

Small-Scale Challenges to the Λ CDM Paradigm

James S. Bullock¹ and Michael Boylan-Kolchin²

¹Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA; email: bullock@uci.edu

²Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA; email: mbk@astro.as.utexas.edu

Annu. Rev. Astron. Astrophys. 2017.
55:343–87

This article's doi:
10.1146/annurev-astro-091916-055313

Copyright © 2017 by Annual Reviews.
All rights reserved

Draft version. Posted with permission
from the Annual Review of Astronomy
and Astrophysics, Volume 55 by Annual
Reviews, <http://www.annualreviews.org>

Keywords

cosmology, dark matter, dwarf galaxies, galaxy formation, Local Group

Abstract

The dark energy plus cold dark matter (Λ CDM) cosmological model has been a demonstrably successful framework for predicting and explaining the large-scale structure of Universe and its evolution with time. Yet on length scales smaller than ~ 1 Mpc and mass scales smaller than $\sim 10^{11} M_{\odot}$, the theory faces a number of challenges. For example, the observed cores of many dark-matter dominated galaxies are both less dense and less cuspy than naively predicted in Λ CDM. The number of small galaxies and dwarf satellites in the Local Group is also far below the predicted count of low-mass dark matter halos and subhalos within similar volumes. These issues underlie the most well-documented problems with Λ CDM: Cusp/Core, Missing Satellites, and Too-Big-to-Fail. The key question is whether a better understanding of baryon physics, dark matter physics, or both will be required to meet these challenges. Other anomalies, including the observed planar and orbital configurations of Local Group satellites and the tight baryonic/dark matter scaling relations obeyed by the galaxy population, have been less thoroughly explored in the context of Λ CDM theory. Future surveys to discover faint, distant dwarf galaxies and to precisely measure their masses and density structure hold promising avenues for testing possible solutions to the small-scale challenges going forward. Observational programs to constrain or discover and characterize the number of truly dark low-mass halos are among the most important, and achievable, goals in this field over the next decade. These efforts will either further verify the Λ CDM paradigm or demand a substantial revision in our understanding of the nature of dark matter.

Contents

1. INTRODUCTION	2
1.1. Preliminaries: how small is a small galaxy?	3
1.2. Overview of the Λ CDM model	5
1.3. Dark matter halos	6
1.4. Dark matter substructure	12
1.5. Linking dark matter halos to galaxies	13
1.6. Connections to particle physics	16
2. OVERVIEW OF PROBLEMS	18
2.1. Missing Satellites	18
2.2. Cusp, Cores, and Excess Mass	20
2.3. Too-Big-To-Fail	22
2.4. Satellite Planes	24
2.5. Regularity in the Face of Diversity	26
3. SOLUTIONS	27
3.1. Solutions within Λ CDM	27
3.2. Solutions requiring modifications to Λ CDM	32
4. Current Frontiers	36
4.1. Dwarf galaxy discovery space in the Local Group	36
4.2. Dwarfs beyond the Local Group	37
4.3. Searches for starless dwarfs	37
4.4. Indirect signatures of dark matter	38
4.5. The high-redshift Universe	39
5. Summary and Outlook	40

1. INTRODUCTION

Astrophysical observations ranging from the scale of the horizon ($\sim 15,000$ Mpc) to the typical spacing between galaxies (~ 1 Mpc) are all consistent with a Universe that was seeded by a nearly scale-invariant fluctuation spectrum and that is dominated today by dark energy ($\sim 70\%$) and Cold Dark Matter ($\sim 25\%$), with baryons contributing only $\sim 5\%$ to the energy density (Planck Collaboration et al. 2016; Guo et al. 2016). This cosmological model has provided a compelling backbone to galaxy formation theory, a field that is becoming increasingly successful at reproducing the detailed properties of galaxies, including their counts, clustering, colors, morphologies, and evolution over time (Vogelsberger et al. 2014; Schaye et al. 2015). As described in this review, there are observations below the scale of ~ 1 Mpc that have proven more problematic to understand in the Λ CDM framework. It is not yet clear whether the small-scale issues with Λ CDM will be accommodated by a better understanding of astrophysics or dark matter physics, or if they will require a radical revision of cosmology, but any correct description of our Universe must look very much like Λ CDM on large scales. It is with this in mind that we discuss the small-scale challenges to the current paradigm. For concreteness, we assume that the default Λ CDM cosmology has parameters $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.6727$, $\Omega_m = 0.3156$, $\Omega_\Lambda = 0.6844$, $\Omega_b = 0.04927$, $\sigma_8 = 0.831$, and $n_s = 0.9645$ (Planck Collaboration et al. 2016).

Given the scope of this review, we must sacrifice detailed discussions for a more broad, high-level approach. There are many recent reviews or overview papers that cover, in more

depth, certain aspects of this review. These include [Frenk & White \(2012\)](#), [Peebles \(2012\)](#), and [Primack \(2012\)](#) on the historical context of Λ CDM and some of its basic predictions; [Willman \(2010\)](#) and [McConnachie \(2012\)](#) on searches for and observed properties of dwarf galaxies in the Local Group; [Feng \(2010\)](#), [Porter, Johnson & Graham \(2011\)](#), and [Strigari \(2013\)](#) on the nature of and searches for dark matter; [Kuhlen, Vogelsberger & Angulo \(2012\)](#) on numerical simulations of cosmological structure formation; and [Brooks \(2014\)](#), [Weinberg et al. \(2015\)](#) and [Del Popolo & Le Delliou \(2017\)](#) on small-scale issues in Λ CDM. Additionally, we will not discuss cosmic acceleration (the Λ in Λ CDM) here; that topic is reviewed in [Weinberg et al. \(2013\)](#). Finally, space does not allow us to address the possibility that the challenges facing Λ CDM on small scales reflects a deeper problem in our understanding of gravity. We point the reader to reviews by [Milgrom \(2002\)](#), [Famaey & McGaugh \(2012\)](#), and [McGaugh \(2015\)](#), which compare Modified Newtonian Dynamics (MOND) to Λ CDM and provide further references on this topic.

1.1. Preliminaries: how small is a small galaxy?

This is a review on small-scale challenges to the Λ CDM model. The past ~ 12 years have seen transformative discoveries that have fundamentally altered our understanding of “small scales” – at least in terms of the low-luminosity limit of galaxy formation.

Prior to 2004, the smallest galaxy known was Draco, with a stellar mass of $M_\star \simeq 5 \times 10^5 M_\odot$. Today, we know of galaxies 1000 times less luminous. While essentially all Milky Way satellites discovered before 2004 were found via visual inspection of photographic plates (with the exceptions of the Carina and Sagittarius dwarf spheroidal galaxies), the advent of large-area digital sky surveys with deep exposures and accurate star-galaxy separation algorithms has revolutionized the search for and discovery of faint stellar systems in the Milky Way (see [Willman 2010](#) for a review of the search for faint satellites). The Sloan Digital Sky Survey (SDSS) ushered in this revolution, doubling the number of known Milky Way satellites in the first five years of active searches. The PAndAS survey discovered a similar population of faint dwarfs around M31 ([Richardson et al. 2011](#)). More recently the DES survey has continued this trend ([Koposov et al. 2015](#); [Drlica-Wagner et al. 2015](#)). All told, we know of ~ 50 satellite galaxies of the Milky Way and ~ 30 satellites of M31 today ([McConnachie 2012](#), updated on-line catalog), most of which are fainter than any galaxy known at the turn of the century. They are also extremely dark-matter-dominated, with mass-to-light ratios within their stellar radii exceeding ~ 1000 in some cases ([Walker et al. 2009](#); [Wolf et al. 2010](#)).

Given this upheaval in our understanding of the faint galaxy frontier over the last decade or so, it is worth pausing to clarify some naming conventions. In what follows, the term “dwarf” will refer to galaxies with $M_\star \lesssim 10^9 M_\odot$. We will further subdivide dwarfs into three mass classes: Bright Dwarfs ($M_\star \approx 10^{7-9} M_\odot$), Classical Dwarfs ($M_\star \approx 10^{5-7} M_\odot$), and Ultra-faint Dwarfs ($M_\star \approx 10^{2-5} M_\odot$). Note that another common classification for dwarf galaxies is between dwarf spheroidals (dSphs) and dwarf irregulars (dIrrs). Dwarfs with gas and ongoing star formation are usually labeled dIrr. The term dSph is reserved for dwarfs that lack gas and have no ongoing star formation. Note that the vast majority of field dwarfs (meaning that they are not satellites) are dIrrs. Most dSph galaxies are satellites of larger systems.

Figure 1 illustrates the morphological differences among galaxies that span these stellar mass ranges. From top to bottom we see three dwarfs each that roughly correspond to Bright, Classical, and Ultra-faint Dwarfs, respectively.

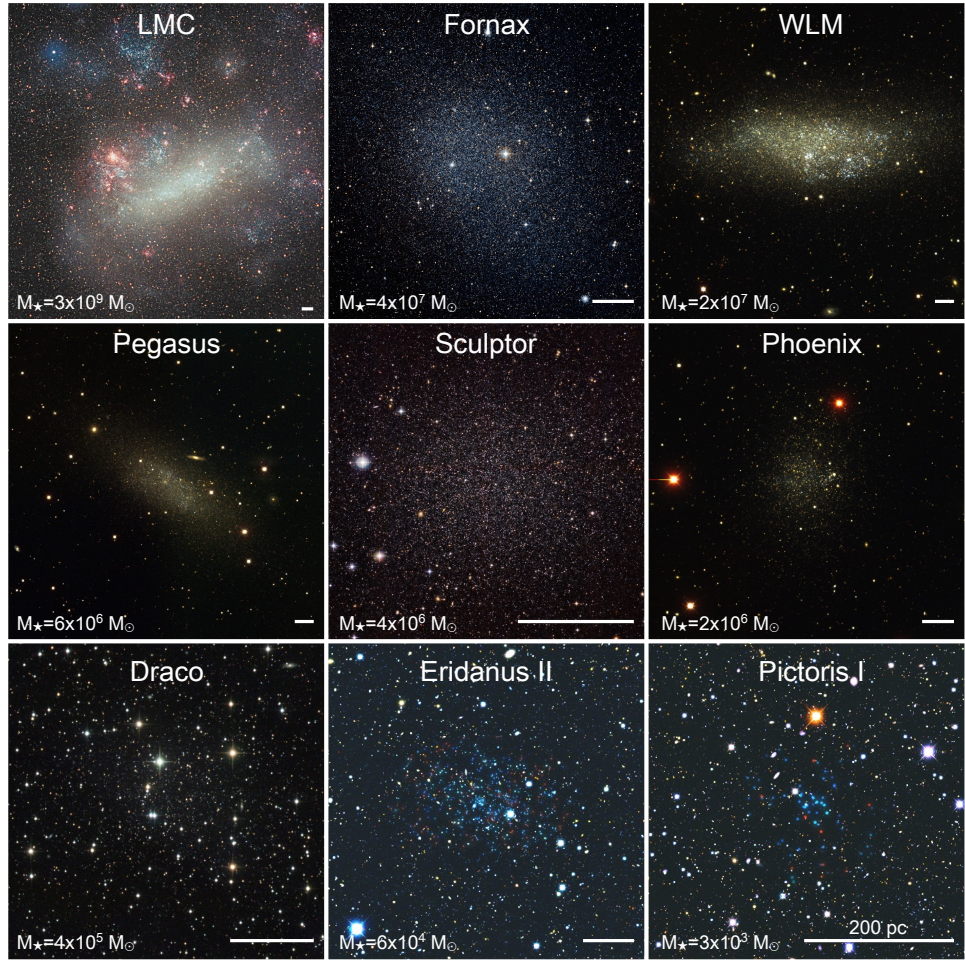


Figure 1

Approaching the threshold of galaxy formation. Shown are images of dwarf galaxies spanning six orders of magnitude in stellar mass. In each panel, the dwarf's stellar mass is listed in the lower-left corner and a scale bar corresponding to 200 pc is shown in the lower-right corner. The LMC, WLM, and Pegasus are dwarf irregular (dIrr) galaxies that have gas and ongoing star formation. The remaining six galaxies shown are gas-free dwarf spheroidal (dSph) galaxies and are not currently forming stars. The faintest galaxies shown here are only detectable in limited volumes around the Milky Way; future surveys may reveal many more such galaxies at greater distances. Image credits: Eckhard Slawik (LMC); ESO/Digitized Sky Survey 2 (Fornax); [Massey et al. \(2007\)](#); WLM, Pegasus, Phoenix); ESO (Sculptor); Mischa Schirmer (Draco), Vasily Belokurov and Sergey Koposov (Eridanus II, Pictoris I).

ADOPTED DWARF GALAXY NAMING CONVENTION

Bright Dwarfs: $M_{\star} \approx 10^{7-9} M_{\odot}$

– the faint galaxy completeness limit for field galaxy surveys

Classical Dwarfs: $M_{\star} \approx 10^{5-7} M_{\odot}$

– the faintest galaxies known prior to SDSS

Ultra-faint Dwarfs: $M_{\star} \approx 10^{2-5} M_{\odot}$

– detected within limited volumes around M31 and the Milky Way

With these definitions in hand, we move to the cosmological model within which we aim to explain the counts, stellar masses, and dark matter content of these dwarfs.

1.2. Overview of the Λ CDM model

The Λ CDM model of cosmology is the culmination of century of work on the physics of structure formation within the framework of general relativity. It also indicates the confluence of particle physics and astrophysics over the past four decades: the particle nature of dark matter directly determines essential properties of non-linear cosmological structure. While the Λ CDM model is phenomenological at present – the actual physics of dark matter and dark energy remain as major theoretical issues – it is highly successful at explaining the large-scale structure of the Universe and basic properties of galaxies that form within dark matter halos.

In the Λ CDM model, cosmic structure is seeded by primordial adiabatic fluctuations and grows by gravitational instability in an expanding background. The primordial power spectrum as a function of wavenumber k is nearly scale-invariant¹, $P(k) \propto k^n$ with $n \simeq 1$. Scales that re-enter the horizon when the Universe is radiation-dominated grow extremely slowly until the epoch of matter domination, leaving a scale-dependent suppression of the primordial power spectrum that goes as k^{-4} at large k . This suppression of power is encapsulated by the “transfer function” $T(k)$, which is defined as the ratio of amplitude of a density perturbation in the post-recombination era to its primordial value as a function of perturbation wavenumber k . This processed power spectrum is the input for structure formation calculations; the dimensionless processed power spectrum, defined by

$$\Delta^2(k, a) = \frac{k^3}{2\pi^2} P(k) T^2(k) d^2(a), \quad (1)$$

therefore rises as k^4 for scales larger than the comoving horizon at matter-radiation equality (corresponding to $k = 0.008 \text{ Mpc}^{-1}$) and is approximately independent of k for scales that re-enter the horizon well before matter-radiation equality. Here, $d(a)$ is the linear growth function, normalized to unity at $a = 1$. The processed $z = 0$ ($a = 1$) linear power spectrum for Λ CDM is shown by the solid line in Figure 2. The asymptotic shape behavior is most easily seen in the bottom panel, which spans the wave number range of cosmological

¹Recent measurements find $n = 0.968 \pm 0.006$ (Planck Collaboration et al. 2016), i.e., small but statistically different from true scale invariance.

interest. For a more complete discussion of primordial fluctuations and the processed power spectrum we recommend that readers consult [Mo, van den Bosch & White \(2010\)](#).

It is useful to associate each wavenumber with a mass scale set by its characteristic length $r_l = \lambda/2 = \pi/k$. In the early Universe, when $\delta \ll 1$, the total amount of matter contained within a sphere of comoving Lagrangian radius r_l at $z = 0$ is

$$M_l = \frac{4\pi}{3} r_l^3 \rho_m = \frac{\Omega_m H_0^2}{2G} r_l^3 \quad (2)$$

$$= 1.71 \times 10^{11} M_\odot \left(\frac{\Omega_m}{0.3} \right) \left(\frac{h}{0.7} \right)^2 \left(\frac{r_l}{1 \text{ Mpc}} \right)^3. \quad (3)$$

The mapping between wave number and mass scale is illustrated by the top and bottom axis in Figure 2. The processed linear power spectrum for Λ CDM shown in the bottom panel (solid line) spans the horizon scale to a typical mass cutoff scale for the most common cold dark matter candidate ($\sim 10^{-6} M_\odot$; see discussion in Section 1.6). A line at $\Delta = 1$ is plotted for reference, showing that fluctuations born on comoving length scales smaller than $r_l \approx 10 h^{-1} \text{ Mpc} \approx 14 \text{ Mpc}$ have gone non-linear today. The top panel is zoomed in on the small scales of relevance for this review (which we define more precisely below). Typical regions on these scales have collapsed into virialized objects today. These collapsed objects – dark matter halos – are the sites of galaxy formation.

1.3. Dark matter halos

1.3.1. Global properties. Soon after overdense regions of the Universe become non-linear, they stop expanding, turn around, and collapse, converting potential energy into kinetic energy in the process. The result is virialized dark matter halos with masses given by

$$M_{\text{vir}} = \frac{4\pi}{3} R_{\text{vir}}^3 \Delta \rho_m, \quad (4)$$

where $\Delta \sim 300$ is the virial over-density parameter, defined here relative to the background matter density. As discussed below, the value of M_{vir} is ultimately a definition that requires some way of defining a halo’s outer edge (R_{vir}). This is done via a choice for Δ . The numerical value for Δ is often chosen to match the over-density one predicts for a virialized dark matter region that has undergone an idealized spherical collapse ([Bryan & Norman 1998](#)), and we will follow that convention here. Note that given a virial mass M_{vir} , the virial radius, R_{vir} , is uniquely defined by Equation 4. Similarly, the virial *velocity*

$$V_{\text{vir}} \equiv \sqrt{\frac{GM_{\text{vir}}}{R_{\text{vir}}}}, \quad (5)$$

is also uniquely defined. The parameters M_{vir} , R_{vir} , and V_{vir} are equivalent mass labels – any one determines the other two, given a specified over-density parameter Δ .

One nice implication of Equation 4 is that a present-day object with virial mass M_{vir} can be associated directly with a linear perturbation with mass M_l . Equating the two gives

$$R_{\text{vir}} = 0.15 \left(\frac{\Delta}{300} \right)^{-1/3} r_l. \quad (6)$$

We see that a collapsed halo of size R_{vir} is approximately 7 times smaller in physical dimension than the comoving linear scale associated with that mass today.

Galaxy Clusters:

$$M_{\text{vir}} \approx 10^{15} M_\odot \\ V_{\text{vir}} \approx 1000 \text{ km s}^{-1}$$

Milky Way:

$$M_{\text{vir}} \approx 10^{12} M_\odot \\ V_{\text{vir}} \approx 100 \text{ km s}^{-1}$$

Smallest Dwarfs:

$$M_{\text{vir}} \approx 10^9 M_\odot \\ V_{\text{vir}} \approx 10 \text{ km s}^{-1}$$

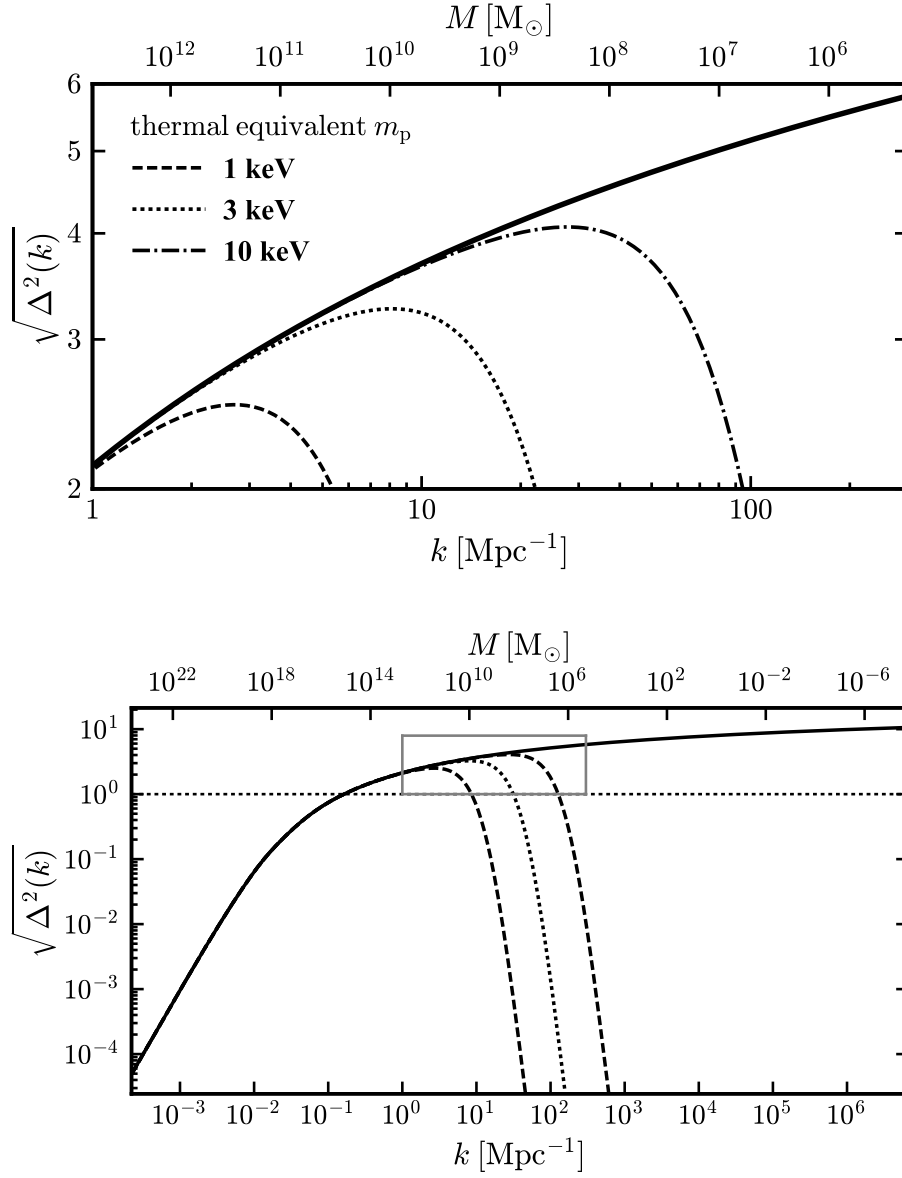


Figure 2

The Λ CDM dimensionless power spectrum (solid lines, Equation 1) plotted as a function of linear wavenumber k (bottom axis) and corresponding linear mass M_l (top axis). The bottom panel spans all physical scales relevant for standard CDM models, from the particle horizon to the free-streaming scale for dark matter composed of standard 100 GeV WIMPs on the far right. The top panel zooms in on the scales of interest for this review, marked with a rectangle in the bottom panel. Known dwarf galaxies are consistent with occupying a relatively narrow 2 decade range of this parameter space – $10^9 - 10^{11} M_\odot$ – even though dwarf galaxies span approximately 7 decades in stellar mass. The effect of WDM models on the power spectrum is illustrated by the dashed, dotted, and dash-dotted lines, which map to the (thermal) WDM particle masses listed. See Section 3.2.1 for a discussion of power suppression in WDM.

With this in mind, Equations 3-6 allow us to self-consistently define “small scales” for both the linear power spectrum and collapsed objects: $M \lesssim 10^{11} M_{\odot}$. As we will discuss, potential problems associated with galaxies inhabiting halos with $V_{\text{vir}} \simeq 50 \text{ km s}^{-1}$ may point to a power spectrum that is non-CDM-like at scales $r_l \lesssim 1 \text{ Mpc}$.

WE DEFINE “SMALL SCALES” AS THOSE SMALLER THAN:

$$M \approx 10^{11} M_{\odot} \leftrightarrow k \approx 3 \text{ Mpc}^{-1} \leftrightarrow r_l \approx 1 \text{ Mpc} \leftrightarrow R_{\text{vir}} \approx 150 \text{ kpc} \leftrightarrow V_{\text{vir}} \approx 50 \text{ km s}^{-1}.$$

As alluded to above, a common point of confusion is that the halo mass definition is subject to the assumed value of Δ , which can vary by a factor of ~ 3 depending on the author. For the spherical collapse definition, $\Delta \simeq 333$ at $z = 0$ (for our fiducial cosmology) and asymptotes to $\Delta = 178$ at high redshift (Bryan & Norman 1998). Another common choice is a fixed $\Delta = 200$ at all z (often labeled M_{200m} in the literature). Finally, some authors prefer to define the virial overdensity as 200 times the critical density, which, according to Equation 4 would mean $\Delta(z) = 200\rho_c(z)/\rho_m(z)$. Such a mass is commonly labeled “ M_{200} ” in the literature. For most purposes (e.g., counting halos), the precise choice does not matter, as long as one is consistent with the definition of halo mass throughout an analysis: every halo has the same center, but its outer radius (and mass contained within that radius) shifts depending on the definition. In what follows, we use the spherical collapse definition ($\Delta = 333$ at $z = 0$) and adhere to the convention of labeling that mass “ M_{vir} ”.

Before moving on, we note that it is also possible (and perhaps even preferable) to give a halo a “mass” label that is directly tied to a physical feature associated with a collapsed dark matter object rather than simply adopting a Δ . More, Diemer & Kravtsov (2015) have advocated the use of a “splash-back” radius, where the density profile shows a sharp break (this typically occurs at $\sim 2R_{\text{vir}}$). Another common choice is to tag halos based not on a mass but on V_{max} , which is the peak value of the circular velocity $V_c(r) = \sqrt{GM(<r)}/r$ as one steps out from the halo center. For any individual halo, the value of V_{max} ($\gtrsim V_{\text{vir}}$) is linked to the internal mass profile or density profile of the system, which is the subject of the next subsection. As discussed below, the ratio $V_{\text{max}}/V_{\text{vir}}$ increases as the halo mass decreases.

1.3.2. Abundance. In principle, the mapping between the initial spectrum of density fluctuations at $z \rightarrow \infty$ and the mass spectrum of collapsed (virialized) dark matter halos at later times could be extremely complicated: as a given scale becomes non-linear, it could affect the collapse of nearby regions or larger scales. In practice, however, the mass spectrum of dark matter halos can be modeled remarkably well with a few simple assumptions. The first of these was taken by Press & Schechter (1974), who assumed that the mass spectrum of collapsed objects could be calculated by extrapolating the overdensity field using linear theory even into the highly non-linear regime and using a spherical collapse model (Gunn & Gott 1972). In the Press-Schechter model, the dark matter halo mass function – the abundance of dark matter halos per unit mass per unit volume at redshift z , often written as $n(M, z)$ – depends only on the rms amplitude of the linear dark matter power spectrum, smoothed using a spherical tophat filter in real space and extrapolated to redshift z using linear theory. Subsequent work has put this formalism on more rigorous mathematical footing (Bond

ROBUST PREDICTIONS FROM CDM-ONLY SIMULATIONS

A defining characteristic of CDM-based hierarchical structure formation is that the smallest scales collapse first – a fact that arises directly from the shape of the power spectrum (Figure 1) and that lies at the heart of many robust predictions for the counts and structure of dark matter halos today. As discussed below, baryonic processes can alter these predictions to various degrees, but pure dark matter simulations have provided a well-defined set of basic predictions used to benchmark the theory.

The dark matter profiles of individual halos are cuspy and dense [Figure 3]

The density profiles of individual Λ CDM halos increase steadily towards small radii, with an overall normalization and detailed shape that reflects the halo’s mass assembly. At fixed mass, early-forming halos tend to be denser than later-forming halos. As with the mass function, both the shape *and* normalization of dark matter halo density structure is predicted by Λ CDM, with a well-quantified prediction for the scatter in halo concentration at fixed mass.

There are many more small halos than large ones [Figure 4]

The comoving number density of dark matter halos rises steeply towards small masses, $dn/dM \propto M^\alpha$ with $\alpha \simeq -1.9$. At large halo masses, counts fall off exponentially above the mass scale that is just going non-linear today. Importantly, both the shape and normalization of the mass function is robustly predicted by the theory.

Substructure is abundant and almost self-similar [Figure 5]

Dark matter halos are filled with substructure, with a mass function that rises as $dN/dm \propto m^{\alpha_s}$ with $\alpha_s \simeq -1.8$ down to the low-mass free-streaming scale ($m \ll 1M_\odot$ for canonical models). Substructure reflects the high-density cores of smaller merging halos that survive the hierarchical assembly process. Substructure counts are nearly self-similar with host mass, with the most massive subhalos seen at $m_{\max} \sim 0.2M_{\text{host}}$.

et al. 1991; Cole 1991; Sheth, Mo & Tormen 2001), and this extended Press-Schechter (EPS) theory yields abundances of dark matter halos that are perhaps surprisingly accurate (see Zentner 2007 for a comprehensive review of EPS theory). This accuracy is tested through comparisons with large-scale numerical simulations.

Simulations and EPS theory both find a universal form for $n(M, z)$: the comoving number density of dark matter halos is a power law with log slope of $\alpha \simeq -1.9$ for $M \ll M^*$ and is exponentially suppressed for $M \gg M^*$, where $M^* = M^*(z)$ is the characteristic mass of fluctuations going non-linear at the redshift z of interest². Importantly, given an initial power spectrum of density fluctuations, it is possible to make highly accurate predictions within Λ CDM for the abundance, clustering, and merger rates of dark matter halos at any cosmic epoch.

1.3.3. Internal structure. Dubinski & Carlberg (1991) were the first to use N -body simu-

²The black line in Figure 4 illustrates the mass function of Λ CDM dark matter halos.

lations to show that the internal structure of a CDM dark matter halo does not follow a simple power-law, but rather bends from a steep outer profile to a mild inner cusp obeying $\rho(r) \sim 1/r$ at small radii. More than twenty years later, simulations have progressed to the point that we now have a fairly robust understanding of the structure of Λ CDM halos and the important factors that govern halo-to-halo variance (e.g., Navarro et al. 2010; Diemer & Kravtsov 2015; Klypin et al. 2016), at least for dark-matter-only simulations.

To first approximation, dark matter halo profiles can be described by a nearly universal form over all masses, with a steep fall-off at large radii transitioning to mildly divergent cusp towards the center. A common way to characterize this is via the NFW functional form (Navarro, Frenk & White 1997), which provides a good (but not perfect) description dark matter profiles:

$$\rho(r) = \frac{4\rho_{-2}}{(r/r_{-2})(1 + r/r_{-2})^2}. \quad (7)$$

Here, r_{-2} is a characteristic radius where the log-slope of the density profile is -2 , marking a transition point from the inner $1/r$ cusp to an outer $1/r^3$ profile. The second parameter, ρ_{-2} , sets the value of $\rho(r)$ at $r = r_{-2}$. In practice, dark matter halos are better described the three-parameter Einasto (1965) profile (Navarro et al. 2004; Gao et al. 2008). However, for the small halos of most concern for this review, NFW fits do almost as well as Einasto in describing the density profiles of halos in simulations (Dutton & Macciò 2014). Given that the NFW form is slightly simpler, we have opted to adopt this approximation for illustrative purposes in this review.

As Equation 7 makes clear, two parameters (e.g., ρ_{-2} and r_{-2}) are required to determine a halo’s NFW density profile. For a fixed halo mass M_{vir} (which fixes R_{vir}), the second parameter is often expressed as the halo concentration: $c = R_{\text{vir}}/r_{-2}$. Together, a $M_{\text{vir}} - c$ combination completely specifies the profile. In the median, and over the mass and redshift regime of interest to this review, halo concentrations increase with decreasing mass and redshift: $c \propto M_{\text{vir}}^{-a} (1+z)^{-1}$, with $a \simeq 0.1$ (Bullock et al. 2001). Though halo concentration correlates with halo mass, there is significant scatter (~ 0.1 dex) about the median at fixed M_{vir} (Jing 2000; Bullock et al. 2001). Some fraction of this scatter is driven by the variation in halo mass accretion history (Wechsler et al. 2002; Ludlow et al. 2016), with early-forming halos having higher concentrations at fixed final virial mass.

The dependence of halo profile on a mass-dependent concentration parameter and the correlation between formation time and concentration at fixed virial mass are caused by the hierarchical build-up of halos in Λ CDM: low-mass halos assemble earlier, when the mean density of the Universe is higher, and therefore have higher concentrations than high-mass halos (e.g., Navarro, Frenk & White 1997; Wechsler et al. 2002). At the very smallest masses, the concentration-mass relation likely flattens, reflecting the shape of the dimensionless power spectrum (see our Figure 1 and the discussion in Ludlow et al. 2016); at the highest masses and redshifts, characteristic of very rare peaks, the trend seems to reverse ($a < 0$; Klypin et al. 2016).

The right panel of Figure 3 summarizes the median NFW density profiles for $z = 0$ halos with masses that span those of large galaxy clusters ($M_{\text{vir}} = 10^{15} M_{\odot}$) to those of the smallest dwarf galaxies ($M_{\text{vir}} = 10^8 M_{\odot}$). We assume the $c - M_{\text{vir}}$ relation from Klypin et al. (2016). These profiles are plotted in physical units (unscaled to the virial radius) in order to emphasize that higher mass halos are denser at every radius than lower mass halos (at least in the median). However, at a fixed small fraction of the virial radius, small halos are slightly denser than larger ones. This is a result of the concentration-mass relation. Under

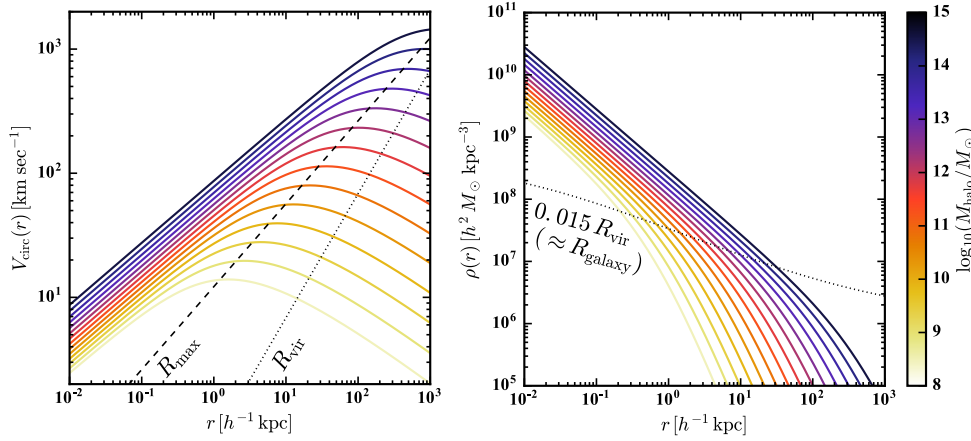


Figure 3

Right: The density profiles of median NFW dark matter halos at $z = 0$ with masses that span galaxy clusters ($M_{\text{vir}} = 10^{15} M_{\odot}$, black) to the approximate HI cooling threshold that is expected to correspond to the smallest dwarf galaxies ($M_{\text{vir}} \approx 10^8 M_{\odot}$, yellow). The lines are color coded by halo virial mass according to the bar on the right and are separated in mass by 0.5 dex. We see that (in the median) massive halos are denser than low-mass halos at a fixed *physical* radius. However, at a fixed small fraction of the virial radius, smaller halos are typically slightly denser than larger halos, reflecting the concentration-mass relation. This is demonstrated by the dotted line which connects $\rho(r)$ evaluated at $r = \epsilon R_{\text{vir}}$ for halos over a range of masses. We have chosen $\epsilon = 0.015$ because this value provides a good match to observed galaxy half-light radii over a wide range of galaxy luminosities under the assumption that galaxies occupy halos according to abundance matching (see Section 1.5 and Figure 6). Interestingly, the characteristic dark matter density at this ‘galaxy radius’ increases only by a factor of ~ 6 over almost seven orders of magnitude in halo virial mass. *Left:* The equivalent circular velocity curves $V_c(r) \equiv \sqrt{GM(<r)}/r$ for the same halos plotted on the right. The dashed line connects the radius R_{max} where the circular velocity is maximum (V_{max}) for each rotation curve. The dotted line tracks the $R_{\text{max}} - V_{\text{vir}}$ relation. The ratio $R_{\text{max}}/R_{\text{vir}}$ decreases towards smaller halos, reflecting the mass-concentration relation. The ratio $V_{\text{max}}/V_{\text{vir}}$ also increases with decreasing concentration.

the ansatz of abundance matching (Section 1.5, Figure 6), galaxy sizes (half-mass radii) track a fixed fraction of their host halo virial radius: $r_{\text{gal}} \simeq 0.015 R_{\text{vir}}$ (Kravtsov 2013). This relation is plotted as a dotted line such that the dotted line intersects each solid line at that $r = 0.015 R_{\text{vir}}$, where R_{vir} is that particular halo’s virial radius. We see that small halos are slightly denser at the typical radii of the galaxies they host than are larger halos. Interestingly, however, the density range is remarkably small, with a local density of dark matter increasing by only a factor of ~ 6 over the full mass range of halos that are expected to host galaxies, from the smallest dwarfs to the largest cD galaxies in the universe.

On the left we show the same halos, now presented in terms of the implied circular velocity curves: $V_c \equiv \sqrt{GM(<r)}/r$. The dotted line in left panel intersects V_{vir} at R_{vir} for each value of M_{vir} . The dashed line does the same for V_{max} and its corresponding radius R_{max} . Higher mass systems, with lower concentrations, typically have $V_{\text{max}} \simeq V_{\text{vir}}$, but for smaller halos the ratio is noticeably different than one and can be as large as ~ 1.5 for high-concentration outliers. Note also that the lowest mass halos have $R_{\text{max}} \ll R_{\text{vir}}$ and thus it is the value of V_{max} (rather than V_{vir}) that is more closely linked to the observable

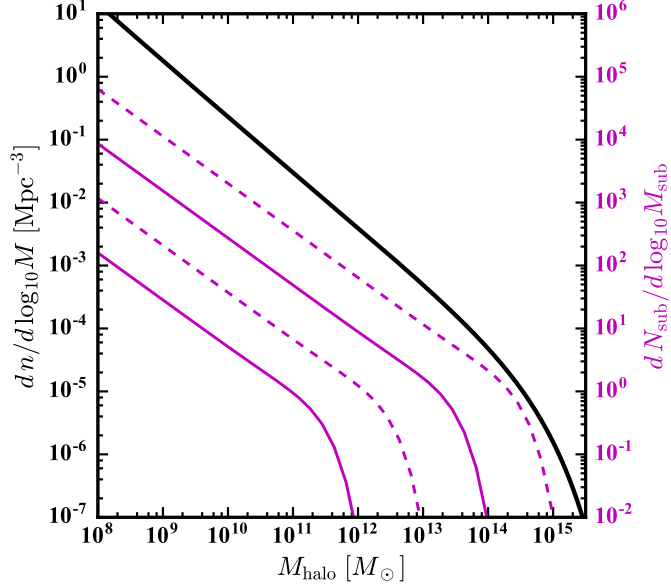


Figure 4

Steep mass functions. The black solid line shows the $z = 0$ dark matter halo mass function ($M_{\text{halo}} = M_{\text{vir}}$) for the full population of halos in the universe as approximated by [Sheth, Mo & Tormen \(2001\)](#). For comparison, the magenta lines show the subhalo mass functions at $z = 0$ (defined as $M_{\text{halo}} = M_{\text{sub}} = M_{\text{peak}}$, see text) at the same redshift for host halos at four characteristic masses ($M_{\text{vir}} = 10^{12}, 10^{13}, 10^{14}$, and $10^{15} M_{\odot}$) with units given along the right-hand axis. Note that the subhalo mass functions are almost self-similar with host mass, roughly shifting to the right by $10\times$ for every decade increase in host mass. The low-mass slope of subhalo mass function is similar than the field halo mass function. Both field and subhalo mass functions are expected to rise steadily to the cutoff scale of the power spectrum, which for fiducial CDM scenarios is $\ll 1 M_{\odot}$.

“flat” region of a galaxy rotation curve. For our “small-scale” mass of $M_{\text{vir}} = 10^{11} M_{\odot}$, typically $V_{\text{max}} \simeq 1.2 V_{\text{vir}} \simeq 60 \text{ km s}^{-1}$.

1.4. Dark matter substructure

It was only just before the turn of the century that N -body simulations set within a cosmological CDM framework were able to robustly resolve the substructure *within* individual dark matter halos ([Ghigna et al. 1998](#); [Klypin et al. 1999a](#)). It soon became clear that the dense centers of small halos are able to survive the hierarchical merging process: dark matter halos should be filled with substructure. Indeed, subhalo counts are nearly self-similar with host halo mass. This was seen as welcome news for cluster-mass halos, as the substructure could be easily identified with cluster galaxies. However, as we will discuss in the next section, the fact that Milky-Way-size halos are filled with substructure is less clearly consistent with what we see around the Galaxy.

Quantifying subhalo counts, however, is not so straightforward. Counting by mass is tricky because the definition of “mass” for an extended distribution orbiting within a

collapsed halo is even more fraught with subjective decisions than virial mass. When a small halo is accreted into a large one, mass is preferentially stripped from the outside. Typically, the standard virial overdensity “edge” is subsumed by the ambient host halo. One option is to compute the mass that is bound to the subhalo, but even these masses vary from halo finder to halo finder. The value of a subhalo’s V_{max} is better defined, and often serves as a good tag for quantifying halos.

Another option is to tag bound subhalos using the maximum virial mass that the halos had at the time they were first accreted³ onto a host, M_{peak} . This is a useful option because stars in a central galaxy belonging to a halo at accretion will be more tightly bound than the dark matter. The resultant satellite’s stellar mass is most certainly more closely related to M_{peak} than the bound dark matter mass that remains at $z = 0$. Moreover, the subsequent mass loss (and even V_{max} evolution) could change depending on the baryonic content of the *host* because of tidal heating and other dynamical effects (D’Onghia et al. 2010). For these reasons, we adopt $M_{\text{sub}} = M_{\text{peak}}$ for illustrative purposes here.

The magenta lines in Figure 4 show the median subhalo mass functions ($M_{\text{sub}} = M_{\text{peak}}$) for four characteristic host halo masses ($M_{\text{vir}} = 10^{12-14} M_{\odot}$) according to the results of Rodríguez-Puebla et al. (2016). These lines are normalized to the right-hand vertical axis. Subhalos are counted only if they exist within the virial radius of the host, which means the counting volume increases as $\propto M_{\text{vir}} \propto R_{\text{vir}}^3$ for these four lines. For comparison, the black line (normalized to the left vertical axis) shows the global halo mass function (as estimated via the fitting function from Sheth, Mo & Tormen 2001). The subhalo mass function rises with a similar (though slightly shallower) slope as the field halo mass function and is also roughly self-similar in host halo mass.

1.5. Linking dark matter halos to galaxies

How do we associate dark matter halos with galaxies? One simple approximation is to assume that each halo is allotted its cosmic share of baryons $f_b = \Omega_b/\Omega_m \approx 0.15$ and that those baryons are converted to stars with some constant efficiency ϵ_* : $M_* = \epsilon_* f_b M_{\text{vir}}$. Unfortunately, as shown in Figure 5, this simple approximation fails miserably. Galaxy stellar masses do not scale linearly with halo mass; the relationship is much more complicated. Indeed, the goal of forward modeling galaxy formation from known physics within the Λ CDM framework is an entire field of its own (galaxy evolution; Somerville & Davé 2015). Though galaxy formation theory has progressed significantly in the last several decades, many problems remain unsolved.

Other than forward modeling galaxy formation, there are two common approaches that give an independent assessment of how galaxies relate to dark matter halos. The first involves matching the observed volume density of galaxies of a given stellar mass (or other observable such as luminosity, velocity width, or baryon mass) to the predicted abundance of halos of a given virial mass. The second way is to measure the mass of the galaxy directly and to infer the dark matter halo properties based on this dynamical estimator.

1.5.1. Abundance matching. As illustrated in Figure 5, the predicted mass function of collapsed dark matter halos has a considerably different normalization and shape than the

³This maximum mass is similar to the virial mass at the time of accretion, though infalling halos can begin losing mass prior to first crossing the virial radius.

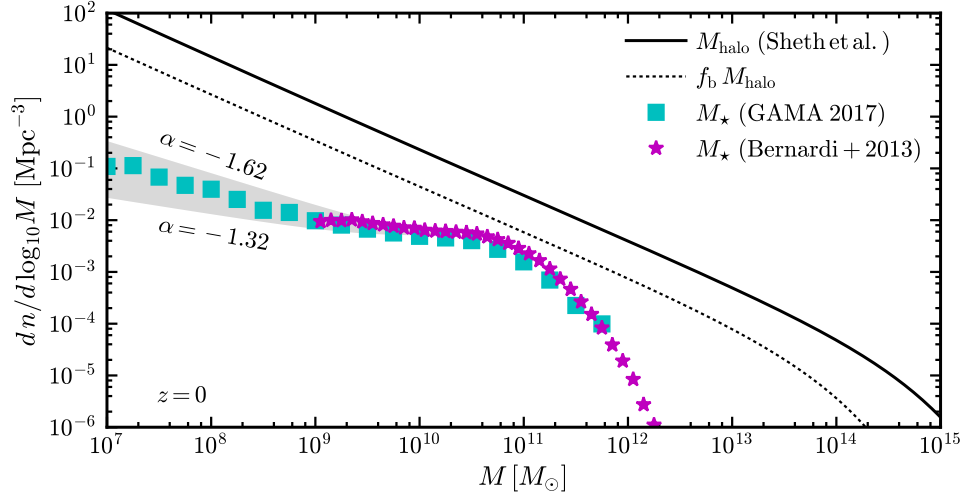


Figure 5

The thick black line shows the global dark matter mass function. The dotted line is shifted to the left by the cosmic baryon fraction for each halo $M_{\text{vir}} \rightarrow f_b M_{\text{vir}}$. This is compared to the observed stellar mass function of galaxies from [Bernardi et al. \(2013\)](#), magenta stars) and [Wright et al. \(2017\)](#), cyan squares). The shaded bands demonstrate a range of faint-end slopes $\alpha_g = -1.62$ to -1.32 . This range of power laws will produce dramatic differences at the scales of the classical Milky Way satellites ($M_* \simeq 10^{5-7} M_\odot$). Pushing large sky surveys down below $10^6 M_\odot$ in stellar mass, where the differences between the power law range shown would exceed a factor of ten, would provide a powerful constraint on our understanding of the low-mass behavior. Until then, this mass regime can only be explored without large completeness corrections in vicinity of the Milky Way.

observed stellar mass function of galaxies. The difference grows dramatically at both large and small masses, with a maximum efficiency of $\epsilon_* \simeq 0.2$ at the stellar mass scale of the Milky Way ($M_* \approx 10^{10.75} M_\odot$). This basic mismatch in shape has been understood since the earliest galaxy formation models set within the dark matter paradigm ([White & Rees 1978](#)) and is generally recognized as one of the primary constraints on feedback-regulated galaxy formation ([White & Frenk 1991](#); [Benson et al. 2003](#); [Somerville & Davé 2015](#)).

At the small masses that most concern this review, dark matter halo counts follow $dn/dM \propto M^\alpha$ with a steep slope $\alpha_{dm} \simeq -1.9$ compared to the observed stellar mass function slope of $\alpha_g = -1.47$ ([Baldry et al. 2012](#), which is consistent with the updated GAMA results shown in Figure 5). Current surveys that cover enough sky to provide a global field stellar mass function reach a completeness limit of $M_* \approx 10^{7.5} M_\odot$. At this mass, galaxy counts are more than two orders of magnitude below the naive baryonic mass function $f_b M_{\text{vir}}$. The shaded band illustrates how the stellar mass function would extrapolate to the faint regime spanning a range of faint-end slopes α that are marginally consistent with observations at the completeness limit.

One clear implication of this comparison is that galaxy formation efficiency (ϵ_*) must vary in a non-linear way as a function of M_{vir} (at least if Λ CDM is the correct underlying model). Perhaps the cleanest way to illustrate this is adopt the simple assumption of Abundance Matching (AM): that galaxies and dark matter halos are related in a one-to-

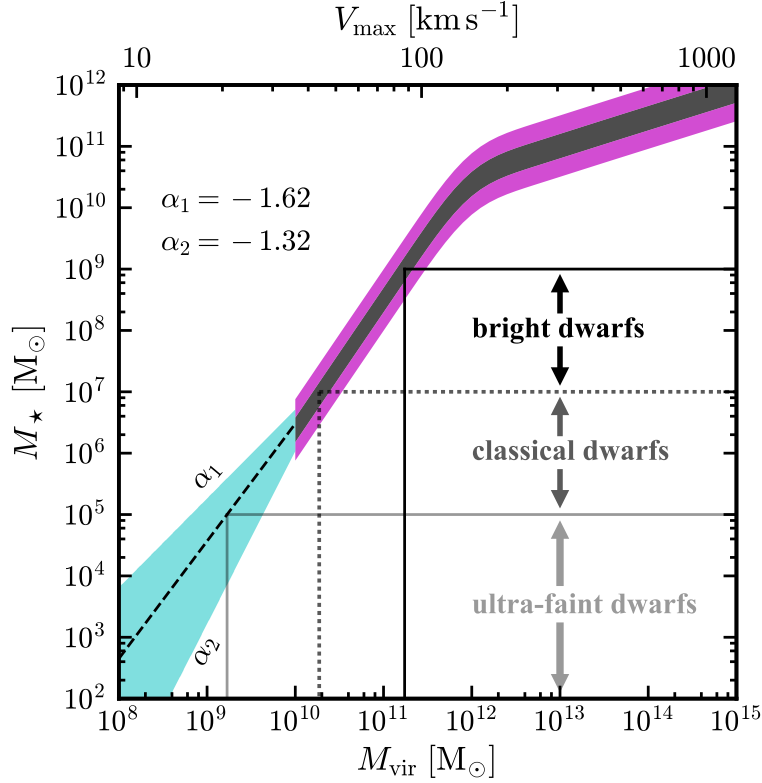


Figure 6

Abundance matching relation from Behroozi et al. (in preparation). Gray (magenta) shows a scatter of 0.2 (0.5) dex about the median relation. The dashed line is power-law extrapolation below the regime where large sky surveys are currently complete. The cyan band shows how the extrapolation would change as the faint-end slope of the galaxy stellar mass function (α) is varied over the same range illustrated by the shaded gray band in Figure 5. Note that the enumeration of $M_* = 10^5 M_\odot$ galaxies could provide a strong discriminator on faint-end slope, as the ± 0.15 range in α shown maps to an order of magnitude difference in the halo mass associated with this galaxy stellar mass and a corresponding factor of ~ 10 shift in the galaxy/halo counts shown in Figure 4.

one way, with the most massive galaxies inhabiting the most massive dark matter halos (Frenk et al. 1988; Kravtsov et al. 2004; Conroy, Wechsler & Kravtsov 2006; Moster et al. 2010; Behroozi, Wechsler & Conroy 2013). The results of such an exercise are presented in Figure 6 (as derived by Behroozi et al., in preparation). The gray band shows the median $M_* - M_{\text{vir}}$ relation with an assumed 0.2 dex of scatter in M_* at fixed M_{vir} . The magenta band expands the scatter to 0.5 dex. This relation is truncated near the completeness limit in Baldry et al. (2012). The central dashed line in Figure 6 shows the median relation that comes from extrapolating the Baldry et al. (2012) mass function with their best-fit $\alpha_g = -1.47$ down to the stellar mass regime of Local Group dwarfs. The cyan band brackets the range for the two other faint-end slopes shown in Figure 5: $\alpha_g = -1.62$ and -1.32 .

Bright Dwarfs:

$$M_\star \approx 10^8 M_\odot$$
$$M_{\text{vir}} \approx 10^{11} M_\odot$$
$$M_\star/M_{\text{vir}} \approx 10^{-3}$$

Classical Dwarfs:

$$M_\star \approx 10^6 M_\odot$$
$$M_{\text{vir}} \approx 10^{10} M_\odot$$
$$M_\star/M_{\text{vir}} \approx 10^{-4}$$

Ultra-faint Dwarfs:

$$M_\star \approx 10^4 M_\odot$$
$$M_{\text{vir}} \approx 10^9 M_\odot$$
$$M_\star/M_{\text{vir}} \approx 10^{-5}$$

Figure 6 allows us to read off the virial mass expectations for galaxies of various sizes. We see that Bright Dwarfs at the limit of detection in large sky surveys ($M_\star \approx 10^8 M_\odot$) are naively associated with $M_{\text{vir}} \approx 10^{11} M_\odot$ halos. Galaxies with stellar masses similar to the Classical Dwarfs at $M_\star \approx 10^6 M_\odot$ are associated with $M_{\text{vir}} \approx 10^{10} M_\odot$ halos. As we will discuss in Section 3, galaxies at this scale with $M_\star/M_{\text{vir}} \approx 10^{-4}$ are at the critical scale where feedback from star formation may not be energetic enough to alter halo density profiles significantly. Finally, Ultra-faint Dwarfs with $M_\star \approx 10^4 M_\odot$, $M_{\text{vir}} \approx 10^9 M_\odot$, and $M_\star/M_{\text{vir}} \approx 10^{-5}$ likely sit at the low-mass extreme of galaxy formation.

1.5.2. Kinematic Measures. An alternative way to connect to the dark matter halo hosting a galaxy is to determine the galaxy’s dark matter mass kinematically. This, of course, can only be done within a central radius probed by the baryons. For the small galaxies of concern for this review, extended mass measurements via weak lensing or hot gas emission is infeasible. Instead, masses (or mass profiles) must be inferred within some inner radius, defined either by the stellar extent of the system for dSphs and/or the outer rotation curves for rotationally-supported gas disks.

Bright dwarfs, especially those in the field, often have gas disks with ordered kinematics. If the gas extends far enough out, rotation curves can be extracted that extend as far as the flat part of the galaxy rotation curve V_{flat} . If care is taken to account for non-trivial velocity dispersions in the mass extraction (e.g., Kuzio de Naray, McGaugh & de Blok 2008), then we can associate $V_{\text{flat}} \approx V_{\text{max}}$.

Owing to the difficulty in detecting them, the faintest galaxies known are all satellites of the Milky Way or M31 and are dSphs. These lack rotating gas components, so rotation curve measurements are impossible. Instead, dSphs are primarily stellar dispersion-supported systems, with masses that are best probed by velocity dispersion measurements obtained star-by-star for the closest dwarfs (e.g., Walker et al. 2009; Simon et al. 2011; Kirby et al. 2014). For systems of this kind, the mass can be measured accurately within the stellar half-light radius (Walker et al. 2009). The mass within the de-projected (3D) half-light radius ($r_{1/2}$) is relatively robust to uncertainties in the stellar velocity anisotropy and is given by $M(< r_{1/2}) = 3 \sigma_\star^2 r_{1/2} / G$, where σ_\star is the measured, luminosity-weighted line-of-sight velocity dispersion (Wolf et al. 2010). This formula is equivalent to saying that the circular velocity at the half-light radius is $V_{1/2} = V_c(r_{1/2}) = \sqrt{3} \sigma_\star$. The value of $V_{1/2}$ ($\leq V_{\text{max}}$) provides a one-point measurement of the host halo’s rotation curve at $r = r_{1/2}$.

1.6. Connections to particle physics

Although the idea of “dark” matter had been around since at least Zwicky (1933), it was not until rotation curve measurements of galaxies in the 1970s revealed the need for significant amounts of non-luminous matter (Freeman 1970; Rubin, Thonnard & Ford 1978; Bosma 1978; Rubin, Ford & Thonnard 1980) that dark matter was taken seriously by the broader astronomical community (and shortly thereafter, it was recognized that dwarf galaxies might serve as sensitive probes of dark matter; Aaronson 1983; Faber & Lin 1983; Lin & Faber 1983). Very quickly, particle physicists realized the potential implications for their discipline as well. Dark matter candidates were grouped into categories based on their effects on structure formation. “Hot” dark matter (HDM) particles remain relativistic until relatively late in the Universe’s evolution and smooth out perturbations even on super-galactic scales; “warm” dark matter (WDM) particles have smaller initial velocities,

become non-relativistic earlier, and suppress perturbations on galactic scales (and smaller); and CDM has negligible thermal velocity and does not suppress structure formation on any scale relevant for galaxy formation. Standard Model neutrinos were initially an attractive (hot) dark matter candidate; by the mid-1980s, however, this possibility had been excluded on the basis of general phase-space arguments (Tremaine & Gunn 1979), the large-scale distribution of galaxies (White, Frenk & Davis 1983), and properties of dwarf galaxies (Lin & Faber 1983). The lack of a suitable Standard Model candidate for particle dark matter has led to significant work on particle physics extensions of the Standard Model. From a cosmology and galaxy formation perspective, the unknown particle nature of dark matter means that cosmologists must make assumptions about dark matter's origins and particle physics properties and then investigate the resulting cosmological implications.

A general class of models that are appealing in their simplicity is that of *thermal relics*. Production and destruction of dark matter particles are in equilibrium so long as the temperature of the Universe kT is larger than the mass of the dark matter particle $m_{\text{DM}}c^2$. At lower temperatures, the abundance is exponentially suppressed, as destruction (via annihilation) dominates over production. At some point, the interaction rate of dark matter particles drops below the Hubble rate, however, and the dark matter particles “freeze out” at a fixed number density (see, e.g., Kolb & Turner 1994; this is also known as chemical decoupling). Amazingly, if the annihilation cross section is typical of weak-scale physics, the resulting freeze-out density of thermal relics with $m \sim 100$ GeV is approximately equal to the observed density of dark matter today (e.g., Jungman, Kamionkowski & Griest 1996). This subset of thermal relics is referred to as *weakly-interacting massive particles (WIMPs)*. The observation that new physics at the weak scale naturally leads to the correct abundance of dark matter in the form of WIMPs is known as the “WIMP miracle” (Feng & Kumar 2008) and has been the basic framework for dark matter over the past 30 years.

WIMPs are not the only viable dark matter candidate, however, and it is important to note that the WIMP miracle could be a red herring. *Axions*, which are particles invoked to explain the strong CP problem of quantum chromodynamics (QCD), and right-handed neutrinos (often called *sterile neutrinos*), which are a minimal extension to the Standard Model of particle physics that can explain the observed baryon asymmetry and why neutrino masses are so small compared to other fermions, are two other hypothetical particles that may be dark matter (among a veritable zoo of additional possibilities; see Feng 2010 for a recent review). While WIMPs, axions, and sterile neutrinos are capable of producing the observed abundance of dark matter in the present-day Universe, they can have very different effects on the mass spectrum of cosmological perturbations.

While the cosmological perturbation spectrum is initially set by physics in the very early universe (inflation in the standard scenario), the microphysics of dark matter affects the evolution of those fluctuations at later times. In the standard WIMP paradigm, the low-mass end of the CDM hierarchy is set by first collisional damping (subsequent to chemical decoupling but prior to kinetic decoupling of the WIMPs), followed by free-streaming (e.g., Hofmann, Schwarz & Stöcker 2001; Bertschinger 2006). For typical 100 GeV WIMP candidates, these processes erase cosmological perturbations with $M \lesssim 10^{-6} M_{\odot}$ (i.e., Earth mass; Green, Hofmann & Schwarz 2004). Free-streaming also sets the low-mass end of the mass spectrum in models where sterile neutrinos decouple from the plasma while relativistic. In this case, the free-streaming scale can be approximated by the (comoving) size of the horizon when the sterile neutrinos become non-relativistic. The comoving horizon size at $z = 10^7$, corresponding to $m \approx 2.5$ keV, is approximately 50 kpc, which is significantly

Cold Dark Matter (CDM):

$$m \sim 100 \text{ GeV}, \\ v_{\text{th}}^{z=0} \approx 0 \text{ km s}^{-1}$$

Warm Dark Matter (WDM):

$$m \sim 1 \text{ keV}, \\ v_{\text{th}}^{z=0} \sim 0.03 \text{ km s}^{-1}$$

Hot Dark Matter (HDM):

$$m \sim 1 \text{ eV}, \\ v_{\text{th}}^{z=0} \sim 30 \text{ km s}^{-1}$$

smaller than the scale derived above for L^* galaxies. keV-scale sterile neutrinos are therefore observationally-viable dark matter candidates (see [Adhikari et al. 2016](#) for a recent, comprehensive review). QCD axions are typically $\sim \mu\text{eV}$ -scale particles but are produced out of thermal equilibrium ([Kawasaki & Nakayama 2013](#)). Their free-streaming scale is significantly smaller than that of a typical WIMP (see Section 3.2.1).

The previous paragraphs have focused on the effects of collisionless damping and free-streaming – direct consequences of the particle nature of dark matter – in the linear regime of structure formation. Dark matter microphysics can also affect the non-linear regime of structure formation. In particular, dark matter self-interactions – scattering between two dark matter particles – will affect the phase space distribution of dark matter. Within observational constraints, dark matter self-interactions could be relevant in the dense centers of dark matter halos. By transferring kinetic energy from high-velocity particles to low-velocity particles, scattering transfers “heat” to the centers of dark matter halos, reducing their central densities and making their velocity distributions nearly isothermal. This would have a direct effect on galaxy formation, as galaxies form within the centers of dark matter halos and the motions of their stars and gas trace the central gravitational potential. These effects are discussed further in Section 3.2.2.

The particle nature of dark matter is therefore reflected in the cosmological perturbation spectrum, in the abundance of collapsed dark matter structures as a function of mass, and in the density and velocity distribution of dark matter in virialized dark matter halos.

2. OVERVIEW OF PROBLEMS

The CDM paradigm as summarized in the previous section emerged among other dark matter variants in the early 1980s ([Peebles 1982](#); [Blumenthal et al. 1984](#); [Davis et al. 1985](#)) with model parameters gradually settling to their current precise state (including Λ) in the wake of overwhelming evidence from large-scale galaxy clustering, supernovae measurements of cosmic acceleration, and cosmic microwave background studies, among other data. The 1990s saw the first N -body simulations to resolve the internal structure of CDM halos on small scales. Almost immediately researchers pinpointed the two most well-known challenges to the theory: the cusp-core problem ([Flores & Primack 1994](#); [Moore 1994](#)) and the missing satellites problem ([Klypin et al. 1999b](#); [Moore et al. 1999](#)). This section discusses these two classic issues from a current perspective goes on to describe a third problem, too-big-to-fail ([Boylan-Kolchin, Bullock & Kaplinghat 2011](#)), which in some sense is a confluence of the first two. Finally, we conclude this section with a more limited discussion of two other challenges faced by ΛCDM on small scales: the apparent planar distributions seen for Local Group satellites and the dynamical scaling relations seen in galaxy populations.

2.1. Missing Satellites

The highest-resolution cosmological simulations of MW-size halos in the ΛCDM paradigm have demonstrated that dark matter (DM) clumps exist at all resolved masses, with no break in the subhalo mass function down to the numerical convergence limit (e.g., [Springel et al. 2008](#); [Kuhlen, Madau & Silk 2009](#); [Stadel et al. 2009](#); [Garrison-Kimmel et al. 2014](#); [Griffen et al. 2016](#)). We expect thousands of subhalos with masses that are (in principle) large enough to have supported molecular cooling ($M_{\text{peak}} \gtrsim 10^7 M_{\odot}$). Meanwhile, only

THREE CHALLENGES TO BASIC Λ CDM PREDICTIONS

There are three classic problems associated with the small-scale predictions for dark matter in the Λ CDM framework. Other anomalies exist, including some that we discuss in this review, but these three are important because 1) they concern basic predictions about dark matter that are fundamental to the hierarchical nature of the theory; and 2) they have received significant attention in the literature.

Missing Satellites and Dwarfs [Figures 4–8]

The observed stellar mass functions of field galaxies and satellite galaxies in the Local Group is much flatter at low masses than predicted dark matter halo mass functions: $dn/dM_\star \propto M_\star^{\alpha_g}$ with $\alpha_g \simeq -1.5$ (vs. $\alpha \simeq -1.9$ for dark matter). The issue is most acute for Galactic satellites, where completeness issues are less of a concern. There are only ~ 50 known galaxies with $M_\star > 300 M_\odot$ within 300 kpc of the Milky Way compared to as many as ~ 1000 dark subhalos (with $M_{\text{sub}} > 10^7 M_\odot$) that could conceivably host galaxies. One solution to this problem is to posit that galaxy formation becomes increasingly inefficient as the halo mass drops. The smallest dark matter halos have simply failed to form stars altogether.

Low-density Cores vs. High-density Cusps [Figure 9]

The central regions of dark-matter dominated galaxies as inferred from rotation curves tend to be both less dense (in normalization) and less cuspy (in inferred density profile slope) than predicted for standard Λ CDM halos (such as those plotted in Figure 3). An important question is whether baryonic feedback alters the structure of dark matter halos.

Too-Big-to-Fail [Figure 10]

The local universe contains too few galaxies with central densities indicative of $M_{\text{vir}} \simeq 10^{10} M_\odot$ halos. Halos of this mass are generally believed to be too massive to have failed to form stars, so the fact that they are missing is hard to understand. The stellar mass associated with this halo mass scale ($M_\star \simeq 10^6 M_\odot$, Figure 6) may be too small for baryonic processes to alter their halo structure (see Figure 13).

~ 50 satellite galaxies down to $\sim 300 M_\odot$ in stars are known to orbit within the virial radius of the Milky Way (Drlica-Wagner et al. 2015). Even though there is real hope that future surveys could bring the census of ultra-faint dwarf galaxies into the hundreds (Tollerud et al. 2008; Hargis, Willman & Peter 2014), it seems unlikely there are thousands of undiscovered dwarf galaxies to this limit within the virial volume of the Milky Way. The current situation is depicted in Figure 7, which shows the dark matter distribution around a Milky Way size galaxy as predicted by a Λ CDM simulation next to a map of the known galaxies of the Milky Way on the same scale.

Given the discussion of abundance matching in Section 1.5 and the associated Figure 6, it is reasonable to expect that dark matter halos become increasingly inefficient at making galaxies at low masses and at some point go completely dark. Physical mass scales of interest in this regard include the mass below which reionization UV feedback likely suppresses gas accretion $M_{\text{vir}} \approx 10^9 M_\odot$ ($V_{\text{max}} \gtrsim 30 \text{ km s}^{-1}$; e.g., Efstathiou 1992; Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002; Bovill & Ricotti 2009; Sawala et al. 2016) and the minimum mass for atomic cooling in the early Universe, $M_{\text{vir}} \approx 10^8 M_\odot$ ($V_{\text{max}} \gtrsim 15 \text{ km s}^{-1}$;

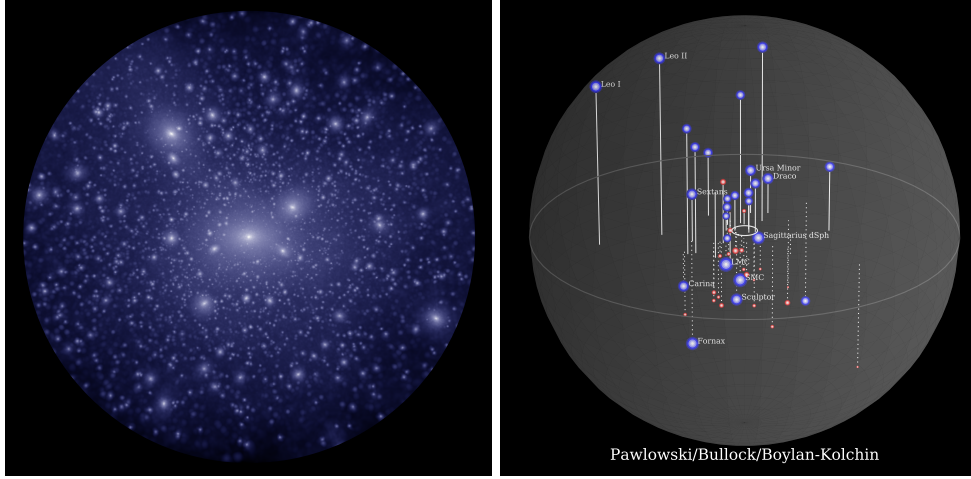


Figure 7

The Missing Satellites Problem: Predicted Λ CDM substructure (left) vs. known Milky Way satellites (right). The image on the left shows the Λ CDM dark matter distribution within a sphere of radius 250 kpc around the center of a Milky-Way size dark matter halo (simulation by V. Robles and T. Kelley in collaboration with the authors). The image on the right (by M. Pawlowski in collaboration with the authors) shows the current census of Milky Way satellite galaxies, with galaxies discovered since 2015 in red. The Galactic disk is represented by a circle of radius 15 kpc at the center and the outer sphere has a radius of 250 kpc. The 11 brightest (classical) Milky Way satellites are labeled by name. Sizes of the symbols are not to scale but are rather proportional to the log of each satellite galaxy’s stellar mass. Currently, there are ~ 50 satellite galaxies of the Milky Way compared to thousands of predicted subhalos with $M_{\text{peak}} \gtrsim 10^7 M_{\odot}$.

see, e.g., [Rees & Ostriker 1977](#)). According to Figure 6, these physical effects are likely to become dominant in the regime of ultra-faint galaxies $M_{\star} \lesssim 10^5 M_{\odot}$.

The question then becomes: can we simply adopt the abundance-matching relation derived from field galaxies to “solve” the Missing Satellites Problem down to the scale of the classical MW satellites (i.e., $M_{\text{vir}} \simeq 10^{10} M_{\odot} \leftrightarrow M_{\star} \simeq 10^6 M_{\odot}$)? Figure 8 (modified from [Garrison-Kimmel et al. 2017a](#)) shows that the answer is likely “yes.” Shown in magenta is the cumulative count of Milky Way satellite galaxies within 300 kpc of the Galaxy plotted down to the stellar mass completeness limit within that volume. The shaded band shows the 68% range predicted stellar mass functions from the dark-matter-only ELVIS simulations ([Garrison-Kimmel et al. 2014](#)) combined with the AM relation shown in Figure 6 with zero scatter. The agreement is not perfect, but there is no over-prediction. The dashed lines show how the predicted satellite stellar mass functions would change for different assumed (field galaxy) faint-end slopes in the calculating the AM relation. An important avenue going forward will be to push these comparisons down to the ultra-faint regime, where strong baryonic feedback effects are expected to begin shutting down galaxy formation altogether.

2.2. Cusp, Cores, and Excess Mass

As discussed in Section 1, Λ CDM simulations that include only dark matter predict that dark matter halos should have density profiles that rise steeply at small radius $\rho(r) \propto r^{-\gamma}$, with $\gamma \simeq 0.8 - 1.4$ over the radii of interest for small galaxies ([Navarro et al. 2010](#)). This is

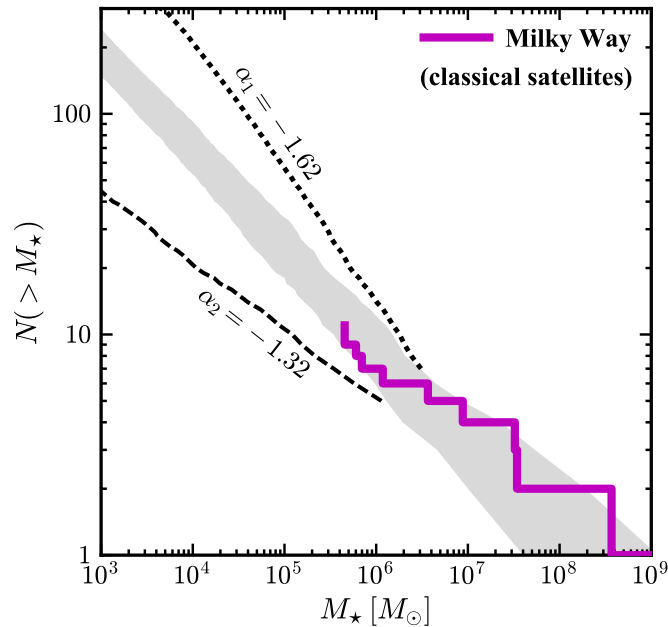


Figure 8

“Solving” the Missing Satellites Problem with abundance matching. The cumulative count of dwarf galaxies around the Milky Way (magenta) plotted down to completeness limits from Garrison-Kimmel et al. (2017a). The gray shaded region shows the predicted stellar mass function from the dark-matter-only ELVIS simulations (Garrison-Kimmel et al. 2014) combined with the fiducial AM relation shown in Figure 6, assuming zero scatter. If the faint end slope of the stellar mass function is shallower (dashed) or steeper (dotted), the predicted abundance of satellites with $M_\star > 10^4 M_\odot$ throughout the Milky Way’s virial volume differs by a factor of 10. Local Group counts can therefore serve as strong constraints on galaxy formation models.

in contrast to many (though not all) low-mass dark-matter-dominated galaxies with well-measured rotation curves, which prefer fits with constant-density cores ($\gamma \approx 0 - 0.5$; e.g., McGaugh, Rubin & de Blok 2001; Marchesini et al. 2002; Simon et al. 2005; de Blok et al. 2008; Kuzio de Naray, McGaugh & de Blok 2008). A related issue is that fiducial Λ CDM simulations predict more dark matter in the central regions of galaxies than is measured for the galaxies that they should host according to AM. This “central density problem” is an issue of normalization and exists independent of the precise slope of the central density profile (Alam, Bullock & Weinberg 2002; Oman et al. 2015). While these problems are in principle distinct issues, as the second refers to a tension in total cumulative mass and the first is an issue with the derivative, it is likely that they point to a common tension. Dark-matter-only Λ CDM halos are too dense and too cuspy in their centers compared to many observed galaxies.

Figure 9 summarizes the basic problem. Shown as a dashed line is the typical circular velocity curve predicted for an NFW Λ CDM dark matter halo with $V_{\text{max}} \approx 40 \text{ km s}^{-1}$ compared to the observed rotation curves for two galaxies with the same asymptotic velocity from Oh et al. (2015). The observed rotation curves rise much more slowly than the Λ CDM expectation, reflecting core densities that are lower and more core-like than the fiducial

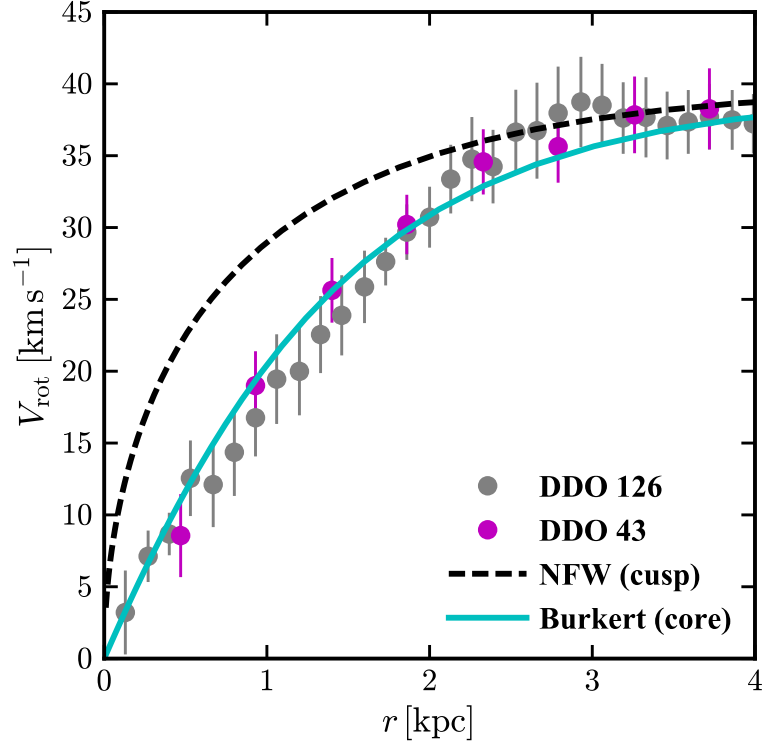


Figure 9

The Cusp-Core problem. The dashed line shows the naive Λ CDM expectation (NFW, from dark-matter-only simulations) for a typical rotation curve of a $V_{\text{max}} \approx 40 \text{ km s}^{-1}$ galaxy. This rotation curve rises quickly, reflecting a central density profile that rises as a cusp with $\rho \propto 1/r$. The data points show the rotation curves of two example galaxies of this size from the LITTLE THINGS survey (Oh et al. 2015), which are more slowly rising and better fit by a density profile with a constant density core (Burkert 1995, cyan line).

prediction.

2.3. Too-Big-To-Fail

As discussed above, a straightforward and natural solution to the missing satellites problem within Λ CDM is to assign the known Milky Way satellites to the largest dark matter subhalos (where largest is in terms of either present-day mass or peak mass) and attribute the lack of observed galaxies in the remaining smaller subhalos to galaxy formation physics. As pointed out by Boylan-Kolchin, Bullock & Kaplinghat (2011), this solution makes a testable prediction: the inferred central masses of Milky Way satellites should be consistent with the central masses of the most massive subhalos in Λ CDM simulations of Milky Way-mass halos. Their comparison of observed central masses to Λ CDM predictions from the Aquarius (Springel et al. 2008) and Via Lactea II (Diemand et al. 2008) simulations revealed that the most massive Λ CDM subhalos were systematically too centrally dense to host the bright Milky Way satellites (Boylan-Kolchin, Bullock & Kaplinghat 2011, 2012).

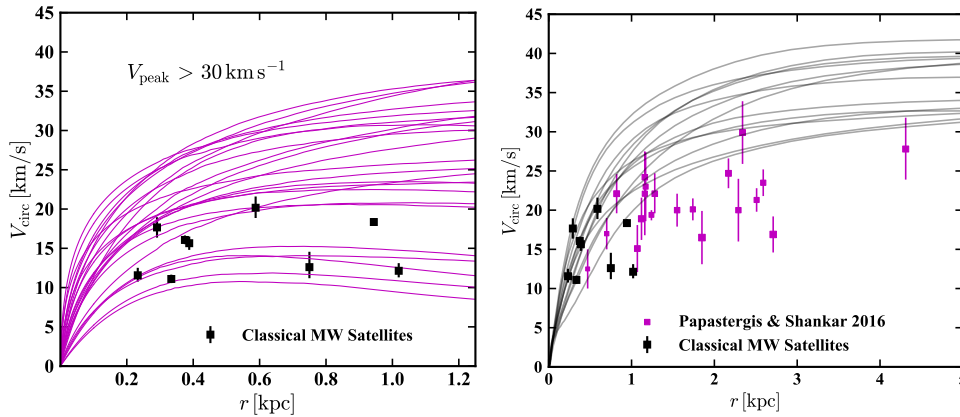


Figure 10

The Too-Big-to-Fail Problem. *Left:* Data points show the circular velocities of classical Milky Way satellite galaxies with $M_\star \simeq 10^{5-7} M_\odot$ measured at their half-light radii $r_{1/2}$. The magenta lines show the circular velocity curves of subhalos from one of the (dark matter only) Aquarius simulations. These are specifically the subhalos of a Milky Way-size host that have peak maximum circular velocities $V_{\text{max}} > 30 \text{ km s}^{-1}$ at some point in their histories. Halos that are this massive are likely resistant to strong star formation suppression by reionization and thus naively too big to have failed to form stars (modified from [Boylan-Kolchin, Bullock & Kaplinghat 2012](#)). The existence of a large population of such satellites with greater central masses than any of the Milky Way’s dwarf spheroidals is the original Too-Big-to-Fail problem. *Right:* The same problem – a mismatch between central masses of simulated dark matter systems and observed galaxies – persists for field dwarfs (magenta points), indicating it is not a satellite-specific process (modified from [Papastergis & Ponomareva 2017](#)). The field galaxies shown all have stellar masses in the range $5.75 \leq \log_{10}(M_\star/M_\odot) \leq 7.5$. The gray curves are predictions for Λ CDM halos from the fully self-consistent hydrodynamic simulations of [Fitts et al. \(2016\)](#) that span the same stellar mass range in the simulations as the observed galaxies.

While there are subhalos with central masses comparable to the Milky Way satellites, these subhalos were never among the ~ 10 most massive (Figure 10). Why would galaxies fail to form in the most massive subhalos, yet form in dark matter satellites of lower mass? The most massive satellites should be “too big to fail” at forming galaxies if the lower-mass satellites are capable of doing so (thus the origin of the name of this problem). In short, while the *number* of massive subhalos in dark-matter-only simulations matches the number of classical dwarfs observed (see Figure 8), the *central densities* of these simulated dwarfs are higher than the central densities observed in the real galaxies (see Figure 10).

While too-big-to-fail was originally identified for satellites of the Milky Way, it was subsequently found to exist in Andromeda ([Tollerud, Boylan-Kolchin & Bullock 2014](#)) and field galaxies in the Local Group (those outside the virial radius of the Milky Way and M31; [Kirby et al. 2014](#)). Similar discrepancies were also pointed out for more isolated low-mass galaxies, first based on HI rotation curve data ([Ferrero et al. 2012](#)) and subsequently using velocity width measurements ([Papastergis et al. 2015](#); [Papastergis & Shankar 2016](#)). This version of too-big-to-fail in the field is also manifested in the velocity function of field galaxies⁴ ([Zavala et al. 2009](#); [Klypin et al. 2015](#); [Trujillo-Gomez et al. 2016](#); [Schneider](#)

⁴We note that the mismatch between the observed and predicted velocity function can also be

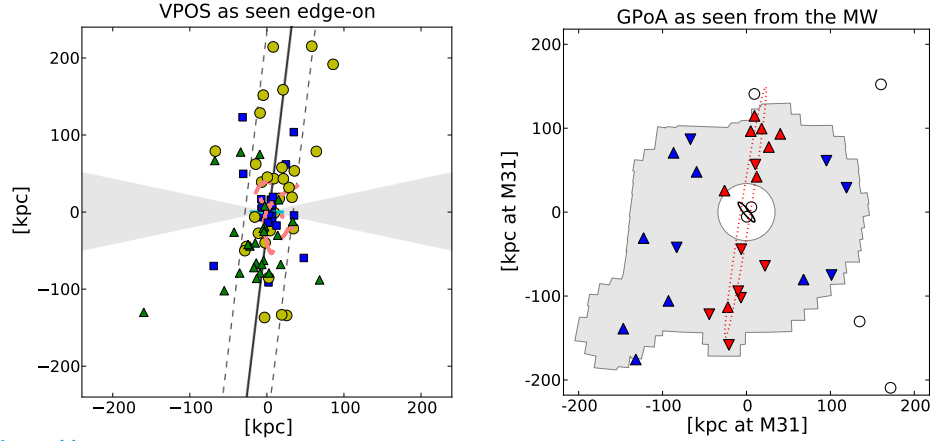


Figure 11

Planes of Satellites. *Left*: Edge-on view of the satellite distribution around the Milky Way (updated from Pawlowski, McGaugh & Jerjen 2015) with the satellite galaxies in yellow, young halo globular clusters and star clusters in blue, and all other newly-discovered objects (unconfirmed dwarf galaxies or star clusters) are shown as green triangles. The red lines in the center dictate the position and orientation of streams in the MW halo. The gray wedges span 24 degrees about the plane of the MW disk, where satellite discovery might be obscured by the Galaxy. *Right*: The satellite distribution around Andromeda (modified by M. Pawlowski from Ibata et al. 2013) where the red points are satellites belonging to the identified kinematic plane. Triangles pointing up are receding relative to M31. Triangles pointing down are approaching.

et al. 2016, though see Macciò et al. 2016 and Brooks et al. 2017 for arguments that no discrepancy exists). The generic observation in the low-redshift Universe, then, is that the inferred central masses of galaxies with $10^5 \lesssim M_*/M_\odot \lesssim 10^8$ are $\sim 50\%$ smaller than expected from dissipationless Λ CDM simulations.

The too-big-to-fail and core/cusp problems would be naturally connected if low-mass galaxies generically have dark matter cores, as this would reduce their central densities relative to CDM expectations⁵. However, the problems are, in principle, separate: one could imagine galaxies that have large constant-density cores yet still with too much central mass relative to CDM predictions (solving the core/cusp problem but not too-big-to-fail), or having cuspy profiles with overall lower density amplitudes than CDM (solving too-big-to-fail but not core/cusp).

2.4. Satellite Planes

Kunkel & Demers (1976) and Lynden-Bell (1976) pointed out that satellite galaxies appeared to lie in a polar great circle around the Milky Way. Insofar as this cannot be explained in a theory of structure formation, this observation pre-dates all other small-scale

interpreted as a “missing dwarfs” problem if one considers the discrepancy as one in numbers at fixed V_{halo} . We believe, however, that the more plausible interpretation is a discrepancy in V_{halo} at fixed number density.

⁵For a sense of the problem, the amount of mass that would need to be removed to alleviate the issue on classical dwarf scales is $\sim 10^7 M_\odot$ within ~ 300 pc

structure issues in the Local Group by approximately two decades. The anisotropic distribution of Galactic satellites received scant attention until a decade ago, when [Kroupa, Theis & Boily \(2005\)](#) argued that it proved that satellite galaxies cannot be related to dark matter substructures (and thereby constituted another crisis for CDM). Kroupa et al. examined classical, pre-SDSS dwarf galaxies in and around the Milky Way and found that the observed distribution was strongly non-spherical. From this analysis, based on the distribution of angles between the normal of the best-fitting plane of dwarfs and the position vector of each MW satellite in the Galacto-centric reference frame, Kroupa et al. argued that “the mismatch between the number and spatial distribution of MW dwarves compared to the theoretical distribution challenges the claim that the MW dwarves are cosmological sub-structures that ought to populate the MW halo.”

This claim was quickly disputed by [Zentner et al. \(2005\)](#), who investigated the spatial distribution of dark matter subhalos in simulated CDM halos and determined that it was highly inconsistent with a spherical distribution. They found that the planar distribution of MW satellites was marginally consistent with being a random sample of the subhalo distributions in their simulations, and furthermore, the distribution of satellites they considered likely to be luminous (corresponding to the more massive subhalos) was even more consistent with observations. A similar result was obtained at roughly the same time by [Kang et al. \(2005\)](#). Slightly later, [Metz, Kroupa & Jerjen \(2007\)](#) argued that the distribution of MW satellite galaxies was inconsistent, at the 99.5% level, with isotropic or prolate substructure distributions (as might be expected in Λ CDM).

Related analysis of Milky Way satellite objects has further supported the idea that the configuration is highly unusual compared to Λ CDM subhalo distributions ([Pawlowski, Pflamm-Altenburg & Kroupa 2012](#)), with the 3D motions of satellites suggesting that there is a preferred orbital pole aligned perpendicular to the observed spatial plane ([Pawlowski & Kroupa 2013](#)). The left hand side of Figure 11 shows the current distribution of satellites (galaxies and star clusters) around the Milky Way looking edge-on at the planar configuration. Note that the disk of the Milky Way could, in principle, bias discoveries away from the MW disk axis, but it is not obvious that the orbital poles would be biased by this effect. Taken together, the orbital poles and spatial configuration of MW satellites is highly unusual for a randomly drawn sample of Λ CDM subhalos ([Pawlowski et al. 2015](#)).

As shown in the right-hand panel of Figure 11, the M31 satellite galaxies also show evidence for having a disk-like configuration ([Metz, Kroupa & Jerjen 2007](#)). Following the discovery of new M31 satellites and the characterization of their velocities, [Conn et al. \(2013\)](#) and [Ibata et al. \(2013\)](#) presented evidence that 15 of 27 Andromeda dwarf galaxies indeed lie in a thin plane, and further, that the southern satellites are mostly approaching us with respect to M31, while the northern satellites are mostly receding (as coded by the direction of the red triangles in Figure 11). This suggests that the plane could be rotationally supported. Our view of this plane is essentially edge-on, meaning we have excellent knowledge of in-plane motions and essentially no knowledge of velocities perpendicular to the plane. Nevertheless, even a transient plane of this kind would be exceedingly rare for Λ CDM subhalos (e.g., [Ahmed, Brooks & Christensen 2017](#)).

Work in a similar vein has argued for the existence of planar structures in the Centaurus A group ([Tully et al. 2015](#)) and for rotationally-supported systems of satellites in a statistical sample of galaxies from the SDSS ([Ibata et al. 2015](#)). [Libeskind et al. \(2015\)](#) have used Λ CDM simulations to suggest that some alignment of satellite systems in the local Universe may be naturally explained by the ambient shear field, though they cannot explain thin

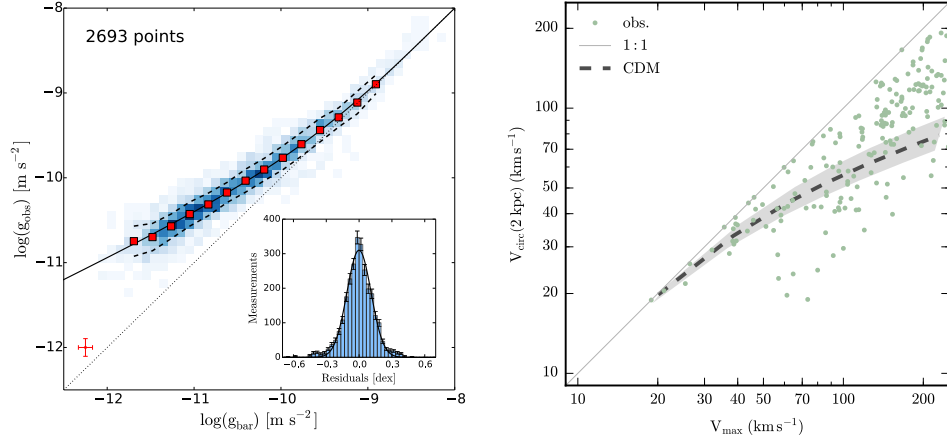


Figure 12

Regularity vs. Diversity. *Left:* The radial acceleration relation from [McGaugh, Lelli & Schombert \(2016\)](#), slightly modified) showing the centripetal acceleration observed in rotation curves, $g_{\text{obs}} = V^2/r$, plotted versus the expected acceleration from observed baryons g_{bar} for 2700 individual data points from 153 galaxy rotation curves. Large squares show the mean and the dashed line lines show the rms width. *Right:* Green points show the circular velocities of observed galaxies measured at 2 kpc as a function of V_{max} from [Oman et al. \(2015\)](#) as re-created by [Creasey et al. \(2017\)](#). For comparison, the gray band shows expectations from dark matter only Λ CDM simulations. There is much more scatter at fixed V_{max} than predicted by the simulations. Note that the galaxies used in the RAR in left-hand panel have V_{max} values that span the range shown on the right. The tightness of the acceleration relation is remarkable (consistent with zero scatter given observational error, red cross), especially given the variation in central densities seen on the right.

planes this way. Importantly, [Phillips et al. \(2015\)](#) have re-analyzed the SDSS data and argued that it is not consistent with a ubiquitous co-rotating satellite population and rather more likely a statistical fluctuation. More data that enables a statistical sample of hosts down to fainter satellites will be needed to determine whether the configurations seen in the Local Group are common.

2.5. Regularity in the Face of Diversity

Among the more puzzling aspects of galaxy phenomenology in the context of Λ CDM are the tight scaling relations between dynamical properties and baryonic properties, even within systems that are dark matter dominated. One well-known example of this is the baryonic Tully-Fisher relation ([McGaugh 2012](#)), which shows a remarkably tight connection between the total baryonic mass of a galaxy (gas plus stars) and its circular velocity $V_{\text{flat}} (\simeq V_{\text{max}})$: $M_b \propto V_{\text{flat}}^4$. Understanding this correlation with Λ CDM models requires care for the low-mass galaxies of most concern in this review ([Brook, Santos-Santos & Stinson 2016](#)).

A generalization of the baryonic Tully-Fisher relation known as the radial acceleration relation (RAR) was recently introduced by [McGaugh, Lelli & Schombert \(2016\)](#). Plotted in left-hand Figure 12, the RAR shows a tight correlation between the radial acceleration traced by rotation curves ($g_{\text{obs}} = V^2/r$) and that predicted solely by the observed dis-

tribution of baryons (g_{bar})⁶. The upper right “high-acceleration” portion of the relation correspond to baryon-dominated regions of (mostly large) galaxies. Here the relation tracks the one-to-one line, as it must. However, rotation curve points begin to peel away from the line, towards an acceleration larger than what can be explained by the baryons alone below a characteristic acceleration of $a_0 \simeq 10^{-10} \text{ m s}^{-2}$. It is this additional acceleration that we attribute to dark matter. The outer parts of large galaxies contribute to this region, as do virtually all parts of small galaxies. It is surprising, however, that the dark matter contribution in the low-acceleration regime tracks the baryonic distribution so closely, particularly in light of the diversity in galaxy rotation curves seen among galaxies of at a fixed V_{flat} , as we now discuss.

The right-hand panel of Figure 12 illustrates the diversity in rotation curve shapes seen from galaxy to galaxy. Shown is a slightly modified version of a figure introduced by [Oman et al. \(2015\)](#) and recreated by [Creasey et al. \(2017\)](#). Each data point corresponds to a single galaxy rotation curve. The horizontal axis shows the observed value of V_{flat} ($\approx V_{\text{max}}$) for each galaxy and the vertical axis plots the value of the circular velocity at 2 kpc from the galaxy center. Note that at fixed V_{flat} , galaxies demonstrate a huge diversity in central densities. Remarkably, this diversity is apparently correlated with the baryonic content in such a way as to drive the tight relation seen on the left. The gray band in the right panel shows the expected relationship between V_{max} and $V_c(2\text{kpc})$ for halos in ΛCDM dark-matter-only simulations. Clearly, the real galaxies demonstrate much more diversity than is naively predicted.

The real challenge, as we see it, is to understand how galaxies can have so much diversity in their rotation curve shapes compared to naive ΛCDM expectations while also yielding tight correlations with baryonic content. The fact that there is a tight correlation with *baryonic* mass and not stellar mass (which presumably correlates more closely with total feedback energy) makes the question all the more interesting.

3. SOLUTIONS

3.1. Solutions within ΛCDM

In this subsection, we explore some of the most popular and promising solutions to the problems discussed above. We take as our starting point the basic ΛCDM model plus reionization, i.e., we take it as a fundamental prediction of ΛCDM that the heating of the intergalactic medium to $\sim 10^4 \text{ K}$ by cosmic reionization will suppress galaxy formation in halos with virial temperatures below $\sim 10^4 \text{ K}$ (or equivalently, with $V_{\text{vir}} \lesssim 20 \text{ km s}^{-1}$) at $z \lesssim 6$.

3.1.1. Feedback-induced cores. Many of the most advanced hydrodynamic simulations today have shown that it is possible for baryonic feedback to erase the central cusps shown in the density profiles in Figure 3 and produce core-like density profiles as inferred from rotation curves such as those shown in Figure 9 ([Mashchenko, Wadsley & Couchman 2008](#); [Pontzen & Governato 2012](#); [Madau, Shen & Governato 2014](#); [Oñorbe et al. 2015](#); [Read, Agertz & Collins 2016](#)). One key prediction is that the effect of core creation will vary with

⁶This type of relation is what is generally expected in MOND, though the precise shape of the relation depends on the MOND interpolation function assumed (see [McGaugh, Lelli & Schombert 2016](#) for a brief discussion).

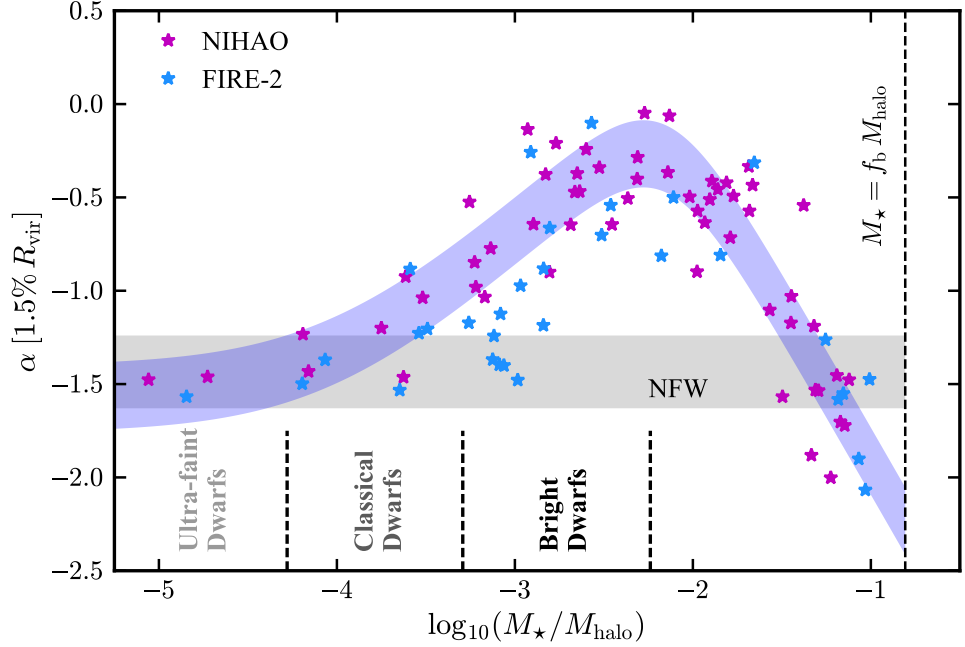


Figure 13

The impact of baryonic feedback on the inner profiles of dark matter halos. Plotted is the inner dark matter density slope α at $r = 0.015 R_{\text{vir}}$ as a function of M_{\star}/M_{vir} for simulated galaxies at $z = 0$. Larger values of $\alpha \approx 0$ imply core profiles, while lower values of $\alpha \lesssim 0.8$ imply cusps. The shaded gray band shows the expected range of dark matter profile slopes for NFW profiles as derived from dark-matter-only simulations (including concentration scatter). The filled magenta stars and shaded purple band (to guide the eye) show the predicted inner density slopes from the NIHAO cosmological hydrodynamic simulations by Tollet et al. (2016). The cyan stars are a similar prediction from an entirely different suite of simulations from the FIRE-2 simulations (Fitts et al. 2016; Hopkins et al. 2017, Chan et al., in preparation). Note that at dark matter core formation peaks in efficiency at $M_{\star}/M_{\text{vir}} \approx 0.005$, in the regime of the brightest dwarfs. Both simulations find that for $M_{\star}/M_{\text{vir}} \lesssim 10^{-4}$, the impact of baryonic feedback is negligible. This critical ratio below which core formation via stellar feedback is difficult corresponds to the regime of classical dwarfs and ultra-faint dwarfs.

the mass in stars formed (Governato et al. 2012; Di Cintio et al. 2014). If galaxies form enough stars, there will be enough supernovae energy to redistribute dark matter and create significant cores. If too many baryons end up in stars, however, the excess central mass can compensate and drag dark matter back in. At the other extreme, if too few stars are formed, there will not be enough energy in supernovae to alter halo density structure and the resultant dark matter distribution will resemble dark-matter-only simulations. While the possible importance of supernova-driven blowouts for the central dark matter structure of dwarf galaxies was already appreciated by Navarro, Eke & Frenk (1996) and Gnedin & Zhao (2002), an important recent development is the understanding that even low-level star formation over an extended period can drive gravitational potential fluctuations that lead to dark matter core formation.

This general behavior is illustrated in Figure 13, which shows the impact of baryonic

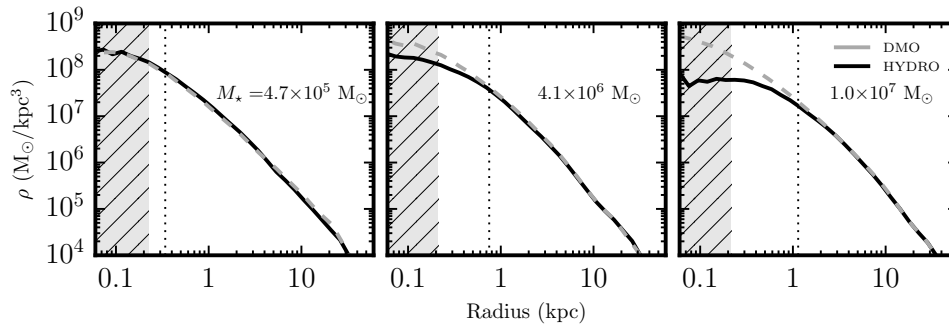


Figure 14

Dark matter density profiles from full hydrodynamic FIRE-2 simulations (Fitts et al. 2016). Shown are three different galaxy halos, each at mass $M_{\text{vir}} \approx 10^{10} M_{\odot}$. Solid lines show the hydro runs while the dashed show the same halos run with dark matter only. The hatched band at the left of each panel marks the region where numerical relaxation may artificially modify density profiles and the vertical dotted line shows the half-light radius of the galaxy formed. The stellar mass of the galaxy formed increases from left to right: $M_{\star} \approx 5 \times 10^5$, 4×10^6 , and $10^7 M_{\odot}$, respectively. As M_{\star} increases, so does the effect of feedback. The smallest galaxy has no effect on the density structure of its host halo.

feedback on the inner slopes of dark matter halos α measured at 1 – 2% of the halo virial radii. Core-like density profiles have $\alpha \rightarrow 0$. The magenta stars show results from the NIHAO hydrodynamic simulations as a function of M_{\star}/M_{vir} , the ratio of stellar mass to the total halo mass (Tollet et al. 2016). The cyan stars show results from an entirely different set of simulations from the FIRE-2 collaboration (Wetzel et al. 2016; Fitts et al. 2016; Garrison-Kimmel et al. 2017b, Chan et al., in preparation). The shaded gray band shows the expected slopes for NFW halos with the predicted range of concentrations from dark-matter-only simulations. We see that both sets of simulations find core formation to be fairly efficient $M_{\star}/M_{\text{vir}} \approx 0.005$. This “peak core formation” ratio maps to $M_{\star} \simeq 10^{8-9} M_{\odot}$, corresponding to the brightest dwarfs. At ratios below $M_{\star}/M_{\text{vir}} \approx 10^{-4}$, however, the impact of baryonic feedback is negligible. The ratio below which core formation is difficult corresponds to $M_{\star} \approx 10^6 M_{\odot}$ – the mass-range of interest for the too-big-to-fail problem.

The effect of feedback on density profile shapes as a function of stellar mass is further illustrated in Figure 14. Here we show simulation results from Fitts et al. (2016) for three galaxies (from a cosmological sample of fourteen), all formed in halos with $M_{\text{vir}}(z=0) \approx 10^{10} M_{\odot}$ using the FIRE-2 galaxy formation prescriptions (Hopkins et al. 2014 and in preparation). The dark matter density profiles of the resultant hydrodynamical runs are shown as solid black lines in each panel, with stellar mass labeled, increasing from left to right. The dashed lines in each panel show dark-matter-only versions of the same halos. We see that only in the run that forms the most stars ($M_{\star} \approx 10^7 M_{\odot}$, $M_{\star}/M_{\text{vir}} \approx 10^{-3}$) does the feedback produce a large core. Being conservative, for systems with $M_{\star}/M_{\text{vir}} \lesssim 10^{-4}$, feedback is likely to be ineffective in altering dark matter profiles significantly compared to dark-matter-only simulations.

SCALE WHERE FEEDBACK BECOMES INEFFECTIVE IN PRODUCING CORES

$$M_{\star}/M_{\text{vir}} \approx 10^{-4} \leftrightarrow M_{\star} \approx 10^6 M_{\odot} \leftrightarrow M_{\text{vir}} \approx 10^{10} M_{\odot}$$

It is important to note that while many independent groups are now obtaining similar results in cosmological simulations of dwarf galaxies (Governato et al. 2012; Munshi et al. 2013; Madau, Shen & Governato 2014; Chan et al. 2015; Oñorbe et al. 2015; Tollet et al. 2016; Fitts et al. 2016) – indicating a threshold mass of $M_{\star} \sim 10^6 M_{\odot}$ or $M_{\text{vir}} \sim 10^{10} M_{\odot}$ – this is *not* an ab initio Λ CDM prediction, and it depends on various adopted parameters in galaxy formation modeling. For example, Sawala et al. (2016) do not obtain cores in their simulations of dwarf galaxies, yet they still produce systems that match observations well owing to a combination of feedback effects that lower central densities of satellites (thereby avoiding the too-big-to-fail problem). On the other hand, the very high resolution, non-cosmological simulations presented in Read, Agertz & Collins (2016) produce cores in galaxies of *all* stellar masses. We note that Read et al.’s galaxies have somewhat higher M_{\star} at a given M_{vir} than the cosmological runs described cited above; this leads to additional feedback energy per unit dark matter mass, likely explaining the differences with cosmological simulations and pointing to a testable prediction for dwarf galaxies’ M_{\star}/M_{vir} .

3.1.2. Resolving too-big-to-fail. The baryon-induced cores described in Section 3.1.1 have their origins in stellar feedback. The existence of such cores for galaxies above the critical mass scale of $M_{\star} \approx 10^6 M_{\odot}$ would explain why \sim half of the classical Milky Way dwarfs – those above this mass – have low observed densities. However, about half of the MW’s classical dwarfs have $M_{\star} < 10^6 M_{\odot}$, so the scenario described in Section 3.1.1 does not explain these systems’ low central masses. Several other mechanisms exist to reconcile Λ CDM with the internal structure of low-mass halos, however.

Interactions between satellites and the Milky Way – tidal stripping, disk shocking, and ram pressure stripping – all act as additional forms of feedback that can reduce the central masses of satellites. Many numerical simulations of galaxy formation point to the importance of such interactions (which are generally absent in dark-matter-only simulations⁷), and these environmental influences are often invoked in explaining too-big-to-fail (e.g., Zolotov et al. 2012; Arraki et al. 2014; Brooks & Zolotov 2014; Brook & Di Cintio 2015; Wetzel et al. 2016; Tomozeiu, Mayer & Quinn 2016; Sawala et al. 2016; Dutton et al. 2016). In many of these papers, environmental effects are limited to 1-2 virial radii from the host galaxy. Several Local Group galaxies reside at greater distances. While only a handful of these systems have $M_{\star} < 10^6 M_{\odot}$ (most are $M_{\star} \sim 10^7 M_{\odot}$), these galaxies provide an initial test of the importance of external feedback: if environmental factors are key in setting the central densities of low-mass systems, satellites should differ systematically from field galaxies. The results of Kirby et al. (2014) find no such difference; further progress will likely have to await the discovery of fainter systems an larger optical telescopes to provide spectroscopic samples for performing dynamical analyses. Other forms of feedback may

⁷We note that capturing these effects is extremely demanding numerically, and it is not clear that any published cosmological hydrodynamical simulation of a Milky Way-size system can resolve the mass within 300 – 500 pc of satellite galaxies with the accuracy required to address this issue.

persist to larger distances. For example, [Benítez-Llambay et al. \(2013\)](#) note that “cosmic web stripping” (ram pressure from large-scale filaments or pancakes) may be important in dwarf galaxy evolution.

None of these solutions would explain too-big-to-fail in isolated field dwarfs. However, a number of factors could influence the conversion between observed HI line widths and the underlying gravitational potential, complicating the interpretation of systematically low densities (for a discussion of some of these issues, see [Papastergis & Ponomareva 2017](#)). Some examples are: (1) gas may not have the radial extent necessary to reach the maximum of the dark matter halo rotation curve; (2) the contribution of non-rotational support (pressure from turbulent motions) may be non-negligible and not correctly handled; and (3) determinations of inclinations angles of galaxies may be systematically wrong. [Macciò et al. \(2016\)](#) find good agreement between their simulations and the observed abundance of field dwarf galaxies in large part because the gas distributions in the simulated dwarfs do not extend to the peak of the dark matter rotation curve (see also [Kormendy & Freeman 2016](#) for a similar conclusion reached via different considerations). A more complete understanding of observational samples and very careful comparisons between observations and simulations are crucial for quantifying the magnitude of any discrepancies between observations and theory.

3.1.3. Explaining planes. Even prior to the [Ibata et al. \(2013\)](#) result on the potential rotationally-supported plane in M31, multiple groups continued to study the observed distribution of satellite galaxies, their orbits, and the consistency of these with Λ CDM. [Libeskind et al. \(2009\)](#) and [Lovell et al. \(2011\)](#) argued that the MW satellite configuration and orbital distribution are consistent with predictions from Λ CDM simulations, while [Metz, Kroupa & Jerjen \(2009\)](#) and [Pawlowski, Pflamm-Altenburg & Kroupa \(2012\)](#) argued that evidence of a serious discrepancy had only become stronger. A major point of disagreement was whether or not filamentary accretion within Λ CDM is sufficient to explain satellite orbits. Given that SDSS only surveyed about 1/3 of the northern sky (centered on the North Galactic Pole, thereby focusing on the portion of the sky where the polar plane was claimed to lie), areal coverage was a serious concern when trying to understand the significance of the polar distribution of satellites. DES has mitigated this concern somewhat, but it is also surveying near the polar plane. [Pawlowski \(2016\)](#) has recently argued that incomplete sky coverage is *not* the driving factor in assessing phase-space alignments in the Milky Way; future surveys with coverage nearer the Galactic plane should definitively test this assertion.

Following [Ibata et al. \(2013\)](#), the question of whether the M31 configuration (right-hand side of Figure 11) is expected in Λ CDM also became a topic of substantial interest. The general consensus of work rooted in Λ CDM is that planes qualitatively similar (though not as thin) as the M31 plane are not particularly uncommon in Λ CDM simulations, but that these planes are not rotationally-supported structures (e.g., [Bahl & Baumgardt 2014](#); [Gillet et al. 2015](#); [Buck, Dutton & Macciò 2016](#)). Since we view the M31 plane almost perfectly edge-on, proper motions of dwarf galaxies in the plane would provide a clean test of its nature. Should this plane turn out to be rotationally supported, it would be *extremely* difficult to explain with our current understanding of the Λ CDM model. These proper motions may be possible with a combination of *Hubble* and *James Webb Space Telescope* data. [Skillman et al. \(2017\)](#) presented preliminary observations of three plane and three non-plane galaxies, finding no obvious differences between the two sets of galaxies. Future observations of this sort could help shed light on the M31 plane and its nature.

3.1.4. Explaining the radial acceleration relation. Almost immediately after [McGaugh, Lelli & Schombert \(2016\)](#) published their RAR relation paper, [Keller & Wadsley \(2017\)](#) responded by demonstrating that a similar relation can be obtained using Λ CDM hydrodynamic simulations of disk galaxies. Importantly, however, the systems simulated did not include low-mass galaxies, which are dark-matter-dominated throughout. The smallest galaxies are the ones with low acceleration in their centers as well as in their outer parts, and they remain the most puzzling to explain (see [Milgrom 2016](#) for a discussion related to this issue).

More recently, [Navarro et al. \(2016\)](#) have argued that Λ CDM can naturally produce an acceleration relation similar to that shown in Figure 12. A particularly compelling section of their argument follows directly from abundance-matching (Figure 6): the most massive disk galaxies that exist are not expected to be in halos much larger than $5 \times 10^{12} M_{\odot}$. This sets a maximum acceleration scale ($\sim 10^{-10} \text{ m s}^{-2}$) above which any observed acceleration *must* track the baryonic acceleration. The implication is that any mass-discrepancy attributable to dark matter will only begin to appear at accelerations below this scale. Stated slightly differently, any successful model of galaxy formation set within a Λ CDM context *must* produce a relation that begins to peel above the one-to-one only below the characteristic scale observed.

It remains to be seen whether the absolute normalization and shape of the RAR in the low-acceleration regime can be reproduced in Λ CDM simulations that span the full range of galaxy types that are observed to obey the RAR. As stated previously, these same simulations must also simultaneously reproduce the observed diversity of galaxies at fixed V_{max} that is seen in the data (e.g., as shown in the right-panel of Figure 12).

3.2. Solutions requiring modifications to Λ CDM

3.2.1. Modifying linear theory predictions. As discussed in Section 1.6, the dominant impact of dark matter particle nature on the linear theory power spectrum for CDM models is in the high- k cut-off (see labeled curves in Figure 2). This cut-off is set by the free-streaming or collisional damping scale associated with CDM and is of order 1 comoving pc (corresponding to perturbations of $10^{-6} M_{\odot}$) for canonical WIMPs ([Green, Hofmann & Schwarz 2004](#)) or 0.001 comoving pc (corresponding to $10^{-15} M_{\odot}$) for a $m \approx 10 \mu\text{eV}$ QCD axion ([Nambu & Sasaki 1990](#)). In these models, the dark matter halo hierarchy should therefore extend 18 to 27 orders of magnitude below the mass scale of the Milky Way ($10^{12} M_{\odot}$; see Fig. 2).

A variety of dark matter models result in a truncation of linear perturbations at much larger masses, however. For example, WDM models have an effective free-streaming length λ_{fs} that scales inversely with particle mass ([Bode, Ostriker & Turok 2001](#); [Viel et al. 2005](#)); in the Planck ([2016](#)) cosmology, this relation is approximately

$$\lambda_{\text{fs}} = 33 \left(\frac{m_{\text{WDM}}}{1 \text{ keV}} \right)^{-1.11} \text{ kpc} \quad (8)$$

and the corresponding free-streaming mass is

$$\begin{aligned} M_{\text{fs}} &= \frac{4\pi}{3} \rho_{\text{m}} \left(\frac{\lambda_{\text{fs}}}{2} \right)^3 \\ &= 2 \times 10^7 \left(\frac{m_{\text{WDM}}}{1 \text{ keV}} \right)^{-3.33} M_{\odot}. \end{aligned} \quad (9)$$

The effects of power spectrum truncation are not limited to the free-streaming scale, however: power is substantially suppressed for significantly larger scales (smaller wavenumbers

k). A characterization of the scale at which power is significantly affected is given by the half-mode scale $k_{\text{hm}} = 2\pi/\lambda_{\text{hm}}$, where the transfer function is reduced by 50% relative to CDM. The half-mode wavelength λ_{hm} is approximately fourteen times larger than the free-streaming length (Schneider et al. 2012), meaning that structure below $\sim 5 \times 10^{10} M_{\odot}$ is significantly different from CDM in a 1 keV thermal dark matter model:

$$M_{\text{hm}} = 5.5 \times 10^{10} \left(\frac{m_{\text{WDM}}}{1 \text{ keV}} \right)^{-3.33} M_{\odot}. \quad (10)$$

Examples of power suppression for several thermal WDM models are shown by the dashed, dotted, and dash-dotted lines in Fig. 2.

The lack of small-scale power in models with warm (or hot) dark matter is a testable prediction. As the free-streaming length is increased and higher-mass dark matter substructure is erased, the expected number of dark matter satellites inside of a Milky Way-mass halo decreases. The observed number of dark-matter-dominated satellites sets a lower limit on the number of subhalos within the Milky Way, and therefore, a lower limit on the warm dark matter particle mass. Polisensky & Ricotti (2011) find this constraint is $m > 2.3 \text{ keV}$ (95% confidence) while Lovell et al. (2014) find $m > 1.6 \text{ keV}$; these differences come from slightly different cosmologies, assumptions about the mass of the Milky Way’s dark matter halo, and modeling of completeness limits for satellite detections.

It is important to note that particle mass and the free-streaming scale are *not* uniquely related: the free-streaming scale depends on the particle production mechanism and is set by the momentum distribution of the dark matter particles. For example, a resonantly-produced sterile neutrino can have a much “cooler” momentum distribution than a particle of the same mass that is produced by a process in thermal equilibrium (Shi & Fuller 1999). Constraints therefore must be computed separately for each production mechanism (Merle & Schneider 2015; Venumadhav et al. 2016). As an example, the effects of Dodelson-Widrow (1994) sterile neutrinos, which are produced through non-resonant oscillations from active neutrinos, can be matched to effects of thermal relics via the following relation:

$$m(\nu_s) = 3.9 \text{ keV} \left(\frac{m_{\text{thermal}}}{1 \text{ keV}} \right)^{1.294} \left(\frac{\Omega_{\text{DM}} h^2}{0.1225} \right)^{-1/3} \quad (11)$$

(Abazajian 2006; Bozek et al. 2016).

The effects of power spectrum suppression are not limited to pure number counts of dark matter halos: since cosmological structure form hierarchically, the erasure of small perturbations affects the collapse of more massive objects. The primary result of this effect is to delay the assembly of halos of a given mass relative to the case of no power spectrum suppression. Since the central densities of dark matter halos reflect the density of the Universe at the time of their formation, models with reduced small-scale power also result in shallower central gravitational potentials at fixed total mass for halos within 2-3 dex of the free-streaming mass. This effect is highlighted in the lower-middle panel of Figure 15. It compares V_{max} values for a CDM simulation and a WDM simulation that assumes a thermal-equivalent mass of 2 keV but is otherwise identical to the CDM simulation. Massive halos ($V_{\text{max}} \gtrsim 50 \text{ km s}^{-1}$) have identical structure; at lower masses, WDM halos have systematically lower V_{max} values than their CDM counterparts. This effect comes from a reduction of V_{max} for a given halo in the WDM runs, *not* from there being fewer objects. The reduction in central density due to power spectrum suppression for halos near or just below the half-mode mass (but significantly more massive than the

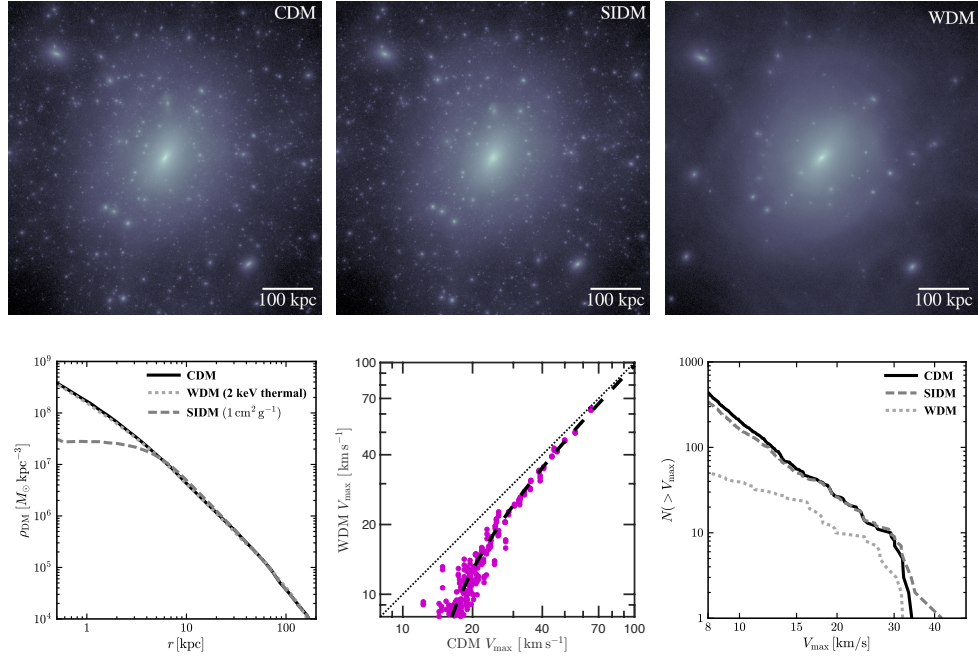


Figure 15

Dark matter phenomenology in the halo of the Milky Way. The three images in the upper row show the same Milky-Way-size dark matter halo simulated with CDM, SIDM ($\sigma/m = 1 \text{ cm}^2/\text{g}$), and WDM (a Shi-Fuller resonant model with a thermal equivalent mass of 2 keV). The left panel in the bottom row shows the dark matter density profiles of the same three halos while the bottom-right panel shows the subhalo velocity functions for each as well. The middle panel on the bottom shows that while the host halos have virtually identical density structure in WDM and CDM, individual subhalos identified in both simulations smaller V_{max} values in WDM (Bozek et al. 2016). This effect can explain the bulk of the differences seen in the V_{max} functions (bottom right panel). Note that SIDM does not reduce the abundance of substructure (unless the power spectrum is truncated) but it does naturally produce large constant-density cores in the dark matter distribution. WDM does not produce large constant-density cores at Milky Way-mass scales but does result in fewer subhalos near the free-streaming mass and reduces V_{max} of a given subhalo (through reduced concentration) near the half-mode mass ($M_{\text{halo}} \lesssim 10^{10} M_{\odot}$ for the plotted 2 keV thermal equivalent model).

free-streaming mass) is how WDM can solve the too-big-to-fail problem (Anderhalden et al. 2013).

3.2.2. Modifying non-linear predictions. The non-linear evolution of CDM is described by the Collisionless Boltzmann equation. Gravitational interactions are the only ones that are relevant for CDM particles, and these interactions operate in the mean field limit (that is, gravitational interactions between individual DM particles are negligible compared to interactions between a dark matter particle and the large-scale gravitational potential). The question of how strong the constraints are on non-gravitational interactions between individual dark matter particles is therefore crucial for evaluating non-CDM models.

There has been long-standing interest in models that involve dark matter self-

interactions (Carlson, Machacek & Hall 1992; Spergel & Steinhardt 2000). In its simplest form, self-interacting dark matter (SIDM, sometimes called collisional dark matter) is characterized by an energy-exchange interaction cross section σ . The mean free path λ of dark matter particles is then $\lambda = (n\sigma)^{-1}$, where n is the local number density of dark matter particles. Since the mass of the dark matter particle is not known, it is often useful to express the mean free path as $(\rho\sigma/m)^{-1}$ and to quantify self-interactions in terms of the cross section per unit particle mass, σ/m . If $\lambda(r)/r \ll 1$ at radius r from the center of a dark matter halo, many scattering events occur per local dynamical time and SIDM acts like a fluid, with conductive transport of heat. In the opposite regime, $\lambda(r)/r \gg 1$, particles are unlikely to scatter over a local dynamical time and SIDM is effectively an optically thin (rarefied) gas, with elastic scattering between dark matter particles. Most work in recent years has been far from the fluid limit.

As originally envisioned by Spergel & Steinhardt (2000) in the context of solving the missing satellites and cusp/core problems, the mean free path for self-interactions is of order $1 \text{ kpc} \lesssim \lambda \lesssim 1 \text{ Mpc}$ at densities characteristic of the Milky Way's dark matter halo (0.4 GeV/cm^3 ; Read 2014), leading to self-interaction cross sections of $400 \gtrsim \sigma/m \gtrsim 0.4 \text{ cm}^2/\text{g}$ ($800 \gtrsim \sigma/m \gtrsim 0.8 \text{ barn/GeV}$). This scale ($\sim \text{barn/GeV}$) is enormous in particle physics terms – it is comparable to the cross-section for neutron-neutron scattering – yet it remains difficult to exclude observationally. It is important to emphasize that the dark matter particle self-interaction strength can, in principle, be completely decoupled from the dark matter's interaction strength with Standard Model particles and thus standard direct detection constraints offer no absolute model-independent limits on σ/m for the dark matter. Astrophysical constraints are therefore essential for understanding dark matter physics.

Though the SIDM cross section estimates put forth by Spergel & Steinhardt (2000) were based on analytic arguments, the interaction scale they proposed to alleviate the cusp/core problem does overlap (at the low end) with more modern results based on fully self-consistent cosmological simulations. Several groups have now run cosmological simulations with dark matter self-interactions and have found that models with $\sigma/m \approx 0.5 - 10 \text{ cm}^2/\text{g}$ produce dark matter cores in dwarf galaxies with sizes $\sim 0.3 - 1.5 \text{ kpc}$ and central densities $2 - 0.2 \times 10^8 M_\odot \text{ kpc}^{-3} = 7.4 - 0.74 \text{ GeV cm}^{-3}$ that can alleviate the cusp/core and too-big-to-fail problems discussed above (e.g., Vogelsberger, Zavala & Loeb 2012; Peter et al. 2013; Fry et al. 2015; Elbert et al. 2015). SIDM does not, however, significantly alleviate the missing satellites problem, as the substructure counts in SIDM simulations are almost identical to those in CDM simulations (Rocha et al. 2013; see Figure 15).

One important constraint on possible SIDM models comes from galaxy clusters. The high central dark matter densities observed in clusters exclude SIDM models with $\sigma/m \gtrsim 0.5 \text{ cm}^2/\text{g}$, though SIDM with $\sigma/m \simeq 0.1 \text{ cm}^2/\text{g}$ may be preferred over CDM (e.g., Kaplinghat, Tulin & Yu 2016; Elbert et al. 2016). This means that in order for SIDM to alleviate the small-scale problems that arise in standard CDM and also match constraints seen on the galaxy cluster scale, it needs to have a velocity-dependent cross section $\sigma(v)$ that decreases as the rms speed of dark matter particles involved in the scattering rises from the scale of dwarfs ($v \sim 10 \text{ km s}^{-1}$) to galaxy clusters ($v \sim 1000 \text{ km s}^{-1}$). Velocity-dependent scattering cross sections are not uncommon among Standard Model particles.

Figure 15 shows the results of three high-resolution cosmological simulations (performed by V. Robles, T. Kelley, and B. Bozek in collaboration with the authors) of the same Milky Way mass halo done with CDM, SIDM ($\sigma/m = 1 \text{ cm}^2/\text{g}$), and WDM (a 7 keV resonant model, with thermal-equivalent mass of 2 keV). The images show density maps spanning

600 kpc. It is clear that while WDM produces many fewer subhalos than CDM, the SIDM model yields a subhalo distribution that is very similar to CDM, with only slightly less substructure near the halo core, which itself is slightly lower density than the CDM case.

These visual impressions are quantified in the bottom three panels, which show the main halo density profiles (left) and the subhalo V_{max} functions for all three simulations (right). The middle panel shows the relationship between the V_{max} values of individual halos identified in both CDM and WDM simulations (Bozek et al. 2016). The left panel shows clearly that SIDM produces a large, constant-density core in the main halo, while the WDM profile is almost identical to the CDM case. However, for mass scales close to the half-mode suppression mass of the WDM model ($M_{\text{halo}} \lesssim 10^{10} M_{\odot}$ for this case), the density structure is affected significantly. This effect accounts for most of the difference seen in the right panel: WDM subhalos have V_{max} values that are greatly reduced compared to their CDM counterparts, meaning there is a V_{max} -dependent shift *leftward* at fixed number (i.e., subhalos at this mass scale are not being destroyed, which would result in a reduction in number at fixed V_{max}).

Finally, we conclude by noting that it is possible to write down SIDM models that have both truncated power spectra and significant self-interactions. Such models produce results that are a hybrid between traditional WDM and SIDM with scale-invariant spectra (e.g. Cyr-Racine et al. 2016; Vogelsberger et al. 2016). Specifically, it is possible to modify dark matter in such a way that it produces both fewer subhalos (owing to power spectra effects) and constant density cores (owing to particle self-interactions) and thus solve the substructure problem and core/cusp problem simultaneously without appealing to strong baryonic feedback.

4. Current Frontiers

4.1. Dwarf galaxy discovery space in the Local Group

The tremendous progress in identifying and characterizing faint stellar systems in the Local Group has led to a variety of new questions. For one, these discoveries have blurred what was previously a clear difference between dwarf galaxies and star clusters, leading to the question, “what is a galaxy?” (Willman & Strader 2012). DES has identified several new satellite galaxies, many of which appear to be clustered around the Large Magellanic Cloud (LMC; Drlaca-Wagner et al. 2015). The putative association of these satellites with the LMC is intriguing (Jethwa, Erkal & Belokurov 2016; Sales et al. 2017), as the nearly self-similar nature of dark matter substructure implies that the LMC – which is likely to be hosted by a halo of $M_{\text{peak}} \sim 10^{11} M_{\odot}$ (Boylan-Kolchin et al. 2010) – could itself contain multiple dark matter satellites above the mass threshold required for galaxy formation. Satellites of the LMC and even fainter dwarfs will be attractive targets for ongoing and future observations to test basic predictions of Λ CDM (Wheeler et al. 2015).

The 800 pound gorilla in the dwarf discovery landscape is the Large Synoptic Survey Telescope (LSST). Currently under construction and set to begin operations in 2022, LSST has the potential to expand dwarf galaxy discovery space substantially: by the end of the survey, co-added LSST data will be sensitive to galaxies ten times more distant (at fixed luminosity) than SDSS, or equivalently, LSST will be able to detect galaxies that are one hundred times fainter than SDSS at the same distance. This means that LSST should be complete for galaxies with $L_{\star} \gtrsim 2 \times 10^3 L_{\odot}$ within ~ 1 Mpc of the Galaxy, dramatically increasing the census of very faint galaxies beyond ~ 100 kpc from the Earth.

One of the unique features of LSST data sets will be the ability to explore the properties of low-mass, *isolated* dark matter halos (i.e., those that have not interacted with a more massive system such as the Milky Way), thereby separating out the effects of environment from internal feedback and dark matter physics. Given the predictions discussed in Sec. 3.1.1, any new discoveries with $M_\star \lesssim 10^6 M_\odot$ at ~ 1 Mpc from the Milky Way and M31 will be attractive targets for discriminating between baryonic feedback and dark matter physics. At this distance, spectrographs on 10m-class telescopes will not be sufficient to measure kinematics of resolved stars; planned 30m-class telescopes will be uniquely suited to this task.

In addition to hosting surviving satellites, galactic halos also act as a graveyard for satellite galaxies that have been disrupted through tidal interactions. These disrupted satellites can form long-lived tidal streams; more generally, the stars from these satellites are part of a galaxy’s stellar halo (which may also encompass stars from globular clusters or other sources). Efforts are underway to disentangle disrupted satellites from other stars in the Milky Way halo via chemistry and kinematics (see [Bland-Hawthorn & Gerhard 2016](#) for a recent review).

4.2. Dwarfs beyond the Local Group

An alternate avenue to probing deeper within the Local Group is to search for low-mass galaxies further away (but still in the very local Universe). The Dark Energy Camera (DECam) and Subaru/Hyper Suprime-Cam are being used by several groups to search for very faint companions in a variety of systems (from NGC 3109, itself a dwarf galaxy at ~ 1.3 Mpc, to Centaurus A, a relatively massive elliptical galaxy at ≈ 3.8 Mpc ([Sand et al. 2015b](#); [Crnojević et al. 2016](#); [Carlin et al. 2016](#)). Searches for the *gaseous* components of galaxies that would otherwise be missed by surveys have also proven fruitful, with a number of individual discoveries ([Giovanelli et al. 2013](#); [Sand et al. 2015a](#); [Tollerud et al. 2016](#)).

Recently, the rediscovery of ultra-diffuse dwarf galaxies ([Impey, Bothun & Malin 1988](#); [Dalcanton et al. 1997](#); [Koda et al. 2015](#); [van Dokkum et al. 2015](#)) has led to significant interest in these odd systems, which have sizes comparable to the Milky Way but luminosities comparable to bright dwarf galaxies. Ultra-diffuse dwarfs have been discovered predominantly in galaxy clusters, but if similar systems – perhaps with even lower luminosities – exist near the Local Group, they could have escaped detection. Understanding the formation and evolution of ultra-diffuse dwarfs, as well as their dark matter content and connection to the broader galaxy population, has the potential to alter our current understanding of faint stellar systems.

4.3. Searches for starless dwarfs

Very low mass dark matter halos *must* be starless, should they exist. Detecting starless halos would represent a strong confirmation of the Λ CDM model (and would place stringent constraints on the possible solutions to problems covered in this review); accordingly, astronomers and physicists are exploring a variety of possibilities for detecting such halos.

A promising technique for inferring the presence of the predicted population of low-mass, dark substructure within the Milky Way is through subhalos’ effects on very cold low velocity dispersion stellar streams ([Ibata et al. 2002](#); [Carlberg 2009](#); [Yoon, Johnston & Hogg 2011](#)). Dark matter substructure passing through a stream will perturb the orbits of the stars, creating gaps and bunches in the stream. Although many physical phenomena

may produce similar effects, and the very existence of gaps themselves remains a matter of debate, large samples of cold streams would likely provide the means to test the abundance of low-mass ($M_{\text{vir}} \sim 10^{5-6} M_{\odot}$) substructure in the Milky Way. We note that the streams from disrupting satellite galaxies discussed above are not suitable for this technique, as they are produced with large enough stellar velocity dispersions that subhalos' effects will go undetectable. Blind surveys for HI gas provide yet another path to searching for starless (or extremely faint) substructure in the very nearby Universe. Some ultra-compact high-velocity clouds (UCHVCs) may be gas-bearing “mini-halos” that are devoid of stars (e.g., Blitz et al. 1999).

Most of the probes we have discussed so far rely on electromagnetic signatures of dark matter. Gravitational lensing is unique in that it is sensitive to *mass* alone, potentially providing a different window into low-mass dark matter halos. Vegetti et al. (2010, 2012) have detected two relatively low-mass dark matter subhalos within lensed galaxies using this technique. The galaxies are at cosmological distances, making it difficult to identify any stellar component associated with the subhalos; Vegetti et al. quote upper limits on the luminosities of detected subhalos of $\sim 5 \times 10^{6-7} L_{\odot}$, comparable to classical dwarfs in the Local Group. The inferred dynamical masses are much higher, however: within 300 pc, Milky Way satellites all have $M_{300} \approx 10^7 M_{\odot}$ (Strigari et al. 2008), while the detected subhalos have $M_{300} \approx (1 - 10) \times 10^8 M_{\odot}$. It remains to be seen whether this is related to the lens modeling or if the substructure in lensing galaxies is fundamentally different from that in the Local Group.

More recently, ALMA has emerged as a promising tool for detecting dark matter halo substructure via spatially-resolved spectroscopy of lensed galaxies. This technique was discussed in Hezaveh et al. (2013), and recently, a subhalo with a total mass of $\sim 10^9 M_{\odot}$ within ~ 1 kpc was detected with ALMA (Hezaveh et al. 2016). At present, the detected substructure is significantly more massive than the hosts of dwarf galaxies in the Local Group: the velocity dispersion of the substructure is $\sigma_{\text{DM}} \sim 30 \text{ km s}^{-1}$ as opposed to $\sigma_{\star} \approx 5 - 10 \text{ km s}^{-1}$ for Local Group dwarf satellites. This value of σ_{DM} is indicative of a galaxy similar to the Small Magellanic Cloud, which has $M_{\star} \sim 5 \times 10^8 M_{\odot}$ and $M_{\text{vir}} \sim (5 - 10) \times 10^{10} M_{\odot}$. The discovery of additional lens systems, and the enhanced resolution and sensitivity of ALMA in its completed configuration, promise to reveal lower-mass substructure, perhaps down to scales similar to Local Group satellites but at cosmological distances and in very different host galaxies.

4.4. Indirect signatures of dark matter

If dark matter is indeed a standard WIMP, two dark matter particles can annihilate into Standard Model particles with electromagnetic signatures. This process is exceedingly rare, on average; as discussed in Section 1.6, the freeze-out of dark matter annihilations is what sets the relic density of dark matter in the WIMP paradigm. Nevertheless, the annihilation rate is proportional to the local value ρ_{DM}^2 , meaning that the centers of dark matter halos are potential sites for annihilations. While the brightest source of such annihilations in the sky should be the Galactic Center, foregrounds make unambiguous detection of annihilating dark matter toward the Galaxy challenging. Dwarf spheroidal galaxies have somewhat lower predicted annihilation fluxes owing both to their greater distances and lower masses, but they have the significant advantage of being free of foreground contamination. The *Fermi* γ -ray telescope has surveyed MW dwarfs extensively, with no conclusive evidence for dark

matter annihilation products. The upper limits on combined dwarf data from *Fermi* are already placing moderate tension on the most basic “WIMP miracle” predictions for the annihilation cross section for wimps with $m \lesssim 100$ GeV (Ackermann et al. 2015). Searches for annihilation from starless dark matter subhalos within the Milky Way via the *Fermi* point source catalog have not yielded any detections to date (Calore et al. 2016).

On cosmic scales, dark matter annihilations may contribute to the extragalactic gamma-ray background (Zavala, Springel & Boylan-Kolchin 2010). The expected contributions of dark matter depend sensitively on the spectrum of dark matter halos and subhalos, as well as the relation between concentration and mass for very low mass systems. These relations can be estimated by a variety of methods (though generally not simulated directly, owing to the enormous range of scales that contribute), with uncertainties being grouped into a “boost factor” that describes unresolved annihilations.

If dark matter is a sterile neutrino rather than a WIMP-like particle, self-annihilation will not be seen. Sterile neutrinos decay radiatively to an active neutrino and a photon, however; for all of the relevant sterile neutrino parameter space, this decay is effectively at rest and a clean signature is therefore a spectral line at half the rest mass energy of the dark matter particle, $E_\gamma = m_{\text{DM}}/2$. While there is no *a priori* expectation for the mass of the sterile neutrino, arguments from Section 3.2.1 point to $E_\gamma \gtrsim 1$ keV, so searches in the soft X-ray band are constraining. The most promising recent result in this field is the detection of a previously unknown X-ray line near 3.51 keV in the spectra of individual galaxy clusters, stacked galaxy clusters, and the halo of M31 (Bulbul et al. 2014; Boyarsky et al. 2014). X-ray observations and satellite counts in M31 rule out an oscillation (Dodelson & Widrow 1994) origin for this line if it indeed originates from sterile neutrino dark matter (Horiuchi et al. 2014), leaving heavy scalar decay and possibly resonant conversion as possible production mechanisms (Merle & Schneider 2015). A definitive test of the origin of the 3.5 keV line was expected from the *Hitomi* satellite, as it had the requisite energy resolution to see the thermal broadening of the line due to virial motions (i.e., the line width from a halo with mass M_{vir} should be $\sim V_{\text{vir}}/c$). With *Hitomi*’s untimely demise, tests of the line’s origin may have to wait for *Athena*.

4.5. The high-redshift Universe

While studies of low-mass dark matter halos are most easily conducted in the very nearby Universe owing to the faintness of the galaxies they host, there are avenues at higher redshifts that may provide alternate windows in to the spectrum of density perturbations. One potentially powerful probe at $z \sim 2 - 6$ is the *Lyman- α forest* of absorption lines produced by neutral hydrogen in the intergalactic medium between us and high-redshift quasars (see McQuinn 2016 for a recent review and further details). This hydrogen probes the density field in the quasi-linear regime (i.e., it is in perturbations that are just starting to collapse) and can constrain the dark matter power spectrum to wavenumbers as large as $k \sim 10 h \text{ Mpc}^{-1}$. Any model that reduces the power on this scale relative to ΛCDM expectations will predict different absorption patterns. In particular, WDM will suppress power on these scales.

Viel et al. (2013) used Lyman- α flux power spectra from 25 quasar sightlines to constrain the mass of thermal relic WDM particle to $m_{\text{WDM,th}} > 3.3 \text{ keV}$ at 95% confidence. This translates into a density perturbation spectrum that must be very close to ΛCDM down to $M \sim 10^8 M_\odot$ (Schneider et al. 2012) and would rule out the possibility that free-

The challenge of detecting “empty” dark matter halos

The detection of abundant, baryon-free, low-mass dark matter halos would be an unambiguous validation of the particle dark matter paradigm, would strongly constrain particle physics models, and would eliminate many of the dark matter candidates for the origin of the small-scale issues described in this review. Why is this such a challenging task?

The answer lies in the densities of low-mass dark matter halos compared to other astrophysical objects. From Equation (4), the average density within a halo’s virial radius is 200 times the cosmic matter density. For the most abundant low-mass halos in standard WIMP models – those just above the free-streaming scale of $\sim 10^{-6} M_{\odot}$ – the virial radius is approximately 0.1 pc. This is the equivalent of the mass of the Earth spread over a distance that is *significantly* larger than the Solar System (the mean distance between Pluto and the Sun is $\sim 2 \times 10^{-4}$ pc). Even the lowest-mass, earliest-collapsing CDM structures are incredibly diffuse compared to typical astrophysical objects. Although there may be $\mathcal{O}(10^{17})$ Earth-mass dark matter subhalos within the Milky Way’s ≈ 300 kpc virial radius, detecting them is a daunting challenge.

streaming has direct relevance for the scales of classical dwarfs (and larger-mass systems). The potential complication with this interpretation is the relationship between density and temperature in the intergalactic medium, as pressure or thermal motions can mimic the effects of dark matter free-streaming.

Counts of galaxies in the high-redshift Universe also trace the spectrum of collapsed density perturbations at low masses, albeit in a non-trivial manner. The mere existence of galaxies at high redshift places an upper limit on the free-streaming length of dark matter (so long as all galaxies form within dark matter halos) in much the same way that the existence of substructure in the local Universe does (Schultz et al. 2014). Menci et al. (2016) have placed limits on the masses of thermal relic WDM particles of 2.4 keV (2.1 keV) at 68% (95%) confidence based on the detection of a single galaxy in the *Hubble* Frontier Fields at $z \sim 6$ with absolute UV magnitude of $M_{\text{UV}} = -12.5$ (Livermore, Finkelstein & Lotz 2017). While this stated constraint is very strong, and the technique is promising, correctly modeling faint, high-redshift galaxies – particularly lensed ones – can be very challenging. Furthermore, the true redshift of the galaxy can only be localized to $\Delta z \sim 1$; the rapid evolution of the halo mass function at high redshift further complicates constraints. With the upcoming *James Webb Space Telescope*, the high-redshift frontier will be pushed fainter and to higher redshifts, raising the possibility of placing strong constraints on the free-streaming length of dark matter through structures in the early Universe.

5. Summary and Outlook

Small-scale structure sits at the nexus of astrophysics, particle physics, and cosmology. Within the standard Λ CDM model, most properties of small-scale structure can be modeled with high precision in the limit that baryonic physics is unimportant. And yet, the level of agreement between theory and observations remains remarkably hard to assess, in large part because of hard-to-model effects of baryonic physics on first-principles predictions. Given the stakes – absent direct detection of dark matter on Earth, indirect evidence from

astrophysics provides the strongest clues to dark matter’s nature – it is essential to take potential discrepancies seriously and to explore all avenues for their resolution.

We have discussed three main classes of problems in this review: (1) counts and (2) densities of low-mass objects, and (3) tight scaling relations between the dark and luminous components of galaxies. All of these issues may have their origin in baryonic physics, but they may also point to the need for a phenomenological theory that goes beyond Λ CDM. Understanding which of these two options is correct is pressing for both astrophysics and particle physics.

In our opinion, the search for abundant dark matter halos with inferred virial masses substantially lower than the expected threshold of galaxy formation ($M_{\text{vir}} \sim 10^8 M_{\odot}$) is the most urgent calling in this field today. The existence of these structures is an unambiguous prediction of all WIMP-based dark matter models (though it is not unique to WIMP models), and confirmation of the existence of dark matter halos with $M \sim 10^6 M_{\odot}$ or less would strongly constrain particle physics of dark matter and effectively rule out any role of dark matter free-streaming in galaxy formation. Here, too, accurate predictions for the number of expected dark subhalos will require an honest accounting of baryon physics – specifically the destructive effects of central galaxies themselves (e.g., [Garrison-Kimmel et al. 2017b](#)). Of nearly equal importance is characterizing the central dark matter density structure of very faint ($M_{\star} \lesssim 10^6 M_{\odot}$) galaxies, as a prediction of many recent high-resolution cosmological simulations within the Λ CDM paradigm is that stellar feedback from galaxies below this threshold mass should not modify their host dark matter halos’ cuspy density profile shape. The detection of ubiquitous cores in very low-mass galaxies therefore has the potential to falsify the Λ CDM paradigm.

While some of the tests of the paradigm are clear, their implementation is difficult. Dark matter substructure is extremely diffuse compared to baryonic matter, making its detection highly challenging. The smallest galaxies have very few stars to base accurate dynamical studies upon. Nevertheless, a variety of independent probes of the small-scale structure of dark matter are now feasible, and the LSST era will likely provide a watershed for our understanding of the nature of dark matter and the threshold of galaxy formation. It is not far-fetched to think that improved astrophysical data, theoretical understanding, and numerical simulations will provide a definitive test of Λ CDM within the next decade, even without the direct detection of particle dark matter on Earth.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

It is a pleasure to thank our collaborators and colleagues for helpful discussions and for making important contributions to our perspectives on this topic. We specifically thank Peter Behroozi, Brandon Bozek, Peter Creasey, Sandy Faber, Alex Fitts, Shea Garrison-Kimmel, Andrea Macciò, Stacy McGaugh, Se-Heon Oh, Manolis Papastergis, Marcel Pawlowski, Victor Robles, Laura Sales, Eduardo Tolley, Mark Vogelsberger, and Hai-Bo Yu for feedback and help in preparing the figures. MBK acknowledges support from The University of Texas at Austin, from NSF grant AST-1517226, and from NASA grants HST-AR-12836,

HST-AR-13888, HST-AR-13896, and HST-AR-14282 from the Space Telescope Science Institute (STScI), which is operated by AURA, Inc., under NASA contract NAS5-26555. JSB was supported by NSF grant AST-1518291 and by NASA through HST theory grants (programs AR-13921, AR-13888, and AR-14282) awarded by STScI. This work used computational resources granted by the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575 and ACI-1053575. Resources supporting this work were also provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center.

LITERATURE CITED

- Aaronson M. 1983. *ApJ* 266:L11–L15
- Abazajian K. 2006. *Phys. Rev. D* 73:063506
- Ackermann M, Albert A, Anderson B, Atwood WB, Baldini L, et al. 2015. *Phys. Rev. Lett.* 115:231301
- Adhikari R, Agostini M, Ky NA, Araki T, Archidiacono M, et al. 2016. *arXiv:1602.04816 [astro-ph]*
- Ahmed SH, Brooks AM, Christensen CR. 2017. *MNRAS* 466:3119–3132
- Alam SMK, Bullock JS, Weinberg DH. 2002. *ApJ* 572:34–40
- Anderhalden D, Schneider A, Macciò AV, Diemand J, Bertone G. 2013. *JCAP* 3:014
- Arraki KS, Klypin A, More S, Trujillo-Gomez S. 2014. *MNRAS* 438:1466–1482
- Bahl H, Baumgardt H. 2014. *MNRAS* 438:2916–2923
- Baldry IK, Driver SP, Loveday J, Taylor EN, Kelvin LS, et al. 2012. *MNRAS* 421:621–634
- Behroozi PS, Wechsler RH, Conroy C. 2013. *ApJ* 762:L31
- Benítez-Llambay A, Navarro JF, Abadi MG, Gottlöber S, Yepes G, et al. 2013. *ApJ* 763:L41
- Benson AJ, Bower RG, Frenk CS, Lacey CG, Baugh CM, Cole S. 2003. *ApJ* 599:38–49
- Benson AJ, Lacey CG, Baugh CM, Cole S, Frenk CS. 2002. *MNRAS* 333:156–176
- Bernardi M, Meert A, Sheth RK, Vikram V, Huertas-Company M, et al. 2013. *MNRAS* 436:697–704
- Bertschinger E. 2006. *Phys. Rev. D* 74:063509
- Bland-Hawthorn J, Gerhard O. 2016. *ARA&A* 54:529–596
- Blitz L, Spergel DN, Teuben PJ, Hartmann D, Burton WB. 1999. *ApJ* 514:818–843
- Blumenthal GR, Faber SM, Primack JR, Rees MJ. 1984. *Nature* 311:517–525
- Bode P, Ostriker JP, Turok N. 2001. *ApJ* 556:93–107
- Bond JR, Cole S, Efstathiou G, Kaiser N. 1991. *ApJ* 379:440–460
- Bosma A. 1978. *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types*. Ph.D. thesis, PhD Thesis, Groningen Univ., (1978)
- Bovill MS, Ricotti M. 2009. *ApJ* 693:1859–1870
- Boyarsky A, Ruchayskiy O, Iakubovskiy D, Franse J. 2014. *Phys. Rev. Lett.* 113:251301
- Boylan-Kolchin M, Bullock JS, Kaplinghat M. 2011. *MNRAS* 415:L40–L44
- Boylan-Kolchin M, Bullock JS, Kaplinghat M. 2012. *MNRAS* 422:1203–1218
- Boylan-Kolchin M, Springel V, White SDM, Jenkins A. 2010. *MNRAS* 406:896–912
- Bozek B, Boylan-Kolchin M, Horiuchi S, Garrison-Kimmel S, Abazajian K, Bullock JS. 2016. *MNRAS* 459:1489–1504
- Brook CB, Di Cintio A. 2015. *MNRAS* 450:3920–3934
- Brook CB, Santos-Santos I, Stinson G. 2016. *MNRAS* 459:638–645
- Brooks A. 2014. *Annalen der Physik* 526:294–308
- Brooks AM, Papastergis E, Christensen CR, Governato F, Stilp A, et al. 2017. *arXiv:1701.07835 [astro-ph]*
- Brooks AM, Zolotov A. 2014. *ApJ* 786:87
- Bryan GL, Norman ML. 1998. *ApJ* 495:80–99

- Buck T, Dutton AA, Macciò AV. 2016. MNRAS
- Bulbul E, Markevitch M, Foster A, Smith RK, Loewenstein M, Randall SW. 2014. ApJ 789:13
- Bullock JS, Kolatt TS, Sigad Y, Somerville RS, Kravtsov AV, et al. 2001. MNRAS 321:559–575
- Bullock JS, Kravtsov AV, Weinberg DH. 2000. ApJ 539:517–521
- Burkert A. 1995. ApJ 447:L25–L28
- Calore F, De Romeri V, Di Mauro M, Donato F, Marinacci F. 2016. *arXiv:1611.03503 [astro-ph]*
- Carlberg RG. 2009. ApJ 705:L223–L226
- Carlin JL, Sand DJ, Price P, Willman B, Karunakaran A, et al. 2016. ApJ 828:L5
- Carlson ED, Machacek ME, Hall LJ. 1992. ApJ 398:43–52
- Chan TK, Kereš D, Oñorbe J, Hopkins PF, Muratov AL, et al. 2015. MNRAS 454:2981–3001
- Cole S. 1991. ApJ 367:45–53
- Conn AR, Lewis GF, Ibata RA, Parker QA, Zucker DB, et al. 2013. ApJ 766:120
- Conroy C, Wechsler RH, Kravtsov AV. 2006. ApJ 647:201–214
- Creasey P, Sameie O, Sales LV, Yu HB, Vogelsberger M, Zavala J. 2017. MNRAS 468:2283–2295
- Crnojević D, Sand DJ, Zaritsky D, Spekkens K, Willman B, Hargis JR. 2016. ApJ 824:L14
- Cyr-Racine FY, Sigurdson K, Zavala J, Bringmann T, Vogelsberger M, Pfrommer C. 2016. Phys. Rev. D 93:123527
- Dalcanton JJ, Spergel DN, Gunn JE, Schmidt M, Schneider DP. 1997. AJ 114:635–654
- Davis M, Efstathiou G, Frenk CS, White SDM. 1985. ApJ 292:371–394
- de Blok WJG, Walter F, Brinks E, Trachternach C, Oh SH, Kennicutt Jr. RC. 2008. AJ 136:2648–2719
- Del Popolo A, Le Delliou M. 2017. *Galaxies* 5:17
- Di Cintio A, Brook CB, Macciò AV, Stinson GS, Knebe A, et al. 2014. MNRAS 437:415–423
- Diemand J, Kuhlen M, Madau P, Zemp M, Moore B, et al. 2008. Nature 454:735–738
- Diemer B, Kravtsov AV. 2015. ApJ 799:108
- Dodelson S, Widrow LM. 1994. Phys. Rev. Lett. 72:17–20
- D’Onghia E, Springel V, Hernquist L, Keres D. 2010. ApJ 709:1138–1147
- Drlica-Wagner A, Bechtol K, Rykoff ES, Luque E, Queiroz A, et al. 2015. ApJ 813:109
- Dubinski J, Carlberg RG. 1991. ApJ 378:496–503
- Dutton AA, Macciò AV. 2014. MNRAS 441:3359–3374
- Dutton AA, Macciò AV, Frings J, Wang L, Stinson GS, et al. 2016. MNRAS 457:L74–L78
- Efstathiou G. 1992. MNRAS 256:43P–47P
- Einasto J. 1965. *Trudy Astrofizicheskogo Instituta Alma-Ata* 5:87–100
- Elbert OD, Bullock JS, Garrison-Kimmel S, Rocha M, Oñorbe J, Peter AHG. 2015. MNRAS 453:29–37
- Elbert OD, Bullock JS, Kaplinghat M, Garrison-Kimmel S, Graus AS, Rocha M. 2016. *arXiv:1609.08626 [astro-ph]*
- Faber SM, Lin DNC. 1983. ApJ 266:L17–L20
- Famaey B, McGaugh SS. 2012. *Living Reviews in Relativity* 15:10
- Feng JL. 2010. ARA&A 48:495–545
- Feng JL, Kumar J. 2008. *Physical Review Letters* 101:231301
- Ferrero I, Abadi MG, Navarro JF, Sales LV, Gurovich S. 2012. MNRAS 425:2817–2823
- Fitts A, Boylan-Kolchin M, Elbert OD, Bullock JS, Hopkins PF, et al. 2016. *arXiv:1611.02281 [astro-ph]*
- Flores RA, Primack JR. 1994. ApJ 427:L1–L4
- Freeman KC. 1970. ApJ 160:811
- Frenk CS, White SDM. 2012. *Annalen der Physik* 524:507–534
- Frenk CS, White SDM, Davis M, Efstathiou G. 1988. ApJ 327:507–525
- Fry AB, Governato F, Pontzen A, Quinn T, Tremmel M, et al. 2015. MNRAS 452:1468–1479
- Gao L, Navarro JF, Cole S, Frenk CS, White SDM, et al. 2008. MNRAS 387:536–544
- Garrison-Kimmel S, Boylan-Kolchin M, Bullock JS, Lee K. 2014. MNRAS 438:2578–2596

- Garrison-Kimmel S, Bullock JS, Boylan-Kolchin M, Bardwell E. 2017a. MNRAS 464:3108–3120
- Garrison-Kimmel S, Wetzel AR, Bullock JS, Hopkins PF, Boylan-Kolchin M, et al. 2017b. *arXiv:1701.03792 [astro-ph]*
- Ghigna S, Moore B, Governato F, Lake G, Quinn T, Stadel J. 1998. MNRAS 300:146–162
- Gillet N, Ocvirk P, Aubert D, Knebe A, Libeskind N, et al. 2015. ApJ 800:34
- Giovanelli R, Haynes MP, Adams EAK, Cannon JM, Rhode KL, et al. 2013. AJ 146:15
- Gnedin OY, Zhao H. 2002. MNRAS 333:299–306
- Governato F, Zolotov A, Pontzen A, Christensen C, Oh SH, et al. 2012. MNRAS 422:1231–1240
- Green AM, Hofmann S, Schwarz DJ. 2004. MNRAS 353:L23–L27
- Griffen BF, Ji AP, Dooley GA, Gómez FA, Vogelsberger M, et al. 2016. ApJ 818:10
- Gunn JE, Gott JRI. 1972. ApJ 176:1
- Guo H, Zheng Z, Behroozi PS, Zehavi I, Chuang CH, et al. 2016. MNRAS 459:3040–3058
- Hargis JR, Willman B, Peter AHG. 2014. ApJ 795:L13
- Hezaveh Y, Dalal N, Holder G, Kuhlen M, Marrone D, et al. 2013. ApJ 767:9
- Hezaveh YD, Dalal N, Marrone DP, Mao YY, Morningstar W, et al. 2016. ApJ 823:37
- Hofmann S, Schwarz DJ, Stöcker H. 2001. Phys. Rev. D 64:083507
- Hopkins PF, Kereš D, Oñorbe J, Faucher-Giguère CA, Quataert E, et al. 2014. MNRAS 445:581–603
- Hopkins PF, Wetzel A, Keres D, Faucher-Giguere CA, Quataert E, et al. 2017. *arXiv:1702.06148 [astro-ph]*
- Horiuchi S, Humphrey PJ, Oñorbe J, Abazajian KN, Kaplinghat M, Garrison-Kimmel S. 2014. Phys. Rev. D 89:025017
- Ibata RA, Famaey B, Lewis GF, Ibata NG, Martin N. 2015. ApJ 805:67
- Ibata RA, Lewis GF, Conn AR, Irwin MJ, McConnachie AW, et al. 2013. Nature 493:62–65
- Ibata RA, Lewis GF, Irwin MJ, Quinn T. 2002. MNRAS 332:915–920
- Impey C, Bothun G, Malin D. 1988. ApJ 330:634–660
- Jethwa P, Erkal D, Belokurov V. 2016. MNRAS 461:2212–2233
- Jing YP. 2000. ApJ 535:30–36
- Jungman G, Kamionkowski M, Griest K. 1996. Phys. Rep. 267:195–373
- Kang X, Mao S, Gao L, Jing YP. 2005. A&A 437:383–388
- Kaplinghat M, Tulin S, Yu HB. 2016. *Physical Review Letters* 116:041302
- Kawasaki M, Nakayama K. 2013. *Annual Review of Nuclear and Particle Science* 63:69–95
- Keller BW, Wadsley JW. 2017. ApJ 835:L17
- Kirby EN, Bullock JS, Boylan-Kolchin M, Kaplinghat M, Cohen JG. 2014. MNRAS 439:1015–1027
- Klypin A, Gottlöber S, Kravtsov AV, Khokhlov AM. 1999a. ApJ 516:530–551
- Klypin A, Karachentsev I, Makarov D, Nasonova O. 2015. MNRAS 454:1798–1810
- Klypin A, Kravtsov AV, Valenzuela O, Prada F. 1999b. ApJ 522:82–92
- Klypin A, Yepes G, Gottlöber S, Prada F, Heß S. 2016. MNRAS 457:4340–4359
- Koda J, Yagi M, Yamanoi H, Komiyama Y. 2015. ApJ 807:L2
- Kolb EW, Turner MS. 1994. *The Early Universe*. vol. 69 of *Frontiers in Physics*. Boulder, CO: Westview Press
- Koposov SE, Belokurov V, Torrealba G, Evans NW. 2015. ApJ 805:130
- Kormendy J, Freeman KC. 2016. ApJ 817:84
- Kravtsov AV. 2013. ApJ 764:L31
- Kravtsov AV, Berlind AA, Wechsler RH, Klypin AA, Gottlöber S, et al. 2004. ApJ 609:35–49
- Kroupa P, Theis C, Boily CM. 2005. A&A 431:517–521
- Kuhlen M, Madau P, Silk J. 2009. *Science* 325:970–
- Kuhlen M, Vogelsberger M, Angulo R. 2012. *Physics of the Dark Universe* 1:50–93
- Kunkel WE, Demers S. 1976. *The Magellanic Plane*. In *The Galaxy and the Local Group*, eds. RJ Dickens, JE Perry, FG Smith, IR King, vol. 182 of *Royal Greenwich Observatory Bulletins*
- Kuzio de Naray R, McGaugh SS, de Blok WJG. 2008. ApJ 676:920–943
- Libeskind NI, Frenk CS, Cole S, Jenkins A, Helly JC. 2009. MNRAS 399:550–558

- Libeskind NI, Hoffman Y, Tully RB, Courtois HM, Pomarède D, et al. 2015. MNRAS 452:1052–1059
- Lin DNC, Faber SM. 1983. ApJ 266:L21–L25
- Livermore RC, Finkelstein SL, Lotz JM. 2017. ApJ 835:113
- Lovell MR, Eke VR, Frenk CS, Jenkins A. 2011. MNRAS 413:3013–3021
- Lovell MR, Frenk CS, Eke VR, Jenkins A, Gao L, Theuns T. 2014. MNRAS 439:300–317
- Ludlow AD, Bose S, Angulo RE, Wang L, Hellwing WA, et al. 2016. MNRAS 460:1214–1232
- Lynden-Bell D. 1976. MNRAS 174:695–710
- Macciò AV, Udrescu SM, Dutton AA, Obreja A, Wang L, et al. 2016. MNRAS 463:L69–L73
- Madau P, Shen S, Governato F. 2014. ApJ 789:L17
- Marchesini D, D’Onghia E, Chincarini G, Firmani C, Conconi P, et al. 2002. ApJ 575:801–813
- Mashchenko S, Wadsley J, Couchman HMP. 2008. *Science* 319:174–
- Massey P, Olsen KAG, Hodge PW, Jacoby GH, McNeill RT, et al. 2007. AJ 133:2393–2417
- McConnachie AW. 2012. AJ 144:4
- McGaugh SS. 2012. AJ 143:40
- McGaugh SS. 2015. *Canadian Journal of Physics* 93:250–259
- McGaugh SS, Lelli F, Schombert JM. 2016. *Physical Review Letters* 117:201101
- McGaugh SS, Rubin VC, de Blok WJG. 2001. AJ 122:2381–2395
- McQuinn M. 2016. ARA&A 54:313–362
- Menci N, Grazian A, Castellano M, Sanchez NG. 2016. ApJ 825:L1
- Merle A, Schneider A. 2015. *Physics Letters B* 749:283–288
- Metz M, Kroupa P, Jerjen H. 2007. MNRAS 374:1125–1145
- Metz M, Kroupa P, Jerjen H. 2009. MNRAS 394:2223–2228
- Milgrom M. 2002. *New Astron. Rev.* 46:741–753
- Milgrom M. 2016. *arXiv:1610.07538 [astro-ph]*
- Mo H, van den Bosch FC, White S. 2010. *Galaxy Formation and Evolution*
- Moore B. 1994. *Nature* 370:629–631
- Moore B, Ghigna S, Governato F, Lake G, Quinn T, et al. 1999. ApJ 524:L19–L22
- More S, Diemer B, Kravtsov AV. 2015. ApJ 810:36
- Moster BP, Somerville RS, Maulbetsch C, van den Bosch FC, Macciò AV, et al. 2010. ApJ 710:903–923
- Munshi F, Governato F, Brooks AM, Christensen C, Shen S, et al. 2013. ApJ 766:56
- Nambu Y, Sasaki M. 1990. *Phys. Rev. D* 42:3918–3924
- Navarro JF, Benítez-Llambay A, Fattahi A, Frenk CS, Ludlow AD, et al. 2016. *arXiv:1612.06329 [astro-ph]*
- Navarro JF, Eke VR, Frenk CS. 1996. MNRAS 283:L72–L78
- Navarro JF, Frenk CS, White SDM. 1997. ApJ 490:493–508
- Navarro JF, Hayashi E, Power C, Jenkins AR, Frenk CS, et al. 2004. MNRAS 349:1039–1051
- Navarro JF, Ludlow A, Springel V, Wang J, Vogelsberger M, et al. 2010. MNRAS 402:21–34
- Oñorbe J, Boylan-Kolchin M, Bullock JS, Hopkins PF, Kereš D, et al. 2015. MNRAS 454:2092–2106
- Oh SH, Hunter DA, Brinks E, Elmegreen BG, Schruba A, et al. 2015. AJ 149:180
- Oman KA, Navarro JF, Fattahi A, Frenk CS, Sawala T, et al. 2015. MNRAS 452:3650–3665
- Papastergis E, Giovanelli R, Haynes MP, Shankar F. 2015. A&A 574:A113
- Papastergis E, Ponomareva AA. 2017. A&A 601:A1
- Papastergis E, Shankar F. 2016. A&A 591:A58
- Pawlowski MS. 2016. MNRAS 456:448–458
- Pawlowski MS, Famaey B, Merritt D, Kroupa P. 2015. ApJ 815:19
- Pawlowski MS, Kroupa P. 2013. MNRAS 435:2116–2131
- Pawlowski MS, McGaugh SS, Jerjen H. 2015. MNRAS 453:1047–1061
- Pawlowski MS, Pflamm-Altenburg J, Kroupa P. 2012. MNRAS 423:1109–1126
- Peebles PJE. 1982. ApJ 263:L1–L5
- Peebles PJE. 2012. ARA&A 50:1–28

- Peter AHG, Rocha M, Bullock JS, Kaplinghat M. 2013. MNRAS 430:105–120
- Phillips JI, Cooper MC, Bullock JS, Boylan-Kolchin M. 2015. MNRAS 453:3839–3847
- Planck Collaboration, Ade PAR, Aghanim N, Arnaud M, Ashdown M, et al. 2016. A&A 594:A13
- Polisensky E, Ricotti M. 2011. Phys. Rev. D 83:043506
- Pontzen A, Governato F. 2012. MNRAS 421:3464–3471
- Porter TA, Johnson RP, Graham PW. 2011. ARA&A 49:155–194
- Press WH, Schechter P. 1974. ApJ 187:425–438
- Primack JR. 2012. *Annalen der Physik* 524:535–544
- Read JI. 2014. *Journal of Physics G Nuclear Physics* 41:063101
- Read JI, Agertz O, Collins MLM. 2016. MNRAS 459:2573–2590
- Rees MJ, Ostriker JP. 1977. MNRAS 179:541–559
- Richardson JC, Irwin MJ, McConnachie AW, Martin NF, Dotter AL, et al. 2011. ApJ 732:76
- Rocha M, Peter AHG, Bullock JS, Kaplinghat M, Garrison-Kimmel S, et al. 2013. MNRAS 430:81–104
- Rodríguez-Puebla A, Behroozi P, Primack J, Klypin A, Lee C, Hellinger D. 2016. MNRAS 462:893–916
- Rubin VC, Ford Jr. WK, Thonnard N. 1980. ApJ 238:471–487
- Rubin VC, Thonnard N, Ford Jr. WK. 1978. ApJ 225:L107–L111
- Sales LV, Navarro JF, Kallivayalil N, Frenk CS. 2017. MNRAS 465:1879–1888
- Sand DJ, Crnojević D, Bennet P, Willman B, Hargis J, et al. 2015a. ApJ 806:95
- Sand DJ, Spekkens K, Crnojević D, Hargis JR, Willman B, et al. 2015b. ApJ 812:L13
- Sawala T, Frenk CS, Fattahi A, Navarro JF, Bower RG, et al. 2016. MNRAS 457:1931–1943
- Schaye J, Crain RA, Bower RG, Furlong M, Schaller M, et al. 2015. MNRAS 446:521–554
- Schneider A, Smith RE, Macciò AV, Moore B. 2012. MNRAS 424:684–698
- Schneider A, Trujillo-Gomez S, Papastergis E, Reed DS, Lake G. 2016. *arXiv:1611.09362 [astro-ph]*
- Schultz C, Oñorbe J, Abazajian KN, Bullock JS. 2014. MNRAS 442:1597–1609
- Sheth RK, Mo HJ, Tormen G. 2001. MNRAS 323:1–12
- Shi X, Fuller GM. 1999. Phys. Rev. Lett. 82:2832–2835
- Simon JD, Bolatto AD, Leroy A, Blitz L, Gates EL. 2005. ApJ 621:757–776
- Simon JD, Geha M, Minor QE, Martinez GD, Kirby EN, et al. 2011. ApJ 733:46
- Skillman ED, Monelli M, Weisz DR, Hidalgo SL, Aparicio A, et al. 2017. ApJ 837:102
- Somerville RS, Davé R. 2015. ARA&A 53:51–113
- Spergel DN, Steinhardt PJ. 2000. *Physical Review Letters* 84:3760–3763
- Springel V, Wang J, Vogelsberger M, Ludlow A, Jenkins A, et al. 2008. MNRAS 391:1685–1711
- Stadel J, Potter D, Moore B, Diemand J, Madau P, et al. 2009. MNRAS 398:L21–L25
- Strigari LE. 2013. Phys. Rep. 531:1–88
- Strigari LE, Bullock JS, Kaplinghat M, Simon JD, Geha M, et al. 2008. Nature 454:1096–1097
- Tollerud EJ, Boylan-Kolchin M, Bullock JS. 2014. MNRAS 440:3511–3519
- Tollerud EJ, Bullock JS, Strigari LE, Willman B. 2008. ApJ 688:277–289
- Tollerud EJ, Geha MC, Grcevich J, Putman ME, Weisz DR, Dolphin AE. 2016. ApJ 827:89
- Tollet E, Macciò AV, Dutton AA, Stinson GS, Wang L, et al. 2016. MNRAS 456:3542–3552
- Tomozeiu M, Mayer L, Quinn T. 2016. ApJ 827:L15
- Tremaine S, Gunn JE. 1979. Phys. Rev. Lett. 42:407–410
- Trujillo-Gomez S, Schneider A, Papastergis E, Reed DS, Lake G. 2016. *arXiv:1610.09335 [astro-ph]*
- Tully RB, Libeskind NI, Karachentsev ID, Karachentseva VE, Rizzi L, Shaya EJ. 2015. ApJ 802:L25
- van Dokkum PG, Abraham R, Merritt A, Zhang J, Geha M, Conroy C. 2015. ApJ 798:L45
- Vegetti S, Koopmans LVE, Bolton A, Treu T, Gavazzi R. 2010. MNRAS 408:1969–1981
- Vegetti S, Lagattuta DJ, McKean JP, Auger MW, Fassnacht CD, Koopmans LVE. 2012. Nature 481:341–343
- Venumadhav T, Cyr-Racine FY, Abazajian KN, Hirata CM. 2016. Phys. Rev. D 94:043515
- Viel M, Becker GD, Bolton JS, Haehnelt MG. 2013. Phys. Rev. D 88:043502

- Viel M, Lesgourgues J, Haehnelt MG, Matarrese S, Riotto A. 2005. *Phys. Rev. D* 71:063534
- Vogelsberger M, Genel S, Springel V, Torrey P, Sijacki D, et al. 2014. *MNRAS* 444:1518–1547
- Vogelsberger M, Zavala J, Cyr-Racine FY, Pfrommer C, Bringmann T, Sigurdson K. 2016. *MNRAS* 460:1399–1416
- Vogelsberger M, Zavala J, Loeb A. 2012. *MNRAS* 423:3740–3752
- Walker MG, Mateo M, Olszewski EW, Peñarrubia J, Wyn Evans N, Gilmore G. 2009. *ApJ* 704:1274–1287
- Wechsler RH, Bullock JS, Primack JR, Kravtsov AV, Dekel A. 2002. *ApJ* 568:52–70
- Weinberg DH, Bullock JS, Governato F, Kuzio de Naray R, Peter AHG. 2015. *Proceedings of the National Academy of Science* 112:12249–12255
- Weinberg DH, Mortonson MJ, Eisenstein DJ, Hirata C, Riess AG, Rozo E. 2013. *Phys. Rep.* 530:87–255
- Wetzel AR, Hopkins PF, Kim Jh, Faucher-Giguère CA, Kereš D, Quataert E. 2016. *ApJ* 827:L23
- Wheeler C, Oñorbe J, Bullock JS, Boylan-Kolchin M, Elbert OD, et al. 2015. *MNRAS* 453:1305–1316
- White SDM, Frenk CS. 1991. *ApJ* 379:52–79
- White SDM, Frenk CS, Davis M. 1983. *ApJ* 274:L1–L5
- White SDM, Rees MJ. 1978. *MNRAS* 183:341–358
- Willman B. 2010. *Advances in Astronomy* 2010:285454
- Willman B, Strader J. 2012. *AJ* 144:76
- Wolf J, Martinez GD, Bullock JS, Kaplinghat M, Geha M, et al. 2010. *MNRAS* 406:1220–1237
- Wright AH, Robotham ASG, Driver SP, Alpaslan M, Andrews SK, et al. 2017. *MNRAS* 470:283–302
- Yoon JH, Johnston KV, Hogg DW. 2011. *ApJ* 731:58
- Zavala J, Jing YP, Faltenbacher A, Yepes G, Hoffman Y, et al. 2009. *ApJ* 700:1779–1793
- Zavala J, Springel V, Boylan-Kolchin M. 2010. *MNRAS* 405:593–612
- Zentner AR. 2007. *International Journal of Modern Physics D* 16:763–815
- Zentner AR, Kravtsov AV, Gnedin OY, Klypin AA. 2005. *ApJ* 629:219–232
- Zolotov A, Brooks AM, Willman B, Governato F, Pontzen A, et al. 2012. *ApJ* 761:71
- Zwicky F. 1933. *Helvetica Physica Acta* 6:110–127