Snowmass2021 Cosmic Frontier White Paper: Dark Matter Physics from Halo Measurements

Keith Bechtol\textsuperscript{1}, Simon Birrer\textsuperscript{2,3,4}, Francis-Yan Cyr-Racine\textsuperscript{5}, Katelin Schutz\textsuperscript{6}, Susmita Adhikari\textsuperscript{7,8}, Arka Banerjee\textsuperscript{7,9}, Simeon Bird\textsuperscript{10}, Nikita Blinov\textsuperscript{11}, Kimberly K. Boddy\textsuperscript{12}, Celine Boehm\textsuperscript{13}, Kevin Bundy\textsuperscript{14,15}, Malte Buschmann\textsuperscript{16}, Sukanya Chakrabarti\textsuperscript{17}, David Curtin\textsuperscript{18}, Liang Dai\textsuperscript{19}, Alex Drlica-Wagner\textsuperscript{8,9,20}, Cora Dvorkin\textsuperscript{21}, Adrienne L. Erickcek\textsuperscript{22}, Daniel Gilman\textsuperscript{23}, Saniya Heeba\textsuperscript{6}, Stacy Kim\textsuperscript{24}, Vid Iršič\textsuperscript{25,26}, Alexie Leauthaud\textsuperscript{14}, Mark Lovell\textsuperscript{27}, Zarija Lučić\textsuperscript{28}, Yao-Yuan Mao\textsuperscript{29}, Sidney Man\textsuperscript{2,3}, Andrea Mitridate\textsuperscript{30}, Philip Mocz\textsuperscript{31}, Julian B. Muñoz\textsuperscript{22}, Ethan O. Nadler\textsuperscript{3,34}, Annika H. G. Peter\textsuperscript{35}, Adrian Price-Whelan\textsuperscript{36}, Andrew Robertson\textsuperscript{37}, Nashwan Sabil\textsuperscript{38}, Neelima Sehgal\textsuperscript{39}, Nora Shipp\textsuperscript{8,20}, Joshua D. Simon\textsuperscript{33}, Rajeev Singh\textsuperscript{42}, Ken Van Tilburg\textsuperscript{35,40}, Risa H. Wechsler\textsuperscript{2,3,4}, Axel Widmark\textsuperscript{11}, and Hai-Bo Yu\textsuperscript{10}

\textsuperscript{1}Physics Department, University of Wisconsin-Madison, Madison, WI 53706, USA
\textsuperscript{2}Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA
\textsuperscript{3}Department of Physics, Stanford University, Stanford, CA 94305, USA
\textsuperscript{4}SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
\textsuperscript{5}Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87106, USA
\textsuperscript{6}Department of Physics & McGill Space Institute, McGill University, Montréal, QC H3A 2T8, Canada
\textsuperscript{7}Department of Physics, Indian Institute of Science Education and Research, Pashan, Pune 411008, India
\textsuperscript{8}Department of Astronomy and Astrophysics, University of Wisconsin-Madison, Madison, WI 53706, USA
\textsuperscript{9}Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
\textsuperscript{10}Department of Physics and Astronomy, University of California, Santa Cruz, CA 95064, USA
\textsuperscript{11}Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 5C2, Canada
\textsuperscript{12}Department of Physics, The University of Texas at Austin, Austin, TX 78712, USA
\textsuperscript{13}Sydney Consortium for Particle Physics and Cosmology, School of Physics, The University of Sydney, NSW 2006, Australia
\textsuperscript{14}Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
\textsuperscript{15}University of California Observatories, University of California, Santa Cruz, CA 95064, USA
\textsuperscript{16}Department of Physics, Princeton University, Princeton, NJ 08544, USA
\textsuperscript{17}School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA
\textsuperscript{18}Department of Physics, University of Toronto, Toronto ON, M5S 3H4, Canada
\textsuperscript{19}Department of Physics, University of California, Berkeley, CA 94720, USA
\textsuperscript{20}Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
\textsuperscript{21}Department of Physics, Harvard University, Cambridge, MA 02138, USA
\textsuperscript{22}Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA
\textsuperscript{23}Department of Astronomy and Astrophysics, University of Toronto, Toronto ON, M5S 3H4, CA
\textsuperscript{24}Department of Physics, University of Surrey, Guildford, GU2 7XH, UK
\textsuperscript{25}Kavli Institute for Cosmology, University of Cambridge, Cambridge CB3 0HA, UK
\textsuperscript{26}Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK
\textsuperscript{27}Center for Astrophysics and Cosmology, Science Institute, University of Iceland, 107 Reykjavik, Iceland
\textsuperscript{28}Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
\textsuperscript{29}Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA
\textsuperscript{30}Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA
\textsuperscript{31}Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
\textsuperscript{32}Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
\textsuperscript{33}Carnegie Observatories, Pasadena, CA 91101, USA
\textsuperscript{34}Department of Physics & Astronomy, University of Southern California, Los Angeles, CA, 90007, USA
\textsuperscript{35}CCAPP, Department of Physics, Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA
\textsuperscript{36}Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
\textsuperscript{37}Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
\textsuperscript{38}Department of Physics, King’s College London, Strand, London WC2R 2LS, UK
\textsuperscript{39}Physics and Astronomy Department, Stony Brook University, Stony Brook, NY 11794, USA
\textsuperscript{40}Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, NY 10003, USA
\textsuperscript{41}Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, 2200 Copenhagen N, Denmark
\textsuperscript{42}Institute of Nuclear Physics Polish Academy of Sciences, PL-31-342 Kraków, Poland
Abstract

The non-linear process of cosmic structure formation produces gravitationally bound overdensities of dark matter known as halos. The abundances, density profiles, ellipticities, and spins of these halos can be tied to the underlying fundamental particle physics that governs dark matter at microscopic scales. Thus, macroscopic measurements of dark matter halos offer a unique opportunity to determine the underlying properties of dark matter across the vast landscape of dark matter theories. This white paper summarizes the ongoing rapid development of theoretical and experimental methods, as well as new opportunities, to use dark matter halo measurements as a pillar of dark matter physics.

Contents

1 Executive Summary 3

2 Classes of Dark Matter Behavior 5
   2.1 Ab Initio Impact on Structure Formation .......................... 6
      2.1.1 Thermal history .............................................. 6
      2.1.2 Non-thermal history ........................................ 8
      2.1.3 Early departures from radiation domination ................. 9
      2.1.4 Interactions with the SM ................................... 10
   2.2 In Situ Impact on Structure Formation ............................ 12
      2.2.1 Self-interacting dark matter ................................. 12
   2.3 Hybrid Impact on Structure Formation ............................ 14
      2.3.1 Composite dark matter ..................................... 14
      2.3.2 Low-mass bosons ........................................... 16

3 Cluster Scale 17
   3.1 Scale and Objects .............................................. 17
   3.2 Observational Probes ........................................... 17
      3.2.1 Density profiles of clusters .............................. 17
      3.2.2 Merging clusters ........................................... 18
      3.2.3 Brightest cluster galaxy oscillations ...................... 20
      3.2.4 Substructure in massive clusters ......................... 21
   3.3 Theory Connections ............................................. 21

4 Galaxy Scale 21
   4.1 Scale and Objects .............................................. 21
   4.2 Observational Probes ........................................... 21
      4.2.1 Dwarf galaxy lensing ...................................... 22
      4.2.2 Milky Way stellar dynamics ............................... 24
      4.2.3 Galactic rotation curves .................................. 25
      4.2.4 Ultraviolet luminosity function ........................... 27
   4.3 Theory Connections ............................................. 28
1 Executive Summary

The standard collisionless, cold dark matter (CDM) paradigm predicts that dark matter in the late-time Universe is clustered into gravitationally bound overdensities called halos. Dark matter halos are distributed along a roughly power-law mass spectrum extending from galaxy clusters ($10^{15} M_\odot$) to ultra-faint dwarf galaxies ($10^{8} M_\odot$). Furthermore, CDM predicts that dark matter clusters into gravitationally bound halos that are not massive enough to host galaxies, with the mass spectrum of halos extending to sub-solar mass scales. While CDM provides an excellent parametric description of the Universe on large spatial scales, it does not provide any underlying mechanism for the production or properties of dark matter. Many theoretical models describing the particle physics of dark matter predict that the simple collisionless CDM model breaks down at small physical scales. This paper highlights the key science opportunities and observational milestones for exploring the fundamental nature of dark matter using astrophysical measurements of dark matter halos in the next decade.

Key Science Opportunities

1. Sensitivity to a broad range of models: To date, all positive experimental evidence for the existence and properties of dark matter come from astrophysical observations. Measurements of the abundance, density profiles, and spatial distribution of dark matter halos offer sensitivity to an enormous range of dark matter models,
and are complementary to both terrestrial experiments and indirect searches for dark matter annihilation and decay products.

2. **Rapidly advancing phenomenology:** Our understanding of how the fundamental properties of dark matter at a microscopic scale impact structure formation throughout cosmic history is rapidly advancing. There is enormous potential to further develop the detailed phenomenology for a broader range of dark matter models, and to explore new regions of theory space with new and archival data.

3. **Connection between dark matter halos and the early Universe:** The seeds of cosmological structure formation were established in the earliest moments after the Big Bang. As we measure the distribution of dark matter across a broader range of physical scales, we simultaneously learn about the initial conditions of the Universe and probe periods of cosmic history that might be inaccessible by other means.

### Key Experimental Milestones

1. **Precision measurements of galaxy-scale dark matter halos:** Current and near-future facilities will provide a detailed mapping between luminous galaxies and their invisible dark matter halos across 13 billions years of cosmic history and 7 orders of magnitude in dark matter halo mass. Detailed measurements of halo abundances and density profiles will provide a stringent test of the CDM paradigm.

2. **Population studies of completely dark halos:** Within the next decade, several observational techniques will become sensitive to dark matter halos at or below the minimum mass required to host stellar populations. Such completely dark halos offer unique advantages to constrain the microphysical properties of dark matter because their evolution is less affected by baryonic physics. Furthermore, many theoretical models of dark matter predict conspicuous deviations from collisionless CDM in low-mass halos.

3. **Testing dark models that predict enhanced central densities in solar- and planetary-mass-scale halos:** A suite of innovative and ambitious observational techniques can be used to search for solar- and planetary-mass-scale halos via their subtle gravitational effects. The discovery of such low mass halos would extend our knowledge of the matter power spectrum by many orders of magnitude in physical scale and would transform our understanding of both dark matter properties and the early Universe.

This white paper is organized as follows. In Section 2, we enumerate several broad classes of DM behavior that can arise in different theories of DM. The models under discussion are not meant to be exhaustive, and we do not commit to particular theories at the level of a Lagrangian; rather, we attempt to “coarse grain” in theory space to capture the most essential DM properties that can observably affect DM halos, with such properties then mapping back on to specific models. Having established the diverse range of possible DM behaviors, we describe the discovery potential of DM halo and subhalo observations spanning nearly thirty orders of magnitude in halo mass and highlight the complementarity of these scales to different kinds of DM behavior. We group the DM halos and subhalos into the
following categories: cluster scale, galactic scale, the scale below the threshold of galaxy formation, and sub-stellar scale. These correspond to Sections 3 – 6, respectively. Brief concluding remarks follow in Section 7.

2 Classes of Dark Matter Behavior

The Standard Model (SM) of particle physics has been a stunning success, with certain predictions of the theory having been validated out to many decimal places. However, the SM (as it is presently known) lacks a description of dark matter (DM), whose presence can be inferred via its gravitational influence in a wide variety of astrophysical environments. The astrophysical discovery of DM has inspired a concerted multi-disciplinary effort to characterize the microphysical properties of DM. To date, however, all attempts to conclusively pin down the nature of DM have yielded null results, indicating that any non-gravitational behavior (for instance, interactions with particles in the SM) must either be very weak or arise only under specific conditions that have not yet been explored. The growing number of null searches for DM in various experimental settings has inspired fresh theoretical ideas over the past decade since the last Snowmass process. It is now more clear than ever that the landscape of potential DM candidates is unfathomably diverse; for instance, if the DM is comprised of a single constituent (e.g., one particle species), then the allowed range of DM particle mass spans roughly ninety orders of magnitude and is constrained only by astrophysical observations. Many plausible theories of DM are not so simple: there is a very real possibility that DM exists as part of an entire “dark sector,” with additional particles and forces and with a rich structure analogous to the SM.

The complexity of the space of plausible DM theories is difficult to quantify, but it is abundantly clear that a one-size-fits-all approach is not suitable in any empirical attempts to characterize the nature of DM. Halos and subhalos — loosely defined for the purposes of this White Paper as gravitationally self-bound clumps of DM — are the most promising systems for testing many broad classes of DM behavior. Halo observations can characterize the properties of dark matter using gravity, the only force through which DM is known to interact, since halos have large enough accumulations of DM to leave a gravitational imprint on the SM matter that we can observe. Halos can also be incredibly diverse and can appear on different lengthscales, at different times in the history of the Universe, and in different ambient environments, providing a wide range of conditions for testing many different theoretical possibilities.

Given everything that we have inferred about cosmological structure formation and DM halos from observation, we know that on sufficiently large scales greater than $O(\text{Mpc})$ the DM can be described within the paradigm of collisionless, cold DM (CDM) where the DM has negligible interactions and negligible thermal velocity. However, many theories of DM predict deviations from this CDM paradigm that would be manifest in DM halos on a range of scales. We broadly classify these deviations into two categories: (1) ab initio deviations, where early-Universe DM behavior modifies the initial conditions for structure formation and hence alters the formation of DM halos and subhalos and (2) in situ deviations where a relatively high density of DM in the present-day Universe allows certain DM behaviors to become manifest.
In the remainder of this section, we discuss several broad features of DM theories with signatures from \textit{ab initio} DM behavior, \textit{in situ} DM behavior, and \textit{hybrid} models that are affected by both \textit{ab initio} and \textit{in situ} behaviors. This section is not meant to be exhaustive but rather serves to capture the diversity of possible DM characteristics that can affect observations of DM halos.

\section*{2.1 \textit{Ab Initio} Impact on Structure Formation}

\subsection*{2.1.1 Thermal history}

The cosmological clustering of dark matter on the scale of dark matter halos is set in large part by the DM thermal history in the early Universe. CDM predicts an abundance of halos down to arbitrarily small mass scales, but this will not be the case even for the simplest plausible particle physics models of DM. Thermal DM that becomes non-relativistic soon after its relic abundance is set (i.e. via decoupling from the SM plasma) is able to condense into low mass structures. Conversely, if the dark matter retains a relatively large kinetic energy, it will free-stream through low-mass density perturbations with low escape velocities and thus inhibit the formation of halos until the DM speed drops, either via kinetic decoupling or Hubble damping.

If DM is made of weakly interacting massive particles (WIMPs) that are produced while non-relativistic via the freeze-out mechanism, then kinetic decoupling typically happens close to the freeze-out temperature for chemical decoupling, $T_f \sim m_{\text{DM}}/20$. Many supersymmetric theories naturally contain WIMP DM candidates, whose kinetic decoupling from the SM plasma would leave an imprint on structure formation on very small scales. For instance, for a bino-like neutralino with a mass around 100 GeV, the smallest halos that could form would be around an Earth mass \cite{10, 11}; structure on even smaller scales would be erased by the thermal motions of DM prior to decoupling. Thermal freeze-out production is excluded for DM below $\sim 10$ MeV based on observations of the CMB and from BBN \cite{12–18}, so exploring new freeze-out parameter space from structure formation will remain an enticing challenge for the future.

In a scenario where DM decouples from the SM while still relativistic, which is thermal history very similar to that of neutrinos, then the suppression of structure can occur on even larger scales. In this case, which is known as warm DM (WDM), structure can only form once the DM velocity is sufficiently Hubble damped, which suppresses structure on small scales \cite{19–21}. Currently, these effects are being probed primarily on scales near the threshold of galaxy formation, as discussed in Sec. 5. This spatial scale is too large to probe WIMP decoupling, however headway is being made in constraining new WDM parameter space above the keV mass scale and will continue over the next decade up to masses of $\sim 20$ keV \cite{22–36}.

A final thermal history relevant for structure formation is DM freeze-in, where DM is produced non-thermally from SM particles in the thermal bath of the early Universe annihilating and decaying. A classic example of a DM candidate born via freeze-in is the sterile neutrino \cite{37, 38}, but freeze-in is now known to be a more general mechanism that can be realized in a number of different models in the limit of small DM-SM couplings \cite{39–41}. Freeze-in DM below the GeV mass scale is a key benchmark of the low-mass DM direct detec-
Figure 1: The dimensionless linear matter power spectrum extrapolated to $z = 0$. Theoretical predictions are plotted for four models that suppress power: (1) ultra-light axion “fuzzy” dark matter with a mass $m = 10^{-22}$ eV (magenta; [1]), (2) dark matter–baryon interactions with interaction cross section that scales with velocity as $\sigma_0 v^{-4}$ for $\sigma_0 = 10^{-22}$ cm$^2$ (blue; [2]), (3) thermal relic warm dark matter with a mass $m \sim 40$ keV (red; [3]), (4) weakly interacting massive particle dark matter represented by a bino-like neutralino with a mass $m \sim 100$ GeV (black; [4]). Also shown are four models that enhance power on very small scales: (1) early matter domination assuming a reheat temperature of 10 MeV (green; [5]), (2) post-inflationary production of QCD axions dominated by the misalignment mechanism (pink; [6]), (3) vector dark matter produced during inflation assuming an inflationary scale of $10^{14}$ GeV and a DM mass of $10^{-6}$ eV (orange; [7]), and (4) post-inflationary production of axions dominated by strings (cyan; [8]). Note that the position of the power spectrum cutoff and/or enhancement depends on model parameters and is flexible for most cases shown here. Power spectrum measurements on large scales are compiled from [9]. Shaded vertical bands roughly indicate the characteristic kinds of halos formed on each scale, and the horizontal axes indicate wavenumber, halo mass, and the temperature of the Universe when that mode entered the horizon.

...
the exploration of this parameter space is a timely goal for the coming decade.

2.1.2 Non-thermal history

Besides the thermal history of dark matter, there are a range of non-thermal scenarios with intriguing signatures. One of such scenarios involves axion minihalos which occur naturally for any axion as long as the underlying Pecci–Quinn symmetry is broken after inflation. These axion minihalos are seeded purely by self-interactions and topological defects in the radiation era [49–68], where overdensities in the axion field then collapse gravitationally at later times during the matter era [69–76]. The resulting halos are small, with a typical mass of $10^{-10} M_\odot$ and a Solar-System-sized radius. As the axion is a cold dark matter candidate, these minihalos form in addition to the much heavier thermally created halos. Despite their miniscule size, axion minihalos are likely to hinder direct detection experiments as a large amount of mass could be bound inside of them [77–81]. However, several entirely new search strategies have been highlighted recently as well [82–92]. First proper estimates for the mass spectrum and abundance of axion minihalos have been obtained only recently. As this often requires large-scale numerical simulations due to non-linearities in the axion equation of motion, more computational work is needed in the future to confirm and refine these results. Furthermore, these halos are volatile against tidal disruptions within the Milky Way halo, though recent studies have found that a decent amount can survive until the present day [93–97]. More simulation work is needed on all fronts to properly understand axion minihalos and to confirm the feasibility of experimental search strategies.

If the Pecci–Quinn symmetry is broken before inflation, it is still possible for various compact structures to emerge non-gravitationally should the initial misalignment angle be large [98]. For axion masses above $10^{-5}$ eV those compact objects can be numerous enough to disrupt direct detection experiments. This does not only lead to unique new signatures but could also explain the non-observation of axions so far due to the potential lack of unbound mass. For smaller axion masses between $10^{-22}$ eV and $10^{-5}$ eV possible observable signatures include gravitational waves, disruptions of baryonic structures and an accelerated early Universe star formation. Numerical simulations will be required to quantitatively understand these signatures as the fraction of bound mass and the exact halo mass spectrum is not known.

Inflationary particle production provides another compelling non-thermal mechanism for the enhancement of small-scale structure. This is a very minimal and generic way of producing the DM of the Universe since it only requires the gravitational coupling of a stable particle with mass less than the inflationary scale, $H_I$. Particle production occurs because the early-time vacuum state does not correspond to a vacuum state for an observer at late times in an expanding background. The resulting power spectrum of field fluctuations is strongly dependent on the spin of the particle; since the fluctuations of the DM field are uncorrelated with the inflaton, the DM density perturbations are isocurvature. The latter fact is important because, for example, gravitational production of scalar particles $\phi$ leads to a scale-invariant dimensionless power spectrum, $\Delta_\phi^2 = (H_I/2\pi)^2$, which can conflict with CMB isocurvature constraints [99, 100]. In contrast, the inflationary production of a (longitudinal) vector boson, $A_L$, gives rise to a fluctuation spectrum $\Delta_{A_L}^2 = [kH_I/(2\pi m)]^2$ which is negligible at CMB scales (i.e., low-$k$), thereby automatically avoiding isocurvature constraints [7].
Gravitational particle production also predicts the relic abundance as a function of the boson mass and the scale of inflation: $\Omega_{A_L}/\Omega_{cdm} \sim \sqrt{m/(10^{-5}\text{eV})}[H_I/(10^{14}\text{ GeV})^2]^{3/2}$ [7, 101–103]; CMB bounds on $H_I$ imply a lower bound $m \gtrsim 10^{-5}\text{ eV}$ to saturate the relic abundance via this mechanism.\(^1\) The scale-dependent field spectrum produces a density power spectrum that peaks at $k \sim a_* m$ with an $\mathcal{O}(1)$ magnitude, where $a_*$ is the scale factor when $H(a_*) = m$; the density power spectrum falls off as $k^3$ below the peak, and as $k^{-1}$ above. This leads to the gravitational formation of minihalos beginning near the time of matter-radiation equality (MRE). Note that because the power spectrum is enhanced for a wide range of scales (not just at $k \sim a_* m$), and the smallest minihalos formed at MRE are hierarchically assembled into larger and larger minihalos, resulting in a broad minihalo mass function [104].

2.1.3 Early departures from radiation domination

There is a troubling gap in the cosmological record: we do not know how the Universe evolved between the end of inflation and the onset of Big Bang nucleosynthesis (BBN) [105]. While it is customary to assume that the Universe became radiation dominated shortly after inflation ended and remained so for the next 50,000 years, there is no evidence than the expansion history of the early Universe was so straightforward. As long as the Universe was radiation dominated at a temperature of $\sim 5\text{ MeV}$, thermal neutrino production is sufficient to produce the observed abundances of light elements [106] and the observed CMB and matter power spectra [107]. Furthermore, deviations from radiation domination prior to BBN are strongly motivated. After inflation ends, it is generally assumed that the inflaton oscillates and then decays, transferring its energy to a radiative plasma [108]. While a scalar field oscillates around the minimum of a quadratic potential, it has zero pressure, so most inflationary scenarios include an early matter-dominated era (EMDE) between the end of inflation and the decay of the inflaton. After the inflaton decays into radiation, other oscillating scalar fields may come to dominate the energy density of the Universe; this phenomenon is a generic consequence of stabilized moduli in string theories [109]. It is also possible that unstable massive particles could temporarily dominate the Universe prior to BBN, and hidden sector theories of dark matter frequently include EMDEs [110–116]. Another possible deviation from radiation domination is a period of kination, during which the kinetic energy of a scalar field dominates the energy density of the Universe [117–119].

Our ignorance of the pre-BBN expansion history profoundly limits our understanding of the origins of the baryon asymmetry [e.g. 120–122] and the origins of dark matter. If dark matter chemically decouples during kination, $\langle \sigma v \rangle \gtrsim 3 \times 10^{-26}\text{ cm}^3\text{ s}^{-1}$ is required to generate the observed dark matter density [123–126]. Conversely, entropy production during an EMDE dilutes the relic abundance of dark matter, so a smaller annihilation cross section is required to generate the observed dark matter density [120]. However, if the dark matter is also produced nonthermally (as a decay product, for instance), $\langle \sigma v \rangle$ may need to be increased so that any excess dark matter is eliminated [127]. We are therefore faced with a dark degeneracy: we cannot determine the annihilation cross section required to produce the observed dark matter density without knowing the thermal history of the pre-BBN

\(^{1}\)The calculation of the relic abundance is modified if post-inflationary reheating is non-instantaneous [103].
A period of kination or an EMDE also relaxes constraints on axion [120, 128, 129] and neutrino [120, 130, 131] dark matter. Fortunately, sub-solar-mass dark matter microhalos could provide a unique window into the pre-BBN Universe that can be used to fill in this gap in the cosmological record. During an EMDE or a period of kination, subhorizon perturbations in the dark matter density grow much faster than they would during radiation domination [5, 132]. If $T_{RH}$ is the radiation temperature when the Universe became radiation dominated, then an EMDE or a period of kination enhances the abundance of microhalos with masses less than $30M_{\odot} \left( \frac{10\text{ MeV}}{T_{RH}} \right)^3$ [5]. The formation time and size of the smallest microhalos depends on the temperature of the dark matter particles [133–136], the properties of the particles that dominated the energy density during the EMDE [115, 116, 137], and the duration of the EMDE [137, 138]. If dark matter is sufficiently cold and density perturbations remain in the linear regime during the EMDE, most of the dark matter is contained in microhalos at redshifts as high as 200, long before halos are expected to form in standard cosmologies [135, 137, 139]. These early-forming microhalos are extremely dense, and they boost the dark matter annihilation rate by several orders of magnitude. Consequently, the upper bounds on the dark matter annihilation rate established by gamma-ray observations can be used to constrain EMDEs [135, 137, 139, 140]. In the future, EMDEs may also be constrained by pulsar timing arrays [141, 142] and observations of stellar microlensing within galaxy clusters [104], as detailed in Section 6.

2.1.4 Interactions with the SM

While we only have gravitational evidence for dark matter so far, it is entirely plausible that dark matter interacts with SM particles. There have been numerous efforts to search for interactions between particle-like DM and the SM through collider experiments, direct detection, and indirect detection, but no concrete evidence has been found. Cosmological and astrophysical observations offer a variety of avenues to test DM physics in regions of parameter space that both overlap with and are inaccessible to the more traditional searches. One such avenue is to investigate the effects that interactions in the early Universe have on the subsequent formation of DM structure and substructure. In particular, DM elastic scattering with SM relativistic particles (i.e., photons and neutrinos) or nonrelativistic matter (generically referred to as “baryons”) generically hinder the growth of structure and thus reduce DM halo abundances on corresponding scales.

The impact of DM interactions with ordinary particles is two-fold. First, collisions permit heat and momentum exchange between dark matter and the species it interacts with, thereby damping their primordial fluctuations. Second, they alter the free-streaming properties of DM and of the interacting species, either by inducing free streaming of DM particles (mixed damping [143]) or by delaying free streaming. This second effect, important for light DM coupled to relativistic particles (i.e. photons and neutrinos), also affects the clumping of DM and of the interacting species. The overall effect of collisional damping (analogous to Silk damping) usually translates into a reduction of power at progressively smaller scales, thereby leading to a smaller number of small-scale structures both in the Universe and within complex objects such as large galaxies.

The magnitude of the effect can be understood intuitively by an analytical approximation
Dark matter Physics from Halo Measurements

of the damping scale (at least in the simplest case, where the viscosity is dominant), namely:

$$l_{cd}^2 \simeq \sum_i \int_{t_{\text{dec}}(\text{dm} - i)}^{t_{\text{dec}(\text{dm})}} \frac{\rho_i v_i^2}{a(t) \Gamma_i}$$

where $t_{\text{dec(dm)}}$ stands for the dark matter kinetic decoupling time with species $i$, $\rho_i$ is the energy density of species $i$, $v_i$ its velocity and $\Gamma_i$ its interaction rate that is $\Gamma_i = \sum_j \Gamma_{ij}$ and $a(t)$ is the scale-factor. The larger the density $\rho_i$ is and the lighter the species are, the bigger the effect is — unless the dark matter decouples from $i$ quickly. Since the DM decoupling is given that the condition $\sigma_{v_{\text{dm} - i}} n_i \simeq H$ (with $n_i$ the number density of species $i$ and $\sigma_{v_{\text{dm} - i}}$ the elastic scattering cross section between the DM and species $i$), one can immediately see that the damping scale is the largest for the two relativistic species of the SM of Particle Physics, namely the neutrinos and photons which turn out to also have the largest densities $n_i$ and $\rho_i$; unless $\sigma_{v_{\text{dm} - i}}$ is small. In fact studies have shown that, in scenarios where the DM interacts with photons [143–147] or neutrinos [143, 145, 148–151] with a relatively large cross section, the Universe may not contain any small galaxies and, in extreme scenarios where the cross section is too large, the Universe may not even contain enough large galaxies, thereby providing a means to exclude certain interaction strengths, otherwise invisible by more conventional means.

Dark matter–baryon interactions produce a similar effect due to collisional damping [29, 30, 152–156]. Since the baryons are nonrelativistic and their density is relatively small, damping effects are negligible at large scales. At small scales, DM interactions produce a cutoff in the matter power spectrum, relative to $\Lambda$CDM. The cutoff behavior is strikingly similar to that of warm dark matter [29], though there are subtleties related to the velocity dependence of the scattering cross section [154]. In particular, cross sections with a steep power-law velocity dependence recover substantial power after the initial cutoff, akin to dark acoustic oscillations. A small-scale suppression of the matter power spectrum implies the existence of a minimum halo mass. Constraints on the scattering cross section as a function of mass have been placed using current observations of Milky Way substructure [29, 30]. Future observatories will probe even smaller substructures, extending experimental sensitivity to DM interactions [157].

Once the dark matter interactions stopped, small-scale dark matter fluctuations can be further washed out by the dark matter free-streaming. Note that in most scenarios, the dark matter happens to be among the first species to kinetically decouple and thus the first to free-stream. However this is not necessarily true when the dark matter interacts with neutrinos and is relatively light (in the MeV range). There are in fact two interesting situations that are worth studying, namely: i) the neutrinos last interactions occur with the electrons, as it is expected in the Standard Model of Cosmology. In this case, the dark matter fluctuations are erased owing to the dark matter coupling to free-streaming neutrinos — this is called the mixed damping effect in [143]; ii) the neutrinos continue to interact with the dark matter after they stop interacting with electrons. In this case, the neutrinos experience a longer period of collisional damping than in the SM and the start of their free-streaming period is delayed. Both situations lead to interesting effects. However the second case in particular can increase the number of large-scale-structures, which eventually has an implication on the estimates of the cosmological parameters (in particular $N_{\text{eff}}$ and $H_0$).
2.2 **In Situ** Impact on Structure Formation

2.2.1 Self-interacting dark matter

In the prevailing CDM theory, DM is assumed to be cold and collisionless. CDM is extremely successful for the large-scale structure of the Universe, corresponding to distances greater than $O$(Mpc) today. On smaller scales, cosmological simulations have provided several predictions for the structure of CDM halos and abundance of substructures on these scales. However, it remains unclear whether these predictions are all consistent with observations. In particular, DM distributions in galactic halos are much more diverse than predicted in CDM; see [160] for a review on the small-scale issues of CDM.

A promising alternative to CDM is self-interacting DM (SIDM), see [158] for a review. In this scenario, DM particles scatter with each other through $2 \rightarrow 2$ interactions. Collisions conduct heat from the hotter outer to the cooler inner regions of a DM halo, leading to a radical deviation from CDM predictions for the inner halo structure [159, 161–163]. Fig. 3 shows profiles of SIDM velocity dispersions (left), densities (middle) and shapes (right), with (blue) and without (black) DM self-interactions. Compared to CDM, the self-interactions thermalize the inner halo, and produce a shallower density profile and a rounder shape, which can be probed in astrophysical observations over a wide range of galactic mass scales, see Sec. 3 for more details.

Recent studies show that SIDM predicts more diverse DM distributions than CDM in galactic systems. For spiral galaxies in the field, SIDM thermalization ties DM and baryonic distributions together and the inner halo density could be either cored or cuspy, depending on the baryon concentration [164]. This feature helps explain diverse galactic rotation curves of the spirals [165, 166], a long-standing challenge for CDM, see Sec. 4.2.3. For the substructures, the interplay between self-interactions and tidal interactions can change halo properties in a dynamical way, resulting in a large spread in the central density of subhalos [167–172]. For example, SIDM may naturally explain diffuse stellar distributions and low DM contents of newly-discovered ultra-diffuse galaxies [173, 174].

In addition, SIDM predicts gravothermal catastrophe [175], i.e., a phase of gravothermal evolution of the halo at which the central density becomes much higher than its CDM counterpart. For a typical halo, the timescale for the onset of the collapse is longer than the Hubble time. However, the process can be accelerated if the self-interactions have a dissipative component [176, 177], the halo is in the tidal environment [167, 168], or the baryon concentration is high [178, 179], leading to unique signatures in astrophysical observations. Refs. [180–185] further show that SIDM could explain the origin of supermassive black holes via the mechanism of gravothermal collapse.

The figure of merit for the self-interactions is the cross section per unit mass, $\sigma/m$. To have an observable effect on DM halos over cosmological timescales, the required $\sigma/m$ must be at least

$$\sigma/m \sim 1 \text{ cm}^2/\text{g} \approx 2 \text{ barns/GeV} \approx 2 \times 10^{-24} \text{ cm}^2/\text{GeV} \quad (1)$$

in dwarf galaxies, where characteristic velocities of DM particles are $v \lesssim 100 \text{ km/s}$. On the other hand, observations of galaxy clusters put an upper limit of $\sigma/m \lesssim 0.1 \text{ cm}^2/\text{g}$ for $v \sim 1000 \text{ km/s}$ [158, 186]. Thus the self-scattering cross section is necessarily velocity-dependent. From a particle physics perspective, the relevant SIDM cross section is many
Figure 2: A schematic depiction of different in situ DM microphysics effects on the morphology of DM halos and subhalos. See Sections 2.2 and 2.3 for further discussion.
orders of magnitude larger than the typical weak-scale cross section expected for a WIMP. Evidence for self-interactions would therefore point toward a new dark mediator particle that is much lighter than the weak scale [187–194]. The mediator must decay to massless degrees of freedom to avoid over closing the Universe. If it couples to the SM particles, SIDM could produce signals in terrestrial DM detection experiments [195]. On the other hand, if the mediator couples to massless particles in the dark sector, SIDM predicts dark acoustic oscillations, leading to a suppressed matter power spectrum [143, 192, 194, 196, 197].

2.3 Hybrid Impact on Structure Formation

2.3.1 Composite dark matter

For hints that the dark matter may not be described by the simplest model, one need look no further than the Standard Model (SM) of particle physics, which is very far from minimal. The multitudes of SM particles and forces give rise to a large variety of microscopic phenomena, such as composite hadronic and nuclear bound states as well as dissipation, atoms and chemistry. We probably would never have expected the complexity of the SM if we did not live in it. What other strange sectors could be making up a substantial fraction of the Universe’s matter budget? This bottom-up plausibility argument sets the stage for considering dark complexity, the possibility of hidden sectors with particles and forces that approach the SM in complexity, with possibly equally interesting dynamics in the early Universe and today.

Dark complexity and the theories that can be included in this umbrella term, such as composite dark matter [198–208] and atomic dark matter [176, 198, 209–224], are also deeply motivated from a top-down point of view. Many BSM theories postulate the existence of hidden sectors related by discrete symmetries to the SM [225–239]. This includes models of “neutral naturalness” such as the Mirror Twin Higgs that address the hierarchy problem of the Higgs boson mass [225–232, 240] by positing the existence of a hidden sector related to the SM by a $Z_2$ or other discrete symmetry, resulting in analogues of electromagnetism, strong and weak forces as well as various matter states existing in the dark sector.
estingly, while these dark sectors are qualitatively analogous to the SM in their ingredients, their dynamics can be very different owing to different effective values of the corresponding fundamental constants. This can result in a fraction of dark matter being made up of dark baryons, with dark protons, dark electrons, dark photons, and even dark nuclear forces that are similar to their SM counterparts but with possibly different values for masses and interaction strengths. Another solution to the hierarchy problem that may give rise to similar phenomena is $N$-Naturalness [241].

Composite DM [198–208] is a well-motivated DM scenario whereby the hidden sector is some analogue of QCD that could be as simple as some dark fermion(s) charged under a confining dark gauge force. All or part of DM could then be made up of a dark-hadronic bound state, with the details of its possible dynamics in our Universe today ranging from being effectively collisionless to acting like self-interacting dark matter [200, 203, 207, 208, 242] (see Section 2.2.1) or even the nucleonic part of atomic dark matter. As a result, composite DM can realize many rich signatures across the broad range of halo masses discussed in this white paper.

An archetypical scenario that leads to dissipation in DM halos is atomic dark matter [176, 198, 209–224], which postulates that dark matter contains at least two states with different masses and opposite charge under a dark electromagnetism $U(1)_D$ gauge symmetry. Straightforward generalizations include different charge ratios of the “dark nucleon” and the “dark electron,” and the possibility of many dark nucleon species. The latter scenario is especially motivated if atomic dark matter is realized together with composite DM, i.e. if the dark nucleons are actually dark-hadronic bound states which in turn form different nuclei, in analogy to the SM. The dark photon could be detected in future CMB-S4 and CMB-HD measurements [243, 244], and the long-range dark-electromagnetic interaction also causes dark acoustic oscillations of the atomic DM component (see e.g. [217, 230]), which can leave oscillatory deviations in Large Scale Structure matter power spectrum measurements compared to the $\Lambda$CDM expectation. However, by far the most drastic signatures of atomic dark matter and dissipation occur at galactic and smaller scales. They could easily be compatible with cosmological and self-interaction bounds, especially if the dark atoms make up a $\lesssim \mathcal{O}(10\%)$ subcomponent of dark matter [214, 217, 230]. Like their SM counterparts, dark particles could cool and form structure, leading to the formation a dark disk or dark microhalos [214, 215, 220, 245–247], which leave an observable gravitational imprint on visible matter that can be observed for instance using astrometry [248–251], as discussed further in Section 4.2.2.

On even smaller astrophysical scales, collapsing atomic dark matter would form mirror stars [236, 237, 252, 253], possibly emitting dark photons long past their Kelvin–Helmholz time if dark nuclear interactions release energy in their cores. This opens a new window for astrophysical dark matter searches, since it has recently been shown [252, 253] that mirror stars can produce electromagnetic signals that may be visible in optical/IR and X-ray telescopes if the dark photon has a tiny kinetic mixing with the SM photon. Such a small mixing $\epsilon F_{\mu \nu} F_{\mu \nu}^{D}$ is a renormalizable operator in the Lagrangian and generically expected to be produced at some loop level in the complete theory [254], but even without such a coupling, the gravitational waves of mirror neutron star mergers will be observable [255]. In the absence of dark nuclear interactions, atomic dark matter can collapse directly to black holes [245–247] (with very different masses than expected from SM astrophysics), leading to
detectable gravitational waves in their mergers [256, 257]. The Vera Rubin telescope will also be able to detect the microlensing signal of mirror stars in our galaxy [258]. Finally, detection of dissipative atomic dark matter in our stellar neighborhood either via ground-based detectors [231] or anomalous cooling of white dwarfs [259] depends on its distribution in the galaxy and the degree to which it has undergone cooling, recombination, and collapse into a dark disk. A positive detection in any combination of the abovementioned channels would therefore not only reveal the existence of dark matter, but allow us to directly probe details of its distribution within our galaxy, which would reveal aspects of hidden sector dynamics that cannot be probed in other experiments or observations.

2.3.2 Low-mass bosons

One hypothesis postulates that dark matter halos are comprised of ultralight bosons [260–262], which are broadly expected to arise in the context of many string theories [263–267]. This so-called Fuzzy Dark Matter (FDM) model introduces wave dynamics in the dark matter on the scale of the de Broglie wavelength due to the extraordinarily high particle occupation number. Lately, this idea has seen increased attention because bosons with a mass of $\sim 10^{-22}$ eV have a de Broglie wavelength on kiloparsec scales and the resulting “quantum pressure” (which arises due to the macroscopic de Broglie wavelength and not because of additional forces) affects the galaxy-scale distribution of dark matter, which may have an impact on the “small-scale” challenges of the CDM paradigm [268, 269]. Numerical simulations testing the nonlinear structure formation and evolution of this dark matter model have recently become possible [270–279], placing constraints on the particle mass and making predictions for unique small-scale features of the model. The arising quantum pressure suppresses small-scale power in the initial dark matter power spectrum [262, 280], modifies the halo mass function [281, 282], creates extended low-density soliton cores at the centers of dark matter halos [270, 283], exhibits transient density fluctuations on very small scales as a result of wave interference which can dynamically heat stellar streams and clusters [281, 284–286], and induces nontrivial dynamical friction on subhalos [287]. A challenge for the FDM model continues to be whether an ultralight particle mass can simultaneously be consistent with the Lyman-$\alpha$ forest power spectrum extracted from high-redshift quasars [288, 289] as well as the subhalo mass function [30, 282] and still simultaneously explain the core sizes of satellite galaxies [290, 291]. The inclusion of sufficient baryonic physics in a realistic cosmological setting [292–294] will also be necessary in determining the observable effects of FDM. In some cases, these effects can enhance rather than suppress structure beyond the FDM Jeans scale.

There are still many outstanding theoretical challenges concerning the origin and non-gravitational interactions of FDM. A candidate particle for the ultralight boson is the axion, which would arise from a symmetry-breaking needed to solve the strong CP problem [295, 296] (see the discussion in Section 2.1.2). This specific model would also give rise to a weak attractive self-interaction which may further alter the dark matter distribution. Non-gravitational self-interactions can lead to a range of phenomenological consequences including the formation and dynamics of solitons, oscillons, oscillatons, axitons, axion-stars, and vortices in the early and present-day Universe [98, 297–306]. Additionally, due to terms in the potential that are not quadratic in the axion field, the axion energy density does not
necessarily redshift entirely as DM. Current bounds from the Cosmic Microwave Background set an approximate limit on the axion decay constant, \( f_a \gtrsim 10^{13} \text{GeV}(10^{-20} \text{ eV}/m_a) \) [307], independently of the axion production mechanism. All of these effects are just beginning to be explored, and many further theoretical developments can be anticipated in the coming decade.

3 Cluster Scale

3.1 Scale and Objects

Clusters of galaxies are the largest \((M \gtrsim 10^{13.5} M_\odot/h)\) gravitationally bound objects in the Universe, with DM forming most of their mass. Detailed observations of galaxy clusters provided some of the earliest evidence for the existence of DM (see e.g. Ref. [308]), and these objects remain excellent laboratories to probe the collisional properties of DM particles. Perhaps the best known example of such an object is the Bullet Cluster [309–311], for which a clear separation between the hot gas (probed via X-ray emission) and the bulk of the mass of the two merging sub-clusters (probed via gravitational lensing) can be used to constrain the interaction cross section between DM particles at high velocities. Beyond merging clusters, detailed measurements of the density profile of relaxed clusters and of the position of their brightest central galaxy away from the cluster’s center could also detect the presence of central DM cores created by DM self-interaction.

3.2 Observational Probes

In the following, we outline several observational aspects of galaxy clusters that display sensitivity to DM physics.

3.2.1 Density profiles of clusters

Through a combination of data from optical, CMB, and X-ray surveys, it is now possible to accurately measure the stacked weak lensing profiles around massive galaxy clusters \((M \gtrsim 10^{13.5} M_\odot/h)\) [312, 313]. The signal-to-noise is expected to increase significantly in the next decade as the next generation optical (e.g. the Legacy Survey of Space and Time (LSST) at the Rubin Observatory [314]) and microwave (e.g. CMB-S4 [243]) surveys come online. These will allow for measurement errors at the percent level from scales comparable to the scale radius of the clusters out to the cluster outskirts. The lensing signal is sourced by the density profile of all matter, including the dark matter component, of the clusters. Therefore, the precise weak lensing measurements around massive clusters can be used to constrain those models of dark matter physics that can affect the density profiles, especially on scales of \(\sim 0.2 h^{-1}\text{Mpc}\) to \(\sim 2.0 h^{-1}\text{Mpc}\).

Self-interacting dark matter (SIDM) models (see Section 2.2.1) are known to produce characteristic signatures on the density profiles of halos. For interaction cross sections per DM mass of \(\sigma/m \sim 1 \text{ cm}^2/\text{g} \) at the characteristic velocity scale set by the halo mass of interest, the most prominent signature of SIDM is a central core. While the size of the core itself is typically smaller than the scale radius, the radial profile is affected all the way out to
the virial radius due to rearrangement of particles driven out of the core by the interactions. For clusters, therefore, the weak lensing measurements on scales of $\sim 0.2 h^{-1} \text{Mpc}$ to $\sim 2.0 h^{-1} \text{Mpc}$, are a sensitive probe to SIDM cross sections at velocity $v \sim 1000 \text{km/s}$ [315]. Existing data from surveys, e.g. Year 1 data from the Dark Energy Survey (DES) has sufficient statistical power to rule out cross sections above $\sim 2 \text{cm}^2/\text{g}$, similar to that obtained from the study of the Bullet Cluster (see, e.g., Refs. [309, 311]). This is shown in Fig. 4, which is adapted from Ref. [315]. The top panel shows the effects of SIDM on the lensing profile of clusters in the mass range $1 - 2 \times 10^{14} M_\odot/h$. The bottom panel shows the residuals of the SIDM models compared to the CDM model. The light grey shaded region in the bottom panel represents the measurement error on a similar cluster sample from DES Y1 data [312]. The progressively darker bands represent the expected measurement errors from DES Y3 data and LSST at VRO, obtained by rescaling the DES Y1 error bars by an appropriate factor to take into account the increased sky area. These projections are, therefore, conservative in the sense that they do not take into account the greater depth, and therefore more background sources, compared to DES Y1. With LSST, the statistical power is sufficient to constrain the cross section at better than $\sigma/m = 0.5 \text{cm}^2/\text{g}$ at these large velocities.

Since these measurements are determined by the cluster outskirts, they are less affected by baryonic effects than measurements closer to the cluster center. However, in light of the expected statistical constraining power on SIDM models expected in the future, any possible effects of baryonic effects on the size or shape of the signal need to be modeled to reduce the systematic errors associated with the interpretation of these measurements in terms of SIDM cross sections.

### 3.2.2 Merging clusters

If DM particles scatter with one another at astrophysically important rates (i.e. if DM is some form of SIDM, see Section 2.2.1), then two likely scenarios are either that the cross-section is independent of velocity, or that it decreases with increasing velocity. In either of these scenarios merging galaxy clusters are interesting, because of the high relative velocities between DM particles. In the case of a velocity-independent cross-section, the rate of scattering per particle is proportional to the product of the density and the mean pairwise velocity between dark matter particles. The high densities and velocities in merging clusters then mean that these systems would have high rates of SIDM scattering, making them a useful laboratory for these interactions. On the other hand, if DM self-interactions have a velocity-dependent cross-section, merging clusters would then provide complementary information to what can be learnt from lower-mass systems [316], because they measure the cross-section at an otherwise inaccessible velocity scale.

In addition to the interesting velocity scale at which merging clusters probe the SIDM cross section ($v \sim 1000–3000 \text{ km/s}$), another unique feature of merging clusters is that they are systems with a clear directionality. This means that the angular dependence of the SIDM scattering cross-section can be probed [317–319]. This is not the case for more isotropic effects of SIDM, such as core formation at the centre of halos, where SIDM models with different forms of differential cross section can lead to the same evolution of the density profile [318].
The archetypal cluster merger is the Bullet Cluster, in which a low-mass ($\sim 10^{14} M_\odot$) cluster (called “the bullet”) has passed through a more massive ($\sim 10^{15} M_\odot$) cluster, with the collision taking place approximately in the plane of the sky [320]. This system contains a textbook example of a bow shock (visible in X-rays), from which the speed and direction of the collision can be inferred. In addition, strong and weak gravitational lensing can be used to map the distribution of dark matter within the merging cluster system [321, 322], from which it can be inferred that the DM must be at least approximately collisionless in order for it to be leading in front of the (highly collisional) hot gas, which is slowed down by hydrodynamical forces. This statement that DM must be approximately collisionless can be turned into a more formal constraint on the SIDM cross section by considering the optical depth for SIDM interactions as particles from the bullet halo pass through the main halo [309]. The projected DM density in the inner regions of the main halo is approximately $\Sigma = 0.2 \, g \, cm^{-2}$, and the main halo would become optically thick for SIDM interactions if $\tau = \Sigma \sigma/m > 1$. Therefore, requiring that the main halo is not optically thick for SIDM interactions implies $\sigma/m < 5 \, cm^2 \, g^{-1}$. In other words, for cross sections $5 \, cm^2 \, g^{-1}$ or larger, most bullet halo particles would have scattered with a main halo particle during the collision, and the DM’s behaviour would not be approximately collisionless.

The SIDM constraint from the Bullet Cluster was tightened somewhat by comparison with $N$-body simulations of merging clusters with different SIDM cross sections [323], but further attempts to refine this have struggled to lower the upper-limit on the SIDM cross section [311, 317, 324]. Even finding one Bullet Cluster-like system in the observed Universe was fairly unlikely [e.g. 325] so it is highly likely that any merging systems yet to be discovered will be less extreme in terms of their mass and collision velocity, and therefore in terms of the expected size of SIDM effects.

One potential avenue for improving constraints is to study a large number of more minor mergers [326], but an initial attempt at doing this [327] was called into question when individual systems from this sample were analysed in more detail [328], and it was found...
that the offsets between the different components (DM, gas and stars) changed considerably when using more comprehensive data on the same clusters. LSST, Euclid, and Roman will improve the quality of weak and strong lensing constraints across much of the sky, and telescopes such as the Super-pressure Balloon-borne Imaging Telescope (SuperBIT) can provide targeted observations that will allow individual systems to be explored in more detail. Nevertheless, simply improving the quality of lensing measurements of merging systems will not significantly improve constraints on SIDM without some theoretical advances in how to extract DM physics from observed mergers.

### 3.2.3 Brightest cluster galaxy oscillations

As mentioned in Section 3.2.1, SIDM generally predicts that dark matter halos should have constant density cores. In galaxy clusters, the largest bound gravitational objects in the Universe, these cores are expected to be large, \( \sim 100 \) kpc, for a velocity-independent cross section of \( \sigma/m \sim 1 \text{ cm}^2/\text{g} \) [329]. As dynamical friction does not operate in the flat potential of cored profiles, one can imagine that if two clusters of relatively similar mass merge, their galaxies will "slosh," or oscillate, in the core of the merger remnant. In particular, the brightest cluster galaxy (BCG) should oscillate within the merger remnant on an orbit determined by the size of the core. The degree to which BCGs are offset from the centers of relaxed clusters can thus constrain the self-interaction cross section. As the core size is a strong function of cross section, this is a potentially sensitive test of \( \sigma/m \). BCGs are observed to be offset from their parent clusters’ X-ray center by \( \sim 10 \) kpc [330], which implies a cross section of order \( \sim 0.1 \text{ cm}^2/\text{g} \) [331]. A follow-up study of 10 relaxed clusters observed with the Hubble Space Telescope similarly indicated an oscillation amplitude of \( 11.82^{+7.3}_{-3.0} \) kpc from their mass peaks inferred from strong lensing [332].

Several theoretical questions remain. Initial constraints were based on dark-matter-only simulations of idealized equal mass mergers. The inclusion of gas has been shown to produce smaller SIDM cores and reestablish a potential gradient [333]. This can reintroduce dynamical friction and allow the orbits of the BCGs to decay. Simulations exploring non-equal mass mergers are also key. Equal mass mergers are expected to be much rarer than unequal mass mergers, as hinted by the observed BCG offset probability distribution, which is not consistent with that expected for a population dominated by equal mass mergers. In unequal mass mergers, the BCG of the more massive cluster (and presumably the BCG of the merger remnant) would not be significantly perturbed, resulting in smaller-scale oscillations. Simulations also have yet to explore the effect of velocity-dependent cross sections. As such models tend to favor low cross sections at cluster-scale velocities, oscillations are likely small in such models as well.

Observationally, new weak lensing mass maps from upcoming surveys with instruments such as the Roman and Euclid space telescopes, as well as SuperBIT will allow constraints to be improved by a factor of at least two [334]. Positional offsets can be complemented with measurements of oscillations in velocity space, which are also expected to be large, and observable with the Maunakea Spectroscopic Explorer [335].
3.2.4 Substructure in massive clusters

As in the case of the Milky Way, the spatial and velocity distribution of substructure in galaxy clusters is sensitive to DM physics. In the presence of SIDM with a velocity-dependent interaction cross section, the cluster subhalos and the main cluster halo probe different regions of the velocity space. Indeed, while scattering between subhalo DM particles occurs at typical galactic speed \(v \sim 100−300 \text{ km/s}\), collisions between main host and subhalo DM particles occur at the typical velocities found in clusters, which tend to be significantly higher.

In these systems, the interaction cross section at galactic velocities will be responsible for core formation or even the onset of gravothermal collapse [167], while the DM cross section at the higher cluster velocities would impact the ram-pressure stripping (evaporation) that subhalos experience as they move within the cluster halo. Thus, detailed observations of cluster substructure using both weak and strong gravitational lensing [336] could probe the DM cross section over a broad range of velocities [172]. Already, Ref. [337] has found a possible discrepancy (see however Ref. [338]) in the properties of cluster substructure as compared to ΛCDM predictions, which could potentially be explained in the context of SIDM (see e.g. Ref. [171]).

3.3 Theory Connections

As can be inferred from the above discussion, galaxy clusters are excellent probes of the potential collisional nature of DM in the high-velocity regime. Clusters can be thought as high-energy DM particle colliders, hence providing a unique environment to test for any deviations from the standard CDM scenarios at those energies. Galaxy clusters are thus a key probe of SIDM (see Section 2.2.1), as well as composite DM (see Section 2.3.1), which typically have sizable self-interaction.

4 Galaxy Scale

4.1 Scale and Objects

Luminous galaxies span a large range of halo masses \(\sim 10^8 − 10^{13} M_\odot\), making them very versatile laboratories to probe the physical properties of DM. Observations of galactic rotation curves by Rubin, Ford, and Thonnard [339] were instrumental in firmly establishing DM as an integral part of our modern cosmological model. Detailed indirect observations of the DM distribution within galaxies (including within our own) can provide important clues about the fundamental physics of DM. The formation of the first galaxies in the early Universe also holds important information about DM since their abundance is quite sensitive to the low-mass end of the halo mass function.

4.2 Observational Probes

In this section, we outline several probes of dark matter physics on the scales of galaxies. We focus our attention here on “luminous” galaxies (including our own), while ultra-faint dwarfs will be treated in Section 5.
Figure 5: Predicted lensing signal and errors for the $\Delta \Sigma$ profile around dwarf lens galaxies. The left panel corresponds to dwarfs with $M_\star = 10^8 M_\odot$ and the right panel corresponds to dwarfs with $M_\star = 10^9 M_\odot$. Errors correspond to a lens samples selected within a narrow mass range (0.2 dex bin width) and for $0 < z < 0.25$. Predicted diagonal errors are shown for HSC wide survey (blue), Euclid (orange), Roman (red), and Rubin LSST (green). Figure taken from Ref. [341].

4.2.1 Dwarf galaxy lensing

Dwarf galaxies are a unique probe of the nature of dark matter and the interplay between dark matter and baryonic physics. However, rotation curves of dwarf galaxies only constrain the inner regions of dark matter halos (typically a factor of $\sim 10$–$20$ smaller than the actual halo radius). Any “halo mass” estimate from rotation curves is in fact an extrapolation that relies on assumptions about the shape of the dark matter profile (see e.g. Ref. [340]). One of the most powerful ways to directly probe total halo masses out to the halo radius is “galaxy–galaxy lensing.” This is the average weak lensing signal from background “source” galaxies around a sample of foreground “lens” galaxies. Existing weak lensing measurements have been limited to $M_\star > 10^9 M_\odot$. However, recent work has showed that new generation lensing surveys can measure the lensing signal of $M_\star \sim 10^8$–$10^9 M_\odot$ dwarf galaxies with high significance [341]. Figure 5 shows the expected lensing signals for dwarf galaxies from upcoming surveys such as Euclid, Roman, and Rubin. With the number of dwarf lenses available in Rubin LSST, it is even possible to distinguish different density profiles (e.g., cored vs. cuspy) using $M_\star \sim 10^9$–$10^{10} M_\odot$ dwarf galaxies as lenses [Sec. 3.2.1 of Ref. 342]. In combination with measurements of galaxy dynamics, lensing measurements for dwarf galaxies will provide unique and novel constraints on non-CDM theories of dark matter and on feedback models at the dwarf scale. To harness the power of large lensing surveys and to tackle this science goal, however, large samples (tens of thousands) of dwarf galaxies with reliable redshift measurements are needed.

The need for such a large sample precludes directly spectroscopic follow-up [341]. Photometric redshifts from Euclid, Roman, and Rubin are also unlikely to be accurate enough to reliably select large samples of dwarf galaxies in the mass range $M_\star = 10^7$–$10^9 M_\odot$. However, there has been much progress in using machine learning algorithms to improve photomet-
Dark matter Physics from Halo Measurements

ric redshifts, especially in the very low-redshift regime [343, 344]. Combining these novel algorithms with substantial very low-redshift spectroscopic training samples (e.g., SAGA Survey [345]; DESI LOW-Z Secondary Program, E. Darragh-Ford et al. in prep.), we can significantly improve sample selection with photometric redshifts in the Rubin LSST era.

In addition to only using photometric redshifts for sample selection, it will soon be possible to select large samples of star-forming dwarfs using narrow-band imaging targeting their strong H$\alpha$ and O$\text{III}$ emission lines. As an important pathfinder for this promising approach, the Merian Survey\textsuperscript{2} will use the Blanco telescope to image 867 deg$^2$ of the Hyper-Suprime Camera (HSC) Wide field using two custom-made filters probing H$\alpha$ and O$\text{III}$ at $z \sim 0.08$. This survey will provide redshifts to a complete sample of 100,000 star-forming dwarf galaxies (two orders of magnitude times larger than the Sloan Digital Sky Survey (SDSS) and the Galaxy And Mass Accretion (GAMA) survey combined). Combined with deep+high spatial resolution imaging from the HSC, this program will result in the first high S/N weak lensing profile measurements for dwarf galaxies, probing their DM halos out to their virial radii. Comparisons with state-of-the-art, detailed, hydrodynamical simulations will help to fully capture the interplay between DM, feedback, and their roles in the formation and evolution of dwarf galaxies. Merian is an important precursor to a much bigger program that can leverage the Rubin LSST. Custom narrow-band imaging over the LSST footprint affords the exciting opportunity to construct very large samples of dwarf galaxies and to measure their halo profiles with gravitational lensing. This program would provide novel and unique constraints on DM on dwarf scales.

Even better constraints on the DM interaction cross section could be obtained by probing the central density of these dwarf halos in the region where a constant-density core could form in the presence of SIDM. This would require measurements of their central velocity dispersions, $\sigma_*$, which would necessitate a large spectroscopic survey of the Merian-found galaxies. This presents several challenges. First, dwarf galaxies ($M_* < 10^9M_\odot$) are intrinsically faint, making absorption-line dispersion measurements difficult. Next, for the more distant galaxies, 8–10 m class telescopes are required, but fields-of-view and scarce observing time limits the final sample size. Smaller telescopes (2–4 m) offer more access with wider fields, but struggle to reach the required spectroscopic depths.

A way forward is offered by new astrophotonic “spectrometer-on-a-chip” technology coupled with coming wide-field adaptive optics (e.g., ground-layer AO) systems. Such systems are being planned for a variety of telescopes of modest size and should eventually provide a 0.8 deg$^2$ field-of-view with performance at $r$-band wavelengths. The AO systems offer sensitivity gains thanks to improved seeing. In addition, photonic spectrometer arrays hold promise for enabling optimal extraction of turbulent modes, thereby offering further gains in S/N. If realized, these technologies would enable a 2–4 m telescope to achieve S/N $\approx 10$ Å$^{-1}$ at $r_{AB} \approx 21.5$, sufficient for $\sigma_*$ measurements to 20%. An instrument with these capabilities and a 150 multiplex across 0.8 deg$^2$ could observe a 10,000 dwarf sample over $\sim$200 nights. On a 2–4 m class telescope, such an experiment would, at a modest cost, deliver enormous value, with SIDM cross-section constraints improved by several factors to an order-of-magnitude over current constraints.

\textsuperscript{2}https://merian.sites.ucsc.edu
4.2.2 Milky Way stellar dynamics

The gravitational potential of the Milky Way can be measured by studying the dynamics of stars. Such dynamical mass measurements are useful for constraining the local DM density, which is a crucial quantity for direct and some indirect detection experiments [346, 347], as well constraining DM particle models that give rise to sub-structures or have a characteristic larger scale spatial density dependence. The astrometric data that is used for such measurements has improved tremendously in the last few years thanks to the Gaia satellite [348], whose catalogue contains over a billion stars with parallax and proper motion information. At the same time, advances in technology in the coming decade will enable “real-time” Galactic dynamics wherein extremely precise time-series measurements can be used to directly measure Galactic accelerations, and therefore the Galactic potential.

Dynamical mass measurements are typically performed under the assumption of a steady state as well as a few simplifying symmetries (e.g. rotational symmetry or mirror symmetry with respect to the Galactic plane), often reducing the dynamics to only one or two spatial dimensions (e.g. the rotational velocity curve or the vertical oscillation of disk stars). In studies that determine the local DM density, the suite of results from different Gaia-based studies are somewhat discrepant. Moreover, where the results of local analyses (e.g. based on the vertical oscillations of stars within roughly kilo-parsec distance) are found in the range 0.4–0.6 GeV/cm$^3$, while the results of more global analysis (e.g. based on the circular velocity curve) are found in a slightly lower range of 0.3–0.5 GeV/cm$^3$ [349]. The systematic source of these discrepancies could be a mismodeled distribution of baryons or incorrect assumptions of a steady state or symmetry [e.g. 349–351].

This suite of measurements does not only inform us of the local DM density, but can also constrain DM particle models. Analyses of the Milky Way’s rotational velocity curve have been used to place constraints on superfluid dark matter, which can exhibit MOND-like acceleration profiles on smaller spatial scales [352, 353]. Very local analyses of the vertical oscillation of stars have placed limits on a so-called thin dark disk, which would consist of a dark matter subcomponent with strong dissipative self-interactions [214, 215]. Current limits to a thin dark disk’s surface density are roughly 5 M$_\odot$/pc$^2$ [249, 250, 354], such that a thin dark disk is allowed only due to uncertainties associated with the baryonic matter density distribution (at least under the assumption that the thin dark disk is stable and co-planar with the stellar disk).

There is progress in relaxing these assumptions of a steady state and Galactic symmetries, which will be necessary to make further progress in extracting information about DM microphysics from galactic dynamics using Gaia (noting again that Gaia data is precisely what has allowed us to uncover the deviations from these assumptions). For example, the full three-dimensional gravitational potential can analysed with a method using “normalising flows” [355], which is model independent in terms of symmetry assumptions. It is also possible to relax the assumption of a steady state and extract information from the shape of time-varying dynamical structures of the Milky Way; this is exemplified by a new method [354, 356] that infers the gravitational potential using the recently discovered phase-space spiral [357].

Beyond these modeling improvements, it has recently become possible to directly measure the very small Galactic accelerations of stars ($\sim 10$ cm/s/decade) using time-series
Dark matter Physics from Halo Measurements

precision measurements with several different techniques. By measuring a direct acceleration, it is not necessary to assume anything is in equilibrium since accelerations are directly related to gradients of the potential via the Poisson equation. Recent work has shown that spectral measurements of the line-of-sight (radial) velocities (RVs) of individual stars will be sensitive to the local DM density [358–361], for instance using the current generation of high-precision spectrographs like ESPRESSO [362] that have achieved precision RV precision $\sim 10$ cm/s [363]. When extremely large telescopes like the Giant Magellan Telescope begin observations, extreme-precision radial velocity measurements can be carried out across the Galaxy [364]. Line-of-sight accelerations of pulsars measured via their spins [365] and orbital periods [365, 366] may already have hints of the local Galactic acceleration, with improvements possible in the future. It is also now imminently possible to measure the Galactic acceleration directly using eclipse timing, by measuring the small shift in the mid-point of the eclipse time over the past decade induced by the Galactic potential since the Kepler mission [367], using space-based observations from HST, and in the future with JWST and Roman observations. Gaia astrometric observations were also recently analyzed to measure the solar system acceleration [368, 369], which is in agreement with prior measurements by VLBI [370]; given these measurements, it may be possible to use statistically aggregated astrometric accelerations of stars to measure the galactic potential [371], which would be aided further by astrometric missions beyond Gaia like Theia [372] or GaiaNIR [373]. Thus, there is now a wide range of techniques available currently or on the timescale of the next decade that will enable “real-time” Galactic dynamics. In the future, this will allow for precise characterizations of the mass distribution (both the smooth component and dark matter sub-structure) across the entire Galaxy, without assumptions of equilibrium or symmetry. These measurements can constrain DM physics due to the different spatial morphology and clustering of DM substructure in different models of DM as detailed in Section 2.

4.2.3 Galactic rotation curves

The CDM model is extremely successful in explaining the matter distribution of the Universe on large scales. However, it has challenges in matching with observations on galactic scales. Dwarf and low surface brightness galaxies often prefer a cored DM density profile rather than a cuspy one predicted in CDM-only simulations, which was initially coined as the “core vs cuspy” problem, see [375] for a review. However, galaxies are not uniformly cored [376], and the inner rotation curves of spiral galaxies exhibit diverse shapes with a spread larger than predicted in CDM [377]. More recently, the issue has been reframed as the “diversity” problem.

Figure 6 shows circular velocity profiles of spiral galaxies from the SPARC sample [374]. The left panel contains galaxies with low masses and their asymptotical velocities are $V_f \approx 79–91$ km/s, while the right panel contains high-mass ones with $239–315$ km/s. For both cases, the spread in the circular velocity at $\sim 2$ kpc is a factor of $\sim 4$, although the galaxies in each mass bin have similar velocities at further radii.

In CDM, the feedback processes associated with star formation could alter the gravitational potential and change the DM distribution in the inner halo. Simulations show that if the feedback is driven by spatially and temporally fluctuating star formation, a cored CDM halo could be produced [378–380], and hence the predicted rotation curve rises slowly. On
the other hand, the halo remains cuspy if the feedback is driven by smoother star formation [377, 381]. The challenge is to produce both cored and cuspy halos in simulations after adapting a specific feedback model.

Recent studies show that the SIDM model [161, 193] can explain the full range of the diversity [165, 166, 382, 383]. The SIDM model predicts both cored and cuspy profiles, depending on baryon concentration [164]. In DM dominated galaxies, thermalization due to the self-interactions creates large cores and reduces DM densities. In contrast, thermalization leads to denser and smaller cores in more luminous galaxies. Ref. [166] analyzed the rotation curves of 135 galaxies from the SPARC dataset and found that SIDM provides excellent fits, as demonstrated in Fig. 6 (solid curves). The inferred halo parameters are consistent with the concentration-mass relation from cosmological CDM simulations, indicating that SIDM and CDM predictions are consistent on large scales.

Looking forward, there are several promising directions that we can take. The analysis in Ref. [166] assumes a static stellar distribution from observations as an input. This approach is well-justified for the purpose of fitting to the rotation curves at $z = 0$. However, to make further progress, we need to understand the formation history of the stellar distribution. In fact, a key challenge for structure formation simulations in both CDM and SIDM is to reproduce the large spread in the stellar density to be consistent with the observations [383]. To resolve this, simulations may need to adopt different baryonic feedback models depending on the DM properties. In addition, it is important to analyze the rotation-curve data in other DM scenarios beyond CDM, such as warm DM and fuzzy DM. To perform an efficient statistical analysis, we need to develop a semi-analytical approach to model DM distributions with the presence of baryons for those models, as in the case of SIDM.

Furthermore, recent observations have discovered a large number of ultra-diffuse galaxies that exhibit a much more extended baryon distribution [384–387], compared to ordinary
spiral galaxies. Some of them have high baryon contents, but low circular velocities, indicating that they are baryon-dominated and their host halos must have extremely low concentrations \cite{388, 389}. In the near future, high-resolution measurements of stellar and gas kinematics will be available for more ultra-diffuse galaxies, and they will provide an exciting opportunity for testing different DM and feedback models. In the next decade, a collaborative effort among particle physics theorists, simulators and observers is essential for seeking a successful structure formation theory which could provide a uniform explanation to observations of galactic systems on all scales.

4.2.4 Ultraviolet luminosity function

Observations of galaxies at high redshifts provide us with a unique and highly complementary handle on the physics of structure formation during the epoch of cosmic reionization \cite{390}. The abundance of galaxies is measured by the Ultraviolet luminosity function (UV LF), and is controlled by both the astrophysics of galaxy formation and the clustering of DM. The galaxies forming during reionization ($z = 4-10$) have been mapped by the Hubble Space Telescope through a decades-long observational campaign that has resulted in two main catalogs — the Hubble Legacy Fields (HLF) and Hubble Frontier Fields (HFF) — that cover tens of thousands UV-bright galaxies, see e.g. \cite{391, 392, 395–403}. The measured UV LF from Refs. \cite{391, 392} is shown in Fig. 7, together with the best-fit model from Ref. \cite{393}. The top axis shows the corresponding halo masses covered at $z = 6$, defining the scales that we can access with this probe ($M_h \sim 10^{10} - 10^{12} M_\odot$ or, equivalently, $0.1 \text{ Mpc}^{-1} < k < 10 \text{ Mpc}^{-1}$).

From a modeling perspective, several different approaches have been employed to predict the UV LF, ranging from (semi-)analytical methods \cite{404–407}, to more elaborate techniques that are calibrated to simulations \cite{408–411}, and full-scale hydrodynamical simulations \cite{412–416}. A crucial ingredient in constructing the UV LF is the so-called galaxy–halo connection \cite{417}, which encodes the relation between the UV magnitude of the galaxy and the mass of its host halo. The astrophysical parameters that enter this link can be highly degenerate with cosmological parameters, see e.g. \cite{393, 418} for an illustration, making it necessary to carefully
model and marginalize over the unknown astrophysics.

The UV LF has seen many applications that improved our understanding of cosmology at high redshifts. In particular, previous works have exploited this probe to measure $\sigma_8$ [393, 419], and to derive constraints on warm- and fuzzy-DM models [418, 420–425], dynamical dark energy [426], the primordial matter power spectrum [427], as well as the small-scale power spectrum [394] and non-Gaussianities [428, 429] beyond the reach of the Cosmic Microwave Background.

Future advancements with the UV luminosity function will come from both the observational and modeling side. In the former case, upcoming observatories like the James Webb Space Telescope [430] and Roman Space Telescope [431] will cover larger areas and fainter magnitudes than currently possible. Measurements of the faint end of the UV LF will particularly be important for understanding the clustering properties of DM, as such observations cover smaller halo masses. Besides improved data, one of the main challenges will be to better understand the dependence of the cosmological results on the astrophysical modeling. This can be done by studying the galaxy–halo connection in hydrodynamical simulations in order to control the degeneracies between astrophysical and cosmological parameters. Insights from other probes, such as 21-cm maps (see Section 5.2.5), may also prove to be useful in this regard.

4.3 Theory Connections

Detailed measurements of the DM density profile of luminous galaxies provide a window into the in situ impact of DM physics within these objects. Such observations are for instance probing the SIDM cross section at typical galactic velocities (see Section 2.2.1), or the presence of a solitonic core caused by ultra-light bosonic DM (see Section 2.3.2). Furthermore, galactic halo profile measurements are also sensitive to the presence of dissipation in the DM sector (which could happen in a composite DM scenario, see Section 2.3.1) or even to halo energy injection through either DM annihilation or decay (see e.g. Refs. [432–434]). On the other hand, the onset of galaxy formation in the earlier Universe as probed by the UV LF is very much influenced by ab initio DM properties. This makes such observations sensitive to the DM thermal history (see Sections 2.1.1 and 2.1.2), as well as to any nongravitational interactions that DM particles experienced in the early Universe (see Sections 2.1.4 and 2.3.1).

5 Beyond the Threshold of Galaxy Formation

5.1 Scale and Objects

The formation of stars, and hence entire galaxies in the process, requires gas to efficiently cool to further collapse into ever denser regions in the vicinity of radiation coming from the first generation of stars. The limit where atomic hydrogen can cool efficiently is at around virial temperatures of $T_{\text{vir}} \sim 10^4 K$ which corresponds to DM halos at around $\sim 10^{7–8} M_\odot$. Current observations presented in this Section are approaching the limit to empirically measure the minimum halo mass where where stars were formed. At least some halos in the mass range
of $\sim 10^{8-9} M_\odot$ do host stars. These halos can be observed as ultrafaint galaxies. These systems are currently the most DM-dominated systems with a ratio of DM to stellar mass significantly exceeding a factor of a hundred. DM halos below this threshold are expected to be depleted of stars and are effectively invisible. Halos at and below the galaxy formation threshold offer a pristine environment in which baryonic physics is subdominant to the dark matter behavior, and hence are substantially less affected by baryonic systematics.

### 5.2 Observational Probes

In the following we list some of the promising current and future probes that can probe the abundance and inner composition of halos in the mass range of $10^{6-9} M_\odot$.

#### 5.2.1 Ultra-faint galaxies

Ultra-faint dwarf (UFD) galaxies, defined as dark matter-dominated stellar systems with luminosity $L < 10^5 L_\odot$ [435], occupy the smallest dark matter halos that form galaxies, with $10^8 M_\odot \lesssim M_{\text{halo}} \lesssim 10^{10} M_\odot$ [436, 437]. These systems therefore provide crucial information about the astrophysical processes that govern the faint end of galaxy formation and the dark matter physics that impacts the abundance and properties of halos near the galaxy formation threshold.

To date, ultra-faints have exclusively been detected within distances of $\sim 300$ kpc as satellites of the Milky Way. Deep photometric surveys including the Dark Energy Survey have significantly increased the number of known Milky Way satellites in recent years, which now exceeds $\sim 60$ systems [438]. Nevertheless, both empirical and theoretical arguments suggest that our census of ultra-faints—even within the Milky Way’s virial radius—is highly incomplete [439–441]. Forthcoming observational facilities including the Vera C. Rubin Observatory and the Nancy Grace Roman Space Telescope are expected to drastically improve upon current ultra-faint detection capabilities. Specifically, Rubin is forecasted to discover the entire population of UFD Milky Way satellites in the Southern hemisphere outside the Galactic disk and with surface brightness down to $\sim 32$ mag arcsec$^{-2}$ [157, 314]. Excitingly, these detections are expected to extend well beyond distances of $300$ kpc, allowing UFD populations to be characterized throughout much of the Local Volume [442].

Our understanding of the dark matter halos that host ultra-faints has significantly improved over the last decade. In particular, hydrodynamical simulations have started to resolve the formation and evolution of UFDs in a cosmological context, including as satellites of larger hosts [443–448]. In parallel, semi-analytic and empirical galaxy–halo connection models have been combined with observations to infer ultra-faints’ halo properties; in these models, the number of observable ultra-faints suggests that these systems occupy halos with masses down to $\sim 10^8 M_\odot$ (near the “atomic cooling limit”; [436, 437, 449, 450]) or even lower [451, 452]. The fraction of low-mass halos that host UFDs and the scatter in the stellar mass–halo mass relation for these systems are also being studied in order to precisely characterize the properties of halos near the galaxy formation threshold [453–456].

Ultra-faints provide stringent constraints on dark matter particle properties. Specifically, the number of observable UFD Milky Way satellites informs dark matter physics that suppresses the present-day abundance of dark matter halos with masses of $\sim 10^8 M_\odot$, yielding
lower limits on the mass of thermal relic warm dark matter of $\sim 5 \text{ keV}$ [449, 457–466] and on the mass of ultra-light dark matter of $\sim 10^{-21} \text{ eV}$ [464, 467]. Similar analyses constrain the strength of dark matter–Standard Model interactions at the $\sim 10^{-30} \text{ cm}^2$ level [29, 154, 155, 464] and the dark matter particle lifetime at the $\sim 10 \text{ Gyr}$ level [468, 469]. Furthermore, stellar velocity dispersion measurements for brighter dwarf galaxies inform the strength and velocity-dependence of dark matter self-interactions, with sensitivity to cross sections of $\sim 1 \text{ cm}^2 \text{ g}^{-1}$ [470, 471], and combining measurements of UFD internal dynamics from upcoming spectroscopic facilities and giant segmented mirror telescopes [472] with their population statistics will further inform self-interacting dark matter models [473–475].

The forthcoming observations discussed above are expected to increase the sample of known UFDs near the galaxy formation threshold, reducing statistical uncertainties and improving the corresponding dark matter constraints. For example, Rubin observations are forecasted to probe halo masses below $\sim 10^8 \text{ M}_\odot$ [157]. If these halos are occupied by ultra-faints, these observations will provide evidence for the existence of galaxies formed through molecular hydrogen cooling; otherwise, they will determine the halo mass threshold and stochasticity associated with galaxy formation. Thus, combining these measurements with purely gravitational probes of small-scale structure will enable the detection of completely dark halos.

### 5.2.2 Lyman-alpha forest

The intergalactic medium (IGM) is the rarefied gas that fills the vast volumes between the galaxies in the Universe. Physical effects, ranging from the nature of dark matter to the radiation from star-forming galaxies and quasars, set the observable properties of the IGM, making the IGM a powerful probe of both fundamental physics and astrophysics. Baryons in the IGM trace dark matter fluctuations on Mpc scales, while on smaller ($\lesssim 100 \text{ kpc}$ scales) the $T \sim 10^4 \text{ K}$ gas is pressure supported against gravitational collapse. While the IGM is sensitive to the Epoch of Reionization [476–478], complex and poorly understood physical processes related to galaxy formation play only a minor role in determining its structure [479, 480].

Forward modeling the structure of the IGM for a given cosmological and DM scenario is a theoretically well-posed problem, albeit requiring expensive cosmological hydro-dynamical simulations. Despite this apparent simplicity, accurate simulations of the IGM are computationally challenging because of the multi-scale nature of the problem [481]. A spatial resolution of $\approx 20 \text{ kpc}$ is required to resolve the density structure of the IGM, while a relatively large box size is required to capture the large-scale power, to obtain a fair sample of the Universe, and to model the topology of reionization.

These characteristics of the Lyman-\(\alpha\) flux power spectrum have allowed for some of the tightest constraints on the nature of dark matter clustering [482–487]. Current analyses have been shown to be sensitive to the scale of either a suppression [488, 489] or an enhancement [490, 491] of the small scale power. These dark matter constraints come from high-quality observations of a the few currently available ($\sim 100$) high redshift ($z > 4$) quasar spectra obtained with high resolution instruments ($R \sim 80,000$) and high signal-to-noise observations ($S/N \sim 100$). With these spectra, surveys are able to measure scales to wavenumbers of $k \sim 20 \text{ h/Mpc}$ [478, 492]. Notably these surveys are currently limited by
statistical errors. Analysis of the current data make clear that dark matter constraints benefit from access to higher redshifts ($z > 4$). By improving both the theoretical predictions in this regime, as well as increased statistical power, we could push the lower bounds on the dark matter free-streaming scale by several orders of magnitude. Current bounds on the particle mass of various dark matter models include only a few spectra ($\sim 10 - 15$) at the highest redshifts ($z > 4.5$) (e.g. [487, 493, 494]). Current or upcoming surveys will push this number to $\sim 25$ (DESI), $\sim 50$ (SkyMapper) and $\sim 100$ (LSST) at a fixed magnitude limit of $m_g < 18.5$. Each object requires 1 night per quasar, so a survey of 3.5 years can obtain 1000 quasar spectra.

Current lower mass bounds for the WDM model are $m_{WDM} > 2.1 \text{ keV}$ using only high-redshift data and the most simplistic and uninformative priors on the thermal history evolution [482]. Including lower redshift data pushes this bound to $> 3.5 \text{ keV}$, and further including informative thermal history priors leads to $> 5.1 \text{ keV}$. Using a Fisher forecast analysis, at 1000 spectra, the bound is $> 14.4 \text{ keV}$; almost an order of magnitude improvement over the current constraints coming from the equivalent high redshift data set analysis. This constraint maps to a lower bound on the ultra-light axion-like particle mass on the order of $\sim 10^{-21} - 10^{-20} \text{ eV}$. Combined with higher mass constraints from super-massive black hole super-radiance, this could exclude all ultra-light axions of mass $< 10^{-14} \text{ eV}$ as the dominant source of dark matter.

### 5.2.3 Stellar streams

Stellar streams, the tidal remnants of globular clusters and dwarf galaxies, are sensitive tracers of the distribution of matter at a range of scales throughout the Galaxy. At larger scales, stream morphology is sensitive to the global distribution of matter in the Milky Way, with recent studies of streams enabling constraints on the total mass and shape of the Milky Way halo and its massive satellites [495–497]. At smaller scales, dynamically-cold stellar streams, resulting primarily from the disruption of globular clusters, retain information about interactions with small-scale perturbers [498], thus enabling tests of the clustering of dark matter mass scales below the threshold of galaxy formation. Precise observation and modeling of a large population of stellar streams will therefore enable exciting constraints on the minimum dark matter halo mass, as well as measurements of the population of low-mass substructure in the Milky Way halo.

In the simplest case, an interaction between a stellar stream and a dark matter subhalo leads to the formation of a gap in the distribution of stars along the stream. The depth and size of this feature can be used to infer the properties of the perturbing subhalo and the time since interaction [499]. Many Milky Way streams have already been examined for gaps caused by interactions with low-mass subhalos, but at current observational limits it is difficult to unambiguously associate stream density variations with perturbations from subhalos. However, for strong encounters between a subhalo perturber and a thin stream, other features may appear in observed streams that provide even stronger (and less ambiguous) constraints on the properties of the perturber and its orbit. For example, thanks to recent data releases from the Gaia mission, these types of features were recently observed in two previously-known streams: A “spur” of stars has been found emanating from a prominent gap in the GD-1
stream [500] that has led to exciting constraints on the properties of the implied perturber [501, 502], and more recently two streams have been kinematically associated into a single stream with a prominent “break” that may have been caused by another encounter from a different massive perturber [503].

Studies of gaps in stellar streams provide insight into recent individual subhalo interactions. However, stellar streams are expected to experience many impacts over their lifetimes. Strong interactions produce clear gaps, while weaker interactions will lead to subtle variations in stream tracks and densities. All features will be washed out over time due to the internal velocity dispersion of stream member stars. The cumulative effect of a lifetime of perturbations can be studied statistically by measuring the power spectrum of density fluctuations along a stream [504, 505]. Perturbations by dark matter subhalos at various masses, as well as by other Milky Way structures, result in density fluctuations at different scales. Measurements of the density power spectrum can therefore reveal the interaction history of a stellar stream, and the mass distribution of the perturbing subhalo population. Density power spectra have been used to provide evidence for a population of dark subhalos around the Milky Way and constrain a range of dark matter models [506]. Future studies with deeper observations and larger population of streams are projected to constrain the subhalo population down to $10^5 M_\odot$ [505].

Recent large photometric and astrometric surveys have enabled the discovery of more than 60 streams around the Milky Way [507–509]. Upcoming surveys, such as the Rubin Observatory Legacy Survey of Space and Time [LSST; 510] will enable the discovery of streams at larger distances and lower stellar densities, and will improve observations of known streams, reaching fainter member stars and obtaining higher photometric precision. LSST is predicted to enable detections of gaps caused by subhalos with masses down to $10^6$ to $10^7 M_\odot$ for streams with a range of typical surface brightnesses at a distance of 20 kpc [342].

Precise kinematic measurements are also necessary to accurately infer the perturbation history of stellar streams. Kinematic measurements facilitate the separation of stream members from foreground contamination, and enable the dynamical modeling of these systems. The sensitivity of streams to baryonic perturbers varies based on the stream orbit. Precise orbit modeling is therefore necessary to select a clean sample of streams and to model any extraneous sources of perturbation. The Gaia Mission has provided unprecedented proper motion measurements of more than 1 billion stars in our Galaxy [511], and has enabled measurements of the proper motions of Milky Way streams out to $> 50$ kpc [512]. Radial velocities have historically been difficult to measure due to the diffuse, extended nature of these systems. However, by taking advantage of cutting edge photometric and astrometric datasets to efficiently select targets for observation, the Southern Stellar Stream Spectroscopic Survey [S\textsuperscript{5}; 513] has recently provided the first systematic spectroscopic observations of stellar streams, measuring radial velocities and metallicities of $> 20$ stellar streams in the southern hemisphere.

These current and upcoming observations will enable detailed studies of individual perturbations and statistical measurements of the population of low-mass dark matter subhalos around the Milky Way. We can make predictions for the expected number of observable gaps in a given dataset within a particular dark matter mode [514]. For example, with LSST, within $\Lambda$CDM, we expect to observe 17 gaps in the 13 DES stellar streams, and a detection
of fewer than 6 gaps would be inconsistent with ΛCDM at a 99.9% level [342].

5.2.4 Galaxy strong lensing

Strong gravitational lensing provides a unique probe of dark matter structure at the smallest cosmological scales, and thus constrains the particle nature of dark matter itself. By coupling directly to gravity, lensing circumvents luminous tracers of the underlying dark matter, and it is sensitive to mass scales below those of typical dwarf galaxies. Analyses of resolved distortions in extended Einstein rings of galaxy-scale strong lenses have resulted in the detection of substructure and halos along the line of sight with inferred lensing masses of $M_{\text{lens}} \approx 10^{8-9} M_\odot$ with the Hubble Space Telescope, Keck adaptive optics (AO), and ALMA interferometric data [515–521]. The sensitivity of a lensing detection is limited by the angular resolution of the data, and by the structure of the lensed source. Very Long Baseline Array (VLBA) radio observations have the capacity to probe structure power at the $10^7 M_\odot$ scale, albeit for a very small sample of lens systems, and require analysis pipelines that can cope with the large data volume per observational target.

The magnification ratios between adjacent images of multiply-imaged quasars provide an alternate means of probing substructure. Initial studies showed that these data are consistent with CDM at scales of $\gtrsim 10^9 M_\odot$ [522, 523]. Recent progress with newly gathered data and a re-analysis of previously existing data resulted in tighter constraints demonstrating consistency with CDM to scales of $\gtrsim 10^8 M_\odot$ [524, 525]. The latest constraints correspond to a free-streaming length scale of $\sim 10$ kpc, or an equivalent thermal relic particle mass of $m_{\text{th}} > 5$ keV, ruling out warmer mass functions at 2 sigma confidence. A similar study constraining the mass-concentration relation of dark matter halos at $\sim 10^8 M_\odot$ [526] showed consistency with CDM predictions, and the extrapolation of N-body simulation results to the scales probed by lensing. In the context of self-interacting dark matter, mid-IR flux ratios obtained from the James Webb Space telescope will provide enough sensitive to detect populations of core-collapsed subhalos, probing the self-interaction cross section on velocity scales below 30 km s$^{-1}$ [527]. Constraints on halo mass function and concentration-mass relation from strong lensing analyses can also be interpreted in terms of the primordial matter power spectrum on scales $k > 10$ Mpc$^{-1}$, irrespective of a particular dark matter model [528].

Conclusive results at the $10^7 M_\odot$ scale will require an expanded sample size of suitable lens systems, and improved data quality. The Rubin and Roman Observatories will yield up to 10,000 gravitational lenses, including several hundred quadruply-imaged quasars. The order of magnitude expansion of the current strong lens sample size forms the basis upon which strong gravitational lensing probes of dark matter will operate in the next decade.

However, this sample of lenses will not translate into constraints on dark matter at the $10^7 M_\odot$ level on its own; sufficiently precise and sensitive follow-up efforts must be undertaken for a targeted subsample. Gravitational imaging sensitivity will significantly benefit from the improved imaging resolution afforded by next-generation telescopes. The expected number of quadruply-lensed quasars on the sky is sufficient to meet the requirements to measure the mass structure at $\sim 10^7 M_\odot$ [529], but the constraining power of a fixed sample size depends strongly on the measurement precision of the flux ratios. Current state-of-the-art measurements presented [530] achieve 5-6% precision. Considering that most of the quadruply-lensed
quasars to be added to the existing sample will be fainter, enhanced observational capabilities will ensure that lensing reaches the $10^7 M_\odot$ target threshold.

The James Webb Space Telescope (JWST) will enable precise spectroscopic measurement of narrow-line and dusty-tori flux ratios, but may not have the capacity to do so for a required sample of $\sim 100$ quadruply-lensed quasars. In addition, uncertainties in the emitting source structure affect the interpretation of the flux ratios and thus provide a systematic floor unless resolved measurements of a subset of quasars can be achieved. The collecting area of the next generation of ground-based extremely large telescopes (ELTs; US-ELTP for the US has the strongest recommendation of the US Astro2020 Decadal Survey) with adaptive optics will make it possible to precisely measure flux ratios for a sufficiently large sample and allow a statistically precise inference of the structure at $\sim 10^7 M_\odot$. Additionally, in order to fully exploit the future data, new methods have been developed to accelerate the detection of such systems as well as perturbers in the images [531–535] and to analyze the collective perturbation of a population of subhalos [519, 536–541].

In conclusion, exploiting the full power of galaxy-scale strong lensing to constrain the nature of dark matter requires both exquisite data and sophisticated analysis tools. While the latter are being developed on precursor datasets, the former can only be obtained with the resolution and collecting area of extremely large telescopes (ELTs) or next generation VLBI. The analysis of the proposed data will require continued support for the development of methods, computational techniques and resources, as well as a robust interface between strong lensing predictions and cosmological numerical simulations of different dark matter theories.

5.2.5 21-cm at cosmic dawn

The first galaxies that populated our Universe were hosted in dark-matter (DM) halos much smaller in size than their counterparts today, with masses in the range of $M_h \sim 10^5−8 M_\odot$. As a consequence, their abundance can be used to study matter fluctuations with very large wavenumbers ($k \sim 10−100 \text{ Mpc}^{-1}$). The 21-cm line of hydrogen maps the formation of these galaxies during cosmic dawn, at $z \sim 10−20$, and thus provides a trove of information about the small-scale structure of dark matter.

The physics by which 21-cm photons are absorbed [542, 543] from the CMB is as follows. The UV light emitted by the first galaxies couples the hyperfine degrees of freedom of hydrogen to its thermal state through the Wouthuysen-Field (WF) effect [544]. This allows hydrogen to resonantly absorb CMB photons with a 21-cm wavelength, which gives rise to 21-cm absorption (or a negative 21-cm temperature brightness). Remnants from the first stellar formation (likely by X-ray binaries [545]) will later heat up the hydrogen, producing observable 21-cm emission and ending cosmic dawn.

The timing and shape of the 21-cm signal can be used to time the formation of the first galaxies, and thus to test the small-scale structure of DM on halos with $M_h \sim 10^5−8 M_\odot$. Models of non-cold DM that produce a suppression of power at small scales (e.g., warm, fuzzy, or interacting DM) will show a delay in the formation of the first galaxies, and thus of the 21-cm signal. Likewise, an enhancement of power would appear as an earlier 21-cm cosmic-dawn trough, which can also be tested.
Cosmic-dawn 21-cm data comes in two flavors. The first are measurements of the sky-averaged global signal (i.e., the monopole), which are experimentally less costly, requiring only one antenna, but more prone to systematics (e.g., [546]). The first detection by the EDGES collaboration [547] (see, however, Ref. [548] which disfavors it) would indicate the presence of the first galaxies by $z = 20$, when the signal goes into absorption. If interpreted to be cosmological, this signal can be used to place limits on DM at small scales. For instance, the mass of warm DM (WDM) is bound to be above $m_{\text{WDM}} > 6.1$ keV [549] (see also [550, 551]), whereas for fuzzy (i.e., wave-like) DM its mass has to be above $5 \times 10^{-21}$ eV [552]. The EDGES signal can also also be used to measure the running of the scalar tilt at the 0.01 level [553]. Note that this argument is independent of the anomalous EDGES depth [554–558], as it relies just on the timing of the signal and not its amplitude.

The second are measurements of the 21-cm fluctuations, which are performed by interferometers such as HERA [559], LOFAR [560], or the SKA [561]. While these data require more antennas, they are expected to be cleaner of both systematics and foregrounds that plague the global signal [562–564]. Current data just includes upper limits, and thus cannot be used to place limits on the small-scale structure of DM. However, upcoming data on fluctuations will yield a clear picture of structure and galaxy formation at high redshifts [565, 566]. For instance, HERA data will allow for a measurement of the small-scale matter power spectrum up to scales $k \approx 100$ Mpc$^{-1}$ [36], see Fig. 8. As a consequence, it will probe fuzzy or self-interacting DM beyond the reach of the global signal [567, 568].

In summary, the different astrophysical environment of the high-$z$ cosmic-dawn epoch, when the first galaxies formed, makes it an attractive target for understanding dark matter in a regime that is complementary to lower-$z$ observables.

5.2.6 Extreme CMB lensing

The gravitational deflection of the light of the Cosmic Microwave Background (CMB) by intervening structure has been demonstrated to be a powerful way to measure the distribution of matter on large-scales [569], as well as the mass of halos with $M \geq 10^{13}$ $M_\odot$ [570]. Measurements of the weak lensing of the CMB have the potential to probe the distribution
of matter down to much smaller scales, given a CMB facility with sufficient sensitivity and resolution [571]. Using gravitational lensing to probe the matter distribution directly avoids potential uncertainty in the use of baryons to trace the location of dark matter. In addition, using the CMB as a backlight to measure gravitational lensing has the advantage that the CMB has 1) a well-known redshift, 2) well-known intrinsic statistical properties, and 3) is behind all structure out to \( z \sim 1100 \). Moreover, the CMB lensing power spectrum is most sensitive to structure at \( z \sim 2 \); since structure builds up hierarchically that makes it more sensitive to models that suppress small-scale structure than a comparable lower redshift probe.

A CMB lensing measurement that could probe \( k \sim 10 \, h \text{Mpc}^{-1} \) and distinguish between the matter power spectrum predicted by CDM and that predicted by models that can explain observational puzzles of small-scale structure, with at least 5\( \sigma \) significance, requires instrument noise levels in temperature of 0.5 \( \mu \text{K-arcmin} \) and a resolution of \( \sim 15 \) arcseconds over half the sky [244, 571–573] (see Figure 9). Such a measurement, which could be achieved by the proposed CMB-HD experiment [244, 574, 575], would result in a high-resolution map of the projected dark matter distribution over half the sky. The CMB-HD survey would be three times more sensitive and have six times higher resolution than CMB-S4; for comparison, CMB-S4 and precursor CMB experiments can reach \( k \sim 1 \, h \text{Mpc}^{-1} \) given their sensitivity and noise levels.

One challenge facing the interpretation of the small-scale dark matter distribution is the degeneracy between suppression of structure due to baryonic effects and suppression due to alternate models of dark matter. However, the way baryonic effects and dark matter models impact the shape of the CMB lensing power spectrum is sufficiently different that they can potentially be differentiated [571]. Another challenge is that such small-scale CMB lensing measurements at 0.5 \( \mu \text{K-arcmin} \) noise levels gain most of their signal-to-noise ratio from the lensing of CMB temperature maps, as opposed to polarization maps. However, CMB temperature maps have more foreground signals in them, such as the thermal Sunyaev-Zel’dovich (tSZ) effect, the kinetic Sunyaev-Zel’dovich (kSZ) effect, and the Cosmic Infrared Background (CIB). Folding in the tSZ and CIB foregrounds to forecast what a CMB-HD survey could achieve results in distinguishing a benchmark \( 10^{−22} \text{eV} \) FDM model from a CDM model at the 5\( \sigma \) level [572].

Future work, with high-resolution millimeter-wave simulations and machine learning techniques are warranted to explore the full potential of foreground mitigation techniques. For even lower noise levels than discussed above, the small-scale CMB lensing signal gains most
of its signal-to-noise ratio from CMB polarization maps, which are much cleaner in terms of contaminating foregrounds.

### 5.3 Theory Connections

The experiments presented in this section will decisively test the yet-unconfirmed prediction of CDM that there are a large abundance of small-scale collapsed halos below the galaxy formation threshold. Deviations from the CDM predictions at mass scales of \( \sim 10^7 \, M_\odot \), where galaxy formation models have minimal impact on the structure and abundance of dark matter halos, would require physics beyond ΛCDM with a particle physics model of dark matter.

The fact that individual experiments probe different cosmic epochs allows one to disentangle \textit{ab initio} and \textit{in situ} dark matter physics at scales beyond the galaxy formation threshold. A consistent deviation of small-scale structure from CDM directly probes the \textit{ab initio} DM nature and hence such observations are sensitive to the DM thermal history (see Sections 2.1.1 and 2.1.2), as well as to any nongravitational interactions that DM particles experienced in the early Universe (see Sections 2.1.4 and 2.3.1). Deviation from CDM originally solely at late times would directly point towards a \textit{in situ} nature of dark matter and can quantify e.g. SIDM (Section 2.2.1). Hybrid impacts (Section 2.3) can also be constrained.

### 6 Stellar and Planet Mass Scale

#### 6.1 Scale and Objects

The standard cosmological paradigm with (1) a nearly scale-invariant adiabatic primordial density fluctuations and (2) collisionless cold dark matter (CDM) predicts that dark matter is clustered into halos with a power-law mass spectrum extending from massive galaxy clusters \((10^{15} M_\odot)\) down to planet-scale masses.\(^3\) If the dark matter is composed of canonical weakly interacting massive particles (WIMPs) with mass \(\sim 100\) GeV, the minimum halo mass set by the free-streaming length is \(\sim 10^{-6} M_\odot\) [10]. This simple model has proved adequate to explain observations from large-scale structure, down to the least luminous galaxies that reside in halos of \(~10^8 M_\odot\) [30, 449, 577, 578]. Over the next 5–10 years, observations of ultra-faint galaxies, galaxy–galaxy strong lensing, stellar stream perturbations, and 21cm tomography will all probe dark matter halos that are small enough \((10^6 – 10^8 M_\odot)\) to contain few if any stars (see Section 5).

Robust identification of dark matter structures well below the threshold of galaxy formation is a qualitatively new observational frontier that is currently mostly unexplored. Measurements in this regime are generically sensitive to a wide range of models that predict damped or enhanced structure — see the discussion in Sec. 2. While a suppression of DM structure compared to ΛCDM on sub-solar scales is well-motivated in the aforementioned scenarios, all existing and proposed probes of these scales cannot reach even the ΛCDM

---

\(^3\)The absolute lower limit is not known since \(N\)-body simulations suffer from resolution effects at these small scales [576].
level of substructure, so they will be difficult to probe. We therefore focus on scenarios with enhanced structure instead. These models offer the exciting opportunity to study the dark sector directly through its gravitational interactions alone. Enhanced structure originates from either an increased initial density contrast leading to an earlier gravitational collapse, or through DM self-interactions that enable the formation of dense clumps, or some combination of the two channels. Models of the former type include cosmological modifications like EMDE (see Sec. 2.1.3) and inflationary production of vector bosons; in this case the characteristic mass scale of DM substructure is connected to cosmological and microphysical parameters like the duration of EMD [5] or the mass of the DM particle [7]. Cosmological and model-specific constraints then require that the mass of DM subhalos with enhanced density satisfy $M \lesssim 10^{-3} M_\odot$. The characteristic density of these objects is determined by the time of their gravitational collapse which can begin at matter-radiation equality or even earlier (if the initial density contrast is $\gg 1$); in contrast, typical $\Lambda$CDM halos with these masses only form for $z < 30$. Since the matter density scales as $(1 + z)^3$, subhalo density can be orders of magnitude larger than in the standard cosmology. In models where DM self-interactions are important, a much broader range of substructure mass and density is possible. For example, atomic dark matter models can produce clumps with $M \gtrsim 10^{-3} M_\odot$ and densities far larger than $\Lambda$CDM analogues across $10^{-24} < M/M_\odot < 10^9$ [98]. Clearly there is a rich portfolio of models that can be tested by probing substructure on sub-solar mass scales.

6.2 Observational Probes

Below, we highlight several proposed experimental approaches that could achieve sensitivity to stellar- and planet-mass-scale extended dark matter structures by detecting their gravitational influence on a variety of observables.

6.2.1 Precision astrometry

Galactic subhalos can induce detectable proper motions in luminous background sources (e.g., stars and QSOs) due to gravitational lensing [579–581]. Several methods have been identified to extract this signal with the ultimate goal of characterizing the lensing subhalo population, e.g., identification of proper motion outliers due to lensing by compact objects, the use of matched filters to identify extended subhalos from locally correlated proper motions [579, 580], and the spectral decomposition of the astrometric velocity and acceleration fields over large regions of the sky [581]. A proof-of-principle analysis using the matched filter approach and data from \emph{Gaia} DR2 was recently presented [580], demonstrating sensitivity to $\sim 10^8 M_\odot$ subhalos extended at the $\sim$ pc level—much less dense than what established methods relying on photometric microlensing are typically sensitive to. The different proposed methods are complementary and, taken together, will be able to probe a wide variety of scenarios—from NFW subhalos down to $\sim 10^6 M_\odot$ to populations of compact objects—when applied to future astrometric datasets such as those from the proposed space-based ESA \emph{Theia} mission [582], the upcoming NASA Nancy Grace Roman Space telescope,
and radio observations from Square Kilometer Array (SKA).

### 6.2.2 Cluster caustics

Gravitational lensing by galaxy clusters produces caustics in which background objects are magnified by orders of magnitude, and the properties of the images provide a sensitive probe of the mass distribution in the lenses. Observations of background galaxies crossing through such caustics have already been employed to constrain the fraction of DM in compact objects [583], and similar techniques may be used to probe DM substructure. In particular, if the mass distribution of the lens is smooth, images appear in symmetric pairs across critical curves. Thus, asymmetries between these images probe inhomogeneities in the lens, and with detailed modeling, such asymmetries can be related to the abundance of substructure.

Analysis of asymmetries in the caustic-crossing arc SGAS J122651.3+215220 suggests that they are best accounted for by DM subhalos of mass $10^6$–$10^8 M_\odot$ [584].

Time-domain observations of caustic-transiting stars can be sensitive to DM substructure on solar and sub-solar scales [583, 585, 586]. These stars are highly magnified (reaching absolute of magnifications of $|\mu| \sim \mathcal{O}(10^3 - 10^4)$), thereby making them extremely sensitive to fluctuations in the lens convergence field (at the level of $|\mu|^{-1}$) and therefore the underlying DM substructure. The macro-caustic due to the cluster lens is expected to be broken up into a network of micro-caustics by the presence of microlenses (stars, black holes, or other compact objects). When a magnified source star is microlensed by such a compact object, the time-variability of its apparent luminosity enables the star to be identified. The event lasts from days to weeks, during which time the source image is stretched by a factor of $\sim |\mu|$ and moves a distance $\sim |\mu|v_{\text{rel}}\tau$, where $v_{\text{rel}} \sim 1000$ km/s is the relative velocity between source and lens, and $\tau$ is the duration. This means that the image brightness is sensitive to density fluctuations on length scales from $\sim 4 \times 10^4$ AU down to $\sim 5000$ AU ($R_\star/100 R_\odot$), where $R_\star$ is the source radius [587]. The maximum size of lensing convergence fluctuations on these scales, $\Delta\kappa$, is determined by the fraction of DM in subhalos, $f$, and their surface density $\sim M/r^2_\star$ relative to the critical density $\Sigma_{\text{cr}}$. $\Delta\kappa \sim \sqrt{f(M/r^2_\star)/\Sigma_{\text{cr}}}$. This can easily exceed $1/|\mu|$ for interesting examples of DM subhalos. These tiny density fluctuations therefore correspond to $\mathcal{O}(1)$ brightness fluctuations.

Microlensing of highly magnified stars can be used to place constraints on DM substructure both from the observation of the absolute magnitude of the microlensing events, and from the properties of the light curve in an individual event. The former approach has already been used to place bounds on DM fraction in compact objects in Ref. [583]. In Ref. [587] it was argued that fluctuations in the microlensing event lightcurve are directly related to the matter power spectrum at small scales, as described above. This is potentially an extremely powerful technique that is sensitive to non-compact structures which are unable to cause strong lensing effects by themselves (as opposed to, for example, black holes). While the cadence of current observations of extremely-magnified star candidates [588–590] are insufficient to constrain this power spectrum, future measurements can access DM substructure with $M \lesssim 10^{-3} M_\odot$ and densities within an order of magnitude of $\Lambda$CDM substructure. Sensitivity estimates of this technique in the context of several models (based on Ref. [587])

---

39

---

4The critical density $\Sigma_{\text{cr}}$ is the lens density needed to induce strong lensing.
can be found in, e.g., Refs. [96, 98, 104, 184].

The Vera C. Rubin Observatory will detect many more such cluster caustic lens systems. Microlensing events are expected to occur at least once a year for thousands of years [586]; therefore continued high-cadence photometric monitoring of these objects can yield unprecedented insight about the DM distribution on sub-solar scales.

6.2.3 Gravitational microlensing

Gravitational lensing due to the transit of a compact dark matter object near the line of sight to a source star may give rise to one or several lensed images with an enhanced total apparent luminosity. By convention, a microlensing event is defined by a fractional magnification of $\mu = 1.34$. From the non-observation of microlensing events, constraints can be derived on compact dark structures, a technique mostly studied for primordial black holes. Recently, such constraints have generalized to extended dark matter structures, such as subhalos, boson stars, and axion miniclusters [88, 89, 591]. The sensitivity to extended objects depends on their mass profile, with stronger constraints existing for more sharply peaked profiles. Current constraints exist for substructures as large as $10^3 R_\odot$ or about 10 AU for subhalos of $10^{-2} - 10^1 M_\odot$ [88], and reach down to $10^{-11} M_\odot$ for smaller structures. CDM subhalos would have radii of $\sim$pc in this range and are thus currently unconstrained through this probe.

Importantly, the sensitivity of microlensing experiments depends strongly on total observation time and cadence. Longer cadences provide greater sensitivity to more massive objects, and increased observation time increases the overall sensitivity of the survey. Dedicated microlensing studies could also take specific dark matter profiles into account, which typically lead to modified light curves in events.

6.2.4 Gravitational waves

Observations of gravitationally lensed distant quasars and galaxies have a long history of probing the invisible dark matter mass throughout the Universe, with high-resolution imaging of galaxy-scale lenses closing in on the $10^6 - 10^7 M_\odot$ scale. The recent advent of gravitational wave astronomy has the potential to push down the sensitivity to small dark matter halos by several orders of magnitude [592, 593]. Indeed, strongly lensed binary black hole (BBH) merger events detectable in the current LIGO band have periods similar to the Schwarzschild timescale of $\sim 10^2 - 10^3 M_\odot$ compact dark matter clumps. This coincidence of scales leads to a distinctive gravitational wave diffraction pattern that is imprinted on the waveform of the gravitational wave signal. Since the lensing cross section increases significantly with larger source redshift, third generation gravitational wave detectors capable of detecting BBH mergers out to $z_s \sim 2 - 4$ offer the most promising path forward for this technique. Lensed fast radio bursts are expected to be sensitive to low-mass structures via similar diffractive effects [594–596].

6.2.5 Pulsar timing arrays

Recently, pulsar timing observations have been used to directly measure Galactic accelerations [365, 366] and the smooth component of the Galactic potential. In the future, long-term
measurements of the time-of-arrivals of light pulses from arrays of pulsars can be used to provide some of the most sensitive probes of substructure at small scales. Dark matter clumps moving around the Earth-pulsar system imprint characteristic shapes in the timing residuals measured by pulsar timing arrays and can be used to search for small-scale structure in the galaxy by their gravitational influence alone [141, 597–604]. Clumps passing close by to the Earth or pulsar accelerate the astrophysical body, resulting in a Doppler shift in the pulsar frequency. Additionally, if the clump passes through the line of sight it will cause a Shapiro time delay in the signal. This technique is sensitive not only to very compact objects such as black holes, but also to more diffuse objects such as DM minihalos [599, 601, 603, 604].

Standard ΛCDM sub-halos have been shown to be a challenging target for Pulsar Timing Arrays (PTAs) [141], mostly because the local abundance of these objects is drastically reduced by tidal disruption. However, models featuring a population of compact and early-forming sub-halos are expected to be within the reach of near-future PTAs thanks to the reduced impact of tidal effects on these ultra-dense objects.

While the viability of PTAs as probes of DM sub-halos has been established, progress still needs to be done to implement search strategies within the pipelines of PTAs collaborations. Recently a search for sub-halos with $M > 10^{-11} M_{\odot}$ has been implemented by using the analysis code enterprise developed by the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [605]. For these masses, the signal induced by a transiting sub-halos is large enough to be individually resolved, which allows for an easy parametrization of the signal shape. However, PTAs are expected to be sensitive to sub-halos as light as $M \sim 10^{-13} M_{\odot}$. The individual signal induced by these light sub-halos is not large enough to be individually resolved. However, the cumulative signal induced by the entire population could leave a detectable imprint in the timing residuals. The implementation of searches for this signal, which would look like a non-stationary noise in the residuals, is still underway.

### 6.2.6 Lensed fast radio bursts

Fast radio bursts (FRBs) are short and bright electromagnetic emission arriving from cosmological distances (see e.g., [606] for a review). Despite their unknown origins, FRBs are expected to advance our knowledge of dark matter and cosmology over the next decade [607–611]. Current facilities are detecting thousands of these events (for instance, the Canadian HI Mapping Experiment CHIME [612]), enough to perform statistical studies.

Of particular interest is the short temporal structure of FRBs, which allows us to study gravitational microlensing on the time domain (rather than the usual spatial one, e.g., [613]). As a consequence, FRBs have been proposed as powerful targets to search for DM-induced strong lensing [594]. The optical depth for a usual FRB to be lensed by DM, if all of it was composed of point-like lenses (for instance black holes), is $\tau_{\text{lens}} \sim 10^{-2}$. As a consequence, a fiducial observation of $10^4$ FRBs (as expected from CHIME) can test any fraction of DM in compact objects down to 1% of the total DM abundance. Direct microlensing (i.e., searching for a lensed FRB echo) can detect lenses with masses $M_{\text{lens}} \gtrsim 10 M_{\odot}$ [594] (see [614, 615] for current searches), whereas “femtolensing” (where the time delay is shorter and fringes appear in the FRB spectra) can test the range $M_{\text{lens}} \sim 10^{-4} - 10 M_{\odot}$ [595]. Searching for fuzzier lenses (for instance NFW halos) requires more FRBs, as the optical depth is reduced if the mass of the lens is not entirely contained within its Einstein radius. As a consequence,
more futuristic observatories (beyond CHIME or HIRAX) are likely required to measure NFW halo structure with FRBs, given the CDM prediction.

6.3 Theory Connections

For the sub-stellar and sub-Earth halo mass range, the most promising theoretical models to test with current and next-generation facilities are those that predict enhanced central densities relative to CDM halos. The majority of observational probes at this extremely low mass scale utilize gravitational lensing effects that are strongest for compact mass distributions. Already, several of the techniques described here have achieved sensitivity to stellar and planet mass primordial black holes, which can be viewed as a limiting case for the most centrally concentrated halos. In addition to primordial black holes (discussed in more detail in another Snowmass white paper), other particle physics models to target in the next decade include vector DM produced during inflation, post-inflationary production of QCD axions via the misalignment mechanism, and post-inflationary production of axions via strings (Section 2.1.2), as well as non-standard cosmologies with early matter domination (Section 2.1.3).

We advocate for enhanced theoretical, instrumentation, and data analysis efforts to develop experimental techniques to access stellar- and planet-mass-scale extended dark matter structures. To enable the success of these experiments, it is essential for the community to support dedicated theory and simulation efforts, both (1) to model the matter distribution and observables at extremely small scales, and (2) to differentiate between various dark matter and early Universe physics models using detailed and diverse sets of observations. At this stage, we identify multiple experimental approaches be explored in parallel to better quantify their sensitivity, robustness, and long-term potential, and to check for consistency between signals.

7 Conclusion

In this white paper, we have presented a wide-ranging set of astrophysical observations that probe the composition and structure of dark matter halos. Among these diverse probes are those that currently provide the strongest evidence of the existence of dark matter to date, and each of them has exciting potential to push boundaries and enable discoveries in the near future. Importantly, the macroscopic distribution of distribution of dark matter is sensitive to both the fundamental properties of dark matter and its production mechanism in the early Universe.

We have grouped the different probes in four different mass ranges, from the most massive bound structures of galaxy clusters, down to individual galaxies, pushing below the galaxy formation threshold where we expect halos to be depleted of stars and gas, and down to the stellar and planet mass scale. The ensemble of astrophysical experiments can probe the entire mass spectrum where we expect imprints of dark matter in our Universe. The key opportunity lies in the combination of multiple different measurements sensitive to different mass scales. A holistic approach of a combined probe analysis and modeling enables to pinpoint a potential non-gravitational signature of dark matter and directly link the macroscopic
Emerging cosmological structure with the microscopic particle nature of dark matter.

Given that the different individual experiments are also sensitive to different epochs of the Universe, the originating microphysical cause of the effect can be disentangled between \textit{ab initio} where early-Universe DM behavior modifies the initial conditions for structure formation and \textit{in situ} where a relatively high density of DM in the present-day Universe allows certain DM behaviors to become manifest. Overall, the astrophysical dark matter experiment probes an enormous range of dark matter models of which only few specific examples have been mentioned in this white paper. There are more opportunities to develop the phenomenology in terms of the early Universe as well as late-time evolution.

Complementary to the core message of this white paper linking astrophysical observations with dark matter microphysics, the community relies on state-of-the-art facilities and instrumentation together with accurate and precise simulation predictions. These topics are addressed in two accompanying Snowmass white papers.

References


Dark matter Physics from Halo Measurements


48


49


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


72


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


Dark matter Physics from Halo Measurements


