first inspection it is clear that each halo is littered with an abundance of dark matter substructures

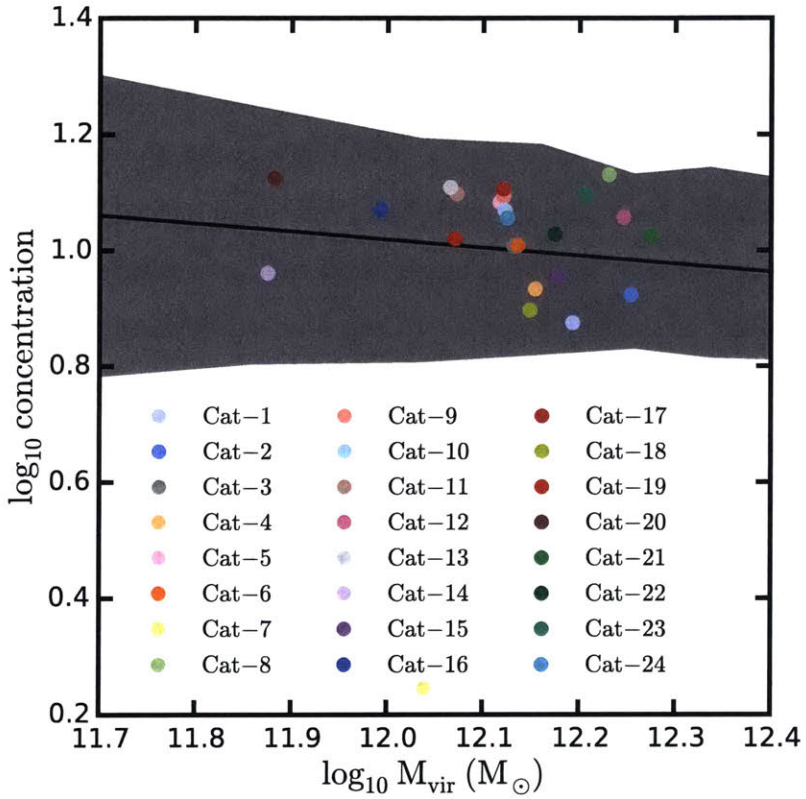

Figure 6-6 Concentration-mass relation for the 24 *Caterpillar* halos relative to those found in the parent simulation of similar mass. The concentration is defined as r_{vir}/r_s. Solid circles are the zoom-in simulations and the black line represents the concentration-mass relation drawn from the parent simulation for relaxed halos. The grey band is the 1σ dispersion in the c–M relation for halos in the parent simulation between $11.5 < \log_{10} \mathrm{M}_\odot < 12.5$.

of varied shapes and sizes. In some cases, there are reasonably large neighbors (e.g., Cat–4, 7, 11, 24). By virtue of our selection criteria these neighbors are no larger than $0.5\times M_{\rm vir}$ of the central host. In any case, in under a Gyr, these SMC/LMC sized systems ($M_{peak} > 10^{11}$ $M_{\odot}$) will likely undergo a major merger with the host galaxy.

In Figure 6-9 we show the mass evolution of each of the halos. As highlighted by the inset which shows the normalized mass evolution, there is a wide variety of formation histories. In our initial catalogue of 24 halos, six halos (Cat–2, 3, 4, 5, 7, 18) have had major mergers since $z = 1$. The halos going above a normalized mass ratio of 1.0 have had a halo pass through them relatively recently which momentarily gives them extra mass such that it is larger than their $z = 0$ mass (e.g, Cat–2). This indicates that many of the halos are yet to reach an equilibrium state.

Adopting the same criteria as Neto et al. (2007) we assess whether the host halos are relaxed. If their substructure mass fraction is below 0.1, their normalized offset between the center of mass of the halo (i.e., computed using all particles within $R_{\rm vir}$) and the potential center ($x_{\rm off} \equiv |r_c - r_{\rm cm}|/R_{\rm vir}$, $r_c \equiv$ center of the potential well, $r_c \equiv$ center of mass) is below 0.07 and their virial ratio ($2T/|U|$) is below 1.35, then the host is considered relaxed. In Table 6.4 we provide the relaxed state of the halo. Many of the halos are in fact unrelaxed under this definition which is by design – we are sampling a wide range of assembly histories and so halos with recent merger events that prevent the halos from being fully virialized naturally make up part of our sample.

6.3.3 Host Halo Profiles

In Figure 6-10 we plot the spherically averaged halo profiles for each of our 24 simulated halos. We draw the measured density profile as a thick line of a given color and continue the fit beyond the smallest radius possible set by Power et al. (2003) as a vertical black dashed line. We truncate each fit at this radius. There is a clear diversity in the profile shapes owing in part to the assembly histories of each halo. Halos which have undergone a recent major merger whose substructure mass fractions

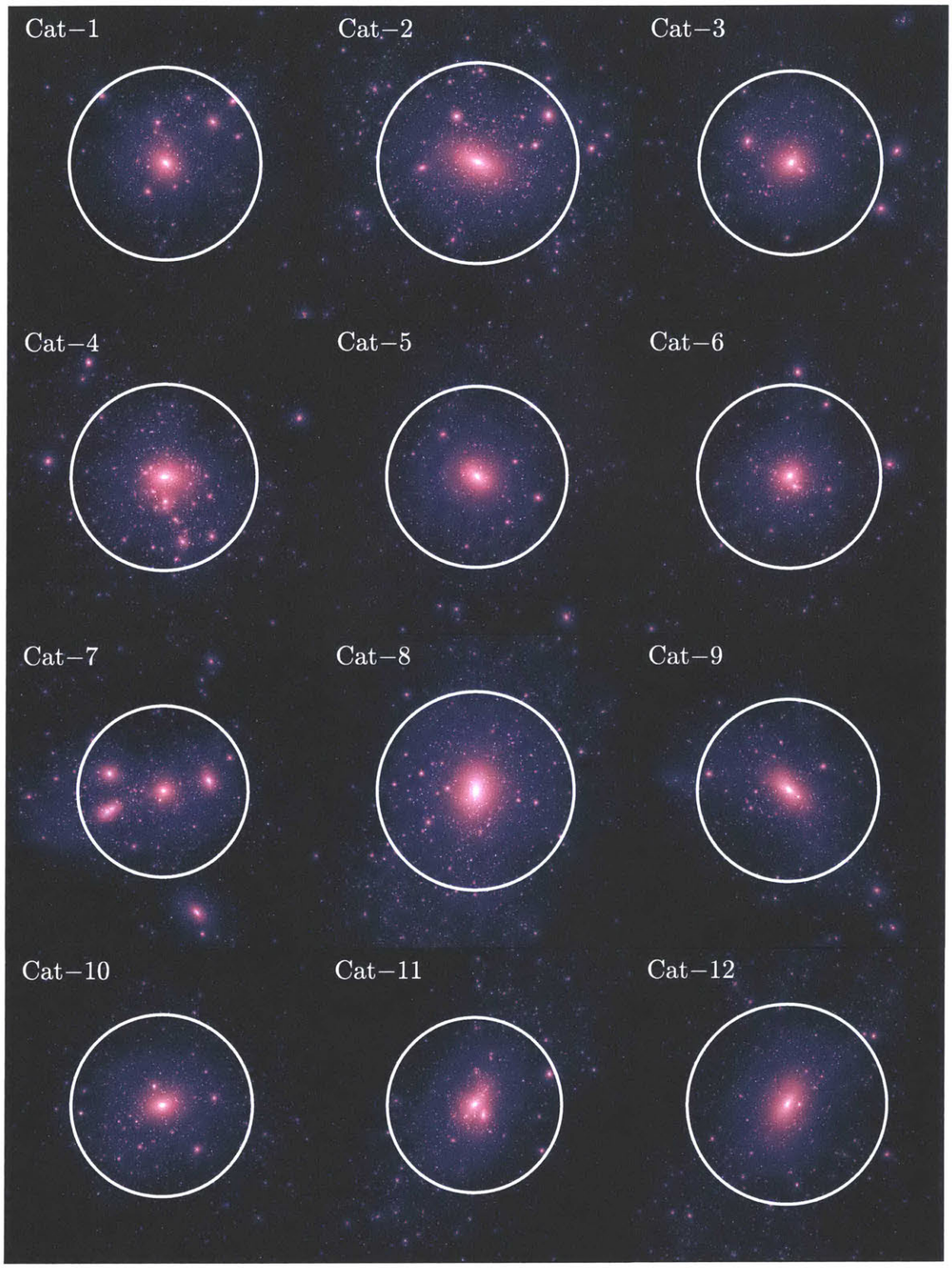

Figure 6-7 Projected dark matter density at $z = 0$ of the first 12 *Caterpillar* halos with a box width of 1 Mpc. The image brightness is proportional to the logarithm of the dark matter density squared. Higher resolution images and animations are available at `www.caterpillarproject.org`.

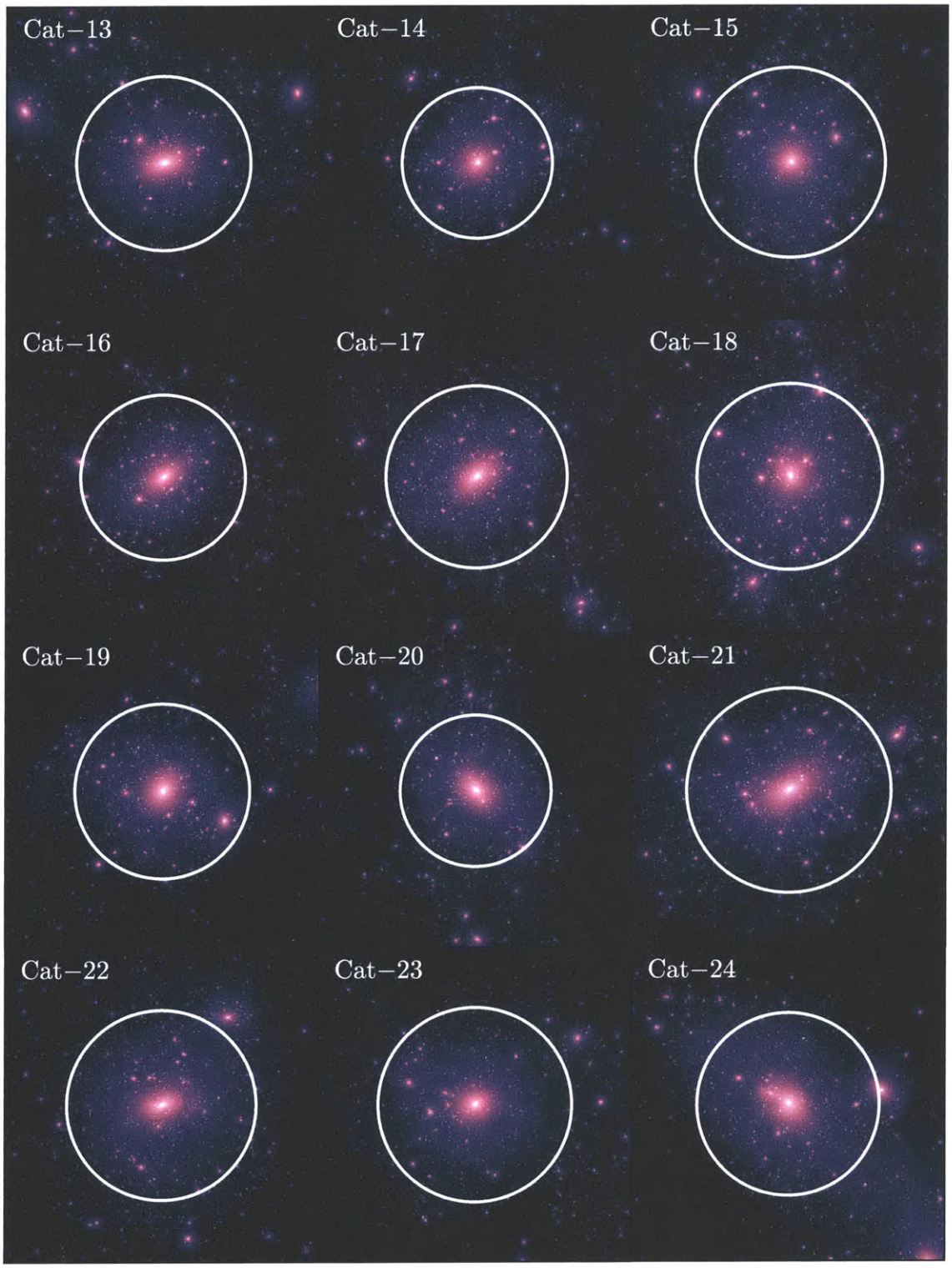

Figure 6-8 Projected dark matter density at $z = 0$ of each of the second set of 12 *Caterpilar* halos with a box width of 1 Mpc. The image brightness is proportional to the logarithm of the dark matter density squared. Higher resolution images and animations are available at `www.caterpillarproject.org`.

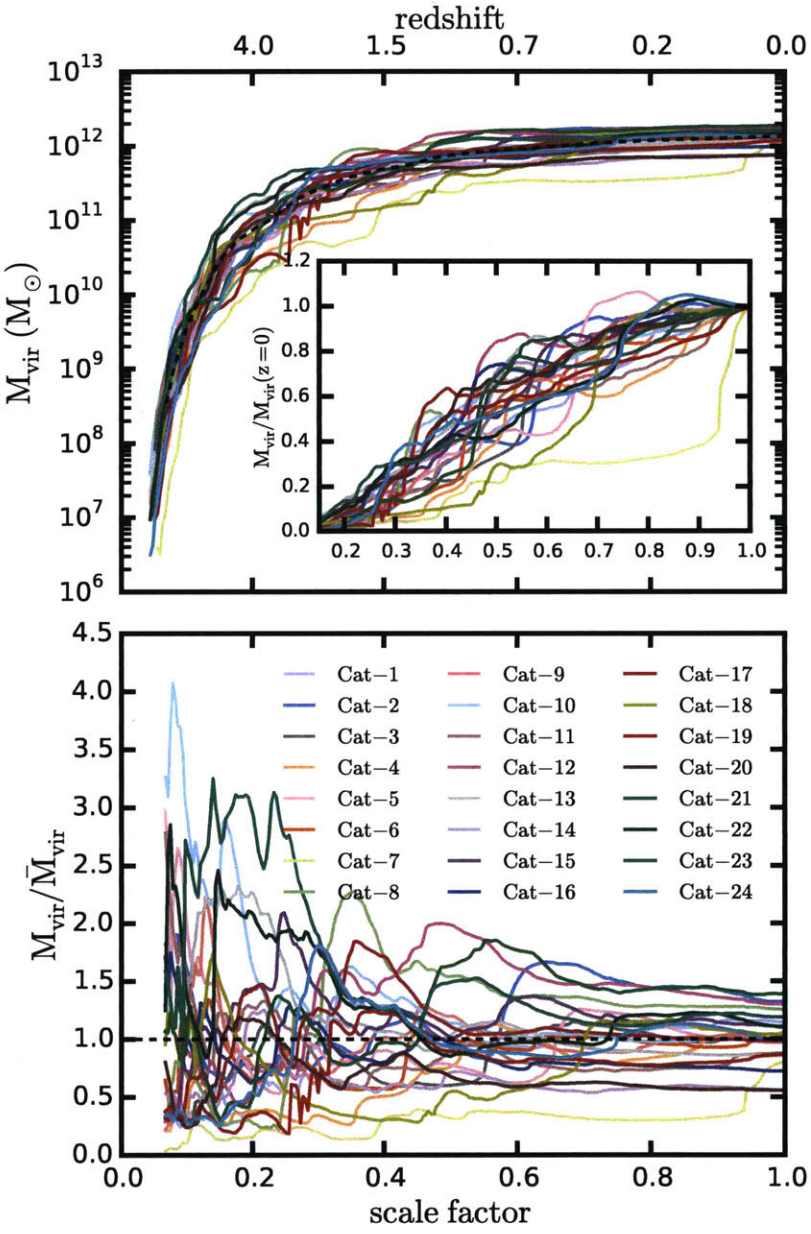

Figure 6-9 Mass evolution of the first 24 *Caterpillar* halos. The top panel shows each halo evolution along with the mean (black dashed) evolution for each of the halos. The inset panel shows the mass evolution normalized to the halo mass at $z = 0$. In the lower panel we show the mass evolution divided by the mean evolution to enhance individual features for each halo. We sample a diverse assembly history, from extremely quiet through to extremely violent (redshift of last major merger, $z \sim 0.07$).

are higher than average are primarily dominated by a single subhalo (e.g, Cat–7 has a large subhalo at 200 kpc). The fitting formula we have used to describe the mass profile of our simulated halos follow the method of Navarro et al. (2010) and is given by the following Einasto form (over all particles within the virial radius):

$$\log[\rho(\mathrm{r})/\rho_{-2}] = (-2/\alpha)\,[(\mathrm{r}/\mathrm{r}_{-2})^{\alpha} - 1]. \tag{6.1}$$

The r_{-2} is the scale length of the halo which can be obtained without resorting to a particular fitting formula. We compute the logarithm of the slope profile and identify where a low-order polynomial fit to it intersects the isothermal value ($\gamma = 2$). Unlike the Navarro-Frenk-White profile (NFW) the peak parameter in the Einasto profile, α, is allowed to vary and thus provides a third parameter for the fitting formula. The best fitting parameters are found by minimizing the deviation between model and simulation at each bin. Specifically we minimize the function Q^2, defined as:

$$\mathrm{Q}^2 = \frac{1}{\mathrm{N_{bins}}} \sum_{\mathrm{i=1}}^{\mathrm{N_{bins}}} (\ln(\rho_{\mathrm{i}}) - \ln(\rho_{\mathrm{i}}^{\mathrm{model}}))^2. \tag{6.2}$$

In this manner we find a function which clearly illustrates the deviation of the simulated and model profiles. In Table 6.4 we show our minimum Q^2 parameter (Q_{min}), characteristic scale radius r_{-2}, and their corresponding densities for each halo. For our relaxed halos, Q_{min} for our Einasto fits are 0.027 $\pm$ 0.010 indicating reasonable agreement between the simulated and model Einasto profiles. This is better than our NFW profile fits for which we obtain $Q_{min} = 0.055 \pm 0.020$. The peak parameters for our Einasto fits are 0.169 $\pm$ 0.023 which is comparable to those of the *Aquarius* halos ($\alpha = 0.145 - 0.173$) studied in Navarro et al. (2010). For halos which significantly deviate from the mean, it is important to note that those halos are not relaxed and so by definition will not provide meaningful Einasto/NFW fits. A more detailed study of the halo density profiles are reserved for future work.

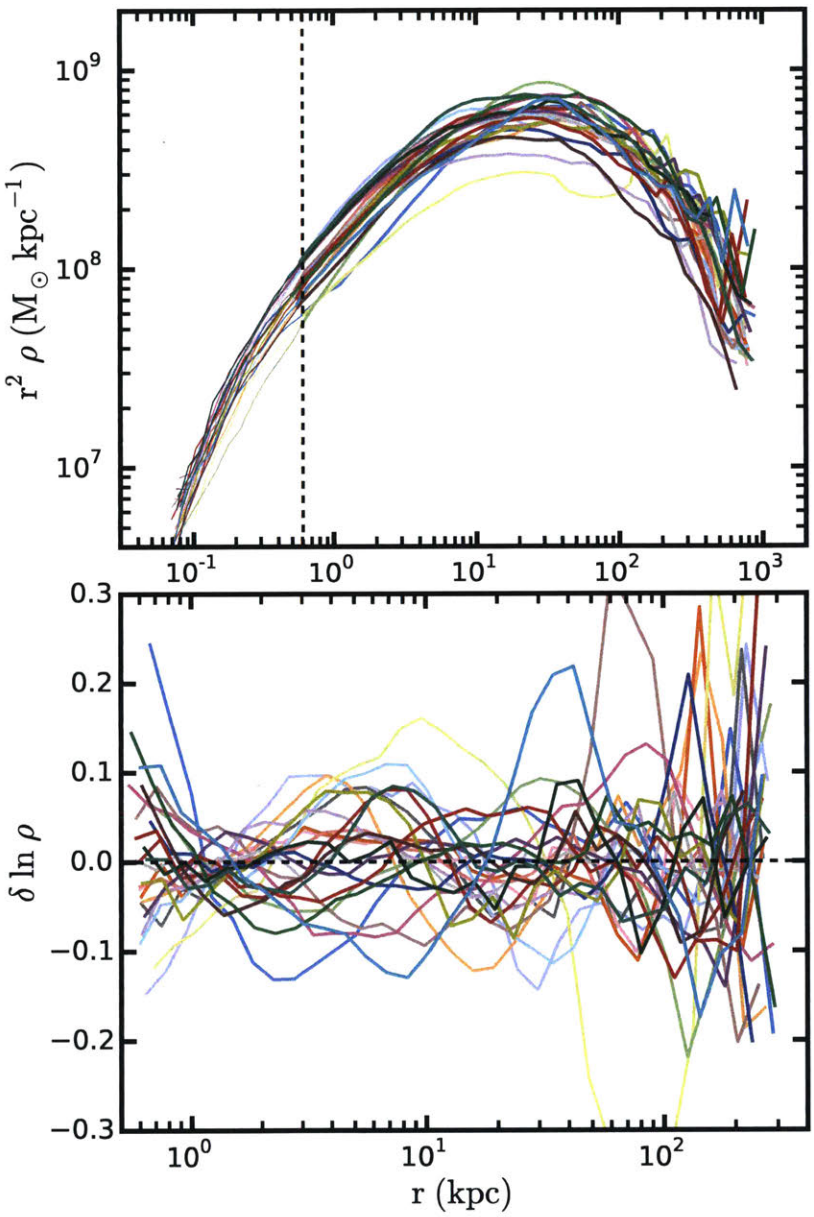

Figure 6-10 *Upper panel*: Normalized halo profiles for each of the host halos. Relaxed halos are better fit by Einasto profiles ($\alpha \sim 0.16$) over NFW profiles though for halos that are unrelaxed, they are unable to be properly fit by NFW or Einasto profiles which is in agreement with Neto et al. (2007) and Navarro et al. (2010). This inability stems purely from their definition, which *a priori* assumes that the halos are in virial equilibrium which is clearly not the case for halos which have undergone a recent merger. The dashed black line indicates the Power et al. (2003) resolution limit as set by our softening length (also represented by a thinner density profile line). *Lower panel*: The log of the ratio between each of the Einasto fits and the data between the Power et al. (2003) radius and the virial radius.

Table 6.4 The relaxed nature and Einasto profile parameters for the first 24 *Caterpillar* halos. For comparison, Q_{min} fits for NFW halo profiles are also listed.

Halo Name	Relaxed [a]	ρ_{-2} [c]	r_{-2} [d] ($\times 10^5$)	α [b]	Q_{min} Ein.	Q_{min} NFW
Cat-1	✗	9.929	27.182	0.128	0.039	0.103
Cat-2	✓	3.531	46.846	0.185	0.040	0.033
Cat-3	✓	11.604	25.309	0.151	0.028	0.067
Cat-4	✗	5.504	34.962	0.154	0.039	0.080
Cat-5	✓	13.305	24.152	0.167	0.018	0.050
Cat-6	✗	9.254	28.299	0.164	0.027	0.058
Cat-7	✗	0.482	90.528	0.075	0.082	0.162
Cat-8	✓	8.027	34.555	0.236	0.033	0.036
Cat-9	✓	12.159	25.543	0.186	0.011	0.034
Cat-10	✓	14.482	23.492	0.168	0.029	0.049
Cat-11	✗	10.617	25.450	0.173	0.049	0.071
Cat-12	✓	8.897	31.157	0.160	0.030	0.062
Cat-13	✓	11.842	25.010	0.187	0.009	0.036
Cat-14	✓	10.455	21.347	0.139	0.018	0.078
Cat-15	✓	8.905	28.744	0.139	0.025	0.078
Cat-16	✓	10.445	24.162	0.166	0.026	0.056
Cat-17	✓	11.283	26.799	0.195	0.019	0.026
Cat-18	✗	6.998	31.012	0.150	0.023	0.082
Cat-19	✗	9.079	26.881	0.164	0.032	0.054
Cat-20	✓	11.415	22.158	0.199	0.018	0.020
Cat-21	✓	6.682	36.486	0.175	0.025	0.043
Cat-22	✓	10.857	26.957	0.156	0.017	0.063
Cat-23	✓	13.486	26.024	0.172	0.018	0.039
Cat-24	✓	7.040	31.734	0.181	0.044	0.050
Mean*	-	9.817	28.446	0.169	0.027	0.055
$\pm 1\sigma$	-	2.6106	5.5677	0.023	0.010	0.020

[a]Relaxed criteria is based on that of Neto et al. (2007). If the substructure mass fraction is below 0.1, their center of mass displacement ($x_{off} = |r_c - r_{cm}|/R_{vir}$) is below 0.07 and their virial ratio ($2T/|U|$) is below 1.35, then the host is considered relaxed.

[b]Einasto slope parameter of the form $\rho(r) \propto \exp(-Ar^{\alpha})$.

[c]The density at the characteristic radius in units of: $10^{10}h^2$ $M_{\odot}$ kpc^{-3}.

[d]The characteristic radius or 'peak' radius of the $r^2\rho$ profile in units of: h^{-1} kpc.

*Means and deviations were calculated over all halos *except* Cat-7 as it has undergone a very recent major merger.

6.3.4 Subhalo Properties

In Figure 6-11a we show the cumulative abundance of subhalos as a function of their maximum circular velocity for each *Caterpillar* halo. Since we achieve excellent convergence (see Appendix A), we reliably resolve halos with circular velocities of $\sim$4 km s^{-1}, which is crucial for identifying the sites of first star formation. At the high $V_{\rm max}$ end we find a variety of different sized subhalos for each host. Some hosts have only one 10 km s^{-1} subhalo whereas another host halo has a large 70 km s^{-1} subhalo within the virial radius. Between 5-20 km s^{-1} all halos are very similar in their $V_{\rm max}$ function slopes within a slight offset owing to normalization stemming from the differences in host halo mass. At low $V_{\rm max}$ values ($\sim$ 3 km s^{-1}) we begin to lose completeness of our host subhalo sample due to lack of resolution. We additionally include in this Figure the $V_{\rm max}$ function for subhalos at infall (i.e., when a subhalo first crosses the virial radius of the host). Since dynamical friction affects the highest mass subhalos the fastest, the biggest difference in the functions occurs at the high-mass end whereby several LMC sized systems ($M_{\rm peak} > 10^{11}\ M_{\odot}$) have been destroyed (over a time scale of 1 – 2 Gyrs) between infall and $z = 0$. These large LMC sized-systems at infall can host anywhere from 4 – 30% of the Milky Way sized halo's subhalos at $z = 0$ depending on their orbit and infall time (Griffen et al. in prep.). In solid black we also plot the *Aquarius* Aq-A2 halo from Springel et al. (2008) (using the same version of ROCKSTAR that we used for the *Caterpillar* halos). We find the differences in the cosmology ($\sigma_8 = 0.9$) and the slightly higher resolution of *Aquarius* leads to systematic differences in subhalo abundance.

In Figure 6-11b, we show the subhalo mass functions for each of the halos. Our results are best fit by the power law $\mathrm{d}N/\mathrm{d}M \propto M^{-1.88\pm0.10}$, which is less steep than that found in the *Aquarius* halos of Springel et al. (2008). This slope is the best fit over the ranges $10^5 - 10^8$ M$_{\odot}$. We do observe a scatter in the subhalo abundances. This can be explained by the subtle concentration-subhalo-abundance relation whereby for fixed halo mass, there are more (less) subhalos belonging to hosts which are less (more) concentrated (e.g, Zentner et al. 2005, Watson et al. 2011, Mao et al. 2015). Indeed,

we find halos which are more concentrated (see Figure 6-6) have lower normalizations than those less concentrated at fixed $V_{\rm max}$. This is simply because halos which are more concentrated have formed earlier and so subhalos have spent substantially longer undergoing dynamical disruption within the host compared to similar sized subhalos orbiting less concentrated hosts.

In Figure 6-11c we show the subhalo radial mass fraction which indicates high variability in the contribution to the total halo mass from substructure as a function of galactocentric distance. For example, at $0.07R_{\rm vir}$, the total mass contributing to the host halo mass from substructure varies by a factor of 3 when normalized by mass. At $R_{\rm vir}$ our substructure mass fraction varies by $\sim$ 10% (see Table 6.3 for exact fractions). Cat-7 has a large component of the halo mass in substructure at low radii because it has recently undergone a major merger (z = 0.03). In Figure 6-11d we plot the normalized number of subhalos as a function of radius scaled by the virial radius of the host. We find the scatter in the number of subhalos as a function of galactocentric distance is a factor of 3 across all halos except within the inner 10% of the host halo where we are subject to noise in the halo finding produced by ROCKSTAR. Again, in solid black we also plot the *Aquarius* Aq-A2 halo from Springel et al. (2008). We find the differences in the cosmology ($\sigma_8 = 0.9$) and in particular the slightly higher resolution of *Aquarius* leads to this systematic difference in the subhalo number density.

6.3.5 Too Big To Fail

We also examine halos which are massive enough to form stars but have no luminous counterpart in the nearby Universe (i.e., *the too big to fail* problem, hereafter TBTF, Boylan-Kolchin et al. 2011). To do this we select halos with $V_{\rm peak} > 30$ km s^{-1} which are subhalos large enough to retain substantial gas in the presence of an ionizing background and therefore theoretically should form stars. We follow the same definition as in Garrison-Kimmel et al. (2014a) to count two classes of halos. Strong massive failures are too dense to host any of the currently known bright MW classical dwarf spherioidals (dSph) galaxies. Massive failures (MFs) include all strong massive fail-

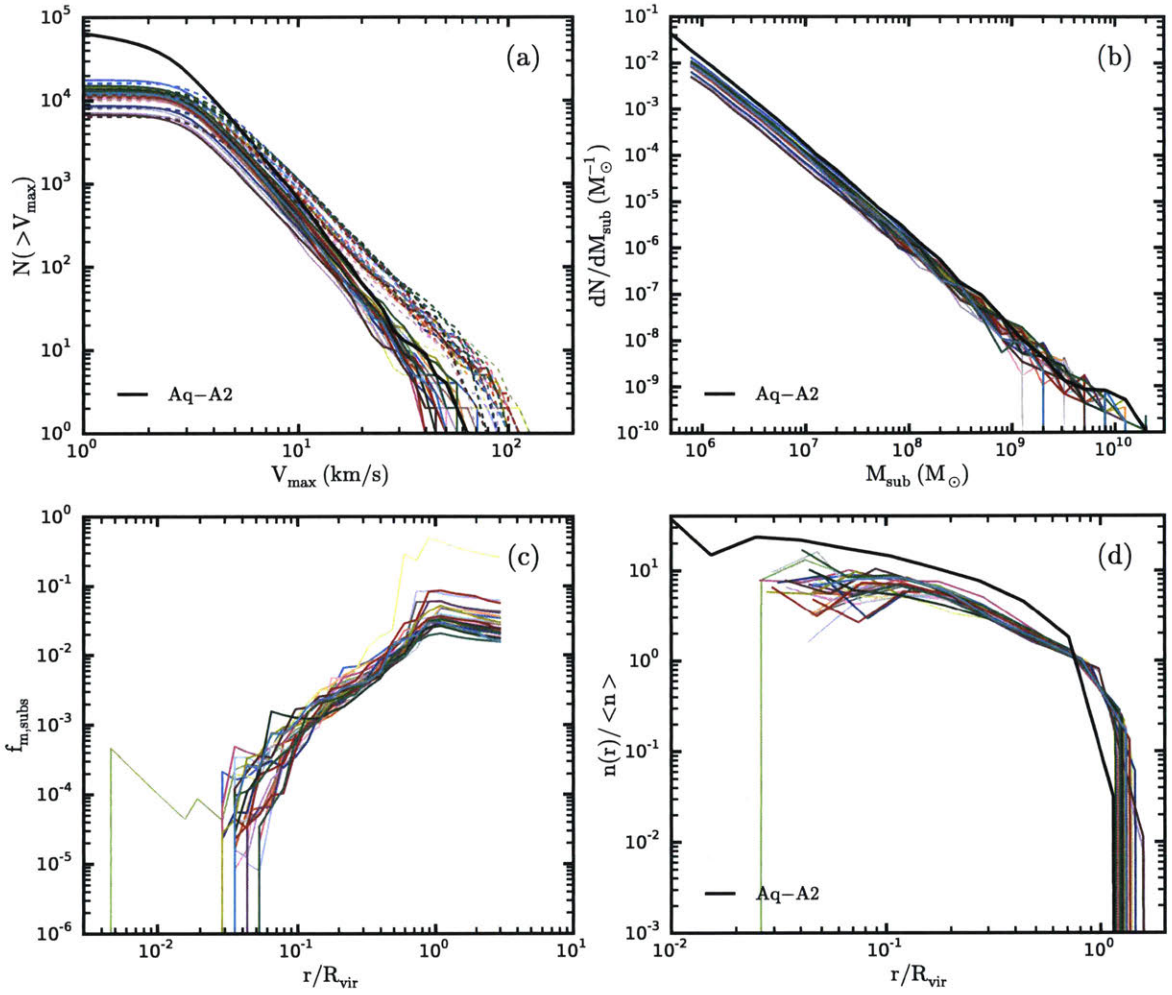

Figure 6-11 (a) The cumulative abundance of subhalos as a function of their maximum circular velocity ($V_{\rm max}$) for each of the *Caterpillar* halos. The solid lines represent subhalos at $z = 0$ and the dashed lines represent halos at infall (i.e., when they first cross the virial radius of the host). Many of the SMC/LMC sized systems ($M_{\rm peak} > 10^{11}$ M$_\odot$) are destroyed by $z = 0$, though some large subhalos do survive. The *Caterpillar* suite is complete down to $\sim$4 km s^{-1}. See our convergence study in the Appendix A. The solid black line is that of the Aquarius-A halo at a similar resolution (level-2, Springel et al. 2008). (b) The subhalo mass functions for each of the host halos. When normalized by mass, the range $10^{-6} - 10^{-4}$ M$_\odot$ there is small scatter in the subhalo abundances as all of our hosts are very similar in mass. Over the ranges 10^{6-8} M$_\odot$ we obtain a median mass function slope of -1.88 $\pm$ 0.1, slightly shallower than Aquarius ($\alpha \sim 1.90$). As we move to higher and lower mass regimes, there are differing abundances of large and small subhalos for each host. There are a number of systems with $\sim 10^{9-10}$ M$_\odot$ halos within the virial radius of the host making them possible Large/Small Magellanic Cloud analogues. Again, the solid black line is that of the Aquarius-A. (c) The subhalo mass fraction as a function of radius scaled by the virial radius of the host. The substructure mass is distributed similarly in nearly all halos with the exception of Cat–7 which has undergone a major merger. There is more variability in the substructure mass fraction at low radii owing to the on-going disruption most prevalent in the inner most dense regions of each host. (d) The normalized number of subhalos as function of radius scaled by the virial radius of the host. The number of subhalos as a function of distance is extremely self-similar across all halos except within the inner 10% of the virial radius of each host. This is in agreement with the findings of Springel et al. (2008) when factoring in their slightly higher resolution of the *Aquarius* suite and use of the observationally disfavoured high value of σ_8 ($\sigma_8 = 0.9$).

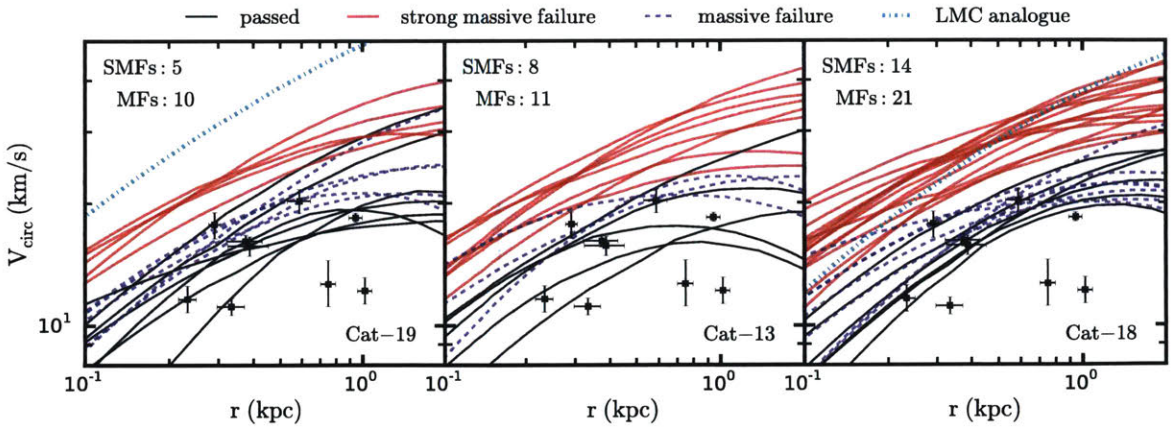

Figure 6-12 The rotation curves for all subhalos identified with a peak maximum circular velocity $V_{\rm peak}$, above 30 km s^{-1} within 300 kpc of three different *Caterpillar* hosts. The black squares indicate observational constraints from dwarf galaxies from Wolf et al. (2010). We only include those observed systems whose luminosities are above 2×10^5 L$_{\odot,V}$ (i.e., showing only classical dwarfs and excluding ultra-faint dwarf galaxies). Dashed-dot cyan lines are LMC analogues (i.e., $V_{\rm max} > 60$ km s^{-1}), blue dashed lines are massive failures (MFs) and red solid lines are strong massive failures (SMFs). Thin black solid lines are subhalos pass the check and have at least one observed dwarf with a comparable circular velocity (i.e., one halo passes through the circular velocity of a known classical dwarf determined by Wolf et al. 2010).

ures (SMFs) as well as all massive subhalos which have densities consistent with the high-density dSphs (i.e., Draco and Ursa Minor) but can not be associated with them without allowing a single dwarf galaxy to be hosted by multiple halos (i.e., assuming every observable dSph galaxy is hosted by exactly *one* halo). Most subhalos in the range of $V_{\rm max} = 25 - 30$ km s^{-1} could host a low-density dwarf and as such are not defined as massive failures.

In Figure 6-12, we plot a sample of the rotation curves for three different *Caterpillar* halos (Cat-19, Cat-13 and Cat-18). We adopt the Boylan-Kolchin et al. (2012) Einasto correction for $R < 291$ pc, which differ from the *ELVIS* profile fits in that they extrapolate their entire profile from $R_{\rm max}$ with various density profile shapes. Black squares depict circular velocities of the classical dwarf galaxies (with luminosities above 2×10^5 L$_{\odot,V}$), as measured by Wolf et al. (2010). Dashed-dot cyan lines are LMC analogues (i.e., $V_{\rm max} > 60$ km s^{-1} which are excluded from our failure analysis), blue dashed lines are massive failures and red solid lines are strong massive failures.

Thin black solid lines are subhalos which pass the test of having at least one observed dwarf with a comparable circular velocity, i.e., a circular profile goes through one of the observed dwarf galaxy data points. The cumulative number of profiles above and below the observed classical dwarfs for Cat-19 are 10 MFs, 5 SMFs and one LMC analogue. Similarly we find Cat-13 has 11 MFs and 8 SMFs. Cat-18 has the most failures of any halo with 21 MFs and 14 SMFs with one LMC analogue.

In Figure 6-13 we plot the number of strong and massive failures across all Caterpillar halos. Specifically we plot the fraction of hosts with fewer than N MFs and N SMFs within 300 kpc of each host as a function of N (black lines). Averaging over the entire Caterpillar sample (excluding Cat-7 as it has recently had a massive major merger), we predict 8 ± 3 (1σ) SMFs and 16 ± 5 (1σ) MFs within 300 kpc. If the Milky Way were well described by such an average we would expect to have these failures.

For comparison, we plot the *ELVIS* MF and SMF counts (red lines). The lower resolution in these simulations requires an extrapolation of the velocity profile from $V_{\rm max}$ and $R_{\rm max}$ using an analytic Einasto profile. Qualitatively, both simulation suites agree that there are a significant number of both MFs and SMFs. Quantitatively, there are several differences, which we now describe. The *Caterpillar* suite has many more MFs. This is due to our ability to better resolve high $V_{\rm peak}$ subhalos that have been tidally stripped. In particular, the iterative unbinding procedure described in Section 6.2.5 removes the need for the ROCKSTAR `unbound_threshold` parameter (Behroozi et al., 2013b). We can simulate the effect of the standard `unbound_threshold` $= 0.5$ cut by removing halos whose bound mass is less than 50% of their mass prior to unbinding. The MF counts with this cut are shown in Figure 6-13 by the blue lines, which are very similar to the *ELVIS* MF counts.

The *Caterpillar* suite also has significantly fewer SMFs compared to *ELVIS*. This discrepancy is likely due to the fact that we have measured rather than extrapolated the subhalo density profiles. Variations in the Einasto shape parameter (α) greatly affect the massive failure count (see Figure 4 of Garrison-Kimmel et al. 2014a). The Einasto fits to our density profiles have α typically closer to 0.2, which

is less discrepant.

Whilst TBTF is a prevalent problem in pure N-body simulations, many authors have indicated the tension between the circular velocities of observed classical dwarfs and simulated subhalos can be alleviated with the addition of supernovae feedback and ram pressure stripping (e.g., Pontzen and Governato 2012, Zolotov et al. 2012, Arraki et al. 2014, Brooks et al. 2013, Del Popolo et al. 2014, Gritschneder and Lin 2013, Elbert et al. 2015, Maxwell et al. 2015) or by making dark matter self-interacting (e.g., Vogelsberger et al. 2012, Zavala et al. 2013). Our results are within 1σ of the number of failures found by Garrison-Kimmel et al. (2014a) (i.e., 12 massive failures within 300 kpc), even when using a better density profile estimation.

6.4 Conclusions

In this work we have presented the first results of the *Caterpillar* simulation project, whose goal is to better understand the formation of Milky Way-sized galaxies and their satellite companions at both high and low redshift. We have carried out 24 initial simulations in a Planck based ΛCDM cosmology. Although the total halo number will increase to 60 – 70 shortly, these first 24 halos provide us with an exquisite initial set of data to achieve our first set of science goals. In our approach, we have taken exceptional care to validate our numerical techniques. We quadruple the current number of halos available in the literature at this extremely high mass and temporal resolution, allowing for detailed statistical studies of the assembly of Milky Way-sized galaxies. We additionally have adjusted our simulation parameters to be more inclusive of potential scientific questions not yet studied in simulations of this size (i.e., decreasing the temporal resolution to $\sim$5 Myrs/snapshot and increasing the volume resolved by high-resolution particles to 1–2 Mpc). The results presented above demonstrate our data quality and give initial clues at how halo properties vary across large numbers of realizations. Our initial key results can be summarized as follows:

1. Key halo properties such as the halo profile, mass functions and substructure

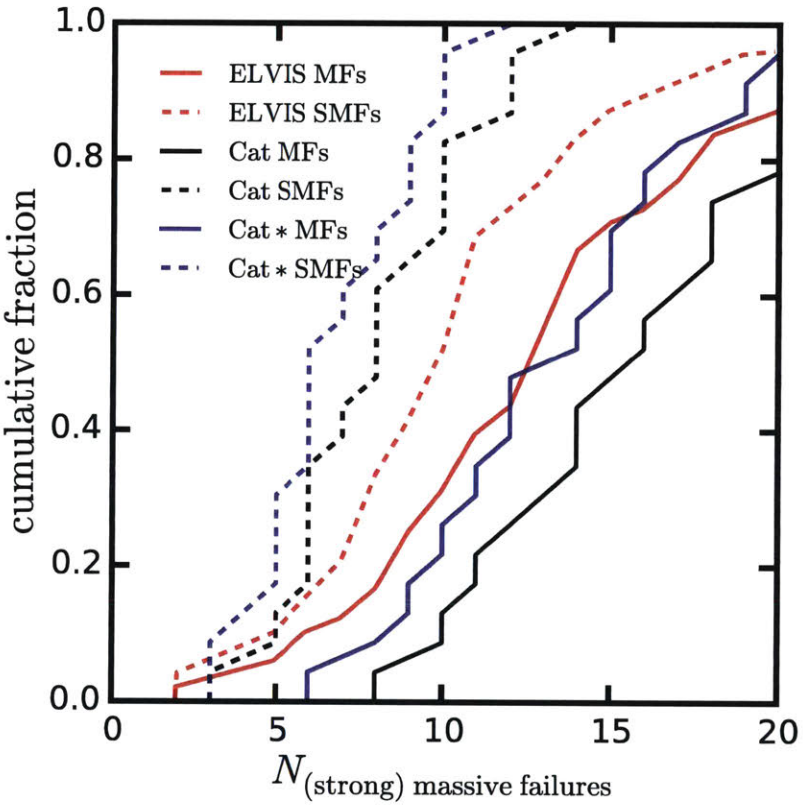

Figure 6-13 Cumulative fraction of massive failures (MFs) and strong massive failures (SMFs) across all *Caterpillar* halos (excluding Cat-7 which has undergone a recent massive major merger). We predict 8 ± 3 (1σ) strong massive failures and 16 ± 5 (1σ) massive failures within 300 kpc of the Milky Way. We also include the result from Garrison-Kimmel et al. (2014a) (i.e, using $\alpha = 0.15$ Einasto profiles) as a reference. The black lines are all failures which are detected by our version of ROCKSTAR which includes iterative unbinding. The blue lines are all failures found using a cut which mimics ROCKSTAR without iterative unbinding (see Section 6.3.5 text for details). This indicates a fraction of failures are undergoing tidal disruption which may have been unaccounted for in the *ELVIS* subhalo catalogues.

fractions are intimately connected to each halo's overall assembly history. Halos which have undergone recent major mergers have profiles which are poorly fit by either the NFW or Einasto profile. For those halos which are well fit by Einasto profiles, they have peak α values of 0.169 ± 0.023. Excluding the Cat-7 halo, we find a $Q_{min} = 0.027 \pm 0.010$ indicating reasonable agreement with Einasto fits of the *Aquarius* halos.

2. The abundance of dark matter subhalos remains relatively similar across our sample when normalized to host halo. As such, our halo mass functions are best fit by a simple power law, $\mathrm{d}N/\mathrm{d}M \propto M^{-1.88\pm0.10}$. The scatter in the normalizations of the mass functions is due to the concentration-subhalo abundance relation for fixed halo mass (i.e., our more concentrated halos exhibit lower normalizations for fixed M_{host}).

3. Regarding TBTF, dividing halos into two categories of massive failures and strong massive failures we predict 8 ± 3 (1σ) strong massive failures and 16 ± 5 (1σ) massive failures within 300 kpc of the Milky Way.

4. Iterative unbinding in ROCKSTAR must be included to properly recover all bound subhalos this resolution. We recover 52 halos above 10^8 $\mathrm{M}_\odot$ across a sample of 13 *Caterpillar* halos ($\sim$4 per host halo) using iterative unbinding which would have otherwise been lost using traditional ROCKSTAR.

This paper outlines the data products of the *Caterpillar* simulations and sets the foundation of many upcoming in-depth studies of the Local Group. Through our statistical approach to the assembly of Milky Way-sized halos we will gain a more fundamental insight into the origin and formation of the Galaxy, its similar sized cousins and their respective satellites.

Appendix A: Convergence Study

In Figure 6-14 we plot the halo profiles (Cat-2) and maximum circular velocity functions (Cat-9) at all our resolutions. We find our halos are well converged down to

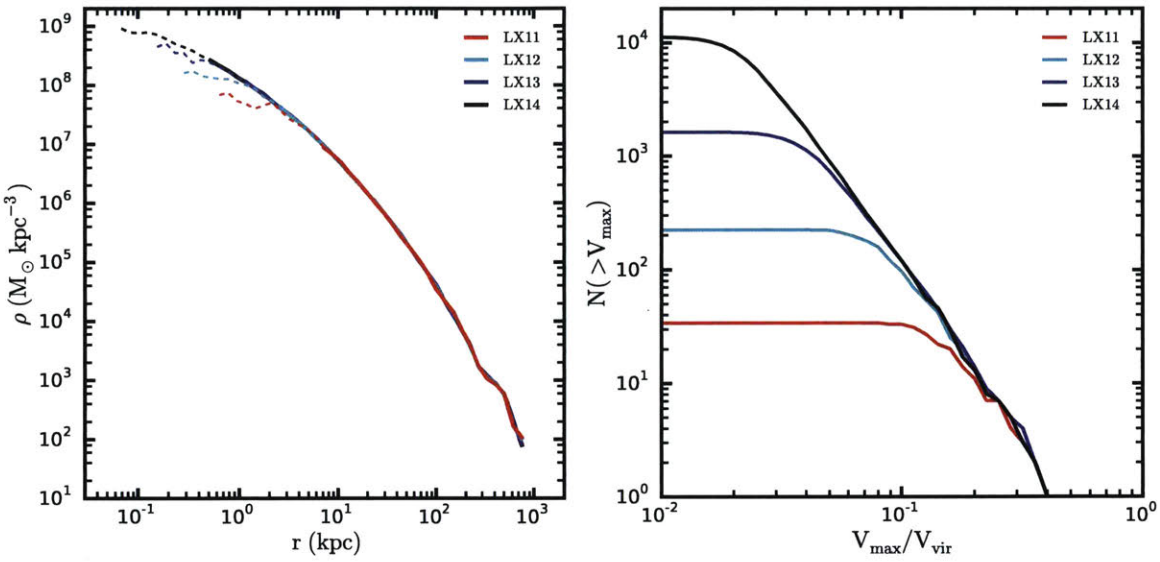

Figure 6-14 Convergence test for the Cat-2 and Cat-9 halos illustrating convergent halo profiles and $V_{\rm max}$ functions respectively. Panel (a) normalized halo profiles where the thick line represents the density profile above the radius defined by Power et al. (2003). The thin dashed line of the same color is the profile extrapolated down to the softening length. Panel (b) the cumulative abundance of subhalos as a function of their maximum circular velocity ($V_{\rm max}$) . These halos are representative of the convergence we achieve for all *relaxed Caterpillar* halos.

$\sim$0.2% of $R_{\rm vir}$. In the case of the $V_{\rm max}$ functions, we find we are converged down to $\sim$4 km s^{-1} at our highest resolution. When normalized to the host halo virial velocity, the halos are in excellent agreement with one another. Halos were re-simulated at successively higher and higher mass and spatial resolution from the initial parent volume. In each instance, care was taken to ensure all halo properties were numerically converged (provided that quantity was not resolution limited, e.g. halo shape). In Tables 6.5, 6.6, 6.7, and 6.8 we show the same quantities as in Table 6.3 from the text but this time include the lower resolution halo properties.

Table 6.5 Halo properties for the first set of 6 *Caterpillar* halos at each resolution.

Name	Geometry	$nR_{\rm vir}$	LX	$M_{\rm vir}$ ($M_\odot$)	$R_{\rm vir}$ (kpc)	c	$V_{\rm max}$ (km s^{-1})	$R_{\rm max}$ (kpc)	$z_{0.5}$	z_{lmm}	$f_{\rm m,subs}$	b/c	c/a	$R_{\rm res}$ (Mpc)
Cat-1	EA	5	11	1.579	307.690	7.762	172.293	35.049	0.881	2.118	0.161	0.810	0.828	1.391
			12	1.560	306.491	7.494	171.060	34.292	0.894	2.118	0.171	0.824	0.869	1.248
			13	1.560	306.458	7.647	170.707	36.451	0.894	2.157	0.197	0.842	0.883	1.138
			14	1.559	306.378	7.492	169.756	34.083	0.894	2.157	0.207	0.841	0.869	0.998
Cat-2	EB	4	11	1.807	321.876	7.621	176.924	64.931	0.742	0.719	0.092	0.596	0.724	1.577
			12	1.782	320.357	8.575	179.069	53.856	0.742	0.719	0.112	0.607	0.716	1.522
			13	1.792	320.970	8.382	178.753	54.360	0.742	0.731	0.137	0.643	0.731	1.480
			14	1.791	320.907	8.374	178.851	55.268	0.742	0.731	0.148	0.636	0.719	1.463
Cat-3	EB	4	11	1.343	291.538	10.763	175.066	26.554	0.802	0.790	0.079	0.961	0.971	1.966
			12	1.355	292.387	10.489	175.142	29.083	0.802	0.790	0.100	0.868	0.915	1.926
			13	1.355	292.400	10.523	172.946	31.565	0.802	0.802	0.117	0.850	0.905	1.906
			14	1.354	292.300	10.170	172.440	31.701	0.802	0.802	0.136	0.865	0.927	1.894
Cat-4	EB	4	11	1.503	302.676	8.308	169.309	91.132	0.894	0.908	0.128	0.749	0.825	1.791
			12	1.415	296.632	9.208	168.170	85.989	0.922	0.922	0.120	0.681	0.762	1.594
			13	1.434	298.009	8.434	164.999	59.225	0.936	0.922	0.156	0.673	0.743	1.561
			14	1.424	297.295	8.573	164.344	53.466	0.936	0.922	0.175	0.671	0.739	1.531
Cat-5	EB	4	11	1.306	288.846	11.897	173.913	33.844	0.584	0.519	0.025	0.551	0.835	1.676
			12	1.318	289.714	11.896	174.223	36.356	0.574	0.519	0.041	0.547	0.765	1.657
			13	1.314	289.450	12.324	176.818	29.346	0.574	0.519	0.055	0.556	0.825	1.617
			14	1.309	289.079	12.108	176.399	32.103	0.564	0.510	0.069	0.552	0.815	1.608
Cat-6	EB	4	11	1.371	293.516	10.373	172.100	32.944	1.144	1.275	0.094	0.495	0.534	1.708
			12	1.347	291.848	10.522	172.873	30.622	1.161	1.275	0.116	0.496	0.525	1.495
			13	1.366	293.186	10.086	170.858	32.794	1.161	1.295	0.138	0.510	0.529	1.300
			14	1.363	292.946	10.196	171.647	33.632	1.161	1.295	0.153	0.508	0.528	1.295

Notes: The resolution details for each refinement level (i.e. 11, 12, 13, 14) can be found in Table 6.2 and the geometry definitions in Table 6.1.

Table 6.6 Halo properties for the second set of 6 *Caterpillar* halos at each resolution.

Name	Geometry	$nR_{\rm vir}$	LX	$M_{\rm vir}$ ($M_\odot$)	$R_{\rm vir}$ (kpc)	c	$V_{\rm max}$ (km s^{-1})	$R_{\rm max}$ (kpc)	$z_{0.5}$	z_{lmm}	$f_{\rm m,subs}$	b/c	c/a	$R_{\rm res}$ (Mpc)
Cat-7	EB	4	11	1.142	276.168	2.513	139.055	140.859	0.065	0.057	0.693	0.191	0.301	1.756
			12	1.111	273.686	2.487	136.803	145.085	0.074	0.057	0.615	0.170	0.288	1.520
			13	1.091	272.009	1.674	133.574	162.291	0.065	0.036	0.693	0.168	0.235	1.510
			14	1.092	272.099	1.757	134.148	157.438	0.070	0.032	0.735	0.151	0.207	1.477
Cat-8	EB	4	11	1.729	317.150	13.081	198.577	46.800	1.541	2.195	0.032	0.602	0.768	1.690
			12	1.716	316.337	13.154	198.229	39.671	1.315	2.195	0.053	0.594	0.775	1.597
			13	1.701	315.450	13.340	197.637	39.810	1.516	2.235	0.066	0.599	0.791	1.550
			14	1.702	315.466	13.507	198.564	40.819	1.516	2.235	0.078	0.605	0.787	1.540
Cat-9	EB	4	11	1.330	290.616	12.568	177.522	32.309	1.236	1.217	0.050	0.493	0.762	2.383
			12	1.331	290.654	11.616	175.047	27.903	1.236	1.236	0.070	0.486	0.754	2.101
			13	1.329	290.538	12.132	176.808	30.297	1.255	1.236	0.085	0.500	0.754	1.833
			14	1.322	289.987	12.401	177.414	30.336	1.255	1.236	0.094	0.513	0.762	2.080
Cat-10	EB	4	11	1.319	289.809	11.902	175.553	41.894	1.699	2.010	0.052	0.561	0.709	1.983
			12	1.332	290.764	11.439	174.479	29.806	1.516	2.010	0.069	0.551	0.679	1.870
			13	1.328	290.477	11.714	175.124	25.839	1.644	2.010	0.088	0.559	0.703	1.740
			14	1.323	290.119	11.714	174.989	39.721	1.644	2.010	0.103	0.559	0.703	1.775
Cat-11	EB	4	11	1.194	280.361	10.551	165.980	62.881	1.059	4.368	0.175	0.527	0.719	1.490
			12	1.196	280.471	10.044	163.290	70.202	1.043	4.368	0.200	0.525	0.703	1.408
			13	1.190	280.043	12.272	173.893	45.727	1.059	1.644	0.199	0.590	0.868	1.192
			14	1.179	279.187	12.522	172.723	53.187	1.059	4.368	0.215	0.597	0.867	1.135
Cat-12	EA	5	11	1.786	320.627	11.723	191.564	59.256	1.336	2.542	0.034	0.592	0.724	1.664
			12	1.749	318.388	11.824	192.085	56.859	1.336	2.542	0.042	0.572	0.686	1.342
			13	1.767	319.441	11.663	191.320	49.435	1.336	2.542	0.062	0.571	0.703	1.239
			14	1.763	319.209	11.402	191.259	52.717	1.336	9.616	0.073	0.584	0.645	1.162

Notes: The resolution details for each refinement level (i.e. 11, 12, 13, 14) can be found in Table 6.2 and the geometry definitions in Table 6.1.

Table 6.7 Halo properties for the third set of 6 *Caterpillar* halos at each resolution.

Name	Geometry	$nR_{\rm vir}$	LX	$M_{\rm vir}$ ($M_\odot$)	$R_{\rm vir}$ (kpc)	c	$V_{\rm max}$ (km/s)	$R_{\rm max}$ (kpc)	$z_{0.5}$	z_{lmm}	$f_{\rm m,subs}$	b/c	c/a	$R_{\rm res}$ (Mpc)
Cat-13	EB	4	11	1.168	278.303	12.664	169.603	31.214	1.180	11.092	0.042	0.595	0.652	2.069
			12	1.171	278.509	12.979	170.750	31.408	1.161	14.748	0.063	0.575	0.634	1.742
			13	1.163	277.896	13.052	171.892	34.163	1.161	15.750	0.073	0.580	0.655	1.634
			14	1.164	277.938	12.850	171.222	33.757	1.161	11.092	0.090	0.578	0.645	1.566
Cat-14	EC	4	11	0.744	239.430	9.526	137.580	42.772	1.180	4.155	0.060	0.714	0.851	2.516
			12	0.757	240.865	8.854	136.512	27.875	1.144	4.258	0.086	0.709	0.849	2.301
			13	0.754	240.529	9.148	137.266	44.395	1.144	4.258	0.097	0.694	0.842	2.234
			14	0.750	240.119	9.135	137.437	26.660	1.144	4.258	0.113	0.705	0.859	2.178
Cat-15	EX	5	11	1.501	302.562	8.950	173.834	31.210	1.144	3.165	0.072	0.897	0.912	1.669
			12	1.497	302.281	9.223	174.792	33.832	1.144	3.165	0.089	0.897	0.926	1.630
			13	1.504	302.755	9.077	174.431	36.520	1.144	3.165	0.111	0.837	0.861	1.597
			14	1.505	302.787	8.983	174.124	37.043	1.144	3.165	0.126	0.849	0.877	1.119
Cat-16	EB	4	11	0.993	263.614	10.997	154.748	42.280	1.315	3.165	0.053	0.567	0.765	1.406
			12	0.976	262.082	12.099	156.589	28.820	1.315	3.165	0.072	0.593	0.791	1.393
			13	0.980	262.447	11.888	156.193	29.498	1.315	3.165	0.088	0.597	0.766	1.384
			14	0.982	262.608	11.737	155.362	28.768	1.315	3.165	0.106	0.618	0.792	0.671
Cat-17	EX	4	11	1.311	289.204	13.216	178.671	38.818	1.846	1.943	0.038	0.646	0.794	1.525
			12	1.314	289.456	12.906	178.676	39.713	1.846	1.943	0.057	0.680	0.875	1.427
			13	1.329	290.487	12.505	178.763	38.717	1.846	1.976	0.084	0.657	0.863	1.333
			14	1.319	289.800	12.765	179.056	38.329	1.846	1.976	0.093	0.664	0.881	1.299
Cat-18	EX	4	11	1.428	297.536	7.909	167.184	32.058	0.451	0.427	0.100	0.677	0.847	1.491
			12	1.414	296.559	7.861	164.702	48.041	0.459	0.412	0.123	0.720	0.840	1.397
			13	1.400	295.596	7.823	165.164	40.766	0.493	0.435	0.141	0.622	0.712	1.228
			14	1.407	296.099	7.887	163.920	57.217	0.493	0.435	0.159	0.676	0.816	0.397

Notes: The resolution details for each refinement level (i.e. 11, 12, 13, 14) can be found in Table 6.2 and the geometry definitions in Table 6.1.

Table 6.8 Halo properties for the fourth set of 6 *Caterpillar* halos at each resolution.

Name	Geometry	$nR_{\rm vir}$	LX	$M_{\rm vir}$ ($M_\odot$)	$R_{\rm vir}$ (kpc)	c	$V_{\rm max}$ (km/s)	$R_{\rm max}$ (kpc)	$z_{0.5}$	z_{lmm}	$f_{\rm m,subs}$	b/c	c/a	$R_{\rm res}$ (Mpc)
Cat-19	EX	5	11	1.179	279.143	10.467	164.816	34.292	1.566	2.693	0.113	0.640	0.857	1.933
			12	1.174	278.788	10.158	163.679	34.514	1.566	2.693	0.132	0.668	0.919	1.861
			13	1.177	279.002	10.139	163.868	30.433	1.541	2.118	0.149	0.672	0.933	1.800
			14	1.174	278.770	10.468	164.726	29.112	1.541	2.118	0.169	0.664	0.937	1.712
Cat-20	BB	4	11	0.765	241.720	13.409	150.030	25.189	1.516	5.588	0.045	0.608	0.743	1.677
			12	0.756	240.683	13.443	148.881	27.312	1.541	5.588	0.053	0.634	0.775	1.521
			13	0.761	241.208	13.456	149.682	29.340	1.516	5.761	0.084	0.613	0.752	1.377
			14	0.763	241.484	13.324	149.672	30.417	1.492	5.427	0.099	0.601	0.733	1.311
Cat-21	EX	4	11	1.865	325.250	11.820	193.253	42.842	1.144	1.198	0.042	0.456	0.584	1.551
			12	1.876	325.890	10.950	191.015	54.116	1.109	1.161	0.075	0.475	0.637	1.426
			13	1.889	326.663	10.465	189.607	57.507	1.126	1.198	0.103	0.472	0.590	1.342
			14	1.881	326.206	10.618	190.683	50.954	1.126	1.198	0.118	0.482	0.611	1.453
Cat-22	EX	5	11	1.560	306.489	9.356	177.811	33.807	0.828	5.940	0.044	0.496	0.643	2.003
			12	1.594	308.677	9.799	181.703	44.919	0.790	5.940	0.052	0.461	0.637	1.903
			13	1.497	302.257	10.655	180.773	37.743	0.854	5.940	0.068	0.518	0.695	1.837
			14	1.495	302.116	10.666	180.647	35.860	0.841	29.488	0.080	0.512	0.694	1.744
Cat-23	EX	4	11	1.608	309.596	11.989	189.267	33.023	1.180	10.062	0.051	0.635	0.845	1.623
			12	1.604	309.328	12.865	191.457	32.232	1.180	9.616	0.071	0.589	0.729	1.236
			13	1.613	309.926	12.135	190.191	31.524	1.161	9.616	0.080	0.602	0.763	1.245
			14	1.607	309.524	12.489	190.705	32.421	1.161	9.616	0.094	0.607	0.784	1.207
Cat-24	EB	4	11	1.329	290.537	11.152	174.259	43.136	1.217	2.801	0.038	0.651	0.705	1.260
			12	1.323	290.054	11.326	175.088	48.435	1.217	2.801	0.052	0.645	0.674	1.396
			13	1.335	290.969	11.490	177.313	34.438	1.144	2.801	0.077	0.675	0.721	1.190
			14	1.334	290.866	11.378	176.911	36.800	1.144	3.608	0.090	0.689	0.734	1.102

Notes: The resolution details for each refinement level (i.e. 11, 12, 13, 14) can be found in Table 6.2 and the geometry definitions in Table 6.1.

Chapter 7

Tracing the first stars and galaxies of the Milky Way

The content of this chapter was submitted to the Monthly Notices of the Royal Astronomical Society *on November 2nd, 2016. It has an arXiv location of* `https://arxiv.org/abs/1611.00759`. *The authors are Brendan F. Griffen, Gregory A. Dooley, Alexander P. Ji, Brian W. O'Shea, Facundo A. Gómez, and Anna Frebel.*

Disclaimer

The majority of the writing and analysis in this chapter was conducted by Dr. Brendan Griffen. Alex Ji and I worked alongside Brendan providing frequent feedback and writing pieces of analysis code. Foremost among these was writing and running the code to traverse the halo merger trees in order to identify the formation of the first minihalos and atomic cooling halos, as well as to track them to the present day. Moreover, I provided code to extract properties of these halos at times of evolutionary interest, and wrote and ran the code to tag and trace the most centrally bound particles in halos. I additionally contributed reviews and edits of the text in the paper.

Acknowledgements

BG would like to thank Paul Hsi for assistance with the compute cluster at MKI. He would also like to thank Bhaskar Agarwal, Andy Casey and Joss Bland-Hawthorn for helpful discussions. The authors thank Oliver Hahn for making the initial conditions code, MUSIC, publicly available. The authors also thank Volker Springel for making GADGET-2 publicly available and for providing a version of GADGET-3/GADGET-4 for our use. The authors thank Peter Behroozi for making ROCKSTAR and CONSISTENT-TREES publicly available and additionally thank him for technical support in modifying ROCKSTAR.

Support for this work was provided by XSEDE through the grants (TG-AST120022, TG-AST110038). BG and AF acknowledges support of the compute cluster of the Astrophysics Division which was built with support from the Kavli Investment Fund administered by the MIT Kavli Institute for Astrophysics and Space Research. GD acknowledges support by NSF Grant 1122374. BWO and FG were supported through the NSF Office of Cyberinfrastructure by grant PHY-0941373 and by the Michigan State University Institute for Cyber-Enabled Research (ICER). BWO was supported in part by NSF grant PHY 08-22648 (Physics Frontiers Center/Joint Institute for Nuclear Astrophysics) and NSF Grant PHY-1430152 (JINA Center for the Evolution of the Elements), and by NASA through grants NNX12AC98G, NNX15AP39G, and Hubble Theory Grants HST-AR-13261.01-A and HST-AR-14315.001-A. AF acknowledges support from the Silverman (1968) Family Career Development professorship.

Abstract

We use 30 high-resolution dark matter halos of the *Caterpillar* simulation suite to probe the first stars and galaxies of Milky Way-mass systems. We quantify the environment of the high-z progenitors of the Milky Way and connect them to the properties of the host and satellites today. We identify the formation sites of the first generation of Population III (Pop III) stars ($z \sim 25$) and first galaxies ($z \sim 22$) with several different models based on a minimum halo mass including a simple model for Lyman-Werner feedback. Through this method we find approximately 23,000 $\pm$ 5,000 Pop III potentially star-forming sites per Milky Way-mass host, though this number is drastically reduced to $\sim$550 star-forming sites when Lyman-Werner feedback is

included, as it has critical effects at these length scales. The majority of these halos identified form in isolation (96% at $z = 15$) and are not subject to external enrichment by neighboring halos (median separation $\sim$1 pkpc at $z = 15$), though half merge with a system larger than themselves within 1.5 Gyrs. Approximately 55% of the entire population has merged into the host halo by $z = 0$. Using particle tagging, we additionally trace the Pop III remnant population to $z = 0$ and find an order of magnitude scatter in their number density at small (i.e. r $<$ 5 kpc) and large (i.e. r $>$ 50 kpc) galactocentric radii at $z = 0$. Using our large number of realizations, we provide fitting functions for determining the number of progenitor minihalo and atomic cooling halo systems that present-day dwarf galaxies and the Magellanic cloud system might have accreted since their formation. We demonstrate that observed dwarf galaxies with stellar masses below $10^{4.6}$ $M_{\odot}$ are unlikely to have merged with any other star-forming systems.

7.1 Introduction

The epoch of the first stars and first galaxies remains a poorly understood period in the Universe's history, although it is broadly known how the first billion years unfolded. Following recombination (z$\sim$1100), small scale density fluctuations collapsed into dark matter halos containing gaseous material capable of molecular hydrogen (H_2) cooling. Once gas densities in these "minihalos" were sufficiently high, the first stars, Population III (Pop III), were able to form, thus marking the end of the so-called "Dark Ages" (Tegmark et al., 1997).

These Pop III stars were predominantly massive (e.g. Bromm et al. 1999) and thus exploded soon after formation as supernovae (SNe). Their deaths produced vast quantities of ionizing radiation and metals, impacting the conditions for subsequent star formation. The metals may have "cross-polluted" nearby minihalos (e.g. Whalen et al. 2008; Smith et al. 2015), and in the case of pair-instability supernovae (PISNe) enriched vast volumes of the early inter-galactic medium (IGM) (Umeda and Nomoto, 2003; Whalen et al., 2004; Kitayama and Yoshida, 2005).

This enriched and ionized environment set a blueprint for more massive galaxies which assembled soon after (Greif et al., 2007; Wise and Abel, 2008; O'Shea et al., 2015). As structure formation progressed ($z \sim 25$, Greif et al. 2008), these more massive dark matter halos (10^8 $M_{\odot}$) formed with gas that could cool via atomic hydrogen

($T_{\rm vir} \sim 10^4\,{\rm K}$) and are widely regarded as the "first galaxies," heralding the first period of galaxy formation (Bromm and Yoshida, 2011). While remarkable progress has been made in arriving at this broad picture of early structure formation, many of the details of both first star and first galaxy formation and associated chemical and physical processes remain elusive (see Frebel and Norris 2015 for a review).

Observational access to these critical periods is quite limited. The optical depth to reionization derived from the cosmic microwave background provides a global constraint on reionization . Other studies have used very deep images of high-z sources (e.g. Finkelstein et al. (2015); Sobral et al. (2015)) or absorption from high-z quasars (e.g. Becker et al., 2015) to study the brightest objects at $6 \lesssim z \lesssim 10$. Future data from the *James Webb Space Telescope* (Gardner et al., 2006) or 21 cm cosmology (Pritchard and Loeb, 2012) will provide valuable further constraints.

One of the best ways to study this early period is by examining local objects. For instance, star formation histories of local group galaxies, and chemical abundances of metal-poor stars in the Milky Way stellar halo (stellar archaeology) or in its satellite dwarf galaxies (dwarf galaxy archaeology). A key step in interpreting these local observations is developing a principled method of connecting low-z stellar systems to their high-z progenitors.

Cosmological simulations have been used extensively to study the non-linear regime of structure formation, but few have been able to resolve and follow the smallest building blocks, which formed in the early universe, to the present day. Indeed, it is still not well understood how many progenitor systems made up the Milky Way nor do we know where they formed and, perhaps most importantly, where any can be found today if they survived (see Frebel 2010 for a review). This is primarily due to the high redshift universe being studied from only two vantage points, either (a) moderately large nondescript volumes (e.g. Ricotti and Shull 2000; Ishiyama et al. 2013; Ishiyama et al. 2016; Sasaki et al. 2014; O'Shea et al. 2015) or (b) small volumes encapsulating single host halos at extremely high-resolution (e.g. Smith et al. 2015; Stacy et al. 2016). By virtue of the computational cost of carrying out hydrodynamical simulations at the required resolutions to resolve the first stellar systems, these calculations

usually only run to relatively high redshifts ($z \sim 10$), prohibiting our ability to test them against local observational data. Currently, exclusively dark matter-only simulations are capable to resolve both the minihalo progenitors ($z > 15$) of the Milky Way *and* to trace their evolution to $z = 0$ (e.g. Springel et al. 2008; Diemand et al. 2008; Griffen et al. 2016b). Until the required hardware and hydrodynamic codes with the necessary physics become available, minihalos and atomic cooling halos have to be modeled semi-analytically (i.e. using the halo properties derived from halo finders to determine the nature of their gas and stellar content).

Indeed, there have been a number of works which have made attempts to connect the high-z universe to the present day via either semi-analytical methods or direct N-body simulations. All of them, however, suffer from at least one drawback, either (a) they contain no spatial information about where the high-z star forming halos reside today (e.g. Hartwig et al. 2015, Smith et al. 2015), or (b) they do not model the critical influence of Lyman-Werner (LW) feedback on the first stellar systems (e.g. Gao et al. 2010), or (c) they lack statistical power to investigate halo-to-halo scatter (e.g. Tumlinson 2010; Bovill and Ricotti 2011b; Corlies et al. 2013; Ishiyama et al. 2016).

In this paper, we address these three issues directly by exploiting the high-resolution cosmological dark matter simulations of the *Caterpillar Project* (www.caterpillarproject.org, Griffen et al. 2016b). Specifically, this work has the following properties which combined is the first of its kind: (a) spatial information about the formation sites and their subsequent evolution to $z = 0$ (b) a model for LW feedback on the first stellar systems and (c) 30 realizations allowing robust halo-to-halo variations to be studied. We adopt simple models to identify the sites of first star and first galaxy formation and include a toy model for chemical enrichment which allows us to separate halos with metal enrichment driven exogenously (externally) and endogenously (internally). We trace our candidate minihalos and first galaxy halos using their most bound particles to determine where their progenitors are today. This approach connects the high-z star formation processes to surviving stars in low-z environments today (e.g. dwarf galaxies and the halo), probes the building

blocks of the Milky Way's metal-poor stellar halo, assists in the hunt for the surviving relics from a unique period of our Galaxy's assembly history, and informs how differing formation histories of similarly-sized galaxies can affect observable properties of metal-poor stellar populations. This is the first time that first star and first galaxy formation is studied with such a wide sample of simulations within the context of the entire Milky Way host assembly.

In Section 7.2, we describe our numerical simulation suite and method for identifying and tracking dark matter halos. In Section 7.3, we present our model for Pop III and Pop II star formation, including our method of treating the LW background. Our results are presented in Section 7.4 where we detail the clustered nature of the high-z progenitor Milky Way and highlight how this critically impacts the present day abundance of possible surviving stellar populations. We additionally discuss these results in the context of the recent discovery of r-process enhanced metal-poor stars inside one of today's ultra-faint dwarf galaxies. Section 7.5 provides our final concluding remarks and avenues for future work.

7.2 Simulations

We analyze 30 of the dark matter only cosmological halos of the *Caterpillar Project* first presented in Griffen et al. (2016b). Each of the 30 halos in our sample are similar in mass to that of the Milky Way and come from a somewhat isolated environment (no nearby clusters). The halos were identified from a larger parent simulation which followed the growth of structure in a periodic box of comoving length 100 h^{-1} Mpc with 1024^3 particles ($m_p = 1.22 \times 10^7$ M$_\odot$). For the underlying cosmological model we adopt the ΛCDM parameter set characterized by a *Planck* 2013 cosmology given by $\Omega_m = 0.32$, $\Omega_\Lambda = 0.68$, $\Omega_b = 0.05$, $n_s = 0.96$, $\sigma_8 = 0.83$ and Hubble constant, H = 100 h km s^{-1} Mpc^{-1} = 67.11 km s^{-1} Mpc^{-1} (Planck Collaboration et al., 2014). All initial conditions were constructed using MUSIC (Hahn and Abel 2011). We identify dark matter halos via a modified version of ROCKSTAR (Behroozi et al., 2013b) and construct merger trees using CONSISTENT-TREES (Behroozi et al., 2012). ROCKSTAR

assigns virial masses to halos, $M_{\rm vir}$, using the evolution of the virial relation from Bryan and Norman (1998) for our particular cosmology. At $z = 0$, this definition corresponds to an over-density of 104 $\times$ the critical density of the Universe. The temporal resolution is $\sim$5 Myrs/snapshot to $z = 6$ and $\sim$50 Myrs to $z = 0$.

Caterpillar halos are zoom-in re-simulations of a parent volume. Particular care was taken to ensure that we restrict our study to only the high-resolution volume of the Milky Way at $z > 10$ and that no halos were contaminated. Indeed in all simulations of this kind, there will be halos in the catalogues which contain lower resolution particle types, particularly near the fringe of the high-resolution region. These halos have poorly determined virial masses and internal velocity dispersions, so they are excluded from our analysis (<1% of the total halo population on average). None of these contaminated halos end up anywhere near the host of the central Milky Way-mass system at $z = 0$.

The dark matter particle mass of the fiducial *Caterpillar* simulation suite is 2.99 $\times$ 10^4 $M_\odot$, resolving halos with masses of 10^6 $M_\odot$ ($\sim$30 particles). Whilst properties such as the velocity dispersion are not converged at the resolution limit, the total mass of the system is reliably determined Power 2013.

We carried out a convergence check of this assumption (see Appendix A) using an even higher resolution run with a particle mass of 3.73 $\times$ 10^3 $M_\odot$. We find that the total number of systems identified between our fiducial run and our ultra-high-resolution counterpart is convergent.

7.3 Modelling The Sites Of High-Redshift Star Formation

To determine which dark matter halos host stellar material and later accrete into the Milky Way, we must consider the nature of star formation in the early Universe. Here, we take a simple approach to modelling star formation sites based on more detailed theoretical work.

Table 7.1 Properties of the 30 *Caterpillar* halos used in analysis

Name	$M_{\rm vir}{}^a$ ($M_\odot$)	$R_{\rm vir}{}^b$ (kpc)	c^c	$V_{\rm max}{}^d$ (km/s)	$z_{0.5}{}^e$
Cat-1	1.559	306.381	7.491	169.760	0.894
Cat-2	1.791	320.906	8.370	178.850	0.742
Cat-3	1.354	292.307	10.168	172.440	0.802
Cat-4	1.424	297.288	8.574	164.340	0.936
Cat-5	1.309	289.081	12.108	176.400	0.564
Cat-6	1.363	292.947	10.195	171.650	1.161
Cat-8	1.702	315.466	13.505	198.560	1.516
Cat-9	1.322	289.987	12.401	177.410	1.255
Cat-10	1.323	290.116	11.715	174.990	1.644
Cat-12	1.763	319.212	11.401	191.260	1.336
Cat-13	1.164	277.938	12.850	171.220	1.161
Cat-14	0.750	240.119	9.135	137.440	1.144
Cat-15	1.505	302.788	8.984	174.120	1.144
Cat-16	0.982	262.608	11.737	155.360	1.315
Cat-17	1.319	289.800	12.765	179.060	1.846
Cat-18	1.407	296.100	7.886	163.920	0.493
Cat-19	1.174	278.771	10.467	164.730	1.541
Cat-20	0.762	241.387	13.376	149.150	1.492
Cat-21	1.882	326.206	10.618	190.680	1.126
Cat-22	1.495	302.114	10.666	180.650	0.841
Cat-23	1.608	309.525	12.489	190.710	1.161
Cat-24	1.334	290.867	11.378	176.910	1.144
Cat-25	1.648	312.153	12.970	191.690	1.126
Cat-26	1.018	265.828	8.130	147.960	0.555
Cat-27	1.357	292.557	7.035	159.730	0.719
Cat-29	1.594	308.698	10.646	182.810	0.980
Cat-31	1.678	313.967	12.461	191.710	1.516
Cat-33	1.675	313.855	13.322	197.710	1.878
Cat-36	1.974	331.521	10.282	191.890	0.966
Cat-37	1.848	324.250	12.854	197.950	1.492

a: Halo virial mass based on Bryan and Norman (1998).
b: Halo virial radius based on Bryan and Norman (1998).
c: Concentration defined by ratio of the virial radius and the scale radius; $R_{\rm vir}/R_{\rm s}$.
d: Maximum of the halo's circular velocity.
e: Redshift at which half the mass of the host has formed.

Structure formation within ΛCDM proceeded first within small dark matter halos forming at early times and merging into larger halos. There are two periods which are significant for star formation at these early times and they both relate to the cooling mechanisms in metal-poor gas. The first of these periods is when star formation proceeds within dark matter halos of mass $\sim 10^6$ $M_\odot$, in which molecular hydrogen cooling is dominant (e.g. Tegmark et al. 1997). The second important period of star formation occurs when the gas within larger halos of mass $\sim 10^8$ $M_\odot$ are able to cool via atomic line cooling (Oh and Haiman, 2002). In the following two sections we outline how we model these two periods of first generation (Pop III) and second generation (Pop II) star formation.

7.3.1 H_2 Cooling

Pop III stars are by definition metal-free, and as such can only form in a minihalo with sufficient H_2 at the appropriate temperature and density to become gravitationally unstable and collapse (Tegmark et al., 1997; O'Shea and Norman, 2007). We assume the gas is in virial equilibrium with the dark matter halo so we can infer the gas temperature from the dark matter virial mass. The minimum temperature required for H_2 cooling to cause gas collapse (Tegmark et al., 1997) thus corresponds to a minimum halo mass that determines possible sites of Pop III star formation.

We identify halos in our merger tree when they first grow above the minimum threshold for collapse. We additionally ensure that none of the progenitors on any branch that merged into a candidate halo were above the temperature threshold.

A critical feature required of a simulations attempting to identify minihalo candidates is the time between each snapshot used by the halo finder. To estimate whether we might be underestimating the number of candidate halos we compare the free-fall time of gas to our temporal resolution. We estimate that the free fall time of gas is $t_{ff} \sim 0.1 * H(z)$, which for $z = 25$ is $\sim$20 Myr, and $z = 10$ is $\sim$70 Myr. Since our temporal resolution is $\sim$5 Myr between each snapshot, we are not under-counting any halos but we may be *over*-counting. By comparison, *Aquarius* has a temporal resolution at these early times of $\sim$100 Myr which means Gao et al. may have under-

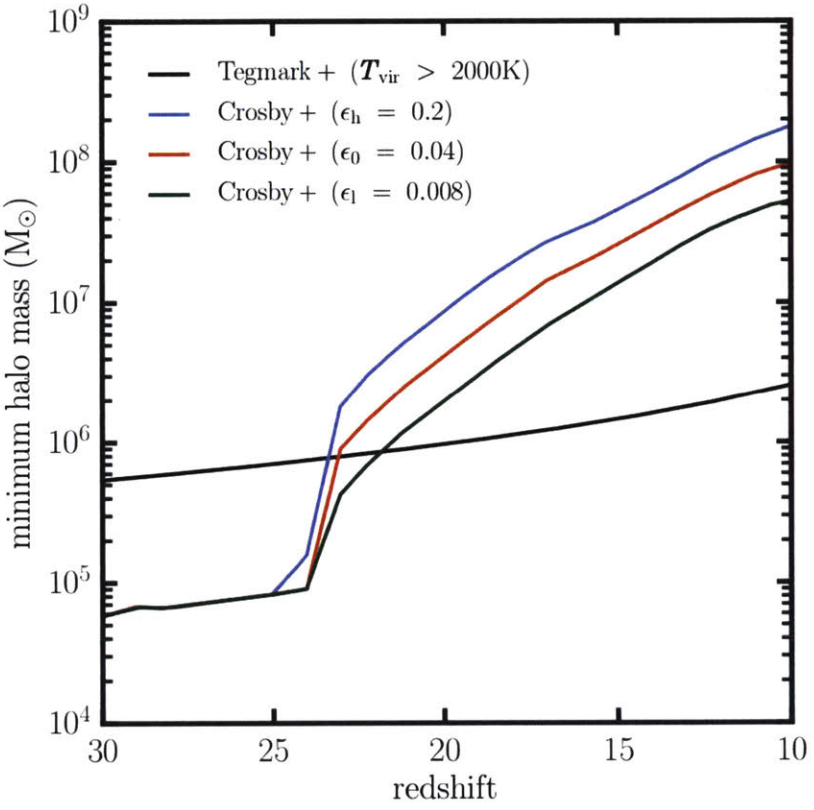

Figure 7-1 Minimum halo mass required for Pop III star-formation to proceed. We adopt two minimum mass thresholds for minihalo formation, one of which contains three variations of same semi-analytic model. Our first model is based on Tegmark et al. (1997). It requires that the H_2 cooling time is less than a Hubble time. For our second model, we interpolate the semi-analytic model of Crosby et al. (2013) which includes LW radiation produced by the first generations of stars in nearby halos at $z > 20$. This results in a raising of the minimum mass depending on the initial mass function adopted. In all models, progenitor halos are also checked to ensure that a candidate halo identified is the first in its history to go above the cooling threshold. The increase in the minimum mass at $z = 24$ is due to the onset of Pop II star formation within the Crosby et al. (2013) model.

counted the number of minihalos forming. If we assume ~50 Myr is approximately the collapse time, and then compare this to the outer panel of Figure 7-6, we find ~5% of minihalos merge within 50 Myrs, and ~10% within 100 Myrs, meaning that we could be over-counting by ~5%, and work using the *Aquarius* simulation will have undercounted by 5%.

7.3.2 LW Feedback

The minimum mass for collapse will be boosted to higher masses with the onset of LW radiation from Pop III stars which will photo-dissociate H_2 via the reaction $H_2 + \gamma_{LW} \rightarrow H + H$, where γ_{LW} is a photon in the LW band of 11.12 – 13.6 eV. We must include this form of feedback in our model if we are to reliably determine which minihalo candidates represent the actual *star-forming* halos at $z > 10$. Accordingly, we model the influence of a LW background via the semi-analytic model constructed by Crosby et al. (2013). We *do not* explicitly calculate the relevant LW flux for each halo's stellar population, but simply adopt the adjusted minimum mass threshold for forming Pop III stars after including LW feedback. The Crosby et al. (2013) model was based on simulations carried out using ENZO, an adaptive mesh refinement + N-body code. We have not carried out any simulation specific to our simulated volume but interpolate the minimum mass threshold they determined. Within their model, they followed 10 chemical species (H, C, N, O, Mg, Ca, Ti, Fe, Co, and Zn) in both the stellar and interstellar medium (ISM) components of every halo. The ISM was treated as a multiphase gas with a central region of dense, cold gas that is capable of forming stars and a hot, diffuse region exterior to the star-forming central region that is incapable of forming stars. For more details see the work of Crosby et al. (2013); Crosby et al. (2013).

In Figure 7-1, we show the minimum host halo virial mass required as determined by Tegmark et al. (1997) in order to cool to its virial temperature via H_2 cooling in the local Hubble time in Crosby et al. (2013) then adopt three star formation efficiencies of $\epsilon = 0.008, 0.04, 0.2$ (hereafter ϵ_l, ϵ_0, ϵ_h) which adjust the minimum mass thresholds for collapse due to differing quantities of LW flux (J_{21}). Crosby et al. (2013) adopted

three different IMFs but since the star formation efficiency drives the Lyman-Werner flux over any particular selection of IMF we adopt parameterizations of three of their models distinguished by their star formation efficiencies only. The minimum mass thresholds we adopt for identifying Pop III star forming regions after including LW feedback are shown in Figure 7-1 (identical to Fig. 6 in Crosby et al. 2013). The increase in the minimum mass threshold is particularly pronounced at z $\sim$ 24 where the onset of Pop II star formation from chemically enriched gas makes Pop II stars the dominant component of the stellar mass.

7.3.3 Population II Star Formation

In this section we describe criteria used for the formation of Pop II stars, which we assume form in the first galaxies.Once the virial temperature of the halo is high enough, atomic line cooling becomes important ($T_{\rm vir} \sim 10^4$ K). These halos are likely the sites of the first galaxies (Bromm and Yoshida, 2011) and as such we refer to all "atomic cooling halos" (ACHs) as first galaxies and vice-versa. The gas inflow rate into these systems largely traces the rate of inflow of the dark matter accretion rate, but this can be suppressed in the presence of an ionizing background. We adopt a simple model of reionization following Bullock and Johnston (2005) whereby we divide atomic cooling halos into three populations based on their maximum circular velocity at the redshift set for reionization (assumed to be instantaneous at $z_{\rm re} = 10$); (1) ACHs with $V_{\rm max}$ (z=10) $> 50\,{\rm km\ s^{-1}}$ are not suppressed, (2) ACHs with $30\,{\rm km\ s^{-1}}$ $< V_{\rm max}$ (z=10) $< 50\,{\rm km\ s^{-1}}$ are partially suppressed (i.e., not all of their cold gas is star forming) and (3) ACHs with $V_{\rm max}$ (z=10) $< 30\,{\rm km\ s^{-1}}$ are completely suppressed (Thoul and Weinberg, 1996).

7.3.4 Simple Chemical Enrichment Model

After the accretion and collapse of cool gas in the central reservoir of a conducive halo, star formation proceeds, with the mass of each star set by the initial mass function (IMF). Some high-mass stars will eventually produce extremely energetic events such

as pair-instability supernovae (PISN), whereby originally bound gas can be nearly entirely ejected (e.g. Whalen et al. 2004, 2008; Kitayama and Yoshida 2005). If any of the proto-Milky Way's star forming regions were extremely clustered, this ejecta could likely pollute neighbouring halos and result in enhanced metal-line cooling spurring on subsequent star formation (e.g. Smith et al. 2015). Detailed modelling of metal-enrichment of the subsequently formed first galaxies in these clustered environments shows that they can become significantly enriched to average metallicities of $Z > 10^{-3}$ $Z_\odot$ (Greif et al., 2010; Safranek-Shrader et al., 2014). This inhomogeneous process can result in large spreads in chemical abundances of two to three orders of magnitude across the host system (Wise and Abel, 2008).

By definition, the first stars form out of chemically pristine gas. However, supernovae spew metals into the IGM (e.g. Madau et al., 2001; Greif et al., 2007; Jeon et al., 2014; Smith et al., 2015; O'Shea et al., 2015), and in some cases they can contaminate nearby minihalos that would otherwise be pristine (e.g. Smith et al., 2015) (probably others). This separates minihalos into *endogenous* and *exogenous* minihalos, i.e. those that are initially unaffected by supernovae, and those that are externally enriched. In principle, this effect reduces the number of minihalos that should be considered as sites of Pop III star formation.

A complete characterization of metal pollution requires a fully hydrodynamic system (e.g. Greif et al., 2010; Wise, 2012; O'Shea et al., 2015), but we can estimate the effect with a simple model based on distances between our halos. We consider a minihalo in our simulation to be exogenous (i.e., polluted) if its center is within the *pollution radius* of any other halo. For minihalos, the pollution radius is the size of a supernova remnant, which we take to be 300 pc for a 10^6 $M_\odot$ halo (Greif et al., 2007; Ritter et al., 2015; Smith et al., 2015). For an atomic cooling halo, the pollution radius is set based on the superbubble created by multiple supernovae associated with extended star formation, which we take to be 3 kpc physical (Madau et al., 2001). We then assume a mass-dependent pollution radius by taking the power law between these two points:

$$R_p = \frac{R}{R_8}\frac{M_{\rm vir}}{M_8}^{\alpha}, \tag{7.1}$$

where R_8 is the pollution radius for a 10^8 $M_\odot$ halo (set to be 3 kpc for the fiducial model), M_8 is 10^8 $M_\odot$ and α is the slope set by the 10^6 $M_\odot$ halo pollution radius. Figure 7-2 illustrates our fiducial model and two alternative normalizations allowing for stronger and weaker feedback.

For simplicity, the pollution radii are assumed to be spherical, instantaneously grow to their maximum size, and instantaneously mix into any matter they encounter. However, detailed hydrodynamic runs find the metal enrichment is inhomogeneous and episodic (Greif et al., 2007; Ritter et al., 2015; Smith et al., 2015), so we expect the number of externally enriched halos is an upper limit. It must be emphasized that we do not expect this simple enrichment prescription to accurately reflect the actual enrichment process of the first stars but to simply provide a broad-stroke model for gaining an understanding of the clustering properties and frequency of externally enriched objects.

7.4 Results

7.4.1 Visual Impression

In Figure 7-3, we show the distribution of molecular cooling and atomic cooling halos across our sample of 30 *Caterpillar* simulation halos. The top five rows shows the distribution of these systems at $z = 10$. We tag the 5% most-bound particles at formation, with H_2 cooling halos ("minihalos") in yellow and atomic cooling halos in red. In this figure, we use the LW feedback model adopting a star formation efficiency of $\epsilon_0 = 0.04$ to identify minihalos (Figure 7-1).

The bottom five rows show the same respective particles at $z = 0$ (image width is 1 physical Mpc in both cases). Halos are only tagged if they form before $z = 10$ as we assume reionization suppresses star formation in all systems at these mass scales. Although, there are stark commonalities between halos at $z = 0$, there are a

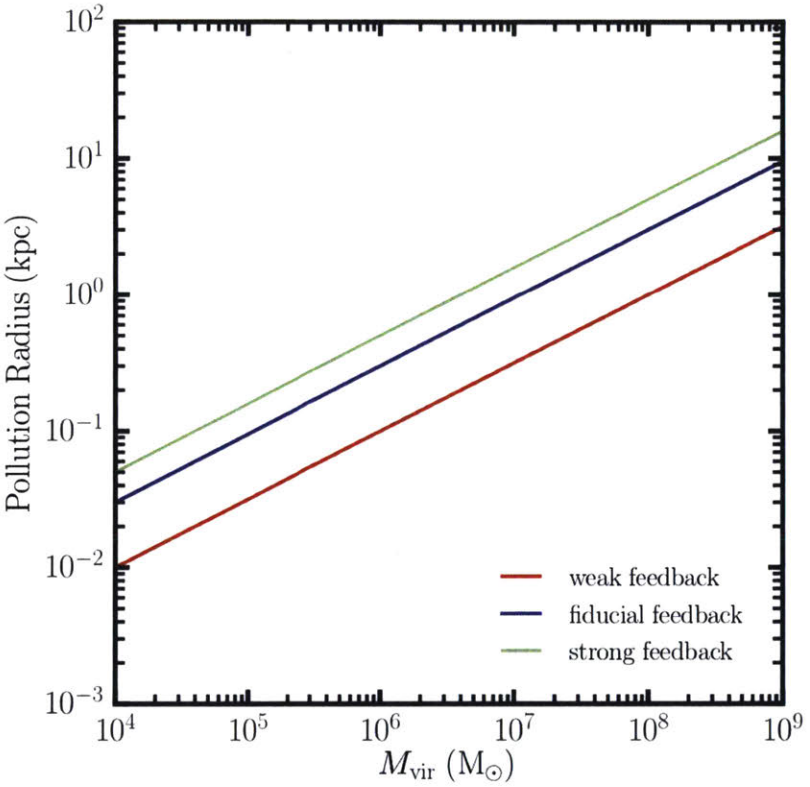

Figure 7-2 Simple chemical enrichment models with varying feedback. Our fiducial model yields a 300 pc (physical) enrichment radius for 10^5 $M_\odot$ halos (Greif et al., 2007; Ritter et al., 2012; Ritter et al., 2015; Smith et al., 2015) and a 3 kpc radius for 10^8 $M_\odot$ halos (Madau et al., 2001). We also adjust our normalization to account for strong feedback cases and weak feedback cases (e.g. for a 10^6 $M_\odot$ halo the enrichment radius varies between 100 pc and 500 pc).

wide variety of Lagrangian geometries at $z = 10$. Some realizations at $z = 10$ (e.g. Cat-2, Cat 9, Cat-36) show high densities of potentially star forming halos whilst other realizations show much more diffuse volumes of potentially star forming halos (e.g. Cat-1, Cat-6, Cat-33). In all cases, satellite systems both inside and outside the virial radius of the host contain potentially ancient stellar systems from the $z > 10$ era.

7.4.2 Progenitors Of The Milky Way

7.4.2.1 Minihalo progenitors of the Milky Way

In Figure 7-4, we plot the cumulative number of minihalos formed over time. We only count the total number of systems which are *accreted* into the central massive host and not those that end up residing in isolated halos at large galactocentric distances from the central host. The first of these Pop III star forming minihalos are identified at $z \sim 26$ and grow in number to approximately ~23,000 total (black line) potential sites assuming the Tegmark et al. (1997) temperature minimum mass criteria ($T_{\rm vir} \sim 2000\,{\rm K}$). The shaded regions for each line indicate 1-σ across all 30 halos in our sample for each of the methods of identification. There is ~20% scatter in the total number at nearly all times but some can be attributed to the fact that larger mass hosts have more progenitors on average ($n/10^{12}\ {\rm M}_{\odot} = 1.08 \times 10^{-8} \pm 0.03 \times 10^{-8}$, where n is the total number of systems).

Further dividing this population into progenitor systems that ultimately end up in the main host halo or in any of the subhalos of the main halo, we find that at $z = 0$ roughly the same number of progenitor minihalos end up in the central host ($45 \pm 11\%$, 10403 ± 2418) as in the host's subhalos ($55 \pm 16\%$, 12746 ± 3568).

Although there are a large number of potential Pop III star forming sites, the first luminous ones to have formed will greatly impact candidate sites for subsequent star formation due to the onset of the LW background. In Figure 7-4, we also show the cumulative number of halos which could have still collapsed in the presence of this LW background. Table 7.2 shows the cumulative number of halos for each population

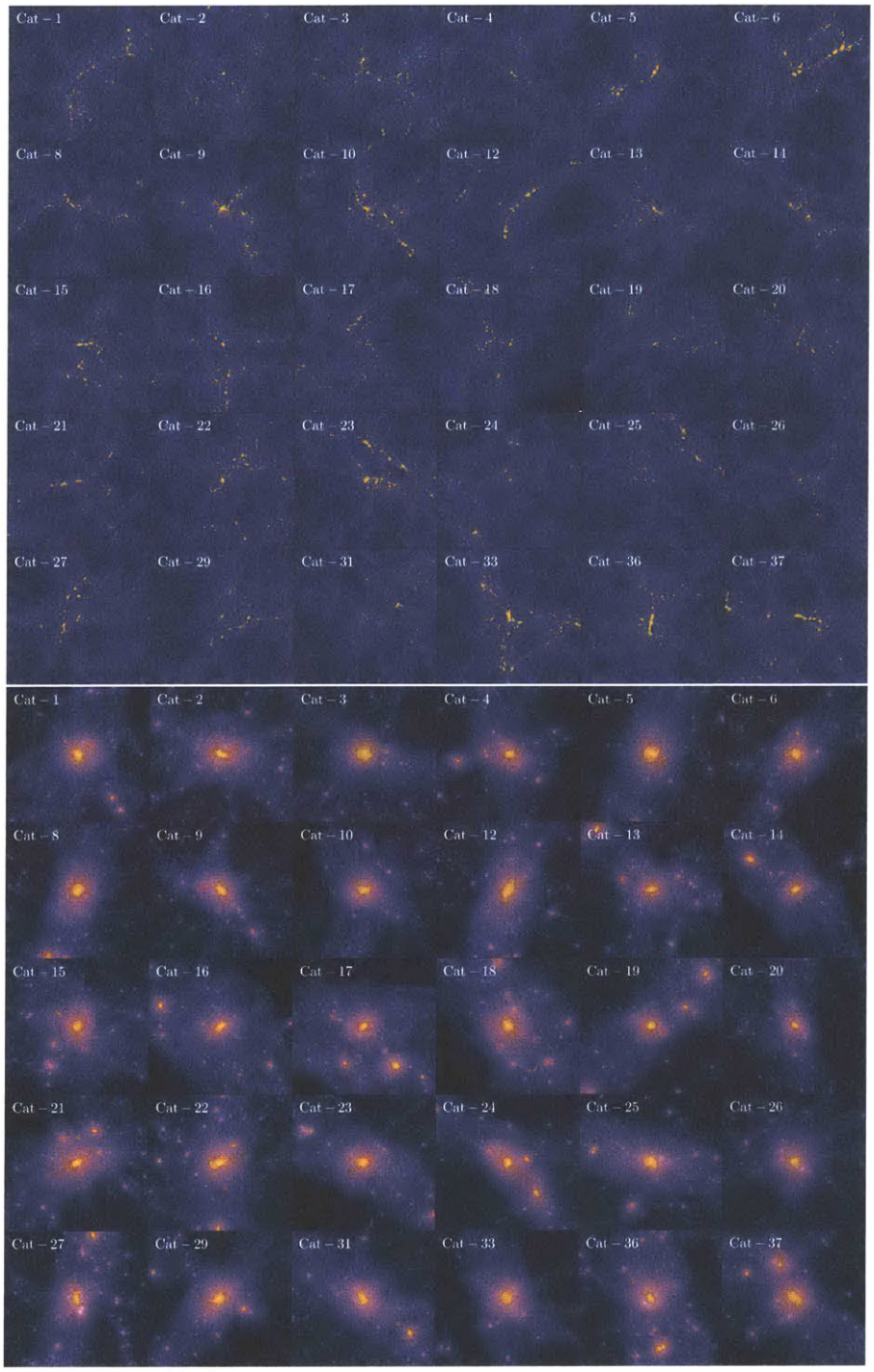

Figure 7-3 The 30 *Caterpillar* halos used in this study illustrate how the underlying dark matter distribution is overlaid with star particles. Particles are tagged as having formed within atomic cooling halos (red) and within molecular line cooling halos using our fiducial star formation efficiency ($\epsilon_0 = 0.04$). Five percent of the most bound particles were tagged for each respective system at formation. This is done purely for visualization purposes. Only halos which satisfy the temperature threshold before $z = 10$ are tagged, as reionization is assumed to suppress star formation at z <10. The top panels shows objects tagged at z = 10 and the bottom panels are the same particles at $z = 0$. The width of the image is 3 physical Mpc.

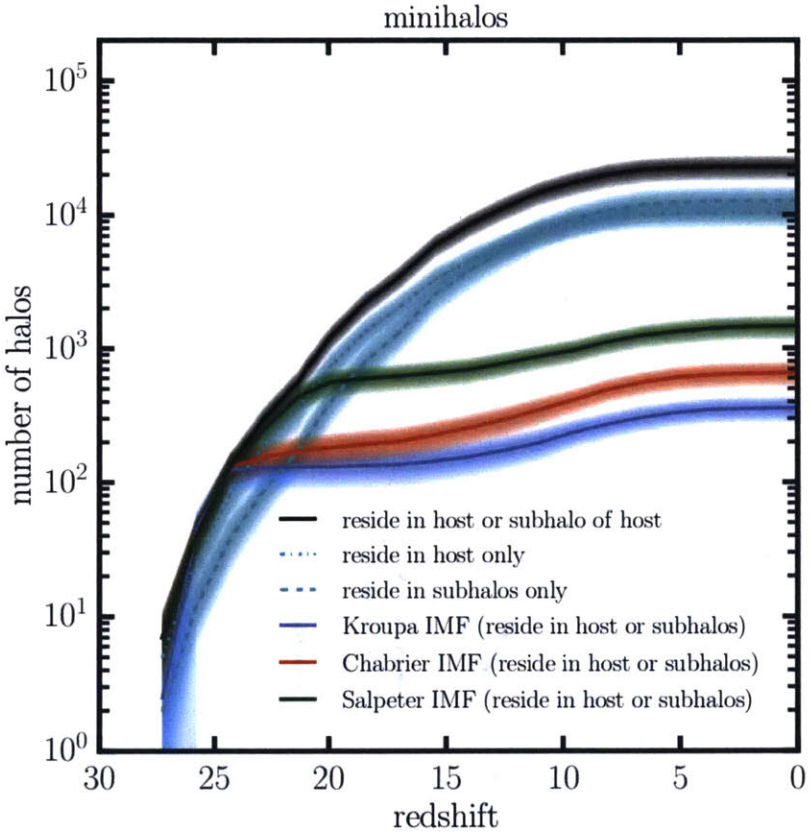

Figure 7-4 The cumulative number of Pop III star formation sites ("minihalos") as a function of time averaged over all 30 *Caterpillar* halos. The onset of the second generation of star formation has a dramatic impact on total minihalo numbers as early as $z \sim 25$, reducing the total number of potential star forming sites by 99.9%, due to LW feedback. Over 90% of all minihalo sites have formed by $z = 10$.

Table 7.2 Number of minihalos across all of the *Caterpillar* halos, broken down by final location at $z = 0$ and by the additional use of different star formation efficiencies (including ± 1-σ variance).

Selection	Number	Fraction
Reside within host or subhalos	22856 ± 4915	1.00 ± 0.22
Reside in host	10403 ± 2418	0.45 ± 0.11
Reside in subhalos	12746 ± 3568	0.55 ± 0.16
incl. LW ($\epsilon_h = 0.2$)	358 ± 82	0.02 ± 0.00
incl. LW ($\epsilon_0 = 0.04$)	653 ± 141	0.03 ± 0.01
incl. LW ($\epsilon_l = 0.008$)	1458 ± 314	0.06 ± 0.01

identified. We find drastic reductions by as much as 98% of potential star forming sites which would have otherwise cooled and collapsed via molecular line cooling in the absence of a LW background. Altering the choice of the star formation efficiency, ϵ, changes the amount of LW flux and consequently the potential number of sites from $\sim$358 ± 82 (1-σ, ϵ_h) to $\sim$1458 ± 314 (1-σ, ϵ_l). Between all cases, a minimum of 94% of the potential number of halos, which are nevertheless later accreted into the central host, are prevented from forming stars. For all three star formation efficiencies, approximately $\sim$50% end up in subhalos and $\sim$50% end up in the primary host by $z = 0$.

7.4.2.2 Atomic cooling halo progenitors of the Milky Way

In Figure 7-5, we plot the total number of halos which satisfy the virial temperature condition ($T_{\rm vir} > 10^4$ K). We divide the population into five categories, three of which are a subset of just one. We only count atomic cooling halos which end up in the central host or in a subhalo of the central host by $z = 0$. Of the subset that accretes into the primary host and subhalos, we further divide them into three groups; (1) halos with $V_{\rm max}$ ($z = 10$) $> 50\,{\rm km\ s^{-1}}$ are not suppressed (green), (2) ACHs with $30\,{\rm km\ s^{-1}} < V_{\rm max}$ ($z = 10$) $< 50\,{\rm km\ s^{-1}}$ are partially suppressed (blue), and (3) halos with $V_{\rm max}$ ($z = 10$) $< 30\,{\rm km\ s^{-1}}$ are completely suppressed (red).

Table 7.3 lists the cumulative number of halos which form in each category. We find that approximately 1793±396 (1-σ) halos within a Milky Way sized system satisfy

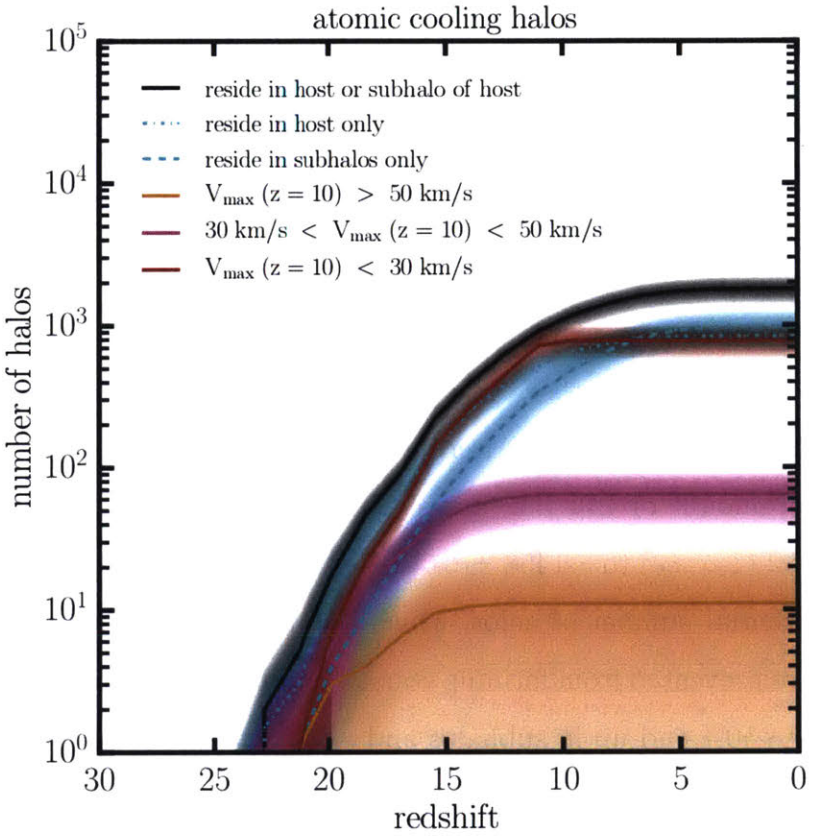

Figure 7-5 Number of atomic cooling halos which reside in either subhalos or halos by $z = 0$. We further divide the population into those which are suppressed, partially suppressed or fully star forming based on their maximum circular velocity of their descendants at $z = 10$. On average, 781 $\pm$ 214 potential atomic cooling halos are suppressed and stop forming stars due to the reionization background. Approximately $\sim$11 survive the reionization era and will continue to form stars provided there exists a supply of cold gas. We find 64 $\pm$ 27 are partially suppressed and will only convert some fraction of their cold gas into stars. Just over half of all atomic cooling halos to have formed reside within the central host (54% $\pm$ 16%) in the present day whilst the remainder (46% $\pm$ 11%) reside in subhalos).

Table 7.3 Number of atomic cooling halos across all of the *Caterpillar* halos broken down by the various models for identification at $z = 0$ (including ± 1-σ variance).

Selection	Number	Fraction
Reside in host or subhalos	1793 ± 396	1.00 ± 0.22
Reside in host	973 ± 290	0.54 ± 0.16
Reside in subhalos	836 ± 206	0.47 ± 0.11
No suppression[a]	11 ± 16	0.01 ± 0.01
Partially suppressed[b]	64 ± 27	0.04 ± 0.02
Fully suppressed[c]	781 ± 214	0.44 ± 0.12

a: $V_{\rm max}(z = 10) > 50\,{\rm km\ s^{-1}}$.
b: $30 < V_{\rm max}(z = 10) \leq 50\,{\rm km\ s^{-1}}$.
c: $V_{\rm max}(z = 10) \leq 30\,{\rm km\ s^{-1}}$. A total of 937 halos per host form after $z = 10$ have $V_{\rm max} < 30$ km/s and are assumed to be suppressed.

the atomic cooling limit and are eventually accreted either into the host itself or its subhalos. As with the minihalos, we find that approximately half ($45 \pm 11\%$) reside within the central host and half ($55 \pm 16\%$) reside within subhalos at the present day. Nearly half the halos that surpass the atomic cooling limit for the first time in their main branch end up within halos below the suppression scale at $z = 10$. We find approximately 11 halos (per host) with $V_{\rm max}(z = 10) > 50\,{\rm km\ s^{-1}}$ at $z = 10$ that will continue to form stars provided there exists a supply of cold gas. Some of these will merge with other halos before being accreted by the central host. These halos, which are not suppressed, combined with any of the partially suppressed ones in the post reionization era (64 ± 27 that will only convert some fraction of their cold gas into stars), could go on to become present day dwarf spheroidal galaxies around the Milky Way.

7.4.3 When were the first stellar systems accreted into the Milky Way?

In Figure 7-6, we show the cumulative distribution function at the time of the first merger for all identified systems which end up within the virial radius of the host

at $z = 0$ (across all *Caterpillar* halos). Approximately 50% of minihalos and atomic cooling halos merge into another halo larger than itself within 1 Gyr. Approximately 22 ± 1 % of all atomic cooling halos never have a merger with another halo larger than itself along its main branch. Similarly, approximately 20% of all minihalos never merge with another host larger than itself along its main branch. For the LW model adopting a high star formation efficiency (ϵ_h) this fell to 19 ± 3 %. For the low star formation efficiency (ϵ_l) and fiducial star formation efficiency (ϵ_0), they both yielded 22 ± 3 %.

In the inner panel of Figure 7-6, we also show the time between formation and accretion for objects which do not merge with anything larger than itself, i.e. the subset of halos amounting to ~20% of halos that have not merged in the outer panel. We find that 50% of all atomic cooling halos and minihalos accrete in the host within 4 Gyrs and 80% are accreted within 8 Gyrs. When compared to the history of an average subhalo at $z = 0$, atomic cooling halos and minihalos systematically cross the virial radius of the central host at earlier times as they were the first halos to form.

7.4.4 Spatial Distribution & Clustering

We investigate the spatial distribution of minihalo and first galaxy progenitors of Milky Way sized systems. In Figure 7-7, we demonstrate the spatial clustering of objects which end up inside subhalos or the central host at $z = 0$. This figure shows the density contours of all systems identified as minihalos (using the Tegmark et al. (1997) prescription) and atomic cooling halos found in a single snapshot corresponding to $z = 10$ across all 30 of our *Caterpillar* halos (the spatial distributions for minihalos identified with LW feedback are the same, see Section 7.5.3). Across all *Caterpillar* halos, we find that objects whose descendants eventually reside within subhalos are much less compactly clustered at high redshift than their counterparts that ultimately end up in the central host.

To get a better understanding of the separation properties of minihalos to their neighbouring star forming halos, we plot in Figure 7-8 how far away star forming halos are located from each of the minihalos. The halos are separated into increasing mass

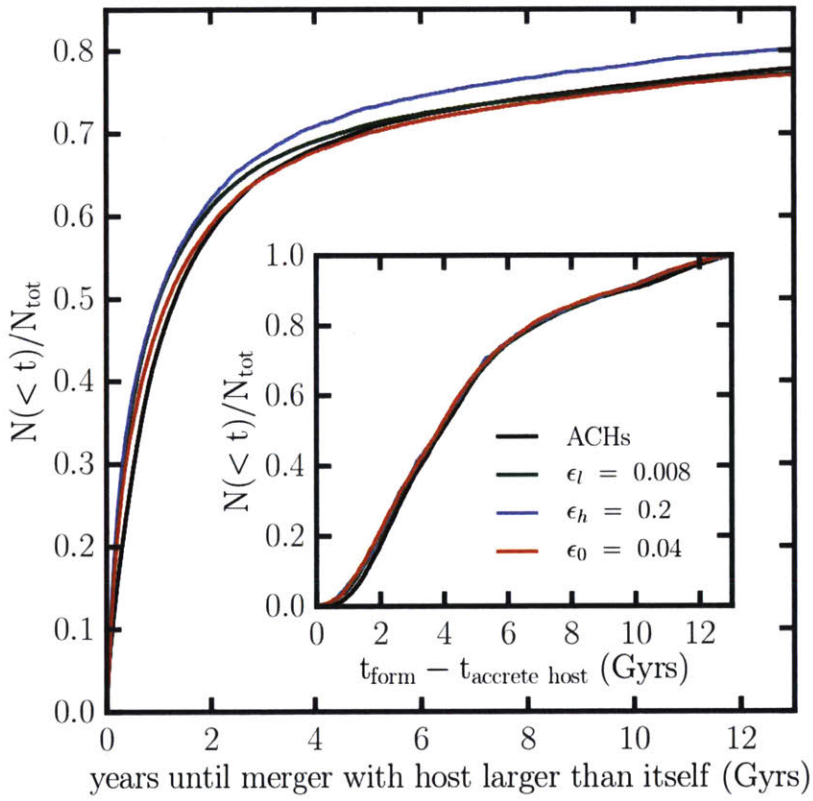

Figure 7-6 *Outer panel:* The cumulative distribution function of the time of the first merger of all atomic cooling halos and minihalos. Approximately 20% of all identified halos do not merge with any other halo larger than itself other than merging with the main host. *Inner panel:* The cumulative distribution function of the time between when halos form and when they enter the virial radius of the central host but have not merged by $z = 0$. Only ~50% of halos enter the host's virial radius within 4 Gyrs, indicating that many systems evolve in isolation for a significant portion of their lifespan.

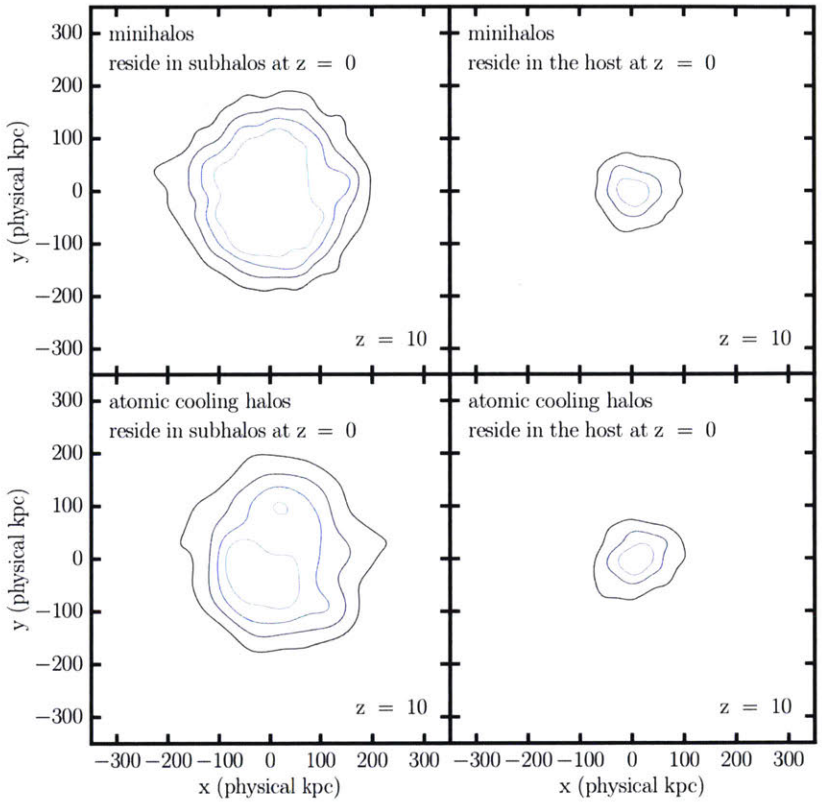

Figure 7-7 Contour plot of the positions of all minihalos and atomic cooling halos relative to the host (physical distance) identified in a single snapshot at $z = 10$. The top rows represent the positions of minihalos while the bottom rows represents the atomic cooling halos. The first column represents systems that reside in subhalos at the present day while the right column represents systems that reside in the central host at the present day. This characteristic spatial correlation between present day environment and formation environment is clear for all times – the initial stellar systems residing in subhalos today were much less clustered at earlier times compared to their counterparts that end up in the central host. This diagram represents the stacked positions for all 30 *Caterpillar* halos in our sample.

bins. In the first panel, we show the median minimum distance of each minihalo to all other star forming 10^6 $M_\odot$ halos. We find most of them are several virial radii away from any minihalo at all times. The median minimum distance to a 10^6 $M_\odot$ star forming halo at z ~20 is ~ 1 kpc (physical) indicating the proto-Milky Way formed in a very clustered environment. There is, however, large scatter in the median minimum separation ranging from 800 pc to 3 kpc at $z = 20$ across each of the *Caterpillar* simulations. The larger neighbouring star forming halos (10^{8-9} $M_\odot$) often have minihalos residing within a few virial radii during the time of their formation. This often leads to minihalos experiencing external chemical enrichment coming from these neighboring halos during their initial fragmentation process. But details depend on the individual case since there is significant scatter of several kiloparsecs of the median minimum separation at z ~20.

7.4.5 Internally & Externally Enriched Fraction

We have shown clearly that there is a spatial preference for progenitors of subhalos when compared to the progenitors of the central host in the high-redshift era of the Milky Way. These spatial biases are expected to manifest themselves in the chemical enrichment history of their respective stellar constituents as systems that reside in the host today come from more clustered environments. We apply our simple chemical enrichment model from Section 7.3.4 to determine what fraction of minihalo progenitors of the Milky Way were likely externally or internally enriched. These processes lead to two classes of systems in the proto-Milky Way era; *endogenous* systems (chemically enriched solely by internal processes) and *exogeneous* systems (enriched by internal *and* external processes).

In Figure 7-9, we show what fraction of the total population are exogenous or endogenous as a function of time for the minihalos identified via the Tegmark et al. (1997) prescription. We further break this population down into minihalos which end up in subhalos and minihalos which end up in the main host. The breakdown of populations is similar in each of the *Caterpillar* halos in the sample. The feedback prescription used in the middle panel is our fiducial model whereby halos with $M_{\rm vir}$

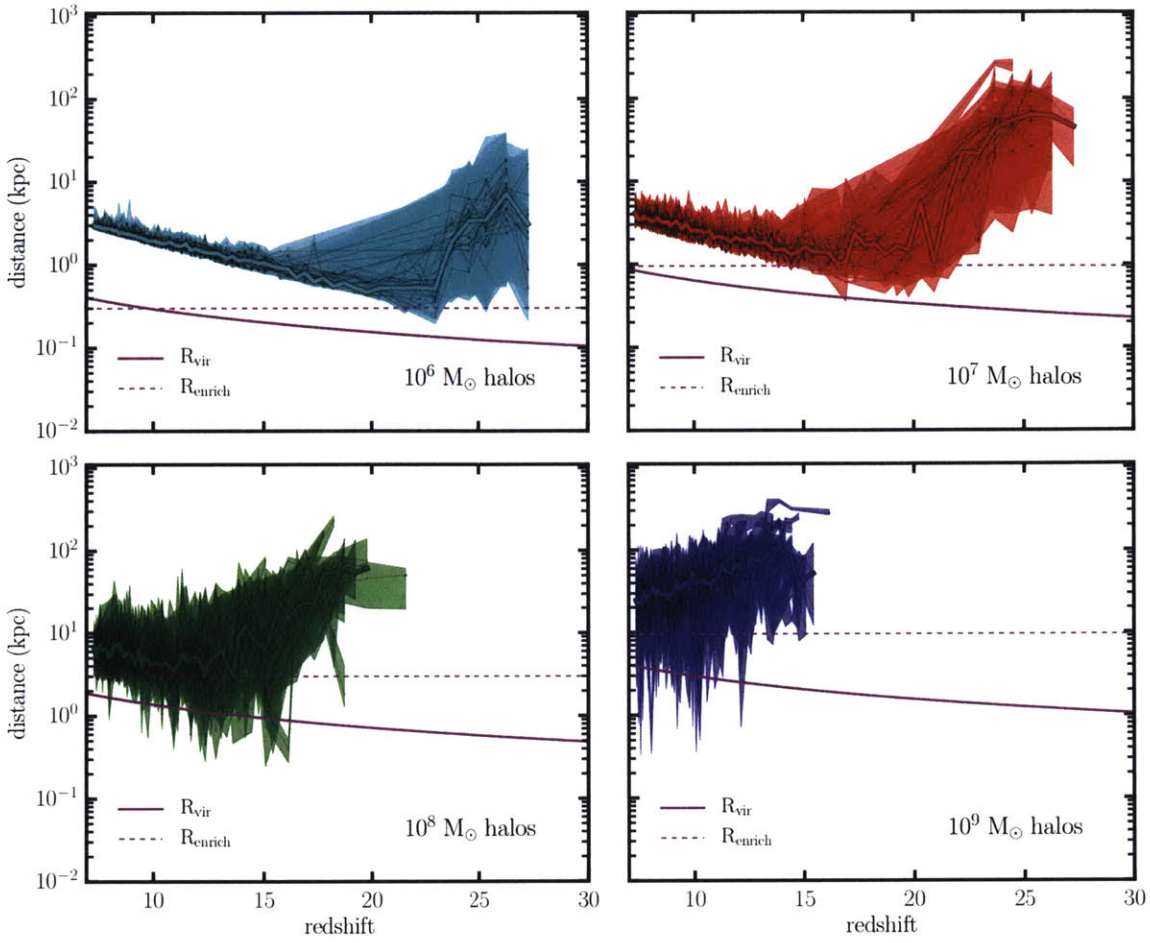

Figure 7-8 Median minimum distance (physical) of every identified minihalo to any different mass halos (star forming) as a function of time. Each minihalo was identified using the minimum mass threshold found by Crosby et al. (2013) which includes LW feedback using our fiducial star formation efficiency ($\epsilon_0 = 0.04$). Each of the thin black lines represent a single *Caterpillar* simulation and the shaded region represents the 1-σ variance. The solid line represents the median of all 30 *Caterpillar* runs. In each panel, the cyan line underneath represents the virial radius of halos with 10^{6-9} M$_\odot$, based on the Bryan and Norman (1998) prescription. The dashed line is the enrichment radius for each of these halos calculated with our fiducial enrichment model. It is clear that the majority of the external enrichment of a given minihalo is driven by neighbouring larger mass halos, e.g. 10^{8-9} M$_\odot$. Each of the neighbouring halos are checked to ensure they are actually star forming by determining if any progenitors contain accreted halos that have satisfied the virial temperature criterion. The median minimum distance is an indicator of the density of star forming halos. Starting at high-z, the density first increases due to a proliferation of galaxy formation. Later, the formation rate of new galaxies declines and the Hubble expansion begins to dominate, leading to a decrease in density (or increase in distance). This turnaround point occurs at higher z for lower mass halos since low mass halos form earlier in the universe than high mass halos.

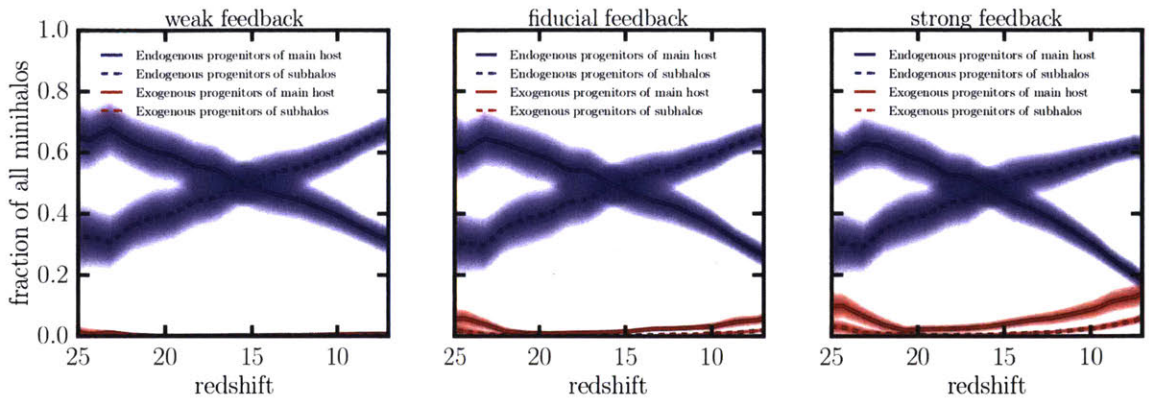

Figure 7-9 Fraction of exogenous (enriched by an external system) and endogenous (enriched by internal processes) progenitor minihalo systems for each of the Milky Way-mass halos in our sample. Each panel represents a different strength of feedback (see Figure 7-2). The total exogenous population varies from $\sim$ 1% of all halos at $z = 7$ in the weak feedback model to 18% in the strong feedback model. Minihalos that form early, near $z = 25$, are more likely to be progenitors of the host since they have more time to be accreted, lose angular momentum, and get disrupted, whereas minihalos forming later, near $z = 7$ are more likely to be progenitors of subhalos since they don't have enough time to be fully disrupted in the host.

$= 10^6$ $M_\odot$ have enrichment radii of 300 pc while halos with $M_{\rm vir} = 10^8$ $M_\odot$ have enrichment radii of 3 kpc. We also show results for the weak and strong feedback models from Figure 7-2.

At $z = 20$, we find that an overwhelming proportion of the minihalo population are endogenous systems, evolving in isolation, for all three feedback models. This continues to later times and only by $z = 7$ do we observe any significant number of minihalos becoming exogenous, or externally enriched. In the strong feedback model, the fraction of exogenous minihalos rises from just 3% at $z = 20$ to 18% at $z = 7$. Using the weak feedback model, merely $\leq$ 1% of minihalos are externally enriched between $z = 20$ and $z = 7$. Meanwhile, the endogenous population flips from being dominated by progenitors of the host at $z = 20$, to being dominated by progenitors of subhalos at $z = 7$. This is due to a bias where halos that form earlier have more time to be pulled into the central host and disrupted by $z = 0$ than halos which form later.

In Table 7.4, we list the breakdown of minihalos for the strong feedback model

Table 7.4 Fraction of halos which are exogenously or endogenously enriched for the fiducial feedback model at different times for 30 *Caterpillar* halos (± 1-σ variance).

Type	$z = 20$	$z = 15$	$z = 10$	$z = 7$
Endogenous progenitors of main host	0.57 ± 0.08	0.45 ± 0.07	0.30 ± 0.05	0.20 ± 0.04
Endogenous progenitors of subhalos	0.40 ± 0.09	0.51 ± 0.08	0.58 ± 0.05	0.62 ± 0.05
Exogenous progenitors of main host	0.02 ± 0.02	0.04 ± 0.02	0.09 ± 0.03	0.13 ± 0.04
Exogenous progenitors of subhalos	0.00 ± 0.01	0.01 ± 0.01	0.03 ± 0.01	0.05 ± 0.02

into raw percentages. When restricting the sample to just progenitors of the host, the fraction of exogenous halos begins approaching that of endogenous halos towards $z = 7$. At $z = 20$, 3% of progenitors of the host are exogenous. By $z = 7$, that number jumps to 40%. When restricting to progenitors of subhalos, the fraction of endogenous systems reaches a much smaller peak of 7% at $z = 7$. This is caused by the spatial biasing in the assembly history of the host. The progenitors of accreted systems which reside within the host in the present day tend to be more centrally clustered in the most over-dense regions, leading to a higher probability that the enrichment bubbles of nearby systems overlap with the surrounding halos. Furthermore, those systems which end up in the present day host are more likely to be externally enriched the later they form due to a combination of spatial clustering and a greater abundance of larger, $10^9 M_\odot$, star forming halos (see Figure 7-8).

7.4.6 Remnants of the first stellar systems in dwarf galaxies

The progenitor halos of the Milky Way can be split into two distinct populations: (1) "halo progenitors" (i.e., those that formed, merged and accreted, subsequently dispersing throughout the stellar halo of the Milky Way) and (2) "dwarf progenitors" (i.e., those that formed, accreted and merged into what are now dwarf galaxies). The progenitor merger tree of each of these two systems will invariably be littered with minihalos and atomic cooling halos.

This presents an opportunity to consider in detail the origin and nature of the observable dwarf satellite galaxies of the Milky Way. Especially their early chemical

composition, and consequently also that of their oldest, most metal-poor stars must have been driven by the total number of high-z minihalos and atomic cooling halos that each dwarf galaxy accreted throughout its evolution.

Since our simulation suite runs until $z = 0$, we can determine how many candidate minihalos and atomic cooling halos have merged with a given dwarf galaxy since its formation. In Figure 7-10, we plot a parameterized fit to the number of progenitors of a given subhalo in the present day for all *Caterpillar* simulations. We relate the number of each respective system (i.e., exogenous and endogenous systems for each definition of minihalo and atomic cooling halo) to the peak mass of present day subhalos via the following form,

$$n_p \;=\; n_0 \, (M_{\rm peak})^{\alpha} \,, \tag{7.2}$$

where n_p is the number of progenitor systems and $M_{\rm peak}$ is the peak mass along the main branch of a given subhalo. n_0 is a normalization quantity. Given this functional form, our best fit estimates are presented in Table 7.5. We also show stellar mass estimates for these systems using the Garrison-Kimmel et al. (2014b) abundance matching prescription as a reference. We find that the number of minihalo progenitors at a fixed $M_{\rm peak}$ (peak mass along the main branch) depends very much on whether LW feedback is included. Without LW feedback, a halo with a peak mass of 10^9 $M_\odot$ ($M_\star \sim 10^4$ $M_\odot$) would have accreted ~30 minihalo progenitors. With LW feedback, a halo with the same peak mass would actually only have accreted ~10 halos at most. This is particularly pronounced at even lower peak masses (e.g., UFDs), where one expects less than one minihalo to have been accreted into the system by $z = 0$ when including LW feedback.

7.4.6.1 Classical dwarfs and the Magellanic Clouds

In Table 7.6, we list the number of progenitor systems that we derived for a sample of nine classical dwarf galaxies. We used the abundance matching prescription of Garrison-Kimmel et al. (2014b) to find the corresponding subhalo peak mass. We

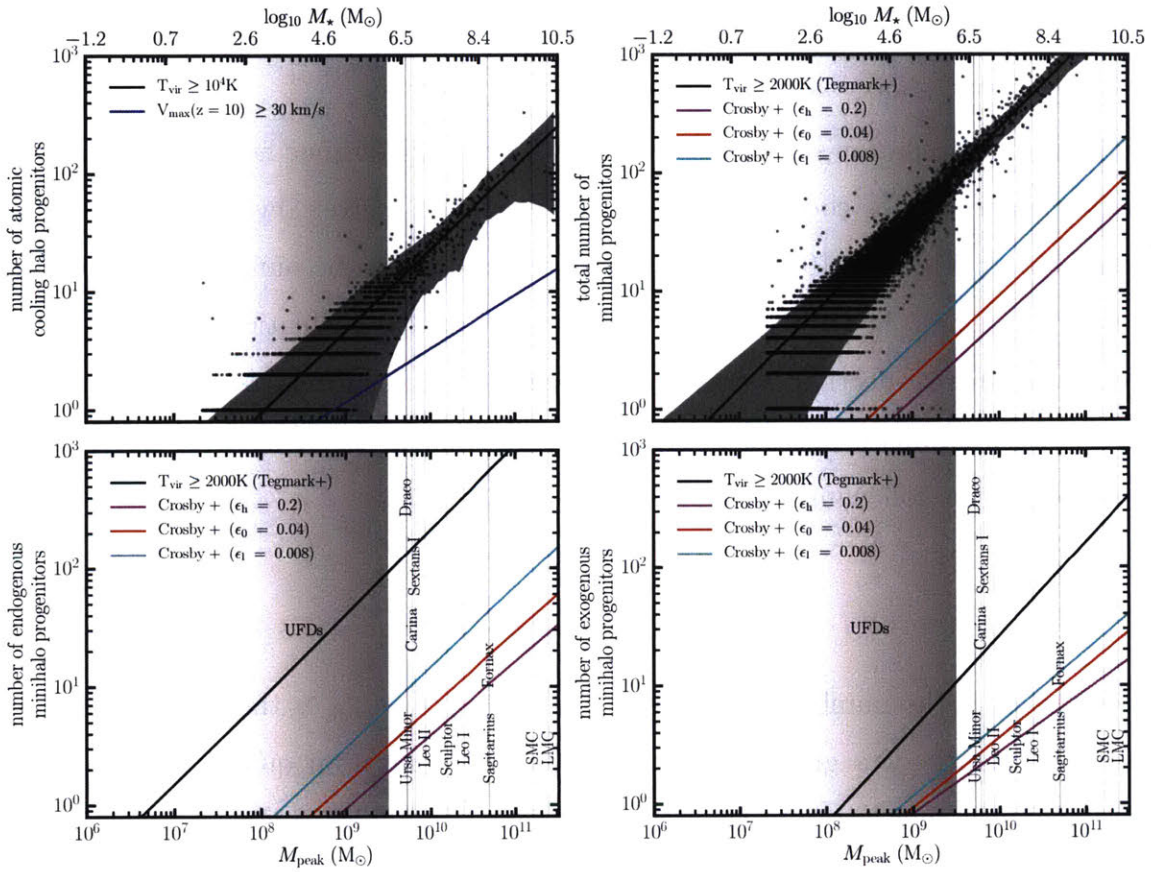

Figure 7-10 Total number of progenitors of a given subhalo (*top left*: atomic cooling halos, *top right*: total, *bottom left*: endogenous minihalos, *bottom right*: exogenous minihalos) as a function of peak subhalo mass. Typical 1-σ variance for each fit (across all 30 *Caterpillar* simulations used in this study) is shown in the top left/right panel (grey band). These are omitted for the other fits for the sake of clarity. The peak mass corresponds to a stellar mass as determined from the abundance matching prescription of Garrison-Kimmel et al. (2014b). As a guide, we have drawn vertical lines corresponding to the stellar mass of each of the observed nine classical dwarf spheroidal galaxies. Although more massive dwarf galaxies tend to have a large number of minihalo progenitors, the total number depends strongly on the inclusion of the LW feedback. This reduction is particularly pronounced for ultra-faint dwarf galaxies, with 90% fewer potential minihalo progenitors. There are slightly fewer ($\sim$10%) endogenous progenitors (i.e. progenitors which have evolved in isolation) at fixed subhalo peak mass. The uncertainty in LW models is similar to the halo-to-halo scatter. As previously stated (e.g. Sawala et al. 2014), estimates of stellar mass based on abundance matching are unreliable for $M_{\rm peak} \leq 10^9$ M$_\odot$. We only estimate the number of progenitors for UFDs (range highlighted in green) by extrapolation, which as such, is speculation.

Table 7.5 Number of progenitors for a given halo at $z = 0$ (± 1-σ variance across halos).

Number Of Progenitors	n_0 ($M_\odot$)	± 1-σ	α	± 1-σ
Minihalos				
Tegmark et al. (1997) ($T_{\rm vir} > 2000\,{\rm K}$)	1.107×10^{-5}	3.436×10^{-7}	0.733	0.001
$\rightarrow$ Endogenous minihalos	1.429×10^{-5}	4.323×10^{-7}	0.717	0.001
$\rightarrow$ Exogenous minihalos	3.296×10^{-7}	1.140×10^{-7}	0.790	0.014
LW w/ high star formation efficiency ($\epsilon_h = 0.2$)*	1.085×10^{-6}	1.808×10^{-7}	0.671	0.007
$\rightarrow$ Endogenous minihalos	2.469×10^{-6}	4.734×10^{-7}	0.620	0.008
$\rightarrow$ Exogenous minihalos	1.695×10^{-5}	4.981×10^{-6}	0.521	0.012
LW w/ fiducial star formation efficiency ($\epsilon_0 = 0.04$)*	1.191×10^{-6}	1.262×10^{-7}	0.688	0.004
$\rightarrow$ Endogenous minihalos	2.733×10^{-6}	3.148×10^{-7}	0.639	0.005
$\rightarrow$ Exogenous minihalos	4.050×10^{-6}	1.020×10^{-6}	0.595	0.010
LW w/ low star formation efficiency ($\epsilon_l = 0.008$)*	1.497×10^{-6}	9.628×10^{-8}	0.708	0.003
$\rightarrow$ Endogenous minihalos	2.431×10^{-6}	1.706×10^{-7}	0.678	0.003
$\rightarrow$ Exogenous minihalos	3.380×10^{-6}	6.301×10^{-7}	0.615	0.008
ACHs	2.678×10^{-6}	1.425×10^{-7}	0.693	0.002
ACHs $V_{\rm max}$ ($z = 10$) ≥ 30 km/s	1.001×10^{-4}	1.733×10^{-4}	0.452	0.070

* based on Crosby et al. (2013).

estimate that ~154 atomic cooling halos were swallowed by the LMC prior to its infall into the Milky Way. Draco by comparison may have accreted only 10 atomic cooling halos by infall. Each column contains the number of estimated accreted minihalos which represent the total number of exogenous or endogenous minihalos which fell into the classical dwarf satellites prior to their own infall. The results are purely based on abundance matching (Garrison-Kimmel et al., 2014b) to infer the subhalo peak mass. Depending on the choice of star formation efficiency, we find that only a handful of star forming minihalos fell into the Draco system (1-3 exogenous systems and 3-5 endogenous systems) prior to infall. Direct treatment of the LW radiation at these early times will provide more solid estimates for the number of progenitor systems in each case.

7.4.6.2 Ultra-faint dwarf galaxies

The population of ancient, very low luminosity "ultra-faint" dwarf (UFD) galaxies in the Milky Way has been studied extensively for their star formation, chemical

composition, and association to Galactic building blocks (see Frebel (2010) for a review). Recently, the Dark Energy Survey unveiled nine new such UFDs (DES Collaboration 2015, Koposov et al. 2015a). Interestingly, these satellites are close to the Large Magellanic Cloud (LMC). What remains to be answered, though, is whether most of the stellar material in such UFDs actually formed in-situ, or whether the dwarfs contain a substantial population of stars accreted from other, possibly chemically distinct, star forming systems.

This idea can in principle be tested with detailed chemical abundances of metal-poor stars that are found in all UFD galaxies. For example in the UFD Reticulum II, seven of nine stars observed are strongly enhanced in heavy r-process elements which already led to the suggestion that this UFD experienced a massive r-process event by either a neutron-star merger or a jet driven supernova (Ji et al., 2016). But the other two stars, which also happen to be the two most metal-poor stars in Reticulum II, display extremely low abundances of those same heavy neutron-capture elements, ([Ba/Fe] < 0, Ji et al. 2016). Furthermore, these nuclei were unlikely produced in an r-process but in some other event or site.

These two groups of nucleosynthetic signatures suggest the following about the nature and evolution of Reticulum II: either a) the stars with low heavy neutron-capture abundances formed within Reticulum II but prior to the r-process enrichment event, or b) they formed in a pocket of low-metallicity gas that was not affected by the r-process enrichment. Importantly, the latter scenario could have occurred in a different, smaller system that was later accreted into Reticulum II.

In general, our results (see Figure 7-10) indicate that it is unlikely that many UFD candidates could have accreted more than a few (endogenous or exogenous) minihalos. The vast majority of potential progenitors were simply unable to form stars due to the H_2 dissociating by the onset of the LW background. Even under the most optimistic of circumstances where we assume a Tegmark et al. (1997) minimum mass threshold for formation and remove our model for the LW background, the most massive of the future ultra-faint dwarf galaxies ($M_{\rm peak} \sim 10^{8.5}$ $M_\odot$) accreted <10 minihalos.

With the inclusion of our fiducial LW model, this number is reduced to only one

minihalo. Thus, the "small system accretion" scenario for Reticulum II is unlikely and very few stars, if any, originate from distinct minihalos. However, larger dSphs like Draco and Ursa Minor are very likely to contain metal-poor stars from multiple progenitor minihalos. Moreover, detailed theoretical modelling of UFDs would greatly assist in this question also by constraining metal mixing and star formation processes to determine the exact origin of potentially different stellar abundance patterns within single UFD systems. Hydrodynamic simulations of UFDs may greatly assist interpretations of chemical abundances in UFDs by further constraining the impact of metal mixing and hierarchical galaxy formation on the exact origin of potentially different stellar abundance patterns within single UFD systems.

Table 7.6 Estimates of the number of progenitors for nine classical dwarf galaxies and Magellanic systems.

Dwarf Galaxy	ACHs		Minihalos											
			Tegmark			$\epsilon_h = 0.2$			$\epsilon_0 = 0.04$			$\epsilon_l = 0.008$		
	−reion	+reion	En.	Ex.	Total	En.	Ex.	Total	En.	Ex.	Total	En.	Ex.	Total
Draco	12	0	120	7	127	1	0	1	3	1	4	8	1	9
Ursa Minor	12	0	120	7	127	1	0	1	3	1	4	8	1	9
Carina	14	0	135	8	143	2	0	2	3	1	4	9	1	10
Sextans I	14	0	143	9	152	2	1	3	4	1	5	9	1	10
Leo II	18	0	178	13	191	2	1	3	5	1	6	11	2	13
Sculptor	29	0	285	27	312	4	2	6	8	3	11	19	4	23
Leo I	41	0	410	48	458	6	3	9	11	4	15	27	6	33
Fornax	71	2	706	111	817	12	6	18	20	9	29	47	12	59
Sagitarrius	72	2	723	115	838	12	6	18	20	9	29	48	13	61
SMC	179	20	1810	473	2283	32	19	51	52	29	81	123	40	163
LMC	246	42	2503	780	3283	46	28	74	72	43	115	171	59	230

Note: "En." represent endogenous systems and "Ex." represent exogenous systems. "−reion" means no reionization included, "+reion" refers to the total number of atomic cooling halos which have $V_{\rm max} \geq 30$ km/s at $z = 10$.

7.4.7 Remnants of the first stellar systems in the Galaxy today

With our 30 high-resolution simulations, we can quantify the halo-to-halo scatter in the remnant population. A full treatment requires more detailed modelling of the stellar mass associated with each remnant, but as a first step we tag the 10 of the most bound particles at $z = 10$ for the minihalos identified with Lyman-Werner feedback at our fiducial star formation efficiency, $\epsilon_0 = 0.04$ (see Figure 7-3) and determine their number density as a function of galactocentric distance.

In Fig. 7-11 we plot these number densities divided by the dark matter density of the host out to the virial radius for each host (black line is the median). This ratio highlights any bias in the remnant distribution relative to the overall density of particles in the dark matter halo. The scatter in the number density at fixed galactocentric distance is an order of magnitude at small radii (e.g. within the bulge) and large radii (i.e. r $>$ 50 kpc) but similar within the halo (i.e. r $<$ 30 kpc). Our scatter agrees qualitatively with the result found by Ishiyama et al. (2016) who used four halos. Additionally, we find different overall means owing to the alternative Lyman-Werner treatment and slightly different tagging method (i.e at formation versus at $z = 10$). Tumlinson (2010) has argued that metal-poor stars in the bulge are most likely to be true relics of Pop II. stars. However, Salvadori et al. (2015) and more recently, Starkenburg et al. (2016) find that the oldest stars populate the innermost region of the Galaxy while the relative contribution of very metal poor stars increases with radius from the Galactic center. Without more detailed modeling, we can not compare directly with these works except to state that our oldest remnants populate all parts of the Galaxy with scatter most pronounced in the bulge and at large radii.

7.5 Conclusions

We have presented a systematic study of the general properties of minihalos and atomic cooling halo progenitors of Milky Way sized systems using 30 cosmological

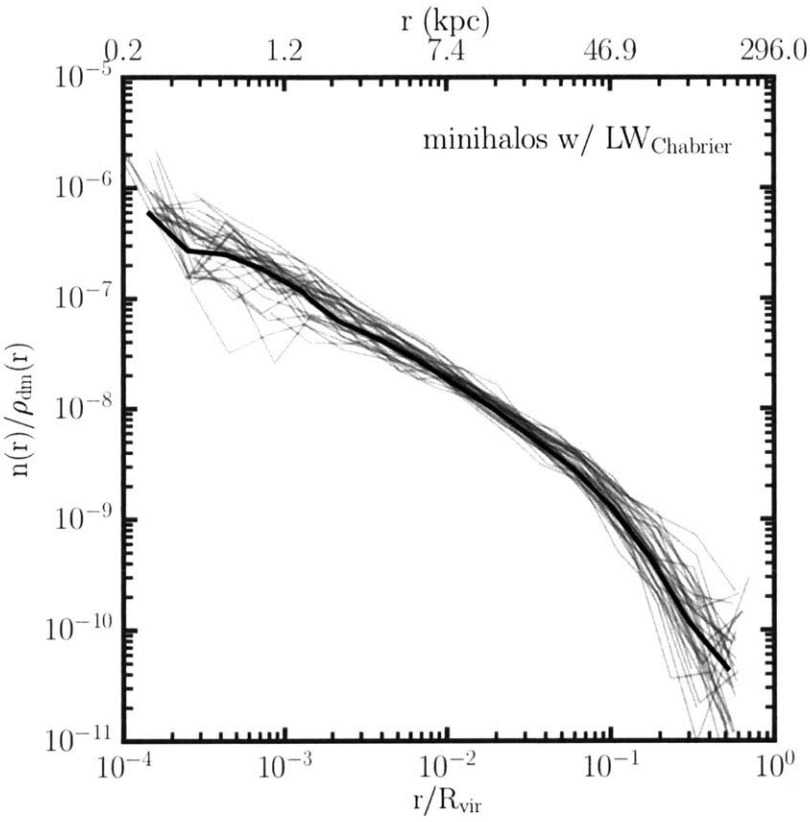

Figure 7-11 Ratio between radial number density profiles of Pop III remnants, $n(r)$, and dark matter mass densities of host halos ($\rho_{\rm dm}(r)$ for 30 *Caterpillar* halos (each individually marked in grey and median marked by thick black line). The top axis represents the bottom axis multiplied by the mean virial radius for all 30 runs (296 kpc). The scatter in the number density at fixed galactocentric distance is quite large at small radii (e.g. an order of magnitude in the bulge) and large radii (i.e. r $>$ 50 kpc) but similar within the halo (i.e. $r < 30$ kpc).

simulations. In our model for first star formation, we include the impact of Lyman-Werner radiation on the earliest stellar systems and determine how the clustering properties of such star-forming systems enriched subsequent generations of stars and galaxies in the Milky Way. Our model and results can be summarized as follows, first with respect to minihalos then atomic cooling halos.

7.5.1 Minihalo Progenitors of Milky Way sized systems

Using a physically motivated minimum mass threshold, we identify all molecular line cooling halos via the cooling threshold of Tegmark et al. (1997). We additionally use the semi-analytic prescriptions of Crosby et al. (2013) for the LW background using three different star formation efficiencies ($\epsilon = 0.008, 0.04, 0.2$). We find the following:

- Without LW feedback, we find 22,856 $\pm$ 4915 progenitor dark matter halos of a Milky way sized host to satisfy the minimum mass threshold required for their molecular hydrogen gas to cool, collapse and form stars.

- With LW feedback, the number of potential star forming minihalo progenitors is significantly reduced (by $\sim$90%) because the radiation raises the minimum mass required to form stars. We find 358$\pm$82/653$\pm$141/1458$\pm$314 (for star formation efficiencies: $\epsilon_l = 0.008$, $\epsilon_0 = 0.04$, $\epsilon_h = 0.2$) minihalos satisfy our requirements to form stars and eventually merge into the host halo.

- By $z = 0$, 55% of all progenitor systems are accreted by the central host and the remainder reside within subhalos of the central host.

- Using a simple chemical enrichment model, we determined what fraction of systems have their chemical composition established by in-situ star formation or by being enriched by neighboring systems. Overwhelmingly, most of the minihalos evolve in isolation without the influence (chemically) of an external halo (i.e. 80-90% of all systems at $z = 7$ are endogenous). For the strong feedback model, we find $\sim$18% of systems are exogenous at $z = 7$ compared

to <7% of systems for the weak and fiducial feedback models. When halos are externally enriched, it is usually by 10^8 $M_\odot$ systems or more massive ones.

- Of the systems which are endogenous, ~50% merge with a system larger than themselves within 1.5 Gyr after formation (Fig. 7-6). This leads to enhanced chemical enrichment, making them only temporarily endogenous systems. Several generations of stars could have formed (and died) between the time of first star formation and the eventual accretion of the system into the main host.

- Star forming minihalos are on average median separated to other star forming 10^7 $M_\odot$ halos by 300 pc at $z = 20$ and 3 kpc at $z = 7$. While we found most systems are internally enriched, a more realistic chemical enrichment model including proper treatment of chemical mixing and non-instantaneous winds may result in an increase in the externally enriched fraction.

- The number of minihalo progenitor systems which have been accreted by a given subhalo halo is best fit via a power law. The number of exogenous progenitors is best fit via the power law, $N_{\rm prog} = 2.97 \times 10^{-4} M_{\rm peak}^{0.4}$. Similarly, the number of endogenous progenitors for a given subhalo is best fit by $N_{\rm prog} = 4.82 \times 10^{-7} M_{\rm peak}^{0.71}$.

- We estimate that there is an order of magnitude scatter in the number density of Pop III remnants at small (i.e. r < 5 kpc) and large galactocentric radii (i.e. r > 50 kpc) across Milky Way-mass halos. The scatter is most minimal at intermediate distances (10 < r < 50 kpc) within the halo.

- We estimate that low luminosity UFD galaxies, such as Reticulum II, have at most one or two star forming minihalo progenitors. Consequently, it highly unlikely that Reticulum II received its r-process enriched material via an external system bringing in chemically enriched stellar material.

- Similarly, we estimate that approximately ~74-230 (~51-163) minihalos were accreted by the proto-LMC (SMC), creating a potentially large number of ultra-

faint satellite systems which could be tidally removed from the LMC during first passage and distributed throughout the Milky Way.

7.5.2 Atomic Cooling Halo Progenitors of Milky Way sized systems

We identified all potentially atomic cooling halos in each simulation and used a simple model of reionization to determine which halos were suppressed, partially suppressed and active in the post-reionization era ($z < 10$). Our results can be summarized as follows:

- There are 1793 $\pm$ 396 atomic cooling halo progenitors per 10^{12} $M_{\odot}$ host (across 30 Milky Way sized systems).

- We find 781 $\pm$ 215 (44 $\pm$ 12%) of these systems do not survive the reionization era and will stop accreting gas and forming new stars (937 additional systems form after $z = 10$ with $V_{\rm max} < 30$ km/s and are suppressed in our model). On average, we also also find that 64 (4 $\pm$ 2%) systems are partially suppressed and 11 (1 $\pm$ 1%) systems are not suppressed at all and will continue to accrete gas and form stars unimpeded by reionization. These will accrete into either larger progenitors and become dwarf galaxies or be disrupted during the accretion onto the primary host.

- By z$\sim$0, 54% of the unsuppressed atomic cooling halo progenitor systems are accreted by the central host and the remainder end up within the subhalos.

- The number of atomic cooling halo progenitor systems of a given subhalo of the host is best fit via the power law, $N_{\rm prog} = 2.69 \times 10^{-7} M_{\rm peak}^{0.69}$.

- Approximately 246 atomic cooling halos were accreted by the LMC prior to infall and $\sim$12 atomic cooling halos were accreted by Draco. Using a simple model for reionization, we find only 42 (0) of these LMC (Draco) progenitor systems have $V_{\rm max}$ ($z = 10$) $\geq$ 30 km/s and will survive the reionization era.

We finally comment that Gao et al. (2010) used the *Aquarius* simulation suite to identify Pop III star forming progenitors. They employed a similar method as ours, though at a lower virial temperature threshold (1100 K). They found $\sim 2 \times 10^4$ Pop III star forming progenitors which agrees well with our estimates of ~23,000. Similarly, they found a mean separation distance of ~ 1 h^{-1} kpc ($z = 10$) which also agrees well with our estimates (~ 3 h^{-1} kpc). They also determine the number of first galaxies (i.e. 10^4 K) to be ~200-300 by $z = 10$. We speculate that this estimate is lower than ours because of the lower temporal resolution used in the *Aquarius* simulation suite (~100 Myrs/snapshot outputs compared to $\sim$ 5 Myr/snapshot outputs in *Caterpillar*). We use a different model for the LW background than the work of Ishiyama et al. (2016) and so it make it difficult to compare numbers directly. Additionally, Gao et al. (2010) also do not provide population statistics which furthermore complicates a detailed comparison of results.

7.5.3 Caveats & Future Work

Our modelling technique is not without drawbacks. Most importantly, we do not resolve the direct collapse of gas, subsequent fragmentation and enrichment *directly* and rely on the assumption that a given halo's temperature is in virial equilibrium with the gas temperature. We additionally assume that the enrichment process proceeds via instantaneous, spherically enriched gas bubbles at a scale set purely by the progenitor host halo mass. It is known from detailed hydrodynamic simulations of single halo systems that star formation proceeds in a much more stochastic manner and that the enrichment process is very unstructured and depends heavily on local environmental conditions. Despite these limitations, we are providing a robust machinery for connecting present day halos with their high-z progenitors, and offer a first glimpse to statistically probe the locations of the first star forming progenitors of Milky Way-mass halos by sampling the largest number of Milky Way halos ever simulated at such high resolution.

The results of this work will invite more direct semi-analytic modelling of the relevant star formation and feedback processes in the future. Moving forward, we

aim to more self-consistently model the formation sites of the first stellar systems and subsequent first galaxies including an enhanced treatment of the relevant radiative processes crucial to regulating each progenitor's assembly history. This modelling will then allow a more detailed understanding of the origin of the chemical make-up of not only the old stellar halo, but also its satellite systems. Only self-consistently modelling of the chemical and dynamical evolution of all of the progenitors of a Milky Way sized host will enable theoretical progress capable of connecting the low-redshift universe to the earliest phases of galaxy formation. Coupling the rich chemical and kinematic data being released by various observational Galactic sky surveys (e.g., GAIA-ESO, Gaia, SkyMapper, GALAH) with advanced modelling of this kind will contribute significantly to the nascent areas of both stellar and dwarf galaxy archaeology.

Appendix

We also carried out the same analysis on a higher resolution halo (LX15, $m_p = 3.73 \times 10^3$ $M_\odot$) which has a particle mass eight times higher than our fiducial run (LX14, $m_p = 2.99 \times 10^4$ $M_\odot$) to check that we identify the same total number of systems. The total number of systems identified as atomic cooling halos and molecular line cooling halos (minihalos) in the two resolution runs of the Cat-9 halo are shown in Figure 7-12. We find good agreement between the runs.

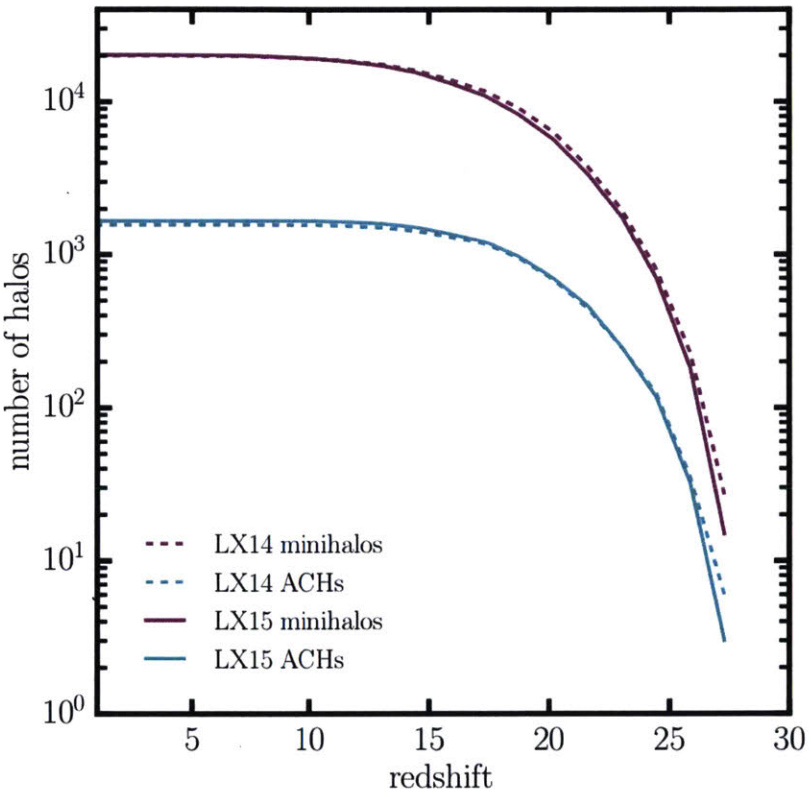

Figure 7-12 The total number of halos identified as molecular line cooling halos (minihalos) and atomic cooling halos in both our fiducial run (LX14, $m_p = 2.99 \times 10^4$ $M_\odot$) and a higher resolution run (LX15, $m_p = 3.73 \times 10^3$ $M_\odot$).

Chapter 8

Conclusion

This thesis consists of a series of investigations on the substructure of galaxies, particularly concerning satellites within the Local Group. Each study is based on analysis of dark matter only simulations of galaxies in a cosmological setting. The main results of each chapter are as follows:

Chapter 2

Although small adjustments to the cosmological parameters σ_8, n_s and Ω_m change the timescale in which halos of a given mass are expected to collapse and virialize, they do not lead to substantial changes in the properties of subhalo distributions for a given host. Increasing any of the parameters will cause a primordial matter overdensity to collapse into a self-bound halo sooner, reach a higher mass by $z = 0$ if it remains isolated, and be accreted and disrupted by a host sooner if it becomes a subhalo. However, the continuous process of subhalo accretion, mass loss, and ultimately destruction leads to a cosmology independent steady state distribution of subhalos within a host when normalized to the mass of the host. This means the subhalo mass function, peak mass function, distribution of accretion times, and radial distribution are unchanged. An exception is that the concentration of subhalos remains increased in cosmologies where objects tend to collapse sooner, which in turn leads to a slightly higher maximum circular velocity of subhalos. Furthermore, the formation time of subhalos remains slightly earlier.

Chapter 3

Self-interacting dark matter (SIDM) reduces the central density of galaxies relative that produced by ordinary cold dark matter, leading to cored density profiles. In this way it offers a resolution to the too big to fail and cusp/core problems. Searching for additional signatures, we find that SIDM causes stars which form in the center of halos to disperse out to larger distances, which leads to an enhanced level of stellar mass-loss in satellite galaxies relative to CDM. Consequently, SIDM suppresses the stellar mass function of satellites and increases the half-light radius of satellites. The level of enhanced stellar stripping can be predicted by the ratio of the subhalo core radius to the tidal radius, which can both be estimated analytically for an SIDM model. Comparing the class of velocity-independent (v-i) and velocity-dependent (v-d) models, tidal stripping is much more efficient at low pericenter orbits in v-d models, and for smaller galaxies, particularly for ultrafaint dwarfs. We further find that subhalo evaporation does not affect MW satellites for viable SIDM models, and that SIDM does not produce easily identifiable global signatures on the stellar halo.

Chapter 4

The number of luminous satellites around any galaxy can be predicted with a methodology developed that applies abundance matching models and a model for reionization to dark matter halo properties inferred from the *Caterpillar* simulation suite. The model fully samples statistical variance, and allows for rapid testing of how all input parameters control resultant predictions. Our baseline models predict around 70 ultrafaint and classical satellites ($M_*^{\rm sat} > 10^3\,{\rm M}_\odot$) around a MW-sized galaxy, consistent with observational expectations and devoid of any missing satellite problem. We also predict the abundance of satellites around specific isolated dwarf galaxies in the Local Group and characterize their expected radial distribution as observed in a line of sight. Even our most conservative model predicts that at least one satellite with $M_* > 10^5\,{\rm M}_\odot$ exists around the five largest host galaxies considered, and that $5-25$ satellites with $M_* > 10^4\,{\rm M}_\odot$ may exist. We advocate for dedicated searches

for these satellites to improve understanding of low mass galaxy formation, galaxy formation in environments with reduced ram-pressure and tidal stripping, the ability of reionization to suppress star formation, and the stellar mass-halo mass relationship for low mass galaxies.

Chapter 5

Using the same methodology as in Chapter 4, we predict the abundance of satellites around isolated galaxies within 8 Mpc of the MW that are comparable in mass to the LMC. We find that a single pointing of a 1.5° diameter field of view camera at most targets should on average contain $1-3$ satellites with $M_* > 10^5\,\mathrm{M}_\odot$. We predict the stellar mass function of satellite galaxies within 50 kpc of the LMC, and find large and intriguing discrepancies with observations. There are 12 known small UFDs ($M_* < 3 \times 10^3\,\mathrm{M}_\odot$) yet no known large UFDs ($M_* > 10^4\,\mathrm{M}_\odot$), and ongoing surveys are expected to increase only the number of known small UFDs. A model that predicts so few large UFDs and so many small UFDs may require large adjustments to currently used models that describe the abundance of classical dwarf satellites. Finding a solution to the puzzle will require, or result in, constraints on the redshift of reionization, the halo size at which star formation first begins, the stellar mass-halo mass relationship, the amount of tidal stripping in UFDs, the mass the MW, and the orbital history of the LMC and SMC. In spite of conflict with the stellar mass function, our model more robustly predicts that within the LMC vicinity, $\sim 47\%$ of satellites originated around the MW, $\sim 33\%$ around the LMC, and $\sim 20\%$ around the SMC.

Chapter 6

The *Caterpillar* project, the largest current suite of dark matter only zoom-in simulations of MW-mass galaxies with a particle resolution of $\sim 10^4\,\mathrm{M}_\odot$/particle is presented. Properties of halos and subhalos are cataloged, convergence tests and other detailed checks of the physical accuracy of the simulations are shown, as well as confirmation of the too big to fail problem. Shortcomings of the ROCKSTAR halo finding

algorithm are identified and corrected for by adding an iterative unbinding procedure. The *Caterpillar* simulations will serve as a valuable resource for research groups to study the diversity of the MW galactic halo and its formation history.

Chapter 7

The formation sites of the first generation of Population III stars and the first galaxies are identified and traced through mergers to the present day. Lyman-Werner radiation eliminates 90% of halos that otherwise would have formed Pop III stars, but still leaves several hundred that eventually become part of a Milky Way-like halo. Overwhelmingly, most of these star forming minihalos evolve in isolation without external chemical enrichment (80-90% of all systems at $z = 7$). However, $\sim$50% of them merge with a system larger than themselves within 1.5 Gyr after formation making them only temporarily purely self-enriched. We predict the number of minihalo and atomic cooling halo progenitor systems accreted by a given subhalo by $z = 0$ and find that low luminosity UFD galaxies, such as Reticulum II, have at most one or two star forming minihalo progenitors, whereas the LMC has $\sim$74-230. Thus the dichotomy of strongly and not at all R-processes enhanced stars in Reticulum II is unlikely to be explained by the merger of stars from multiple minihalos, though still possible.

Future topics to investigate

In the process of researching and writing each chapter, I identify a few areas worthy of future investigation. Regarding Chapter 5, the stellar mass function of LMC vicinity satellites deserves much more attention. Investigations should be made estimating how tidally stripped UFDs should be, and what the effects of the LMC, SMC and MW should be on the orbital dynamics of satellites originally accreted with the LMC and SMC. Predictions for the stellar mass function should be revisited once surveys for satellites within the entire LMC vicinity is completed.

The large number of small UFDs found near the LMC can and should also be connected to predictions from SIDM.Vogelsberger et al. (2016) showed that SIDM

will change the cosmological transfer function and reduce power on small scales, thus reducing the subhalo mass function for halos with $M_{\rm halo} < 10^{11}\,{\rm M}_{\odot}$. The suppression will be particularly strong for $M_{\rm halo} = 10^{8}\,{\rm M}_{\odot}$ halos, which are expected to be hosts of the ultrafaint dwarf galaxies. It would be interesting to apply the methods developed in Chapters 4 and 5 to the reduced subhalo mass function from SIDM. Producing as many UFDs as are found within the MW and near the LMC may require a significantly weaker model of reionization, with a far greater fraction of dark matter subhalos having to host luminous galaxies than in CDM. Given constraints on reionization and $M_{*} - M_{\rm halo}$ relationships, an investigation could lead to constraints on how much the subhalo mass function can be suppressed, and in turn lead to constrains on SIDM.

Lastly, the ability of SIDM to solve too big to fail should be tested against a diversity of simulated halos with differing formation and accretion histories. I have finished running, and started to run several simulations so that such tests can be made.

Bibliography

S J Aarseth. Dynamical evolution of clusters of galaxies, I. *MNRAS*, 126:223, 1963.

J. J. Adams, J. D. Simon, M. H. Fabricius, R. C. E. van den Bosch, J. C. Barentine, R. Bender, K. Gebhardt, G. J. Hill, J. D. Murphy, R. A. Swaters, J. Thomas, and G. van de Ven. Dwarf Galaxy Dark Matter Density Profiles Inferred from Stellar and Gas Kinematics. ApJ, 789:63, July 2014. doi: 10.1088/0004-637X/789/1/63.

Bhaskar Agarwal, Sadegh Khochfar, Jarrett L Johnson, Eyal Neistein, Claudio Dalla Vecchia, and Mario Livio. Ubiquitous seeding of supermassive black holes by direct collapse. *MNRAS*, 425(4):2854–2871, October 2012.

N. C. Amorisco, J. Zavala, and T. J. L. de Boer. Dark Matter Cores in the Fornax and Sculptor Dwarf Galaxies: Joining Halo Assembly and Detailed Star Formation Histories. ApJ, 782:L39, February 2014. doi: 10.1088/2041-8205/782/2/L39.

R. E. Angulo and S. D. M. White. One simulation to fit them all - changing the background parameters of a cosmological N-body simulation. MNRAS, 405:143–154, June 2010. doi: 10.1111/j.1365-2966.2010.16459.x.

G. W. Angus, H. Y. Shan, H. S. Zhao, and B. Famaey. On the Proof of Dark Matter, the Law of Gravity, and the Mass of Neutrinos. ApJ, 654:L13–L16, January 2007. doi: 10.1086/510738.

C. Armendariz-Picon and J. T. Neelakanta. How cold is cold dark matter? J. Cosmology Astropart. Phys., 3:049, March 2014. doi: 10.1088/1475-7516/2014/03/049.

Kenza S Arraki, Anatoly Klypin, Surhud More, and Sebastian Trujillo-Gomez. Effects of baryon removal on the structure of dwarf spheroidal galaxies. *MNRAS*, 438(2): 1466–1482, February 2014.

Vladimir Avila-Reese, Pedro Colin, Stefan Gottlöber, Claudio Firmani, and Christian Maulbetsch. The Dependence on Environment of Cold Dark Matter Halo Properties. *ApJ*, 634(1):51–69, November 2005.

J. S. Bagla, J. Prasad, and N. Khandai. Effects of the size of cosmological N-body simulations on physical quantities - III. Skewness. MNRAS, 395:918–930, May 2009. doi: 10.1111/j.1365-2966.2009.14592.x.

J. Bailin, E. F. Bell, M. Valluri, G. S. Stinson, V. P. Debattista, H. M. P. Couchman, and J. Wadsley. Systematic Problems with Using Dark Matter Simulations to Model Stellar Halos. ApJ, 783:95, March 2014. doi: 10.1088/0004-637X/783/2/95.

C. Barber, E. Starkenburg, J. F. Navarro, A. W. McConnachie, and A. Fattahi. The orbital ellipticity of satellite galaxies and the mass of the Milky Way. MNRAS, 437:959–967, January 2014. doi: 10.1093/mnras/stt1959.

R. Barkana and A. Loeb. The Photoevaporation of Dwarf Galaxies during Reionization. ApJ, 523:54–65, September 1999. doi: 10.1086/307724.

B. Bassett and R. Hlozek. *Baryon acoustic oscillations*, page 246. 2010.

Holger Baumgardt and Junichiro Makino. Dynamical evolution of star clusters in tidal fields. *MNRAS*, 340:227, March 2003.

K. Bechtol, A. Drlica-Wagner, E. Balbinot, A. Pieres, J. D. Simon, B. Yanny, B. Santiago, R. H. Wechsler, J. Frieman, A. R. Walker, P. Williams, E. Rozo, E. S. Rykoff, A. Queiroz, E. Luque, A. Benoit-Lévy, D. Tucker, I. Sevilla, R. A. Gruendl, L. N. da Costa, A. Fausti Neto, M. A. G. Maia, T. Abbott, S. Allam, R. Armstrong, A. H. Bauer, G. M. Bernstein, R. A. Bernstein, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, F. J. Castander, R. Covarrubias, C. B. D'Andrea, D. L. DePoy, S. Desai, H. T. Diehl, T. F. Eifler, J. Estrada, A. E. Evrard, E. Fernandez, D. A. Finley, B. Flaugher, E. Gaztanaga, D. Gerdes, L. Girardi, M. Gladders, D. Gruen, G. Gutierrez, J. Hao, K. Honscheid, B. Jain, D. James, S. Kent, R. Kron, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, H. Lin, M. Makler, M. March, J. Marshall, P. Martini, K. W. Merritt, C. Miller, R. Miquel, J. Mohr, E. Neilsen, R. Nichol, B. Nord, R. Ogando, J. Peoples, D. Petravick, A. A. Plazas, A. K. Romer, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, J. Thaler, D. Thomas, W. Wester, J. Zuntz, and DES Collaboration. Eight New Milky Way Companions Discovered in First-year Dark Energy Survey Data. ApJ, 807:50, July 2015. doi: 10.1088/0004-637X/807/1/50.

George D Becker, James S Bolton, and Adam Lidz. Reionisation and High-Redshift Galaxies: The View from Quasar Absorption Lines. *PASA*, 32:e045, December 2015.

P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. ApJ, 770:57, June 2013a. doi: 10.1088/0004-637X/770/1/57.

P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu. The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. ApJ, 762:109, January 2013b. doi: 10.1088/0004-637X/762/2/109.

P. S. Behroozi, R. H. Wechsler, H.-Y. Wu, M. T. Busha, A. A. Klypin, and J. R. Primack. Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology. ApJ, 763:18, January 2013c. doi: 10.1088/0004-637X/763/1/18.

Peter Behroozi, Alexander Knebe, Frazer R Pearce, Pascal Elahi, Jiaxin Han, Hanni Lux, Yao-Yuan Mao, Stuart I Muldrew, Doug Potter, and Chaichalit Srisawat. Major Mergers Going Notts: Challenges for Modern Halo Finders. *arXiv*, June 2015.

Peter S Behroozi, Risa H Wechsler, Hao-Yi Wu, Michael T Busha, Anatoly A Klypin, and Joel R Primack. GRAVITATIONALLY CONSISTENT HALO CATALOGS AND MERGER TREES FOR PRECISION COSMOLOGY. *ApJ*, 763(1):18, December 2012.

K. Bekki. When was the Large Magellanic Cloud accreted on to the Galaxy? MNRAS, 416:2359–2367, September 2011. doi: 10.1111/j.1365-2966.2011.19211.x.

E. F. Bell, D. B. Zucker, V. Belokurov, S. Sharma, K. V. Johnston, J. S. Bullock, D. W. Hogg, K. Jahnke, J. T. A. de Jong, T. C. Beers, N. W. Evans, E. K. Grebel, Ž. Ivezić, S. E. Koposov, H.-W. Rix, D. P. Schneider, M. Steinmetz, and A. Zolotov. The Accretion Origin of the Milky Way's Stellar Halo. ApJ, 680: 295–311, June 2008. doi: 10.1086/588032.

E. F. Bell, C. T. Slater, and N. F. Martin. Andromeda XXIX: A New Dwarf Spheroidal Galaxy 200 kpc from Andromeda. ApJ, 742:L15, November 2011. doi: 10.1088/2041-8205/742/1/L15.

M. Bellazzini, R. Ibata, L. Monaco, N. Martin, M. J. Irwin, and G. F. Lewis. Detection of the Canis Major galaxy at (l;b) = (244; -8) and in the background of Galactic open clusters. MNRAS, 354:1263–1278, November 2004. doi: 10.1111/j.1365-2966.2004.08283.x.

M. Bellazzini, R. Ibata, N. Martin, G. F. Lewis, B. Conn, and M. J. Irwin. The core of the Canis Major galaxy as traced by red clump stars. MNRAS, 366:865–883, March 2006. doi: 10.1111/j.1365-2966.2005.09973.x.

V. Belokurov, D. B. Zucker, N. W. Evans, M. I. Wilkinson, M. J. Irwin, S. Hodgkin, D. M. Bramich, J. M. Irwin, G. Gilmore, B. Willman, S. Vidrih, H. J. Newberg, R. F. G. Wyse, M. Fellhauer, P. C. Hewett, N. Cole, E. F. Bell, T. C. Beers, C. M. Rockosi, B. Yanny, E. K. Grebel, D. P. Schneider, R. Lupton, J. C. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, A. Nitta, J. A. Smith, and S. A. Snedden. A Faint New Milky Way Satellite in Bootes. ApJ, 647:L111–L114, August 2006. doi: 10.1086/507324.

V. Belokurov, D. B. Zucker, N. W. Evans, J. T. Kleyna, S. Koposov, S. T. Hodgkin, M. J. Irwin, G. Gilmore, M. I. Wilkinson, M. Fellhauer, D. M. Bramich, P. C. Hewett, S. Vidrih, J. T. A. De Jong, J. A. Smith, H.-W. Rix, E. F. Bell, R. F. G. Wyse, H. J. Newberg, P. A. Mayeur, B. Yanny, C. M. Rockosi, O. Y. Gnedin, D. P.

Schneider, T. C. Beers, J. C. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, A. Nitta, and S. A. Snedden. Cats and Dogs, Hair and a Hero: A Quintet of New Milky Way Companions. ApJ, 654: 897–906, January 2007. doi: 10.1086/509718.

V. Belokurov, M. G. Walker, N. W. Evans, D. C. Faria, G. Gilmore, M. J. Irwin, S. Koposov, M. Mateo, E. Olszewski, and D. B. Zucker. Leo V: A Companion of a Companion of the Milky Way Galaxy? ApJ, 686:L83–L86, October 2008. doi: 10.1086/592962.

V. Belokurov, M. G. Walker, N. W. Evans, G. Gilmore, M. J. Irwin, M. Mateo, L. Mayer, E. Olszewski, J. Bechtold, and T. Pickering. The discovery of Segue 2: a prototype of the population of satellites of satellites. MNRAS, 397:1748–1755, August 2009. doi: 10.1111/j.1365-2966.2009.15106.x.

V. Belokurov, M. G. Walker, N. W. Evans, G. Gilmore, M. J. Irwin, D. Just, S. Koposov, M. Mateo, E. Olszewski, L. Watkins, and L. Wyrzykowski. Big Fish, Little Fish: Two New Ultra-faint Satellites of the Milky Way. ApJ, 712:L103–L106, March 2010. doi: 10.1088/2041-8205/712/1/L103.

C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *ArXiv e-prints*, December 2012.

A. J. Benson, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole. The effects of photoionization on galaxy formation - II. Satellite galaxies in the Local Group. MNRAS, 333:177–190, June 2002. doi: 10.1046/j.1365-8711.2002.05388.x.

L. Bergström. Dark matter evidence, particle physics candidates and detection methods. *Annalen der Physik*, 524:479–496, October 2012. doi: 10.1002/andp.201200116.

E. Bertschinger. Multiscale Gaussian Random Fields and Their Application to Cosmological Simulations. ApJS, 137:1–20, November 2001. doi: 10.1086/322526.

G. Besla, N. Kallivayalil, L. Hernquist, B. Robertson, T. J. Cox, R. P. van der Marel, and C. Alcock. Are the Magellanic Clouds on Their First Passage about the Milky Way? ApJ, 668:949–967, October 2007. doi: 10.1086/521385.

J. Binney and S. Tremaine. *Galactic Dynamics*. Princeton Univ. Press, 2nd ed. edition, 2008.

M. S. Bovill and M. Ricotti. Pre-Reionization Fossils, Ultra-Faint Dwarfs, and the Missing Galactic Satellite Problem. ApJ, 693:1859–1870, March 2009. doi: 10.1088/0004-637X/693/2/1859.

M. S. Bovill and M. Ricotti. Where are the Fossils of the First Galaxies? I. Local Volume Maps and Properties of the Undetected Dwarfs. ApJ, 741:17, November 2011a. doi: 10.1088/0004-637X/741/1/17.

M. S. Bovill and M. Ricotti. Where are the Fossils of the First Galaxies? II. True Fossils, Ghost Halos, and the Missing Bright Satellites. ApJ, 741:18, November 2011b. doi: 10.1088/0004-637X/741/1/18.

J. Bovy. Dynamical Modeling of Tidal Streams. ApJ, 795:95, November 2014. doi: 10.1088/0004-637X/795/1/95.

M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. MNRAS, 398:1150–1164, September 2009. doi: 10.1111/j.1365-2966.2009.15191.x.

M. Boylan-Kolchin, V. Springel, S. D. M. White, and A. Jenkins. There's no place like home? Statistics of Milky Way-mass dark matter haloes. MNRAS, 406:896–912, August 2010. doi: 10.1111/j.1365-2966.2010.16774.x.

M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. MNRAS, 415:L40–L44, July 2011. doi: 10.1111/j.1745-3933.2011.01074.x.

M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat. The Milky Way's bright satellites as an apparent failure of ΛCDM. MNRAS, 422:1203–1218, May 2012. doi: 10.1111/j.1365-2966.2012.20695.x.

Michael Boylan-Kolchin, James S Bullock, Sangmo Tony Sohn, Gurtina Besla, and Roeland P van der Marel. THE SPACE MOTION OF LEO I: THE MASS OF THE MILKY WAY'S DARK MATTER HALO. *ApJ*, 768(2):140, April 2013.

R. H. Brandenberger. Introduction to Early Universe Cosmology. *ArXiv e-prints*, March 2011.

M. A. Breddels and A. Helmi. Model comparison of the dark matter profiles of Fornax, Sculptor, Carina and Sextans. A&A, 558:A35, October 2013. doi: 10.1051/0004-6361/201321606.

V. Bromm and N. Yoshida. The First Galaxies. ARA&A, 49:373–407, September 2011. doi: 10.1146/annurev-astro-081710-102608.

Volker Bromm, Paolo S Coppi, and Richard B Larson. Forming the First Stars in the Universe: The Fragmentation of Primordial Gas. *ApJ*, 527(1):L5–L8, December 1999.

C. B. Brook and A. Di Cintio. Expanded haloes, abundance matching and too-big-to-fail in the Local Group. MNRAS, 450:3920–3934, July 2015. doi: 10.1093/mnras/stv864.

C. B. Brook and F. Shankar. A matter of measurement: rotation velocities and the velocity function of dwarf galaxies. MNRAS, 455:3841–3847, February 2016. doi: 10.1093/mnras/stv2550.

C. B. Brook, A. Di Cintio, A. Knebe, S. Gottlöber, Y. Hoffman, G. Yepes, and S. Garrison-Kimmel. The Stellar-to-halo Mass Relation for Local Group Galaxies. ApJ, 784:L14, March 2014. doi: 10.1088/2041-8205/784/1/L14.

A. M. Brooks and A. Zolotov. Why Baryons Matter: The Kinematics of Dwarf Spheroidal Satellites. ApJ, 786:87, May 2014. doi: 10.1088/0004-637X/786/2/87.

A. M. Brooks, M. Kuhlen, A. Zolotov, and D. Hooper. A Baryonic Solution to the Missing Satellites Problem. ApJ, 765:22, March 2013. doi: 10.1088/0004-637X/765/1/22.

T. M. Brown, J. Tumlinson, M. Geha, E. N. Kirby, D. A. VandenBerg, R. R. Muñoz, J. S. Kalirai, J. D. Simon, R. J. Avila, P. Guhathakurta, A. Renzini, and H. C. Ferguson. The Primeval Populations of the Ultra-faint Dwarf Galaxies. ApJ, 753: L21, July 2012. doi: 10.1088/2041-8205/753/1/L21.

T. M. Brown, J. Tumlinson, M. Geha, E. Kirby, D. A. VandenBerg, J. S. Kalirai, J. D. Simon, R. J. Avila, R. R. Munoz, P. Guhathakurta, A. Renzini, H. C. Ferguson, L. C. Vargas, and M. Gennaro. The formation history of the Ultra-Faint Dwarf galaxies . Mem. Soc. Astron. Italiana, 85:493, 2014a.

T. M. Brown, J. Tumlinson, M. Geha, J. D. Simon, L. C. Vargas, D. A. VandenBerg, E. N. Kirby, J. S. Kalirai, R. J. Avila, M. Gennaro, H. C. Ferguson, R. R. Muñoz, P. Guhathakurta, and A. Renzini. The Quenching of the Ultra-faint Dwarf Galaxies in the Reionization Era. ApJ, 796:91, December 2014b. doi: 10.1088/0004-637X/796/2/91.

G. L. Bryan and M. L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. ApJ, 495:80–99, March 1998. doi: 10.1086/305262.

Greg L Bryan, Michael L Norman, Brian W O'Shea, Tom Abel, John H Wise, Matthew J Turk, Daniel R Reynolds, David C Collins, Peng Wang, Samuel W Skillman, Britton Smith, Robert P Harkness, James Bordner, Ji-hoon Kim, Michael Kuhlen, Hao Xu, Nathan Goldbaum, Cameron Hummels, Alexei G Kritsuk, Elizabeth Tasker, Stephen Skory, Christine M Simpson, Oliver Hahn, Jeffrey S Oishi, Geoffrey C So, Fen Zhao, Renyue Cen, Yuan Li, and The Enzo Collaboration. ENZO: AN ADAPTIVE MESH REFINEMENT CODE FOR ASTROPHYSICS. *ApJS*, 211(2):19, March 2014.

Tobias Buck, Andrea V Maccio, and Aaron A Dutton. Evidence for Early Filamentary Accretion from the Andromeda Galaxy's Thin Plane of Satellites. *ApJ*, 809(1):49, August 2015.

M. R. Buckley and P. J. Fox. Dark matter self-interactions and light force carriers. Phys. Rev. D, 81(8):083522, April 2010. doi: 10.1103/PhysRevD.81.083522.

J. S. Bullock and K. V. Johnston. Tracing Galaxy Formation with Stellar Halos. I. Methods. ApJ, 635:931–949, December 2005. doi: 10.1086/497422.

J. S. Bullock, A. V. Kravtsov, and D. H. Weinberg. Reionization and the Abundance of Galactic Satellites. ApJ, 539:517–521, August 2000. doi: 10.1086/309279.

J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. MNRAS, 321:559–575, March 2001. doi: 10.1046/j.1365-8711.2001.04068.x.

A. Burkert. The Structure of Dark Matter Halos in Dwarf Galaxies. ApJ, 447:L25, July 1995. doi: 10.1086/309560.

A. Burkert. The Structure and Evolution of Weakly Self-interacting Cold Dark Matter Halos. ApJ, 534:L143–L146, May 2000. doi: 10.1086/312674.

M. T. Busha, P. J. Marshall, R. H. Wechsler, A. Klypin, and J. Primack. The Mass Distribution and Assembly of the Milky Way from the Properties of the Magellanic Clouds. ApJ, 743:40, December 2011a. doi: 10.1088/0004-637X/743/1/40.

M. T. Busha, R. H. Wechsler, P. S. Behroozi, B. F. Gerke, A. A. Klypin, and J. R. Primack. Statistics of Satellite Galaxies around Milky-Way-like Hosts. ApJ, 743: 117, December 2011b. doi: 10.1088/0004-637X/743/2/117.

F. Capela, M. Pshirkov, and P. Tinyakov. Constraints on primordial black holes as dark matter candidates from capture by neutron stars. Phys. Rev. D, 87(12): 123524, June 2013. doi: 10.1103/PhysRevD.87.123524.

J. L. Carlin, D. J. Sand, P. Price, B. Willman, A. Karunakaran, K. Spekkens, E. F. Bell, J. P. Brodie, D. Crnojević, D. A. Forbes, J. Hargis, E. Kirby, R. Lupton, A. H. G. Peter, A. J. Romanowsky, and J. Strader. First results from the MADCASH Survey: A Faint Dwarf Galaxy Companion to the Low Mass Spiral Galaxy NGC 2403 at 3.2 Mpc. *ArXiv e-prints*, August 2016.

E. D. Carlson, M. E. Machacek, and L. J. Hall. Self-interacting dark matter. ApJ, 398:43–52, October 1992. doi: 10.1086/171833.

G. Carraro. The Milky Way disk. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 57:138–144, August 2015.

G. Castignani, N. Frusciante, D. Vernieri, and P. Salucci. The density profiles of Dark Matter halos in Spiral Galaxies. *ArXiv e-prints*, January 2012.

C.-K. Chang, C.-M. Ko, and T.-H. Peng. Information on the Milky Way from the Two Micron All Sky Survey Whole Sky Star Count: The Structure Parameters. ApJ, 740:34, October 2011. doi: 10.1088/0004-637X/740/1/34.

L. Chen, Q.-G. Huang, and K. Wang. Constraint on the abundance of primordial black holes in dark matter from Planck data. *ArXiv e-prints*, August 2016.

K. Chiboucas, B. A. Jacobs, R. B. Tully, and I. D. Karachentsev. Confirmation of Faint Dwarf Galaxies in the M81 Group. AJ, 146:126, November 2013. doi: 10.1088/0004-6256/146/5/126.

D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter. ApJ, 648: L109–L113, September 2006. doi: 10.1086/508162.

P. Colín, V. Avila-Reese, O. Valenzuela, and C. Firmani. Structure and Subhalo Population of Halos in a Self-interacting Dark Matter Cosmology. ApJ, 581:777–793, December 2002. doi: 10.1086/344259.

M. L. M. Collins, E. J. Tollerud, D. J. Sand, A. Bonaca, B. Willman, and J. Strader. Dynamical evidence for a strong tidal interaction between the Milky Way and its satellite, Leo V. *ArXiv e-prints*, August 2016.

A. P. Cooper, S. Cole, C. S. Frenk, S. D. M. White, J. Helly, A. J. Benson, G. De Lucia, A. Helmi, A. Jenkins, J. F. Navarro, V. Springel, and J. Wang. Galactic stellar haloes in the CDM model. MNRAS, 406:744–766, August 2010. doi: 10.1111/j.1365-2966.2010.16740.x.

Andrew P Cooper, Owen H Parry, Ben Lowing, Shaun Cole, and Carlos Frenk. Formation of In Situ Stellar Haloes in Milky Way-Mass Galaxies. *arXiv*, January 2015.

Lauren Corlies, Kathryn V Johnston, Jason Tumlinson, and Greg Bryan. CHEMICAL ABUNDANCE PATTERNS AND THE EARLY ENVIRONMENT OF DWARF GALAXIES. *ApJ*, 773(2):105, July 2013.

D. Crnojević, D. J. Sand, N. Caldwell, P. Guhathakurta, B. McLeod, A. Seth, J. D. Simon, J. Strader, and E. Toloba. Discovery of a Close Pair of Faint Dwarf Galaxies in the Halo of Centaurus A. ApJ, 795:L35, November 2014. doi: 10.1088/2041-8205/795/2/L35.

D. Crnojević, D. J. Sand, K. Spekkens, N. Caldwell, P. Guhathakurta, B. McLeod, A. Seth, J. D. Simon, J. Strader, and E. Toloba. The Extended Halo of Centaurus A: Uncovering Satellites, Streams, and Substructures. ApJ, 823:19, May 2016. doi: 10.3847/0004-637X/823/1/19.

B. D. Crosby, B. W. O'Shea, T. C. Beers, and J. Tumlinson. Tracing the Evolution of High Redshift Galaxies Using Stellar Abundances. *ArXiv e-prints*, December 2013.

Brian D Crosby, Brian W O'Shea, Britton D Smith, Matthew J Turk, and Oliver Hahn. Population III Star Formation in Large Cosmological Volumes. I. Halo Temporal and Physical Environment. *ApJ*, 773(2):108, August 2013.

F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, and C. Pfrommer. ETHOS - An Effective Theory of Structure Formation: From dark particle physics to the matter distribution of the Universe. *ArXiv e-prints*, December 2015.

R. Davé, D. N. Spergel, P. J. Steinhardt, and B. D. Wandelt. Halo Properties in Cosmological Simulations of Self-interacting Cold Dark Matter. ApJ, 547:574–589, February 2001. doi: 10.1086/318417.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. ApJ, 292:371–394, May 1985. doi: 10.1086/163168.

W. J. G. de Blok. The Core-Cusp Problem. *Advances in Astronomy*, 2010:789293, 2010. doi: 10.1155/2010/789293.

W. J. G. de Blok and S. S. McGaugh. The dark and visible matter content of low surface brightness disc galaxies. MNRAS, 290:533–552, September 1997.

W. J. G. de Blok, S. S. McGaugh, A. Bosma, and V. C. Rubin. Mass Density Profiles of Low Surface Brightness Galaxies. ApJ, 552:L23–L26, May 2001. doi: 10.1086/320262.

J. T. A. de Jong, D. J. Butler, H. W. Rix, A. E. Dolphin, and D. Martínez-Delgado. Modeling the Stellar Populations in the Canis Major Overdensity: The Relation between the Old and Young Populations. ApJ, 662:259–271, June 2007. doi: 10.1086/517967.

Roelof S de Jong, A C Seth, E F Bell, T M Brown, J S Bullock, S Courteau, J J Dalcanton, H C Ferguson, P Goudfrooij, S Holfeltz, C Purcell, D Radburn-Smith, and D Zucker. GHOSTS: The Resolved Stellar Outskirts of Massive Disk Galaxies. *IAU*, 241(S241):503–504, August 2007.

A. A. de Laix, R. J. Scherrer, and R. K. Schaefer. Constraints on Self-interacting Dark Matter. ApJ, 452:495, October 1995. doi: 10.1086/176322.

G. De Lucia and A. Helmi. The Galaxy and its stellar halo: insights on their formation from a hybrid cosmological approach. MNRAS, 391:14–31, November 2008. doi: 10.1111/j.1365-2966.2008.13862.x.

A. J. Deason, A. R. Wetzel, S. Garrison-Kimmel, and V. Belokurov. Satellites of LMC-mass dwarfs: close friendships ruined by Milky Way mass haloes. MNRAS, 453:3568–3574, November 2015. doi: 10.1093/mnras/stv1939.

A Del Popolo, J A S Lima, Júlio C Fabris, and Davi C Rodrigues. A unified solution to the small scale problems of the LambdaCDM model. *J. Cosmol. Astropart. Phys.*, 4(04):021–021, April 2014.

Li-Cai Deng, Heidi Jo Newberg, Chao Liu, Jeffrey L Carlin, Timothy C Beers, Li Chen, Yu-Qin Chen, Norbert Christlieb, Carl J Grillmair, Puragra Guhathakurta, Zhan-Wen Han, Jin-Liang Hou, Hsu-Tai Lee, Sébastien Lépine, Jing Li, Xiao-Wei Liu, Kai-Ke Pan, J A Sellwood, Bo Wang, Hong-Chi Wang, Fan Yang, Brian Yanny, Hao-Tong Zhang, Yue-Yang Zhang, Zheng Zheng, and Zi Zhu. LAMOST Experiment for Galactic Understanding and Exploration (LEGUE) — The survey's science plan. *RAA*, 12(7):735–754, June 2012.

DES Collaboration. Eight New Milky Way Companions Discovered in First-year Dark Energy Survey Data. ApJ, 807:50, July 2015. doi: 10.1088/0004-637X/807/1/50.

J. Diemand, M. Kuhlen, and P. Madau. Dark Matter Substructure and Gamma-Ray Annihilation in the Milky Way Halo. ApJ, 657:262–270, March 2007. doi: 10.1086/510736.

J Diemand, M Kuhlen, P Madau, M Zemp, B Moore, D Potter, and J Stadel. Clumps and streams in the local dark matter distribution. *Nature*, 454(7205):735–738, August 2008.

K. Dolag, S. Borgani, S. Schindler, A. Diaferio, and A. M. Bykov. Simulation Techniques for Cosmological Simulations. Space Sci. Rev., 134:229–268, February 2008. doi: 10.1007/s11214-008-9316-5.

E. D'Onghia and G. Lake. Small Dwarf Galaxies within Larger Dwarfs: Why Some Are Luminous while Most Go Dark. ApJ, 686:L61, October 2008. doi: 10.1086/592995.

G. A. Dooley, B. F. Griffen, P. Zukin, A. P. Ji, M. Vogelsberger, L. E. Hernquist, and A. Frebel. The Effects of Varying Cosmological Parameters on Halo Substructure. ApJ, 786:50, May 2014. doi: 10.1088/0004-637X/786/1/50.

G. A. Dooley, A. H. G. Peter, M. Vogelsberger, J. Zavala, and A. Frebel. Enhanced tidal stripping of satellites in the galactic halo from dark matter self-interactions. MNRAS, 461:710–727, September 2016a. doi: 10.1093/mnras/stw1309.

G. A. Dooley, A. H. G. Peter, T. Yang, B. Willman, B. F. Griffen, and A. Frebel. An observer's guide to the (Local Group) dwarf galaxies: predictions for their own dwarf satellite populations. *ArXiv e-prints*, October 2016b.

M. Drees and G. Gerbier. Mini–Review of Dark Matter: 2012. *ArXiv e-prints*, April 2012.

A. Drlica-Wagner, K. Bechtol, E. S. Rykoff, E. Luque, A. Queiroz, Y.-Y. Mao, R. H. Wechsler, J. D. Simon, B. Santiago, B. Yanny, E. Balbinot, S. Dodelson, A. Fausti Neto, D. J. James, T. S. Li, M. A. G. Maia, J. L. Marshall, A. Pieres, K. Stringer, A. R. Walker, T. M. C. Abbott, F. B. Abdalla, S. Allam, A. Benoit-Lévy, G. M. Bernstein, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, M. Crocce, L. N. da Costa, S. Desai, H. T. Diehl,

J. P. Dietrich, P. Doel, T. F. Eifler, A. E. Evrard, D. A. Finley, B. Flaugher, P. Fosalba, J. Frieman, E. Gaztanaga, D. W. Gerdes, D. Gruen, R. A. Gruendl, G. Gutierrez, K. Honscheid, K. Kuehn, N. Kuropatkin, O. Lahav, P. Martini, R. Miquel, B. Nord, R. Ogando, A. A. Plazas, K. Reil, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Tucker, V. Vikram, W. Wester, Y. Zhang, J. Zuntz, and DES Collaboration. Eight Ultra-faint Galaxy Candidates Discovered in Year Two of the Dark Energy Survey. ApJ, 813:109, November 2015. doi: 10.1088/0004-637X/813/2/109.

A. Drlica-Wagner, K. Bechtol, S. Allam, D. L. Tucker, R. A. Gruendl, M. D. Johnson, A. R. Walker, D. J. James, D. L. Nidever, K. A. G. Olsen, R. H. Wechsler, M. R. L. Cioni, B. C. Conn, K. Kuehn, T. S. Li, Y.-Y. Mao, N. F. Martin, E. Neilsen, N. E. D. Noël, A. Pieres, J. D. Simon, G. S. Stringfellow, R. P. van der Marel, and B. Yanny. An Ultra-Faint Galaxy Candidate Discovered in Early Data from the Magellanic Satellites Survey. *ArXiv e-prints*, September 2016.

A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. MNRAS, 390:L64–L68, October 2008. doi: 10.1111/j.1745-3933.2008.00537.x.

J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, J. L. Weiland, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Likelihoods and Parameters from the WMAP Data. ApJS, 180:306–329, February 2009. doi: 10.1088/0067-0049/180/2/306.

G. Efstathiou. Suppressing the formation of dwarf galaxies via photoionization. MNRAS, 256:43P–47P, May 1992. doi: 10.1093/mnras/256.1.43P.

J. Einasto. 1965. Tr. Inst. Astroz. Alma-Ata, 51, 87.

D. J. Eisenstein and W. Hu. Baryonic Features in the Matter Transfer Function. ApJ, 496:605, March 1998. doi: 10.1086/305424.

O. D. Elbert, J. S. Bullock, S. Garrison-Kimmel, M. Rocha, J. Oñorbe, and A. H. G. Peter. Core formation in dwarf haloes with self-interacting dark matter: no fine-tuning necessary. MNRAS, 453:29–37, October 2015. doi: 10.1093/mnras/stv1470.

A. Emerick, M.-M. Mac Low, J. Grcevich, and A. Gatto. Gas Loss by Ram Pressure Stripping and Internal Feedback from Low-mass Milky Way Satellites. ApJ, 826: 148, August 2016. doi: 10.3847/0004-637X/826/2/148.

R. Errani, J. Peñarrubia, and G. Tormen. Constraining the distribution of dark matter in dwarf spheroidal galaxies with stellar tidal streams. MNRAS, 449:L46–L50, April 2015. doi: 10.1093/mnrasl/slv012.

H. Falcke and S. B. Markoff. Toward the event horizon — the supermassive black hole in the Galactic Center. *Classical and Quantum Gravity*, 30(24):244003, December 2013. doi: 10.1088/0264-9381/30/24/244003.

J. L. Feng, M. Kaplinghat, and H.-B. Yu. Sommerfeld enhancements for thermal relic dark matter. Phys. Rev. D, 82(8):083525, October 2010. doi: 10.1103/PhysRevD.82.083525.

T. Fermi-LAT, DES Collaborations, A. Albert, B. Anderson, K. Bechtol, A. Drlica-Wagner, M. Meyer, M. Sanchez-Conde, L. Strigari, M. Wood, T. M. C. Abbott, F. B. Abdalla, A. Benoit-Levy, G. M. Bernstein, R. A. Bernstein, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, T. F. Eifler, A. E. Evrard, A. Fausti Neto, D. A. Finley, B. Flaugher, P. Fosalba, J. Frieman, D. W. Gerdes, D. A. Goldstein, D. Gruen, R. A. Gruendl, K. Honscheid, D. J. James, S. Kent, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, M. A. G. Maia, M. March, J. L. Marshall, P. Martini, C. J. Miller, R. Miquel, E. Neilsen, B. Nord, R. Ogando, A. A. Plazas, K. Reil, A. K. Romer, E. S. Rykoff, E. Sanchez, B. Santiago, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, V. Vikram, A. R. Walker, and R. H. Wechsler. Searching for Dark Matter Annihilation in Recently Discovered Milky Way Satellites with Fermi-LAT. *ArXiv e-prints*, November 2016.

S. P. Fillingham, M. C. Cooper, A. B. Pace, M. Boylan-Kolchin, J. S. Bullock, S. Garrison-Kimmel, and C. Wheeler. Under pressure: quenching star formation in low-mass satellite galaxies via stripping. MNRAS, 463:1916–1928, December 2016. doi: 10.1093/mnras/stw2131.

Steven L Finkelstein, Russell E Jr Ryan, Casey Papovich, Mark Dickinson, Mimi Song, Rachel S Somerville, Henry C Ferguson, Brett Salmon, Mauro Giavalisco, Anton M Koekemoer, Matthew L N Ashby, Peter Behroozi, Marco Castellano, James S Dunlop, Sandy M Faber, Giovanni G Fazio, Adriano Fontana, Norman A Grogin, Nimish Hathi, Jason Jaacks, Dale D Kocevski, Rachael Livermore, Ross J McLure, Emiliano Merlin, Bahram Mobasher, Jeffrey A Newman, Marc Rafelski, Vithal Tilvi, and S P Willner. The Evolution of the Galaxy Rest-frame Ultraviolet Luminosity Function over the First Two Billion Years. *ApJ*, 810(1):71, September 2015.

C. Firmani, E. D'Onghia, V. Avila-Reese, G. Chincarini, and X. Hernández. Evidence of self-interacting cold dark matter from galactic to galaxy cluster scales. MNRAS, 315:L29–L32, July 2000. doi: 10.1046/j.1365-8711.2000.03555.x.

R. A. Flores and J. R. Primack. Observational and theoretical constraints on singular dark matter halos. ApJ, 427:L1–L4, May 1994. doi: 10.1086/187350.

Andreea S Font, Kathryn V Johnston, James S Bullock, and Brant E Robertson. Phase-Space Distributions of Chemical Abundances in Milky Way–Type Galaxy Halos. *ApJ*, 646(2):886–898, August 2006.

A Frebel. Stellar archaeology: Exploring the Universe with metal-poor stars. *Astron. Nachr.*, 331(5):474–488, May 2010.

A. Frebel and V. Bromm. Chemical Signatures of the First Galaxies: Criteria for One-shot Enrichment. ApJ, 759:115, November 2012. doi: 10.1088/0004-637X/759/2/115.

Anna Frebel and John E Norris. Near-Field Cosmology with Extremely Metal-Poor Stars. *ARA&A*, 53(1):631–688, August 2015.

Anna Frebel, Joshua D Simon, and Evan N Kirby. SEGUE 1: AN UNEVOLVED FOSSIL GALAXY FROM THE EARLY UNIVERSE. *ApJ*, 786(1):74, April 2014.

K C Freeman. The HERMES Project: Reconstructing Galaxy Formation. *Satellites and Tidal Streams*, 458:393, August 2012.

Ken Freeman and Joss Bland-Hawthorn. The New Galaxy: Signatures of Its Formation. *ARA&A*, 40(1):487–537, 2002.

C. S. Frenk and S. D. M. White. Dark matter and cosmic structure. *Ann. Phys.*, 524: 507–534, October 2012. doi: 10.1002/andp.201200212.

A. B. Fry, F. Governato, A. Pontzen, T. Quinn, M. Tremmel, L. Anderson, H. Menon, A. M. Brooks, and J. Wadsley. All about baryons: revisiting SIDM predictions at small halo masses. MNRAS, 452:1468–1479, September 2015. doi: 10.1093/mnras/stv1330.

L. Gao, S. D. M. White, A. Jenkins, F. Stoehr, and V. Springel. The subhalo populations of ΛCDM dark haloes. MNRAS, 355:819–834, December 2004. doi: 10.1111/j.1365-2966.2004.08360.x.

L Gao, Tom Theuns, C S Frenk, A Jenkins, J C Helly, J Navarro, Volker Springel, and S D M White. The earliest stars and their relics in the Milky Way. *MNRAS*, 403(3):1283–1295, April 2010.

Jonathan P Gardner, John C Mather, Mark Clampin, Rene Doyon, Matthew A Greenhouse, Heidi B Hammel, John B Hutchings, Peter Jakobsen, Simon J Lilly, Knox S Long, Jonathan I Lunine, Mark J McCaughrean, Matt Mountain, John Nella, George H Rieke, Marcia J Rieke, Hans-Walter Rix, Eric P Smith, George Sonneborn, Massimo Stiavelli, H S Stockman, Rogier A Windhorst, and Gillian S Wright. The James Webb Space Telescope. *Space Science Reviews*, 123(4):485–606, April 2006.

S. Garrison-Kimmel, M. Rocha, M. Boylan-Kolchin, J. S. Bullock, and J. Lally. Can feedback solve the too-big-to-fail problem? MNRAS, 433:3539–3546, August 2013. doi: 10.1093/mnras/stt984.

S. Garrison-Kimmel, M. Boylan-Kolchin, J. S. Bullock, and E. N. Kirby. Too big to fail in the Local Group. MNRAS, 444:222–236, October 2014a. doi: 10.1093/mnras/stu1477.

S. Garrison-Kimmel, M. Boylan-Kolchin, J. S. Bullock, and K. Lee. ELVIS: Exploring the Local Volume in Simulations. MNRAS, 438:2578–2596, March 2014b. doi: 10.1093/mnras/stt2377.

S. Garrison-Kimmel, J. S. Bullock, M. Boylan-Kolchin, and E. Bardwell. Organized Chaos: Scatter in the relation between stellar mass and halo mass in small galaxies. *ArXiv e-prints*, March 2016.

G. Gentile, P. Salucci, U. Klein, D. Vergani, and P. Kalberla. The cored distribution of dark matter in spiral galaxies. MNRAS, 351:903–922, July 2004. doi: 10.1111/j.1365-2966.2004.07836.x.

C. Giocoli, G. Tormen, R. K. Sheth, and F. C. van den Bosch. The substructure hierarchy in dark matter haloes. *MNRAS*, 404:502–517, May 2010. doi: 10.1111/j.1365-2966.2010.16311.x.

N. Y. Gnedin. Effect of Reionization on Structure Formation in the Universe. ApJ, 542:535–541, October 2000. doi: 10.1086/317042.

O. Y. Gnedin and J. P. Ostriker. Limits on Collisional Dark Matter from Elliptical Galaxies in Clusters. ApJ, 561:61–68, November 2001. doi: 10.1086/323211.

Facundo A Gómez, Amina Helmi, Anthony G A Brown, and Yang-Shyang Li. On the identification of merger debris in the Gaia era. *MNRAS*, 408(2):935–946, October 2010.

Facundo A Gómez, Christopher E Coleman-Smith, Brian W O'Shea, Jason Tumlinson, and Robert L Wolpert. Characterizing the Formation History of Milky Way like Stellar Halos with Model Emulators. *ApJ*, 760(2):112, December 2012.

Facundo A Gómez, Christopher E Coleman-Smith, Brian W O'Shea, Jason Tumlinson, and Robert L Wolpert. Dissecting Galaxy Formation Models with Sensitivity Analysis—a New Approach to Constrain the Milky Way Formation History. *ApJ*, 787(1):20, May 2014.

O. A. Gonzalez and D. Gadotti. The Milky Way Bulge: Observed Properties and a Comparison to External Galaxies. *Galactic Bulges*, 418:199, 2016. doi: 10.1007/978-3-319-19378-6_9.

R. E. González, A. V. Kravtsov, and N. Y. Gnedin. On the Mass of the Local Group. ApJ, 793:91, October 2014. doi: 10.1088/0004-637X/793/2/91.

S. Gottloeber, Y. Hoffman, and G. Yepes. Constrained Local UniversE Simulations (CLUES). *ArXiv e-prints*, May 2010.

F. Governato, C. Brook, L. Mayer, A. Brooks, G. Rhee, J. Wadsley, P. Jonsson, B. Willman, G. Stinson, T. Quinn, and P. Madau. Bulgeless dwarf galaxies and dark matter cores from supernova-driven outflows. Nature, 463:203–206, January 2010. doi: 10.1038/nature08640.

F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen, and J. Wadsley. Cuspy no more: how outflows affect the central dark matter and baryon distribution in Λ cold dark matter galaxies. MNRAS, 422: 1231–1240, May 2012. doi: 10.1111/j.1365-2966.2012.20696.x.

Thomas H Greif, Jarrett L Johnson, Volker Bromm, and Ralf S Klessen. The First Supernova Explosions: Energetics, Feedback, and Chemical Enrichment. *ApJ*, 670 (1):1–14, November 2007.

Thomas H Greif, Jarrett L Johnson, Ralf S Klessen, and Volker Bromm. The first galaxies: assembly, cooling and the onset of turbulence. *MNRAS*, 387(3):1021–1036, July 2008.

Thomas H Greif, Simon C O Glover, Volker Bromm, and Ralf S Klessen. THE FIRST GALAXIES: CHEMICAL ENRICHMENT, MIXING, AND STAR FORMATION. *ApJ*, 716(1):510–520, May 2010.

B. F. Griffen, G. A. Dooley, A. P. Ji, B. W. O'Shea, F. A. Gómez, and A. Frebel. Tracing the first stars and galaxies of the Milky Way. *ArXiv e-prints*, November 2016a.

B. F. Griffen, A. P. Ji, G. A. Dooley, F. A. Gómez, M. Vogelsberger, B. W. O'Shea, and A. Frebel. The Caterpillar Project: A Large Suite of Milky Way Sized Halos. ApJ, 818:10, February 2016b. doi: 10.3847/0004-637X/818/1/10.

Matthias Gritschneder and Douglas N C Lin. EXPLAINING THE OBSERVED VELOCITY DISPERSION OF DWARF GALAXIES BY BARYONIC MASS LOSS DURING THE FIRST COLLAPSE. *ApJ*, 765(1):38, February 2013.

Q. Guo, S. White, C. Li, and M. Boylan-Kolchin. How do galaxies populate dark matter haloes? MNRAS, 404:1111–1120, May 2010. doi: 10.1111/j.1365-2966.2010.16341.x.

Q. Guo, S. White, R. E. Angulo, B. Henriques, G. Lemson, M. Boylan-Kolchin, P. Thomas, and C. Short. Galaxy formation in WMAP1 and WMAP7 cosmologies. MNRAS, 428:1351–1365, January 2013. doi: 10.1093/mnras/sts115.

Qi Guo, Andrew Cooper, Carlos Frenk, John Helly, and Wojciech Hellwing. The Milky Way system in LCDM cosmological simulations. *arXiv*, March 2015.

O. Hahn and T. Abel. Multi-scale initial conditions for cosmological simulations. MNRAS, 415:2101–2121, August 2011. doi: 10.1111/j.1365-2966.2011.18820.x.

J. Han, S. Cole, C. S. Frenk, and Y. Jing. A unified model for the spatial and mass distribution of subhaloes. MNRAS, 457:1208–1223, April 2016. doi: 10.1093/mnras/stv2900.

J. R. Hargis, B. Willman, and A. H. G. Peter. Too Many, Too Few, or Just Right? The Predicted Number and Distribution of Milky Way Dwarf Galaxies. ApJ, 795: L13, November 2014. doi: 10.1088/2041-8205/795/1/L13.

J. Harris and D. Zaritsky. The Star Formation History of the Large Magellanic Cloud. AJ, 138:1243–1260, November 2009. doi: 10.1088/0004-6256/138/5/1243.

Tilman Hartwig, Volker Bromm, Ralf S Klessen, and Simon C O Glover. Constraining the primordial initial mass function with stellar archaeology. *MNRAS*, 447(4):3892–3908, March 2015.

E. Hayashi, J. F. Navarro, J. E. Taylor, J. Stadel, and T. Quinn. The Structural Evolution of Substructure. ApJ, 584:541–558, February 2003. doi: 10.1086/345788.

A. Helmi. The stellar halo of the Galaxy. A&A Rev., 15:145–188, June 2008. doi: 10.1007/s00159-008-0009-6.

A. Helmi, S. D. White, and V. Springel. The phase-space structure of a dark-matter halo: Implications for dark-matter direct detection experiments. Phys. Rev. D, 66 (6):063502, September 2002. doi: 10.1103/PhysRevD.66.063502.

Amina Helmi and P Tim de Zeeuw. Mapping the substructure in the Galactic halo with the next generation of astrometric satellites. *MNRAS*, 319(3):657–665, December 2000.

M Hénon. Sur l'évolution dynamique des amas globulaires. *Annales d'Astrophysique*, 24:369, February 1961.

L. Hernquist. An analytical model for spherical galaxies and bulges. ApJ, 356:359–364, June 1990. doi: 10.1086/168845.

C. R. Higgs, A. W. McConnachie, M. Irwin, N. F. Bate, G. F. Lewis, M. G. Walker, P. Côté, K. Venn, and G. Battaglia. Solo dwarfs I: survey introduction and first results for the Sagittarius dwarf irregular galaxy. MNRAS, 458:1678–1695, May 2016. doi: 10.1093/mnras/stw257.

D. Homma, M. Chiba, S. Okamoto, Y. Komiyama, M. Tanaka, M. Tanaka, M. N. Ishigaki, M. Akiyama, N. Arimoto, J. A. Garmilla, R. H. Lupton, M. A. Strauss, H. Furusawa, S. Miyazaki, H. Murayama, A. J. Nishizawa, M. Takada, T. Usuda, and S.-Y. Wang. A New Milky Way Satellite Discovered In The Subaru/Hyper Suprime-Cam Survey. *ArXiv e-prints*, September 2016.

P. F. Hopkins, D. Kereš, J. Oñorbe, C.-A. Faucher-Giguère, E. Quataert, N. Murray, and J. S. Bullock. Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. MNRAS, 445:581–603, November 2014. doi: 10.1093/mnras/stu1738.

Philip F Hopkins. Accurate, Meshless Methods for Magneto-Hydrodynamics. May 2015.

M. J. Irwin, V. Belokurov, N. W. Evans, E. V. Ryan-Weber, J. T. A. de Jong, S. Koposov, D. B. Zucker, S. T. Hodgkin, G. Gilmore, P. Prema, L. Hebb, A. Begum, M. Fellhauer, P. C. Hewett, R. C. Kennicutt, Jr., M. I. Wilkinson, D. M. Bramich, S. Vidrih, H.-W. Rix, T. C. Beers, J. C. Barentine, H. Brewington, M. Harvanek, J. Krzesinski, D. Long, A. Nitta, and S. A. Snedden. Discovery of an Unusual Dwarf Galaxy in the Outskirts of the Milky Way. ApJ, 656:L13–L16, February 2007. doi: 10.1086/512183.

M. J. Irwin, A. M. N. Ferguson, A. P. Huxor, N. R. Tanvir, R. A. Ibata, and G. F. Lewis. Andromeda XVII: A New Low-Luminosity Satellite of M31. ApJ, 676:L17, March 2008. doi: 10.1086/587100.

T. Ishiyama, K. Sudo, S. Yokoi, K. Hasegawa, N. Tominaga, and H. Susa. Where are the Low-mass Population III Stars? ApJ, 826:9, July 2016. doi: 10.3847/0004-637X/826/1/9.

Tomoaki Ishiyama, Steven Rieder, Junichiro Makino, Simon Portegies Zwart, Derek Groen, Keigo Nitadori, Cees de Laat, Stephen McMillan, Kei Hiraki, and Stefan Harfst. THE COSMOGRID SIMULATION: STATISTICAL PROPERTIES OF SMALL DARK MATTER HALOS. *ApJ*, 767(2):146, April 2013.

I. S. Jang and M. G. Lee. Discovery of an Ultra-faint Dwarf Galaxy in the Intracluster Field of the Virgo Center: A Fossil of the First Galaxies? ApJ, 795:L6, November 2014. doi: 10.1088/2041-8205/795/1/L6.

N Jarosik, C L Bennett, J Dunkley, B Gold, M R Greason, M Halpern, R S Hill, G Hinshaw, A Kogut, E Komatsu, D Larson, M Limon, S S Meyer, M R Nolta, N Odegard, L Page, K M Smith, D. N. Spergel, G S Tucker, J L Weiland, E Wollack, and E L Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results. *ApJS*, 192(2):14, February 2011.

J. H. Jeans. The stability of a spherical nebula. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 199 (312-320):1–53, 1902. ISSN 0264-3952. doi: 10.1098/rsta.1902.0012. URL `http://rsta.royalsocietypublishing.org/content/199/312-320/1`.

A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. MNRAS, 321:372–384, February 2001. doi: 10.1046/j.1365-8711.2001.04029.x.

Myoungwon Jeon, Andreas H Pawlik, Volker Bromm, and Milos Milosavljevic. Recovery from Population III supernova explosions and the onset of second-generation star formation. *MNRAS*, 444(4):3288–3300, November 2014.

P. Jethwa, D. Erkal, and V. Belokurov. A Magellanic origin of the DES dwarfs. MNRAS, 461:2212–2233, September 2016. doi: 10.1093/mnras/stw1343.

A. P. Ji, A. Frebel, A. Chiti, and J. D. Simon. R-process enrichment from a single event in an ancient dwarf galaxy. Nature, 531:610–613, March 2016. doi: 10.1038/nature17425.

Kathryn V Johnston, James S Bullock, Sanjib Sharma, Andreea Font, Brant E Robertson, and Samuel N Leitner. Tracing Galaxy Formation with Stellar Halos. II. Relating Substructure in Phase and Abundance Space to Accretion Histories. *ApJ*, 689(2):936–957, December 2008.

M. Jurić, Ž. Ivezić, A. Brooks, R. H. Lupton, D. Schlegel, D. Finkbeiner, N. Padmanabhan, N. Bond, B. Sesar, C. M. Rockosi, G. R. Knapp, J. E. Gunn, T. Sumi, D. P. Schneider, J. C. Barentine, H. J. Brewington, J. Brinkmann, M. Fukugita, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, E. H. Neilsen, Jr., A. Nitta, S. A. Snedden, and D. G. York. The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution. ApJ, 673:864-914, February 2008. doi: 10.1086/523619.

Nick Kaiser, William Burgett, Ken Chambers, Larry Denneau, Jim Heasley, Robert Jedicke, Eugene Magnier, Jeff Morgan, Peter Onaka, and John Tonry. ¡title¿The Pan-STARRS wide-field optical/NIR imaging survey¡/title¿. In Larry M Stepp, Roberto Gilmozzi, and Helen J Hall, editors, *SPIE Astronomical Telescopes and Instrumentation: Observational Frontiers of Astronomy for the New Decade*, pages 77330E–77330E–14. SPIE, July 2010.

N. Kallivayalil, R. P. van der Marel, G. Besla, J. Anderson, and C. Alcock. Third-epoch Magellanic Cloud Proper Motions. I. Hubble Space Telescope/WFC3 Data and Orbit Implications. ApJ, 764:161, February 2013. doi: 10.1088/0004-637X/764/2/161.

M. Kaplinghat, S. Tulin, and H.-B. Yu. Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters. *Physical Review Letters*, 116(4):041302, January 2016. doi: 10.1103/PhysRevLett.116.041302.

I. D. Karachentsev, D. I. Makarov, and E. I. Kaisina. Updated Nearby Galaxy Catalog. AJ, 145:101, April 2013. doi: 10.1088/0004-6256/145/4/101.

I. D. Karachentsev, L. N. Makarova, R. B. Tully, P.-F. Wu, and A. Y. Kniazev. KK258, a new transition dwarf galaxy neighbouring the Local Group. MNRAS, 443:1281–1290, September 2014. doi: 10.1093/mnras/stu1217.

I. D. Karachentsev, L. N. Makarova, D. I. Makarov, R. B. Tully, and L. Rizzi. A new isolated dSph galaxy near the Local Group. MNRAS, 447:L85–L89, February 2015. doi: 10.1093/mnrasl/slu181.

Igor D Karachentsev, Valentina E Karachentseva, Walter K Huchtmeier, and Dmitry I Makarov. A Catalog of Neighboring Galaxies. *AJ*, 127(4):2031–2068, April 2004.

N. Katz, T. Quinn, E. Bertschinger, and J. M. Gelb. Formation of Quasars at High Redshift. MNRAS, 270:L71, October 1994. doi: 10.1093/mnras/270.1.L71.

S C Keller, B P Schmidt, M S Bessell, P G Conroy, P Francis, A Granlund, E Kowald, A P Oates, T Martin-Jones, T Preston, P Tisserand, A Vaccarella, and M F Waterson. The SkyMapper Telescope and The Southern Sky Survey. *PASA*, 24(01): 1–12, March 2013.

D. Kim and H. Jerjen. Horologium II: A Second Ultra-faint Milky Way Satellite in the Horologium Constellation. ApJ, 808:L39, August 2015. doi: 10.1088/2041-8205/808/2/L39.

D. Kim, H. Jerjen, D. Mackey, G. S. Da Costa, and A. P. Milone. A Heros Dark Horse: Discovery of an Ultra-faint Milky Way Satellite in Pegasus. ApJ, 804:L44, May 2015. doi: 10.1088/2041-8205/804/2/L44.

E. N. Kirby, J. D. Simon, M. Geha, P. Guhathakurta, and A. Frebel. Uncovering Extremely Metal-Poor Stars in the Milky Way's Ultrafaint Dwarf Spheroidal Satellite Galaxies. ApJ, 685:L43, September 2008. doi: 10.1086/592432.

E. N. Kirby, M. Boylan-Kolchin, J. G. Cohen, M. Geha, J. S. Bullock, and M. Kaplinghat. Segue 2: The Least Massive Galaxy. ApJ, 770:16, June 2013a. doi: 10.1088/0004-637X/770/1/16.

E. N. Kirby, J. G. Cohen, P. Guhathakurta, L. Cheng, J. S. Bullock, and A. Gallazzi. The Universal Stellar Mass-Stellar Metallicity Relation for Dwarf Galaxies. ApJ, 779:102, December 2013b. doi: 10.1088/0004-637X/779/2/102.

Tetsu Kitayama and Naoki Yoshida. Supernova Explosions in the Early Universe: Evolution of Radiative Remnants and the Halo Destruction Efficiency. *ApJ*, 630 (2):675–688, September 2005.

A. Klypin, S. Gottlöber, A. V. Kravtsov, and A. M. Khokhlov. Galaxies in N-Body Simulations: Overcoming the Overmerging Problem. ApJ, 516:530–551, May 1999a. doi: 10.1086/307122.

A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada. Where Are the Missing Galactic Satellites? ApJ, 522:82–92, September 1999b. doi: 10.1086/307643.

A. Klypin, I. Karachentsev, D. Makarov, and O. Nasonova. Abundance of field galaxies. MNRAS, 454:1798–1810, December 2015. doi: 10.1093/mnras/stv2040.

A. A. Klypin, S. Trujillo-Gomez, and J. Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. ApJ, 740:102, October 2011. doi: 10.1088/0004-637X/740/2/102.

A. Knebe, S. R. Knollmann, S. I. Muldrew, F. R. Pearce, M. A. Aragon-Calvo, Y. Ascasibar, P. S. Behroozi, D. Ceverino, S. Colombi, J. Diemand, K. Dolag, B. L. Falck, P. Fasel, J. Gardner, S. Gottlöber, C.-H. Hsu, F. Iannuzzi, A. Klypin, Z. Lukić, M. Maciejewski, C. McBride, M. C. Neyrinck, S. Planelles, D. Potter, V. Quilis, Y. Rasera, J. I. Read, P. M. Ricker, F. Roy, V. Springel, J. Stadel, G. Stinson, P. M. Sutter, V. Turchaninov, D. Tweed, G. Yepes, and M. Zemp. Haloes gone MAD: The Halo-Finder Comparison Project. MNRAS, 415:2293–2318, August 2011. doi: 10.1111/j.1365-2966.2011.18858.x.

A. Knebe, F. R. Pearce, H. Lux, Y. Ascasibar, P. Behroozi, J. Casado, C. Corbett Moran, J. Diemand, K. Dolag, R. Dominguez-Tenreiro, P. Elahi, B. Falck, S. Gottloeber, J. Han, A. Klypin, Z. Lukic, M. Maciejewski, C. K. McBride, M. E. Merchan, S. I. Muldrew, M. Neyrinck, J. Onions, S. Planelles, D. Potter, V. Quilis, Y. Rasera, P. M. Ricker, F. Roy, A. N. Ruiz, M. A. Sgro, V. Springel, J. Stadel, P. M. Sutter, D. Tweed, and M. Zemp. Structure Finding in Cosmological Simulations: The State of Affairs. *ArXiv e-prints*, April 2013.

E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. ApJS, 192:18, February 2011. doi: 10.1088/0067-0049/192/2/18.

S. E. Koposov, V. Belokurov, G. Torrealba, and N. W. Evans. Beasts of the Southern Wild : Discovery of nine Ultra Faint satellites in the vicinity of the Magellanic Clouds. ApJ, 805:130, June 2015a. doi: 10.1088/0004-637X/805/2/130.

S. E. Koposov, A. R. Casey, V. Belokurov, J. R. Lewis, G. Gilmore, C. Worley, A. Hourihane, S. Randich, T. Bensby, A. Bragaglia, M. Bergemann, G. Carraro, M. T. Costado, E. Flaccomio, P. Francois, U. Heiter, V. Hill, P. Jofre, C. Lando, A. C. Lanzafame, P. de Laverny, L. Monaco, L. Morbidelli, L. Sbordone, Š. Mikolaitis, and N. Ryde. Kinematics and Chemistry of Recently Discovered Reticulum 2 and Horologium 1 Dwarf Galaxies. ApJ, 811:62, September 2015b. doi: 10.1088/0004-637X/811/1/62.

A. Kravtsov. Dark Matter Substructure and Dwarf Galactic Satellites. *Advances in Astronomy*, 2010:281913, 2010. doi: 10.1155/2010/281913.

A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, S. Gottlöber, B. Allgood, and J. R. Primack. The Dark Side of the Halo Occupation Distribution. ApJ, 609: 35–49, July 2004. doi: 10.1086/420959.

A. H. W. Küpper, K. V. Johnston, S. Mieske, M. L. M. Collins, and E. J. Tollerud. Exploding Satellites – The Tidal Debris of the Ultra-Faint Dwarf Galaxy Hercules. *ArXiv e-prints*, August 2016.

R. Kuzio de Naray and K. Spekkens. Do Baryons Alter the Halos of Low Surface Brightness Galaxies? ApJ, 741:L29, November 2011. doi: 10.1088/2041-8205/741/2/L29.

B. P. M. Laevens, N. F. Martin, E. J. Bernard, E. F. Schlafly, B. Sesar, H.-W. Rix, E. F. Bell, A. M. N. Ferguson, C. T. Slater, W. E. Sweeney, R. F. G. Wyse, A. P. Huxor, W. S. Burgett, K. C. Chambers, P. W. Draper, K. A. Hodapp, N. Kaiser, E. A. Magnier, N. Metcalfe, J. L. Tonry, R. J. Wainscoat, and C. Waters. Sagittarius II, Draco II and Laevens 3: Three New Milky Way Satellites Discovered in the Pan-STARRS 1 3π Survey. ApJ, 813:44, November 2015. doi: 10.1088/0004-637X/813/1/44.

D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M. R. Nolta, C. L. Bennett, B. Gold, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, L. Page, K. M. Smith, D. N. Spergel, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters. ApJS, 192:16, February 2011. doi: 10.1088/0067-0049/192/2/16.

D Larson, J L Weiland, G Hinshaw, and C L Bennett. Comparing Planck and WMAP: Maps, Spectra, and Parameters. *ApJ*, 801(1):9, March 2015.

R. B. Larson. The physics of star formation. *Reports on Progress in Physics*, 66: 1651–1697, October 2003. doi: 10.1088/0034-4885/66/10/R03.

T. Le Bret, A. Pontzen, A. P. Cooper, C. Frenk, A. Zolotov, A. M. Brooks, F. Governato, and O. H. Parry. Particle tagging and its implications for stellar population dynamics. *ArXiv e-prints*, February 2015.

Jounghun Lee. On the Environmental Dependence of Galaxy Properties Established by the Initial Cosmological Conditions. *ApJ*, 644(1):L5–L8, June 2006.

Y. Li, H. J. Mo, and L. Gao. On halo formation times and assembly bias. MNRAS, 389:1419–1426, September 2008. doi: 10.1111/j.1365-2966.2008.13667.x.

Yang-Shyang Li and Simon D M White. Masses for the Local Group and the Milky Way. *MNRAS*, 384(4):1459–1468, March 2008.

T. C. Licquia and J. A. Newman. Improved Estimates of the Milky Way's Stellar Mass and Star Formation Rate from Hierarchical Bayesian Meta-Analysis. ApJ, 806:96, June 2015. doi: 10.1088/0004-637X/806/1/96.

A. Loeb and N. Weiner. Cores in Dwarf Galaxies from Dark Matter with a Yukawa Potential. *Physical Review Letters*, 106(17):171302, April 2011. doi: 10.1103/PhysRevLett.106.171302.

M. López-Corredoira. Galactic warp in the overdensity of the Canis Major region. MNRAS, 369:1911–1915, July 2006. doi: 10.1111/j.1365-2966.2006.10435.x.

M. R. Lovell, C. S. Frenk, V. R. Eke, A. Jenkins, L. Gao, and T. Theuns. The properties of warm dark matter haloes. *ArXiv e-prints*, August 2013.

Y. Lu, A. Benson, Y.-Y. Mao, S. Tonnesen, A. H. G. Peter, A. R. Wetzel, M. Boylan-Kolchin, and R. H. Wechsler. The connection between the host halo and the satellite galaxies of the Milky Way. *ArXiv e-prints*, May 2016.

Z. Lu and H. J. Mo. Star Formation and Quenching of Satellite Galaxies. *ArXiv e-prints*, June 2015.

R. Lunnan, M. Vogelsberger, A. Frebel, L. Hernquist, A. Lidz, and M. Boylan-Kolchin. The Effects of Patchy Reionization on Satellite Galaxies of the Milky Way. ApJ, 746:109, February 2012. doi: 10.1088/0004-637X/746/1/109.

E. Luque, A. Pieres, B. Santiago, B. Yanny, A. K. Vivas, A. Queiroz, A. Drlica-Wagner, E. Morganson, E. Balbinot, J. L. Marshall, T. S. Li, A. Fausti Neto, L. N. da Costa, M. A. G. Maia, K. Bechtol, A. G. Kim, G. M. Bernstein, S. Dodelson, L. Whiteway, H. T. Diehl, D. A. Finley, T. Abbott, F. B. Abdalla, S. Allam, J. Annis, A. Benoit-Lévy, E. Bertin, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, C. E. Cunha, C. B. D'Andrea, S. Desai, P. Doel, A. E. Evrard, B. Flaugher, P. Fosalba, D. W. Gerdes, D. A. Goldstein, D. Gruen, R. A. Gruendl, G. Gutierrez, D. J. James, K. Kuehn, N. Kuropatkin, O. Lahav, P. Martini, R. Miquel, B. Nord, R. Ogando, A. A. Plazas, A. K. Romer, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, and A. R. Walker. The Dark Energy Survey view of the Sagittarius stream: Discovery of two faint stellar system candidates. *ArXiv e-prints*, August 2016.

D. Lynden-Bell. Galaxies and Globular Clusters in High Velocity Hydrogen Streams. In R. J. Dickens, J. E. Perry, F. G. Smith, and I. R. King, editors, *The Galaxy and the Local Group*, volume 182 of *Royal Greenwich Observatory Bulletins*, page 235, 1976.

D. Lynden-Bell. The Ursa Minor dwarf galaxy is a member of the Magellanic Stream. *The Observatory*, 102:7–9, February 1982.

D. Lynden-Bell and R. M. Lynden-Bell. Ghostly streams from the formation of the Galaxy's halo. MNRAS, 275:429–442, July 1995. doi: 10.1093/mnras/275.2.429.

A. V. Macciò, A. A. Dutton, and F. C. van den Bosch. Concentration, spin and shape of dark matter haloes as a function of the cosmological model: WMAP1, WMAP3 and WMAP5 results. MNRAS, 391:1940–1954, December 2008. doi: 10.1111/j.1365-2966.2008.14029.x.

M. E. Machacek, E. D. Carlson, and L. J. Hall. Self-Interacting Dark Matter: An Alternative Scenario? *in Akerlof, C. W., Srednicki, M. A. eds, Annals of the New York Academy of Sciences Vol. 688, Relativistic Astrophysics and Particle*

Cosmology. p. 681, 688:681, October 1993. doi: 10.1111/j.1749-6632.1993.tb43955.x.

P. Madau, M. Kuhlen, J. Diemand, B. Moore, M. Zemp, D. Potter, and J. Stadel. Fossil Remnants of Reionization in the Halo of the Milky Way. ApJ, 689:L41, December 2008. doi: 10.1086/595814.

P. Madau, S. Shen, and F. Governato. Dark Matter Heating and Early Core Formation in Dwarf Galaxies. ApJ, 789:L17, July 2014. doi: 10.1088/2041-8205/789/1/L17.

Piero Madau, Andrea Ferrara, and Martin J Rees. Early Metal Enrichment of the Intergalactic Medium by Pregalactic Outflows. *ApJ*, 555(1):92–105, July 2001.

S. R. Majewski, R. L. Beaton, R. J. Patterson, J. S. Kalirai, M. C. Geha, R. R. Muñoz, M. S. Seigar, P. Guhathakurta, K. M. Gilbert, R. M. Rich, J. S. Bullock, and D. B. Reitzel. Discovery of Andromeda XIV: A Dwarf Spheroidal Dynamical Rogue in the Local Group? ApJ, 670:L9–L12, November 2007. doi: 10.1086/524033.

Steven R Majewski, John C Wilson, Fred Hearty, Ricardo R Schiavon, and Michael F Skrutskie. The Apache Point Observatory Galactic Evolution Experiment (APOGEE) in Sloan Digital Sky Survey III (SDSS-III). *IAU*, 265(S265): 480–481, March 2010.

Y.-Y. Mao, M. Williamson, and R. H. Wechsler. The Dependence of Subhalo Abundance on Halo Concentration. ApJ, 810:21, September 2015. doi: 10.1088/0004-637X/810/1/21.

F. Marinacci, R. Pakmor, V. Springel, and C. M. Simpson. Diffuse gas properties and stellar metallicities in cosmological simulations of disc galaxy formation. MNRAS, 442:3745–3760, August 2014. doi: 10.1093/mnras/stu1136.

N. F. Martin, R. A. Ibata, M. Bellazzini, M. J. Irwin, G. F. Lewis, and W. Dehnen. A dwarf galaxy remnant in Canis Major: the fossil of an in-plane accretion on to the Milky Way. MNRAS, 348:12–23, February 2004. doi: 10.1111/j.1365-2966.2004.07331.x.

N. F. Martin, A. W. McConnachie, M. Irwin, L. M. Widrow, A. M. N. Ferguson, R. A. Ibata, J. Dubinski, A. Babul, S. Chapman, M. Fardal, G. F. Lewis, J. Navarro, and R. M. Rich. PAndAS' CUBS: Discovery of Two New Dwarf Galaxies in the Surroundings of the Andromeda and Triangulum Galaxies. ApJ, 705:758–765, November 2009. doi: 10.1088/0004-637X/705/1/758.

N. F. Martin, D. L. Nidever, G. Besla, K. Olsen, A. R. Walker, A. K. Vivas, R. A. Gruendl, C. C. Kaleida, R. R. Muñoz, R. D. Blum, A. Saha, B. C. Conn, E. F. Bell, Y.-H. Chu, M.-R. L. Cioni, T. J. L. de Boer, C. Gallart, S. Jin, A. Kunder, S. R. Majewski, D. Martinez-Delgado, A. Monachesi, M. Monelli, L. Monteagudo, N. E. D. Noël, E. W. Olszewski, G. S. Stringfellow, R. P. van der Marel, and

D. Zaritsky. Hydra II: A Faint and Compact Milky Way Dwarf Galaxy Found in the Survey of the Magellanic Stellar History. ApJ, 804:L5, May 2015. doi: 10.1088/2041-8205/804/1/L5.

D. Martínez-Delgado, D. J. Butler, H.-W. Rix, V. I. Franco, J. Peñarrubia, E. J. Alfaro, and D. I. Dinescu. The Closest View of a Dwarf Galaxy: New Evidence on the Nature of the Canis Major Overdensity. ApJ, 633:205–209, November 2005. doi: 10.1086/432635.

D. Martínez-Delgado, A. J. Romanowsky, R. J. Gabany, F. Annibali, J. A. Arnold, J. Fliri, S. Zibetti, R. P. van der Marel, H.-W. Rix, T. S. Chonis, J. A. Carballo-Bello, A. Aloisi, A. V. Macciò, J. Gallego-Laborda, J. P. Brodie, and M. R. Merrifield. Dwarfs Gobbling Dwarfs: A Stellar Tidal Stream around NGC 4449 and Hierarchical Galaxy Formation on Small Scales. ApJ, 748:L24, April 2012. doi: 10.1088/2041-8205/748/2/L24.

Aaron J Maxwell, James Wadsley, and H M P Couchman. THE ENERGETICS OF CUSP DESTRUCTION. *ApJ*, 806(2):229, June 2015.

L. Mayer, C. Mastropietro, J. Wadsley, J. Stadel, and B. Moore. Simultaneous ram pressure and tidal stripping; how dwarf spheroidals lost their gas. MNRAS, 369: 1021–1038, July 2006. doi: 10.1111/j.1365-2966.2006.10403.x.

A. W. McConnachie. The Observed Properties of Dwarf Galaxies in and around the Local Group. AJ, 144:4, July 2012. doi: 10.1088/0004-6256/144/1/4.

A. W. McConnachie, A. Huxor, N. F. Martin, M. J. Irwin, S. C. Chapman, G. Fahlman, A. M. N. Ferguson, R. A. Ibata, G. F. Lewis, H. Richer, and N. R. Tanvir. A Trio of New Local Group Galaxies with Extreme Properties. ApJ, 688: 1009-1020, December 2008. doi: 10.1086/591313.

A. W. McConnachie, M. J. Irwin, R. A. Ibata, J. Dubinski, L. M. Widrow, N. F. Martin, P. Côté, A. L. Dotter, J. F. Navarro, A. M. N. Ferguson, T. H. Puzia, G. F. Lewis, A. Babul, P. Barmby, O. Bienaymé, S. C. Chapman, R. Cockcroft, M. L. M. Collins, M. A. Fardal, W. E. Harris, A. Huxor, A. D. Mackey, J. Peñarrubia, R. M. Rich, H. B. Richer, A. Siebert, N. Tanvir, D. Valls-Gabaud, and K. A. Venn. The remnants of galaxy formation from a panoramic survey of the region around M31. Nature, 461:66–69, September 2009. doi: 10.1038/nature08327.

K. B. W. McQuinn, E. D. Skillman, A. Dolphin, J. M. Cannon, J. J. Salzer, K. L. Rhode, E. A. K. Adams, D. Berg, R. Giovanelli, L. Girardi, and M. P. Haynes. Leo P: An Unquenched Very Low-mass Galaxy. ApJ, 812:158, October 2015. doi: 10.1088/0004-637X/812/2/158.

M. Meneghetti and E. Rasia. Reconciling extremely different concentration-mass relations. *ArXiv e-prints*, March 2013.

J. Miralda-Escudé. A Test of the Collisional Dark Matter Hypothesis from Cluster Lensing. ApJ, 564:60–64, January 2002. doi: 10.1086/324138.

A. Moitinho, R. A. Vázquez, G. Carraro, G. Baume, E. E. Giorgi, and W. Lyra. Spiral structure of the third galactic quadrant and the solution to the Canis Major debate. MNRAS, 368:L77–L81, May 2006. doi: 10.1111/j.1745-3933.2006.00163.x.

Y. Momany, S. Zaggia, G. Gilmore, G. Piotto, G. Carraro, L. R. Bedin, and F. de Angeli. Outer structure of the Galactic warp and flare: explaining the Canis Major over-density. A&A, 451:515–538, May 2006. doi: 10.1051/0004-6361:20054081.

Antonela Monachesi, Eric F Bell, David J Radburn-Smith, Marija Vlajić, Roelof S de Jong, Jeremy Bailin, Julianne J Dalcanton, Benne W Holwerda, and David Streich. TESTING GALAXY FORMATION MODELS WITH THE GHOSTS SURVEY: THE COLOR PROFILE OF M81's STELLAR HALO. *ApJ*, 766(2): 106, March 2013.

Antonela Monachesi, Eric F Bell, David Radburn-Smith, Jeremy Bailin, Roelof S de Jong, Benne Holwerda, David Streich, and Grace Silverstein. The GHOSTS survey. II. The diversity of Halo Color and Metallicity Profiles of Massive Disk Galaxies. *arXiv*, July 2015.

B. Moore. Evidence against dissipation-less dark matter from observations of galaxy haloes. Nature, 370:629–631, August 1994. doi: 10.1038/370629a0.

B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, and P. Tozzi. Dark Matter Substructure within Galactic Halos. ApJ, 524:L19–L22, October 1999. doi: 10.1086/312287.

S. More, B. Diemer, and A. V. Kravtsov. The Splashback Radius as a Physical Halo Boundary and the Growth of Halo Mass. ApJ, 810:36, September 2015. doi: 10.1088/0004-637X/810/1/36.

B. P. Moster, R. S. Somerville, C. Maulbetsch, F. C. van den Bosch, A. V. Macciò, T. Naab, and L. Oser. Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift. ApJ, 710:903–923, February 2010. doi: 10.1088/0004-637X/710/2/903.

B. P. Moster, T. Naab, and S. D. M. White. Galactic star formation and accretion histories from matching galaxies to dark matter haloes. MNRAS, 428:3121–3138, February 2013. doi: 10.1093/mnras/sts261.

M. Mouhcine, R. Ibata, and M. Rejkuba. A Panoramic View of the Milky Way Analog NGC 891. ApJ, 714:L12–L15, May 2010. doi: 10.1088/2041-8205/714/1/L12.

S. I. Muldrew, F. R. Pearce, and C. Power. The accuracy of subhalo detection. MNRAS, 410:2617–2624, February 2011. doi: 10.1111/j.1365-2966.2010.17636.x.

F. Munshi, F. Governato, A. M. Brooks, C. Christensen, S. Shen, S. Loebman, B. Moster, T. Quinn, and J. Wadsley. Reproducing the Stellar Mass/Halo Mass Relation in Simulated ΛCDM Galaxies: Theory versus Observational Estimates. ApJ, 766:56, March 2013. doi: 10.1088/0004-637X/766/1/56.

A. L. Muratov, D. Kereš, C.-A. Faucher-Giguère, P. F. Hopkins, E. Quataert, and N. Murray. Gusty, gaseous flows of FIRE: galactic winds in cosmological simulations with explicit stellar feedback. MNRAS, 454:2691–2713, December 2015. doi: 10.1093/mnras/stv2126.

J. F. Navarro and S. D. M. White. Simulations of dissipative galaxy formation in hierarchically clustering universes-2. Dynamics of the baryonic component in galactic haloes. MNRAS, 267:401–412, March 1994. doi: 10.1093/mnras/267.2.401.

J. F. Navarro, V. R. Eke, and C. S. Frenk. The cores of dwarf galaxy haloes. MNRAS, 283:L72–L78, December 1996a. doi: 10.1093/mnras/283.3.72L.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. ApJ, 462:563, May 1996b. doi: 10.1086/177173.

J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJ, 490:493–508, December 1997.

Julio F Navarro, Aaron Ludlow, Volker Springel, Jie Wang, Mark Vogelsberger, Simon D M White, Adrian Jenkins, Carlos S Frenk, and Amina Helmi. The diversity and similarity of simulated cold dark matter haloes. *MNRAS*, 402(1):21–34, February 2010.

A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of Λ CDM halo concentrations. MNRAS, 381:1450–1462, November 2007. doi: 10.1111/j.1365-2966.2007.12381.x.

M. Nichols, J. Colless, M. Colless, and J. Bland-Hawthorn. Accretion of the Magellanic System onto the Galaxy. ApJ, 742:110, December 2011. doi: 10.1088/0004-637X/742/2/110.

S. Nickerson, G. Stinson, H. M. P. Couchman, J. Bailin, and J. Wadsley. Mechanisms of baryon loss for dark satellites in cosmological smoothed particle hydrodynamics simulations. MNRAS, 415:257–270, July 2011. doi: 10.1111/j.1365-2966.2011.18700.x.

M. Niederste-Ostholt, V. Belokurov, N. W. Evans, and J. Peñarrubia. Re-Assembling the Sagittarius Dwarf Galaxy. ApJ, 712:516–526, March 2010. doi: 10.1088/0004-637X/712/1/516.

J. E. Norris, G. Gilmore, R. F. G. Wyse, D. Yong, and A. Frebel. An Extremely Carbon-rich, Extremely Metal-poor Star in the Segue 1 System. ApJ, 722:L104–L109, October 2010. doi: 10.1088/2041-8205/722/1/L104.

J. Oñorbe, M. Boylan-Kolchin, J. S. Bullock, P. F. Hopkins, D. Kereš, C.-A. Faucher-Giguère, E. Quataert, and N. Murray. Forged in FIRE: cusps, cores and baryons in low-mass dwarf galaxies. MNRAS, 454:2092–2106, December 2015. doi: 10.1093/mnras/stv2072.

J. Oñorbe, J. F Hennawi, and Z. Lukić. Self-Consistent Modeling of Reionization in Cosmological Hydrodynamical Simulations. *ArXiv e-prints*, July 2016.

M. Oguri and J. Lee. A realistic model for spatial and mass distributions of dark halo substructures: An analytic approach. MNRAS, 355:120–128, November 2004. doi: 10.1111/j.1365-2966.2004.08304.x.

S.-H. Oh, C. Brook, F. Governato, E. Brinks, L. Mayer, W. J. G. de Blok, A. Brooks, and F. Walter. The Central Slope of Dark Matter Cores in Dwarf Galaxies: Simulations versus THINGS. AJ, 142:24, July 2011. doi: 10.1088/0004-6256/142/1/24.

S Peng Oh and Zoltán Haiman. Second-Generation Objects in the Universe: Radiative Cooling and Collapse of Halos with Virial Temperatures above 10 4K. *ApJ*, 569 (2):558–572, April 2002.

T. Okamoto and C. S. Frenk. The origin of failed subhaloes and the common mass scale of the Milky Way satellite galaxies. MNRAS, 399:L174–L178, October 2009. doi: 10.1111/j.1745-3933.2009.00748.x.

T. Okamoto, L. Gao, and T. Theuns. Mass loss of galaxies due to an ultraviolet background. MNRAS, 390:920–928, November 2008. doi: 10.1111/j.1365-2966.2008.13830.x.

J. Onions, A. Knebe, F. R. Pearce, S. I. Muldrew, H. Lux, S. R. Knollmann, Y. Ascasibar, P. Behroozi, P. Elahi, J. Han, M. Maciejewski, M. E. Merchán, M. Neyrinck, A. N. Ruiz, M. A. Sgró, V. Springel, and D. Tweed. Subhaloes going Notts: the subhalo-finder comparison project. MNRAS, 423:1200–1214, June 2012. doi: 10.1111/j.1365-2966.2012.20947.x.

Jose Onorbe, Shea Garrison-Kimmel, Ariyeh H Maller, James S Bullock, Miguel Rocha, and Oliver Hahn. How to zoom: bias, contamination and Lagrange volumes in multimass cosmological simulations. *MNRAS*, 437(2):1894–1908, January 2014.

B. W. O'Shea, G. Bryan, J. Bordner, M. L. Norman, T. Abel, R. Harkness, and A. Kritsuk. Introducing Enzo, an AMR Cosmology Application. *ArXiv Astrophysics e-prints*, March 2004.

B. W. O'Shea, J. H. Wise, H. Xu, and M. L. Norman. Probing the Ultraviolet Luminosity Function of the Earliest Galaxies with the Renaissance Simulations. ApJ, 807:L12, July 2015. doi: 10.1088/2041-8205/807/1/L12.

Brian W O'Shea and Michael L Norman. Population III Star Formation in a ΛCDM Universe. I. The Effect of Formation Redshift and Environment on Protostellar Accretion Rate. *ApJ*, 654(1):66–92, January 2007.

E. Papastergis, R. Giovanelli, M. P. Haynes, and F. Shankar. Is there a "too big to fail" problem in the field? A&A, 574:A113, February 2015. doi: 10.1051/0004-6361/201424909.

A. H. Pawlik and J. Schaye. Photoheating and supernova feedback amplify each other's effect on the cosmic star formation rate. MNRAS, 396:L46–L50, June 2009. doi: 10.1111/j.1745-3933.2009.00659.x.

J. Peñarrubia, A. J. Benson, M. G. Walker, G. Gilmore, A. W. McConnachie, and L. Mayer. The impact of dark matter cusps and cores on the satellite galaxy population around spiral galaxies. MNRAS, 406:1290–1305, August 2010. doi: 10.1111/j.1365-2966.2010.16762.x.

J. Peñarrubia, A. Pontzen, M. G. Walker, and S. E. Koposov. The Coupling between the Core/Cusp and Missing Satellite Problems. ApJ, 759:L42, November 2012. doi: 10.1088/2041-8205/759/2/L42.

J. Peñarrubia, F. A. Gómez, G. Besla, D. Erkal, and Y.-Z. Ma. A timing constraint on the (total) mass of the Large Magellanic Cloud. MNRAS, 456:L54–L58, February 2016. doi: 10.1093/mnrasl/slv160.

P. J. E. Peebles. *The large-scale structure of the universe.* 1980.

M A C Perryman, K S de Boer, G Gilmore, E Høg, M G Lattanzi, L Lindegren, X Luri, F Mignard, O Pace, and P T de Zeeuw. GAIA: Composition, formation and evolution of the Galaxy. *A&A*, 369(1):339–363, April 2001.

A. H. G. Peter and A. J. Benson. Dark-matter decays and Milky Way satellite galaxies. Phys. Rev. D, 82(12):123521, December 2010. doi: 10.1103/PhysRevD.82.123521.

A. H. G. Peter, M. Rocha, J. S. Bullock, and M. Kaplinghat. Cosmological simulations with self-interacting dark matter - II. Halo shapes versus observations. MNRAS, 430:105–120, March 2013. doi: 10.1093/mnras/sts535.

T Piffl, C Scannapieco, J Binney, M Steinmetz, R D Scholz, M E K Williams, R S de Jong, G Kordopatis, G Matijevič, O Bienaymé, J Bland-Hawthorn, C Boeche, K Freeman, B Gibson, G Gilmore, E K Grebel, A Helmi, U Munari, J F Navarro, Q Parker, W A Reid, G Seabroke, F Watson, R F G Wyse, and T Zwitter. The RAVE survey: the Galactic escape speed and the mass of the Milky Way. *A&A*, 562:A91, February 2014.

A. Pillepich, M. Vogelsberger, A. Deason, V. Rodriguez-Gomez, S. Genel, D. Nelson, P. Torrey, L. V. Sales, F. Marinacci, V. Springel, D. Sijacki, and L. Hernquist. Halo mass and assembly history exposed in the faint outskirts: the stellar and dark matter haloes of Illustris galaxies. MNRAS, 444:237–249, October 2014. doi: 10.1093/mnras/stu1408.

A. Pillepich, P. Madau, and L. Mayer. Building Late-type Spiral Galaxies by In-situ and Ex-situ Star Formation. ApJ, 799:184, February 2015. doi: 10.1088/0004-637X/799/2/184.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. A&A, 571:A16, November 2014. doi: 10.1051/0004-6361/201321591.

Planck Collaboration, R. Adam, N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, A. Bonaldi, L. Bonavera, J. R. Bond, J. Borrill, F. R. Bouchet, M. Bucher, C. Burigana, E. Calabrese, J.-F. Cardoso, J. Carron, H. C. Chiang, L. P. L. Colombo, C. Combet, B. Comis, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, J. Delabrouille, E. Di Valentino, C. Dickinson, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, F. Elsner, T. A. Enßlin, H. K. Eriksen, E. Falgarone, Y. Fantaye, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, A. Gruppuso, J. E. Gudmundsson, F. K. Hansen, G. Helou, S. Henrot-Versillé, D. Herranz, E. Hivon, Z. Huang, S. Ili, A. H. Jaffe, W. C. Jones, E. Keihänen, R. Keskitalo, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, A. Lähteenmäki, J.-M. Lamarre, M. Langer, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. López-Caniego, Y.-Z. Ma, J. F. Macías-Pérez, G. Maggio, A. Mangilli, M. Maris, P. G. Martin, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M.-A. Miville-Deschênes, D. Molinari, A. Moneti, L. Montier, G. Morgante, A. Moss, P. Naselsky, P. Natoli, C. A. Oxborrow, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, L. Patrizii, O. Perdereau, L. Perotto, V. Pettorino, F. Piacentini, S. Plaszczynski, L. Polastri, G. Polenta, J.-L. Puget, J. P. Rachen, B. Racine, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, M. Rossetti, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, G. Sirri, R. Sunyaev, A.-S. Suur-Uski, J. A. Tauber, M. Tenti, L. Toffolatti, M. Tomasi, M. Tristram, T. Trombetti, J. Valiviita, F. Van Tent, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, A. Zacchei, and A. Zonca. Planck intermediate results. XLVII. Planck constraints on reionization history. *ArXiv e-prints*, May 2016a.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13, September 2016b. doi: 10.1051/0004-6361/201525830.

E. Polisensky and M. Ricotti. Massive Milky Way satellites in cold and warm dark

matter: dependence on cosmology. MNRAS, November 2013. doi: 10.1093/mnras/stt2105.

A. Pontzen and F. Governato. How supernova feedback turns dark matter cusps into cores. MNRAS, 421:3464–3471, April 2012. doi: 10.1111/j.1365-2966.2012.20571.x.

A. Pontzen and F. Governato. Cold dark matter heats up. Nature, 506:171–178, February 2014. doi: 10.1038/nature12953.

A. Pontzen, J. I. Read, R. Teyssier, F. Governato, A. Gualandris, N. Roth, and J. Devriendt. Milking the spherical cow - on aspherical dynamics in spherical coordinates. MNRAS, 451:1366–1379, August 2015. doi: 10.1093/mnras/stv1032.

C. Power. Seeking Observable Imprints of Small-Scale Structure on the Properties of Dark Matter Haloes. *PASA*, 30:e053, October 2013. doi: 10.1017/pasa.2013.32.

C. Power and A. Knebe. The impact of box size on the properties of dark matter haloes in cosmological simulations. MNRAS, 370:691–701, August 2006. doi: 10.1111/j.1365-2966.2006.10562.x.

C. Power, G. A. Wynn, A. S. G. Robotham, G. F. Lewis, and M. I. Wilkinson. Stochastic Star Formation Feedback: Mapping Low-Mass Galaxies to Dark Matter Haloes. *ArXiv e-prints*, June 2014.

Chris Power, J F Navarro, A Jenkins, C S Frenk, S D M White, Volker Springel, J Stadel, and T Quinn. The inner structure of CDM haloes - I. A numerical convergence study. *MNRAS*, 338:14–34, 2003.

W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. ApJ, 187:425–438, February 1974. doi: 10.1086/152650.

Jonathan R Pritchard and Abraham Loeb. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, August 2012.

Cinthia Ragone-Figueroa and Manolis Plionis. Environmental influences on the morphology and dynamics of group-sized haloes. *MNRAS*, 377(4):1785–1794, June 2007.

S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradač. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56. ApJ, 679:1173–1180, June 2008. doi: 10.1086/587859.

V. Rashkov, P. Madau, M. Kuhlen, and J. Diemand. On the Assembly of the Milky Way Dwarf Satellites and Their Common Mass Scale. ApJ, 745:142, February 2012. doi: 10.1088/0004-637X/745/2/142.

J. I. Read and G. Gilmore. Mass loss from dwarf spheroidal galaxies: the origins of shallow dark matter cores and exponential surface brightness profiles. MNRAS, 356:107–124, January 2005. doi: 10.1111/j.1365-2966.2004.08424.x.

J. I. Read, M. I. Wilkinson, N. W. Evans, G. Gilmore, and J. T. Kleyna. The importance of tides for the Local Group dwarf spheroidals. MNRAS, 367:387–399, March 2006. doi: 10.1111/j.1365-2966.2005.09959.x.

J. I. Read, O. Agertz, and M. L. M. Collins. Dark matter cores all the way down. MNRAS, 459:2573–2590, July 2016a. doi: 10.1093/mnras/stw713.

J. I. Read, G. Iorio, O. Agertz, and F. Fraternali. The stellar mass-halo mass relation of isolated field dwarfs: a critical test of ΛCDM at the edge of galaxy formation. *ArXiv e-prints*, July 2016b.

M. J. Reid. Is There a Supermassive Black Hole at the Center of the Milky Way? *International Journal of Modern Physics D*, 18:889–910, 2009. doi: 10.1142/S0218271809014820.

R. M. Rich, M. L. M. Collins, C. M. Black, F. A. Longstaff, A. Koch, A. Benson, and D. B. Reitzel. A tidally distorted dwarf galaxy near NGC 4449. Nature, 482: 192–194, February 2012. doi: 10.1038/nature10837.

J. C. Richardson, M. J. Irwin, A. W. McConnachie, N. F. Martin, A. L. Dotter, A. M. N. Ferguson, R. A. Ibata, S. C. Chapman, G. F. Lewis, N. R. Tanvir, and R. M. Rich. PAndAS' Progeny: Extending the M31 Dwarf Galaxy Cabal. ApJ, 732:76, May 2011. doi: 10.1088/0004-637X/732/2/76.

M. Ricotti and J M Shull. Feedback from Galaxy Formation: Escaping Ionizing Radiation from Galaxies at High Redshift. *ApJ*, 542(2):548–558, 2000.

K. Riebe, A. M. Partl, H. Enke, J. Forero-Romero, S. Gottloeber, A. Klypin, G. Lemson, F. Prada, J. R. Primack, M. Steinmetz, and V. Turchaninov. The MultiDark Database: Release of the Bolshoi and MultiDark Cosmological Simulations. *ArXiv e-prints*, August 2011.

J. S. Ritter, A. Sluder, C. Safranek-Shrader, M. Milosavljević, and V. Bromm. Metal transport and chemical heterogeneity in early star forming systems. MNRAS, 451: 1190–1198, August 2015. doi: 10.1093/mnras/stv982.

Jeremy S Ritter, Chalence Safranek-Shrader, Orly Gnat, Milos Milosavljevic, and Volker Bromm. Confined Population III Enrichment and the Prospects for Prompt Second-generation Star Formation. *ApJ*, 761(1):56, December 2012.

M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Oñorbe, and L. A. Moustakas. Cosmological simulations with self-interacting dark matter - I. Constant-density cores and substructure. MNRAS, 430:81–104, March 2013. doi: 10.1093/mnras/sts514.

H. J. Rocha-Pinto, S. R. Majewski, M. F. Skrutskie, R. J. Patterson, H. Nakanishi, R. R. Muñoz, and Y. Sofue. The Dog on the Ship: The Canis Major Dwarf Galaxy as an Outlying Part of the Argo Star System. ApJ, 640:L147–L150, April 2006. doi: 10.1086/503555.

T. A. Roderick, H. Jerjen, A. D. Mackey, and G. S. Da Costa. Stellar Substructures Around the Hercules Dwarf Spheroidal Galaxy. ApJ, 804:134, May 2015. doi: 10.1088/0004-637X/804/2/134.

J. C. Roediger and S. Courteau. On the uncertainties of stellar mass estimates via colour measurements. MNRAS, 452:3209–3225, September 2015. doi: 10.1093/mnras/stv1499.

A. J. Romanowsky, D. Martínez-Delgado, N. F. Martin, G. Morales, Z. G. Jennings, R. J. GaBany, J. P. Brodie, E. K. Grebel, J. Schedler, and M. Sidonio. Satellite accretion in action: a tidally disrupting dwarf spheroidal around the nearby spiral galaxy NGC 253. MNRAS, 457:L103–L107, March 2016. doi: 10.1093/mnrasl/slv207.

A. N. Ruiz, N. D. Padilla, M. J. Domínguez, and S. A. Cora. How accurate is it to update the cosmology of your halo catalogues? MNRAS, 418:2422–2434, December 2011. doi: 10.1111/j.1365-2966.2011.19635.x.

Chalence Safranek-Shrader, Milos Milosavljevic, and Volker Bromm. Formation of the first low-mass stars from cosmological initial conditions. *MNRAS*, 440(1):L76–L80, May 2014.

L. V. Sales, J. F. Navarro, A. P. Cooper, S. D. M. White, C. S. Frenk, and A. Helmi. Clues to the 'Magellanic Galaxy' from cosmological simulations. MNRAS, 418:648–658, November 2011. doi: 10.1111/j.1365-2966.2011.19514.x.

L. V. Sales, W. Wang, S. D. M. White, and J. F. Navarro. Satellites and haloes of dwarf galaxies. MNRAS, 428:573–578, January 2013. doi: 10.1093/mnras/sts054.

L. V. Sales, J. F. Navarro, N. Kallivayalil, and C. S. Frenk. Identifying true satellites of the Magellanic Clouds. *ArXiv e-prints*, May 2016.

P. Salucci, M. I. Wilkinson, M. G. Walker, G. F. Gilmore, E. K. Grebel, A. Koch, C. Frigerio Martins, and R. F. G. Wyse. Dwarf spheroidal galaxy kinematics and spiral galaxy scaling laws. MNRAS, 420:2034–2041, March 2012. doi: 10.1111/j.1365-2966.2011.20144.x.

S. Salvadori and A. Ferrara. Ultra faint dwarfs: probing early cosmic star formation. MNRAS, 395:L6–L10, May 2009. doi: 10.1111/j.1745-3933.2009.00627.x.

S. Salvadori, Á. Skúladóttir, and E. Tolstoy. Carbon-enhanced metal-poor stars in dwarf galaxies. *MNRAS*, 454:1320–1331, December 2015. doi: 10.1093/mnras/stv1969.

D. J. Sand, D. Crnojević, J. Strader, E. Toloba, J. D. Simon, N. Caldwell, P. Guhathakurta, B. McLeod, and A. C. Seth. Discovery of a New Faint Dwarf Galaxy Associated with NGC 253. ApJ, 793:L7, September 2014. doi: 10.1088/2041-8205/793/1/L7.

D. J. Sand, K. Spekkens, D. Crnojević, J. R. Hargis, B. Willman, J. Strader, and C. J. Grillmair. Antlia B: A Faint Dwarf Galaxy Member of the NGC 3109 Association. ApJ, 812:L13, October 2015. doi: 10.1088/2041-8205/812/1/L13.

Mei Sasaki, Paul C Clark, Volker Springel, Ralf S Klessen, and Simon C O Glover. Statistical properties of dark matter mini-haloes at z ¿= 15. *MNRAS*, 442(3): 1942–1955, August 2014.

T. Sawala, C. S. Frenk, R. A. Crain, A. Jenkins, J. Schaye, T. Theuns, and J. Zavala. The abundance of (not just) dark matter haloes. MNRAS, 431:1366–1382, May 2013. doi: 10.1093/mnras/stt259.

T. Sawala, C. S. Frenk, A. Fattahi, J. F. Navarro, R. G. Bower, R. A. Crain, C. Dalla Vecchia, M. Furlong, J. C. Helly, A. Jenkins, K. A. Oman, M. Schaller, J. Schaye, T. Theuns, J. Trayford, and S. D. M. White. Local Group galaxies emerge from the dark. *ArXiv e-prints*, December 2014.

T. Sawala, C. S. Frenk, A. Fattahi, J. F. Navarro, R. G. Bower, R. A. Crain, C. Dalla Vecchia, M. Furlong, A. Jenkins, I. G. McCarthy, Y. Qu, M. Schaller, J. Schaye, and T. Theuns. Bent by baryons: the low-mass galaxy-halo relation. MNRAS, 448: 2941–2947, April 2015. doi: 10.1093/mnras/stu2753.

T. Sawala, C. S. Frenk, A. Fattahi, J. F. Navarro, T. Theuns, R. G. Bower, R. A. Crain, M. Furlong, A. Jenkins, M. Schaller, and J. Schaye. The chosen few: the low-mass haloes that host faint galaxies. MNRAS, 456:85–97, February 2016. doi: 10.1093/mnras/stv2597.

J. Schaye, R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly, A. Jenkins, Y. M. Rosas-Guevara, S. D. M. White, M. Baes, C. M. Booth, P. Camps, J. F. Navarro, Y. Qu, A. Rahmati, T. Sawala, P. A. Thomas, and J. Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. MNRAS, 446: 521–554, January 2015. doi: 10.1093/mnras/stu2058.

R. Schmidt and J. Wambsganss. Limits on MACHOs from microlensing in the double quasar Q0957+561. A&A, 335:379–387, July 1998.

P. R. Shapiro, I. T. Iliev, and A. C. Raga. Photoevaporation of cosmological minihaloes during reionization. MNRAS, 348:753–782, March 2004. doi: 10.1111/j.1365-2966.2004.07364.x.

H. Shapley. Approximate Distance and dimensions of Large Magellanic Cloud. *Harvard College Observatory Bulletin*, 775:1–2, September 1922.

H. Shapley. The Magellanic Clouds, III. The Distance and Linear Dimensions of the Large Cloud. *Harvard College Observatory Circular*, 268:1–4, December 1924.

G. Shattow and A. Loeb. Implications of recent measurements of the Milky Way rotation for the orbit of the Large Magellanic Cloud. MNRAS, 392:L21–L25, January 2009. doi: 10.1111/j.1745-3933.2008.00573.x.

R. K. Sheth and G. Tormen. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. MNRAS, 329:61–75, January 2002. doi: 10.1046/j.1365-8711.2002.04950.x.

J. D. Simon, A. D. Bolatto, A. Leroy, L. Blitz, and E. L. Gates. High-Resolution Measurements of the Halos of Four Dark Matter-Dominated Galaxies: Deviations from a Universal Density Profile. ApJ, 621:757–776, March 2005. doi: 10.1086/427684.

J. D. Simon, A. Drlica-Wagner, T. S. Li, B. Nord, M. Geha, K. Bechtol, E. Balbinot, E. Buckley-Geer, H. Lin, J. Marshall, B. Santiago, L. Strigari, M. Wang, R. H. Wechsler, B. Yanny, T. Abbott, A. H. Bauer, G. M. Bernstein, E. Bertin, D. Brooks, D. L. Burke, D. Capozzi, A. Carnero Rosell, M. Carrasco Kind, C. B. D'Andrea, L. N. da Costa, D. L. DePoy, S. Desai, H. T. Diehl, S. Dodelson, C. E. Cunha, J. Estrada, A. E. Evrard, A. Fausti Neto, E. Fernandez, D. A. Finley, B. Flaugher, J. Frieman, E. Gaztanaga, D. Gerdes, D. Gruen, R. A. Gruendl, K. Honscheid, D. James, S. Kent, K. Kuehn, N. Kuropatkin, O. Lahav, M. A. G. Maia, M. March, P. Martini, C. J. Miller, R. Miquel, R. Ogando, A. K. Romer, A. Roodman, E. S. Rykoff, M. Sako, E. Sanchez, M. Schubnell, I. Sevilla, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, J. Thaler, D. Tucker, V. Vikram, A. R. Walker, W. Wester, and DES Collaboration. Stellar Kinematics and Metallicities in the Ultra-faint Dwarf Galaxy Reticulum II. ApJ, 808:95, July 2015. doi: 10.1088/0004-637X/808/1/95.

J. D. Simon, T. S. Li, A. Drlica-Wagner, K. Bechtol, J. L. Marshall, D. J. James, M. Y. Wang, L. Strigari, E. Balbinot, K. Kuehn, A. R. Walker, T. M. C. Abbott, S. Allam, J. Annis, A. Benoit-Levy, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, D. L. DePoy, S. Desai, P. Doel, E. Fernandez, B. Flaugher, J. Frieman, J. Garcia-Bellido, E. Gaztanaga, D. A. Goldstein, D. Gruen, G. Gutierrez, N. Kuropatkin, M. A. G. Maia, P. Martini, F. Menanteau, C. J. Miller, R. Miquel, E. Neilsen, B. Nord, R. Ogando, A. A. Plazas, A. K. Romer, E. S. Rykoff, E. Sanchez, B. Santiago, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, R. C. Smith, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, L. Whiteway, and B. Yanny. Nearest Neighbor: The Low-Mass Milky Way Satellite Tucana III. *ArXiv e-prints*, October 2016.

E. Sirko. Initial Conditions to Cosmological N-Body Simulations, or, How to Run an Ensemble of Simulations. ApJ, 634:728–743, November 2005. doi: 10.1086/497090.

C. T. Slater, E. F. Bell, and N. F. Martin. Andromeda XXVIII: A Dwarf Galaxy More Than 350 kpc from Andromeda. ApJ, 742:L14, November 2011. doi: 10.1088/2041-8205/742/1/L14.

Britton D Smith, John H Wise, Brian W O'Shea, Michael L Norman, and Sadegh Khochfar. The first Population II stars formed in externally enriched mini-haloes. *MNRAS*, 452(3):2822–2836, July 2015.

Martin C Smith, Gregory R Ruchti, Amina Helmi, Rosemary F G Wyse, J P Fulbright, K C Freeman, J F Navarro, G M Seabroke, M Steinmetz, M Williams, O Bienaymé, J Binney, J Bland-Hawthorn, W Dehnen, B. K. Gibson, G Gilmore, E K Grebel, U Munari, Q A Parker, R D Scholz, A Siebert, F G Watson, and T Zwitter. The RAVE survey: constraining the local Galactic escape speed. *MNRAS*, 379(2):755–772, August 2007.

G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. ApJ, 396:L1–L5, September 1992. doi: 10.1086/186504.

David Sobral, Jorryt Matthee, Behnam Darvish, Daniel Schaerer, Bahram Mobasher, Huub J A Röttgering, Sérgio Santos, and Shoubaneh Hemmati. Evidence for PopIII-like Stellar Populations in the Most Luminous Lyman-α Emitters at the Epoch of Reionization: Spectroscopic Confirmation. *ApJ*, 808(2):139, August 2015.

Sangmo Tony Sohn, Gurtina Besla, Roeland P van der Marel, Michael Boylan-Kolchin, Steven R Majewski, and James S Bullock. THE SPACE MOTION OF LEO I: HUBBLE SPACE TELESCOPEPROPER MOTION AND IMPLIED ORBIT. *ApJ*, 768(2):139, April 2013.

R. S. Somerville. Can Photoionization Squelching Resolve the Substructure Crisis? ApJ, 572:L23–L26, June 2002. doi: 10.1086/341444.

R. S. Somerville and R. Davé. Physical Models of Galaxy Formation in a Cosmological Framework. ARA&A, 53:51–113, August 2015. doi: 10.1146/annurev-astro-082812-140951.

D. N. Spergel and P. J. Steinhardt. Observational Evidence for Self-Interacting Cold Dark Matter. *Physical Review Letters*, 84:3760, April 2000. doi: 10.1103/PhysRevLett.84.3760.

D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page,

G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. ApJS, 148:175–194, September 2003. doi: 10.1086/377226.

D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. ApJS, 170:377–408, June 2007. doi: 10.1086/513700.

V. Springel. The cosmological simulation code GADGET-2. MNRAS, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at [formmu2]z=0. MNRAS, 328:726–750, December 2001. doi: 10.1046/j.1365-8711.2001.04912.x.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. Nature, 435:629–636, June 2005. doi: 10.1038/nature03597.

V. Springel, C. S. Frenk, and S. D. M. White. The large-scale structure of the Universe. Nature, 440:1137–1144, April 2006. doi: 10.1038/nature04805.

V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White. The Aquarius Project: the subhaloes of galactic haloes. MNRAS, 391:1685–1711, December 2008. doi: 10.1111/j.1365-2966.2008.14066.x.

Volker Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *MNRAS*, 401(2):791–851, January 2010.

A. Stacy, V. Bromm, and A. T. Lee. Building up the Population III initial mass function from cosmological initial conditions. MNRAS, 462:1307–1328, October 2016. doi: 10.1093/mnras/stw1728.

J Stadel, D Potter, B Moore, J Diemand, P Madau, M Zemp, M Kuhlen, and V Quilis. Quantifying the heart of darkness with GHALO - a multibillion particle simulation of a galactic halo. *MNRAS*, 398(1):L21–L25, September 2009.

E. Starkenburg, A. Helmi, G. De Lucia, Y.-S. Li, J. F. Navarro, A. S. Font, C. S. Frenk, V. Springel, C. A. Vera-Ciro, and S. D. M. White. The satellites of the Milky Way - insights from semi-analytic modelling in a ΛCDM cosmology. MNRAS, 429: 725–743, February 2013. doi: 10.1093/mnras/sts367.

E. Starkenburg, K. A. Oman, J. F. Navarro, R. A. Crain, A. Fattahi, C. S. Frenk, T. Sawala, and J. Schaye. The oldest and most metal poor stars in the APOSTLE Local Group simulations. *ArXiv e-prints*, September 2016.

M Steinmetz, T Zwitter, A Siebert, F G Watson, K C Freeman, U Munari, R Campbell, M Williams, G M Seabroke, R F G Wyse, Q A Parker, O Bienaymé, S Roeser, B. K. Gibson, G Gilmore, E K Grebel, A Helmi, J F Navarro, D Burton, C J P Cass, J A Dawe, K Fiegert, M Hartley, K S Russell, W Saunders, H Enke, J Bailin, J Binney, J Bland-Hawthorn, C Boeche, W Dehnen, D J Eisenstein, N W Evans, M Fiorucci, J P Fulbright, O Gerhard, U Jauregi, A Kelz, L Mijović, I Minchev, G Parmentier, J Peñarrubia, A C Quillen, M A Read, G Ruchti, R D Scholz, A Siviero, M C Smith, R Sordo, L Veltz, S Vidrih, R von Berlepsch, B J Boyle, and E Schilbach. The Radial Velocity Experiment (RAVE): First Data Release. *AJ*, 132(4):1645–1668, October 2006.

A. Tasitsiomi, A. V. Kravtsov, R. H. Wechsler, and J. R. Primack. Modeling Galaxy-Mass Correlations in Dissipationless Simulations. ApJ, 614:533–546, October 2004. doi: 10.1086/423784.

M. Tegmark, J. Silk, M. J. Rees, A. Blanchard, T. Abel, and F. Palla. How Small Were the First Cosmological Objects? ApJ, 474:1, January 1997. doi: 10.1086/303434.

R. Teyssier, A. Pontzen, Y. Dubois, and J. I. Read. Cusp-core transformations in dwarf galaxies: observational predictions. MNRAS, 429:3068–3078, March 2013. doi: 10.1093/mnras/sts563.

A. A. Thoul and D. H. Weinberg. Hydrodynamic Simulations of Galaxy Formation. II. Photoionization and the Formation of Low-Mass Galaxies. ApJ, 465:608, July 1996. doi: 10.1086/177446.

Anton V Tikhonov and Anatoly Klypin. The emptiness of voids: yet another overabundance problem for the Λ cold dark matter model. *MNRAS*, 395(4):1915–1924, June 2009.

J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. ApJ, 688:709–728, December 2008. doi: 10.1086/591439.

P. Tisserand, L. Le Guillou, C. Afonso, J. N. Albert, J. Andersen, R. Ansari, É. Aubourg, P. Bareyre, J. P. Beaulieu, X. Charlot, C. Coutures, R. Ferlet, P. Fouqué, J. F. Glicenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, C. Hamadache, J. de Kat, T. Lasserre, É. Lesquoy, C. Loup, C. Magneville, J. B. Marquette, É. Maurice, A. Maury, A. Milsztajn, M. Moniez, N. Palanque-Delabrouille, O. Perdereau, Y. R. Rahal, J. Rich, M. Spiro, A. Vidal-Madjar, L. Vigroux, S. Zylberajch, and EROS-2 Collaboration. Limits on the Macho content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. A&A, 469: 387–404, July 2007. doi: 10.1051/0004-6361:20066017.

Erik J Tollerud, Rachael L Beaton, Marla C Geha, James S Bullock, Puragra Guhathakurta, Jason S Kalirai, Steve R Majewski, Evan N Kirby, Karoline M Gilbert, Basilio Yniguez, Richard J Patterson, James C Ostheimer, Jeff Cooke, Claire E Dorman, Abrar Choudhury, and Michael C Cooper. The SPLASH Survey: Spectroscopy of 15 M31 Dwarf Spheroidal Satellite Galaxies. *ApJ*, 752(1):45, June 2012.

E. Toloba, P. Guhathakurta, A. J. Romanowsky, J. P. Brodie, D. Martínez-Delgado, J. A. Arnold, N. Ramachandran, and K. Theakanath. New Spectroscopic Technique Based on Coaddition of Surface Brightness Fluctuations: NGC 4449 and its Stellar Tidal Stream. ApJ, 824:35, June 2016. doi: 10.3847/0004-637X/824/1/35.

Alar Toomre and Juri Toomre. Galactic Bridges and Tails. *ApJ*, 178:623–666, December 1972.

G. Tormen, A. Diaferio, and D. Syer. Survival of substructure within dark matter haloes. MNRAS, 299:728–742, September 1998. doi: 10.1046/j.1365-8711.1998.01775.x.

G. Torrealba, S. E. Koposov, V. Belokurov, and M. Irwin. The feeble giant. Discovery of a large and diffuse Milky Way dwarf galaxy in the constellation of Crater. MNRAS, 459:2370–2378, July 2016a. doi: 10.1093/mnras/stw733.

G. Torrealba, S. E. Koposov, V. Belokurov, M. Irwin, M. Collins, M. Spencer, R. Ibata, M. Mateo, A. Bonaca, and P. Jethwa. At the survey limits: discovery of the Aquarius 2 dwarf galaxy in the VST ATLAS and the SDSS data. *ArXiv e-prints*, May 2016b.

J. Tumlinson. Chemical Evolution in Hierarchical Models of Cosmic Structure. II. The Formation of the Milky Way Stellar Halo and the Distribution of the Oldest Stars. ApJ, 708:1398–1418, January 2010. doi: 10.1088/0004-637X/708/2/1398.

Hideyuki Umeda and Ken'ichi Nomoto. First-generation black-hole-forming supernovae and the metal abundance pattern of a very iron-poor star. *Nature*, 422(6):871–873, April 2003.

U. Ural, M. I. Wilkinson, J. I. Read, and M. G. Walker. A low pre-infall mass for the Carina dwarf galaxy from disequilibrium modelling. *Nature Communications*, 6:7599, July 2015. doi: 10.1038/ncomms8599.

A. Vale and J. P. Ostriker. Linking halo mass to galaxy luminosity. MNRAS, 353:189–200, September 2004. doi: 10.1111/j.1365-2966.2004.08059.x.

A. Vale and J. P. Ostriker. The non-parametric model for linking galaxy luminosity with halo/subhalo mass. MNRAS, 371:1173–1187, September 2006. doi: 10.1111/j.1365-2966.2006.10605.x.

S. van den Bergh. The Early History of Dark Matter. PASP, 111:657–660, June 1999. doi: 10.1086/316369.

S. van den Bergh. Hubble and Shapley - Two Early Giants of Observational Cosmology. JRASC, 105:245, December 2011.

F. C. van den Bosch, G. Tormen, and C. Giocoli. The mass function and average mass-loss rate of dark matter subhaloes. MNRAS, 359:1029–1040, May 2005. doi: 10.1111/j.1365-2966.2005.08964.x.

F. C. van den Bosch, D. Aquino, X. Yang, H. J. Mo, A. Pasquali, D. H. McIntosh, S. M. Weinmann, and X. Kang. The importance of satellite quenching for the build-up of the red sequence of present-day galaxies. MNRAS, 387:79–91, June 2008. doi: 10.1111/j.1365-2966.2008.13230.x.

R. P. van der Marel, D. R. Alves, E. Hardy, and N. B. Suntzeff. New Understanding of Large Magellanic Cloud Structure, Dynamics, and Orbit from Carbon Star Kinematics. AJ, 124:2639–2663, November 2002. doi: 10.1086/343775.

Roeland P van der Marel, Mark Fardal, Gurtina Besla, Rachael L Beaton, Sangmo Tony Sohn, Jay Anderson, Tom Brown, and Puragra Guhathakurta. The M31 Velocity Vector. II. Radial Orbit toward the Milky Way and Implied Local Group Mass. *ApJ*, 753(1):8, July 2012.

I. van Vledder, D. van der Vlugt, B. W. Holwerda, M. A. Kenworthy, R. J. Bouwens, and M. Trenti. The size and shape of the Milky Way disc and halo from M-type brown dwarfs in the BoRG survey. MNRAS, 458:425–437, May 2016. doi: 10.1093/mnras/stw258.

M. Vogelsberger and J. Zavala. Direct detection of self-interacting dark matter. MNRAS, 430:1722–1735, April 2013. doi: 10.1093/mnras/sts712.

M. Vogelsberger, J. Zavala, and A. Loeb. Subhaloes in self-interacting galactic dark matter haloes. MNRAS, 423:3740–3752, July 2012. doi: 10.1111/j.1365-2966.2012.21182.x.

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. MNRAS, 444:1518–1547, October 2014a. doi: 10.1093/mnras/stu1536.

M. Vogelsberger, J. Zavala, C. Simpson, and A. Jenkins. Dwarf galaxies in CDM and SIDM with baryons: observational probes of the nature of dark matter. MNRAS, 444:3684–3698, November 2014b. doi: 10.1093/mnras/stu1713.

M. Vogelsberger, J. Zavala, F.-Y. Cyr-Racine, C. Pfrommer, T. Bringmann, and K. Sigurdson. ETHOS - an effective theory of structure formation: dark matter physics as a possible explanation of the small-scale CDM problems. MNRAS, 460: 1399–1416, August 2016. doi: 10.1093/mnras/stw1076.

M. G. Walker and J. Peñarrubia. A Method for Measuring (Slopes of) the Mass Profiles of Dwarf Spheroidal Galaxies. ApJ, 742:20, November 2011. doi: 10.1088/0004-637X/742/1/20.

M. G. Walker, M. Mateo, E. W. Olszewski, J. I. Bailey, III, S. E. Koposov, V. Belokurov, and N. W. Evans. Magellan/M2FS Spectroscopy of the Reticulum 2 Dwarf Spheroidal Galaxy. ApJ, 808:108, August 2015. doi: 10.1088/0004-637X/808/2/108.

M. G. Walker, M. Mateo, E. W. Olszewski, S. Koposov, V. Belokurov, P. Jethwa, D. L. Nidever, V. Bonnivard, J. I. Bailey, III, E. F. Bell, and S. R. Loebman. Magellan/M2FS Spectroscopy of Tucana 2 and Grus 1. ApJ, 819:53, March 2016. doi: 10.3847/0004-637X/819/1/53.

S. M. Walsh, H. Jerjen, and B. Willman. A Pair of Boötes: A New Milky Way Satellite. ApJ, 662:L83–L86, June 2007. doi: 10.1086/519684.

S. M. Walsh, B. Willman, and H. Jerjen. The Invisibles: A Detection Algorithm to Trace the Faintest Milky Way Satellites. AJ, 137:450–469, January 2009. doi: 10.1088/0004-6256/137/1/450.

J. Wang, G. De Lucia, M. G. Kitzbichler, and S. D. M. White. The dependence of galaxy formation on cosmological parameters: can we distinguish between the WMAP1 and WMAP3 parameter sets? MNRAS, 384:1301–1315, March 2008. doi: 10.1111/j.1365-2966.2007.12797.x.

J. Wang, C. S. Frenk, J. F. Navarro, L. Gao, and T. Sawala. The missing massive satellites of the Milky Way. MNRAS, 424:2715–2721, August 2012. doi: 10.1111/j.1365-2966.2012.21357.x.

L. Wang and Y. P. Jing. Modelling galaxy stellar mass evolution from z ~ 0.8 to today. MNRAS, 402:1796–1806, March 2010. doi: 10.1111/j.1365-2966.2009.16007.x.

W. Wang, J. Han, A. P. Cooper, S. Cole, C. Frenk, and B. Lowing. Estimating the dark matter halo mass of our Milky Way using dynamical tracers. MNRAS, 453:377–400, October 2015. doi: 10.1093/mnras/stv1647.

Douglas F Watson, Andreas A Berlind, and Andrew R Zentner. A Cosmic Coincidence: The Power-law Galaxy Correlation Function. *ApJ*, 738(1):22, September 2011.

R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. ApJ, 568:52–70, March 2002. doi: 10.1086/338765.

A. R. Wetzel, A. J. Deason, and S. Garrison-Kimmel. Satellite Dwarf Galaxies in a Hierarchical Universe: Infall Histories, Group Preprocessing, and Reionization. ApJ, 807:49, July 2015. doi: 10.1088/0004-637X/807/1/49.

A. R. Wetzel, P. F. Hopkins, J.-h. Kim, C.-A. Faucher-Giguere, D. Keres, and E. Quataert. Reconciling dwarf galaxies with LCDM cosmology: Simulating a realistic population of satellites around a Milky Way-mass galaxy. *ArXiv e-prints*, February 2016.

Daniel Whalen, Tom Abel, and Michael L Norman. Radiation Hydrodynamic Evolution of Primordial H iiRegions. *ApJ*, 610(1):14–22, July 2004.

Daniel Whalen, Bob van Veelen, Brian W O'Shea, and Michael L Norman. The Destruction of Cosmological Minihalos by Primordial Supernovae. *ApJ*, 682(1): 49–67, July 2008.

C. Wheeler, J. Oñorbe, J. S. Bullock, M. Boylan-Kolchin, O. D. Elbert, S. Garrison-Kimmel, P. F. Hopkins, and D. Kereš. Sweating the small stuff: simulating dwarf galaxies, ultra-faint dwarf galaxies, and their own tiny satellites. MNRAS, 453: 1305–1316, October 2015. doi: 10.1093/mnras/stv1691.

S D M White. Simulations of merging galaxies. *MNRAS*, 184:185–203, July 1978.

S D M White and M. J. Rees. Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. *MNRAS*, 183:341–358, May 1978.

R. P. C. Wiersma, J. Schaye, and B. D. Smith. The effect of photoionization on the cooling rates of enriched, astrophysical plasmas. MNRAS, 393:99–107, February 2009. doi: 10.1111/j.1365-2966.2008.14191.x.

B. Willman, M. R. Blanton, A. A. West, J. J. Dalcanton, D. W. Hogg, D. P. Schneider, N. Wherry, B. Yanny, and J. Brinkmann. A New Milky Way Companion: Unusual Globular Cluster or Extreme Dwarf Satellite? AJ, 129:2692–2700, June 2005a. doi: 10.1086/430214.

B. Willman, J. J. Dalcanton, D. Martinez-Delgado, A. A. West, M. R. Blanton, D. W. Hogg, J. C. Barentine, H. J. Brewington, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, E. H. Neilsen, Jr., A. Nitta, and S. A. Snedden. A New Milky Way Dwarf Galaxy in Ursa Major. ApJ, 626:L85–L88, June 2005b. doi: 10.1086/431760.

J. H. Wise. First Light: A Brief Review. *ArXiv e-prints*, January 2012.

John H Wise and Tom Abel. Resolving the Formation of Protogalaxies. III. Feedback from the First Stars. *ApJ*, 685(1):40–56, September 2008.

Joe Wolf, Gregory D Martinez, James S Bullock, Manoj Kaplinghat, Marla Geha, Ricardo R Muñoz, Joshua D Simon, and Frank F Avedo. Accurate masses for dispersion-supported galaxies. *MNRAS*, 406:1220–1237, August 2010.

J. S. B. Wyithe and A. Loeb. A suppressed contribution of low-mass galaxies to reionization due to supernova feedback. MNRAS, 428:2741–2754, January 2013. doi: 10.1093/mnras/sts242.

Ł. Wyrzykowski, S. Kozłowski, J. Skowron, A. Udalski, M. K. Szymański, M. Kubiak, G. Pietrzyński, I. Soszyński, O. Szewczyk, K. Ulaczyk, and R. Poleski. The OGLE view of microlensing towards the Magellanic Clouds - III. Ruling out subsolar MACHOs with the OGLE-III LMC data. MNRAS, 413:493–508, May 2011a. doi: 10.1111/j.1365-2966.2010.18150.x.

L. Wyrzykowski, J. Skowron, S. Kozłowski, A. Udalski, M. K. Szymański, M. Kubiak, G. Pietrzyński, I. Soszyński, O. Szewczyk, K. Ulaczyk, R. Poleski, and P. Tisserand. The OGLE view of microlensing towards the Magellanic Clouds - IV. OGLE-III SMC data and final conclusions on MACHOs. MNRAS, 416:2949–2961, October 2011b. doi: 10.1111/j.1365-2966.2011.19243.x.

X X Xue, H W Rix, G Zhao, P Re Fiorentin, T Naab, M Steinmetz, F C van den Bosch, T C Beers, Y S Lee, E F Bell, C Rockosi, B Yanny, H Newberg, R Wilhelm, X Kang, M C Smith, and D P Schneider. The Milky Way's Circular Velocity Curve to 60 kpc and an Estimate of the Dark Matter Halo Mass from the Kinematics of ~2400 SDSS Blue Horizontal-Branch Stars. *ApJ*, 684(2):1143–1158, September 2008.

X. Yang, H. J. Mo, and F. C. van den Bosch. Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies. MNRAS, 339: 1057–1080, March 2003. doi: 10.1046/j.1365-8711.2003.06254.x.

X. Yang, H. J. Mo, F. C. van den Bosch, Y. Zhang, and J. Han. Evolution of the Galaxy-Dark Matter Connection and the Assembly of Galaxies in Dark Matter Halos. ApJ, 752:41, June 2012. doi: 10.1088/0004-637X/752/1/41.

B. Yanny and H. J. Newberg. The Monoceros Ring, and Other Substructure Near the Galactic Plane. In H. J. Newberg and J. L. Carlin, editors, *Astrophysics and Space Science Library*, volume 420 of *Astrophysics and Space Science Library*, page 63, 2016. doi: 10.1007/978-3-319-19336-6_3.

Brian Yanny, Constance Rockosi, Heidi Jo Newberg, Gillian R Knapp, Jennifer K Adelman-McCarthy, Bonnie Alcorn, Sahar Allam, Carlos Allende Prieto, Deokkeun An, Kurt S J Anderson, Scott Anderson, Coryn A L Bailer-Jones, Steve Bastian, Timothy C Beers, Eric Bell, Vasily Belokurov, Dmitry Bizyaev, Norm Blythe, John J Bochanski, William N Boroski, Jarle Brinchmann, J Brinkmann, Howard Brewington, Larry Carey, Kyle M Cudworth, Michael Evans, N W Evans, Evalyn Gates, B T Gänsicke, Bruce Gillespie, Gerald Gilmore, Ada Nebot Gomez-Moran, Eva K Grebel, Jim Greenwell, James E Gunn, Cathy Jordan, Wendell Jordan, Paul Harding, Hugh Harris, John S Hendry, Diana Holder, Inese I Ivans, Željko Ivezić, Sebastian Jester, Jennifer A Johnson, Stephen M Kent, Scot Kleinman, Alexei Kniazev, Jurek Krzesinski, Richard Kron, Nikolay Kuropatkin, Svetlana Lebedeva, Young Sun Lee, R French Leger, Sébastien Lépine, Steve Levine, Huan Lin, Daniel C Long, Craig Loomis, Robert Lupton, Olena Malanushenko, Viktor Malanushenko, Bruce Margon, David Martinez-Delgado, Peregrine McGehee, Dave Monet, Heather L Morrison, Jeffrey A Munn, Eric H Jr Neilsen, Atsuko

Nitta, John E Norris, Dan Oravetz, Russell Owen, Nikhil Padmanabhan, Kaike Pan, R S Peterson, Jeffrey R Pier, Jared Platson, Paola Re Fiorentin, Gordon T Richards, Hans-Walter Rix, David J Schlegel, Donald P Schneider, Matthias R Schreiber, Axel Schwope, Valena Sibley, Audrey Simmons, Stephanie A Snedden, J Allyn Smith, Larry Stark, Fritz Stauffer, M Steinmetz, C Stoughton, Mark SubbaRao, Alex Szalay, Paula Szkody, Aniruddha R Thakar, Thirupathi Sivarani, Douglas Tucker, Alan Uomoto, Dan Vanden Berk, Simon Vidrih, Yogesh Wadadekar, Shannon Watters, Ron Wilhelm, Rosemary F G Wyse, Jean Yarger, and Dan Zucker. SEGUE: A Spectroscopic Survey of 240,000 Stars with g = 14-20. *AJ*, 137 (5):4377–4399, May 2009.

N. Yoshida. Star Formation in the Early Universe. In *38th COSPAR Scientific Assembly*, volume 38 of *COSPAR Meeting*, page 2, 2010.

N. Yoshida, V. Springel, S. D. M. White, and G. Tormen. Collisional Dark Matter and the Structure of Dark Halos. ApJ, 535:L103–L106, June 2000a. doi: 10.1086/312707.

N. Yoshida, V. Springel, S. D. M. White, and G. Tormen. Weakly Self-interacting Dark Matter and the Structure of Dark Halos. ApJ, 544:L87–L90, December 2000b. doi: 10.1086/317306.

C. Yozin and K. Bekki. Local ultra faint dwarves as a product of Galactic processing during a Magellanic group infall. MNRAS, 453:2302–2307, November 2015. doi: 10.1093/mnras/stv1828.

S. Zaroubi. The Epoch of Reionization. In T. Wiklind, B. Mobasher, and V. Bromm, editors, *The First Galaxies*, volume 396 of *Astrophysics and Space Science Library*, page 45, 2013. doi: 10.1007/978-3-642-32362-1_2.

J. Zavala, M. Vogelsberger, and M. G. Walker. Constraining self-interacting dark matter with the Milky Way's dwarf spheroidals. MNRAS, 431:L20–L24, April 2013. doi: 10.1093/mnrasl/sls053.

A. R. Zentner and J. S. Bullock. Halo Substructure and the Power Spectrum. ApJ, 598:49–72, November 2003. doi: 10.1086/378797.

Andrew R Zentner, Andrey V Kravtsov, Oleg Y Gnedin, and Anatoly A Klypin. The Anisotropic Distribution of Galactic Satellites. *ApJ*, 629(1):219–232, August 2005.

D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos. ApJ, 707: 354–369, December 2009. doi: 10.1088/0004-637X/707/1/354.

A. A. Zijlstra and D. Minniti. A Dwarf Irregular Galaxy at the Edge of the Local Group: Stellar Populations and Distance of IC 5152. AJ, 117:1743–1757, April 1999. doi: 10.1086/300802.

A. Zolotov, A. M. Brooks, B. Willman, F. Governato, A. Pontzen, C. Christensen, A. Dekel, T. Quinn, S. Shen, and J. Wadsley. Baryons Matter: Why Luminous Satellite Galaxies have Reduced Central Masses. ApJ, 761:71, December 2012. doi: 10.1088/0004-637X/761/1/71.

D. B. Zucker, A. Y. Kniazev, E. F. Bell, D. Martínez-Delgado, E. K. Grebel, H.-W. Rix, C. M. Rockosi, J. A. Holtzman, R. A. M. Walterbos, J. Annis, D. G. York, Ž. Ivezić, J. Brinkmann, H. Brewington, M. Harvanek, G. Hennessy, S. J. Kleinman, J. Krzesinski, D. Long, P. R. Newman, A. Nitta, and S. A. Snedden. Andromeda IX: A New Dwarf Spheroidal Satellite of M31. ApJ, 612:L121–L124, September 2004. doi: 10.1086/424691.

D. B. Zucker, V. Belokurov, N. W. Evans, J. T. Kleyna, M. J. Irwin, M. I. Wilkinson, M. Fellhauer, D. M. Bramich, G. Gilmore, H. J. Newberg, B. Yanny, J. A. Smith, P. C. Hewett, E. F. Bell, H.-W. Rix, O. Y. Gnedin, S. Vidrih, R. F. G. Wyse, B. Willman, E. K. Grebel, D. P. Schneider, T. C. Beers, A. Y. Kniazev, J. C. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, A. Nitta, and S. A. Snedden. A Curious Milky Way Satellite in Ursa Major. ApJ, 650:L41–L44, October 2006a. doi: 10.1086/508628.

D. B. Zucker, V. Belokurov, N. W. Evans, M. I. Wilkinson, M. J. Irwin, T. Sivarani, S. Hodgkin, D. M. Bramich, J. M. Irwin, G. Gilmore, B. Willman, S. Vidrih, M. Fellhauer, P. C. Hewett, T. C. Beers, E. F. Bell, E. K. Grebel, D. P. Schneider, H. J. Newberg, R. F. G. Wyse, C. M. Rockosi, B. Yanny, R. Lupton, J. A. Smith, J. C. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, A. Nitta, and S. A. Snedden. A New Milky Way Dwarf Satellite in Canes Venatici. ApJ, 643:L103–L106, June 2006b. doi: 10.1086/505216.

D. B. Zucker, A. Y. Kniazev, D. Martínez-Delgado, E. F. Bell, H.-W. Rix, E. K. Grebel, J. A. Holtzman, R. A. M. Walterbos, C. M. Rockosi, D. G. York, J. C. Barentine, H. Brewington, J. Brinkmann, M. Harvanek, S. J. Kleinman, J. Krzesinski, D. Long, E. H. Neilsen, Jr., A. Nitta, and S. A. Snedden. Andromeda X, a New Dwarf Spheroidal Satellite of M31: Photometry. ApJ, 659:L21–L24, April 2007. doi: 10.1086/516748.