Dark matters on the scale of galaxies

Ivan de Martino*,1,2  Sankha S. Chakrabarty,1,2  Valentina Cesare,1,2  Arianna Gallo,1,2  Luisa Ostorero,1,2  and Antonaldo Diaferio1,2

1 Dipartimento di Fisica, Università di Torino, Via P. Giuria 1, I-10125 Torino, Italy
2 Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy

The cold dark matter model successfully explains both the emergence and evolution of cosmic structures on large scales and, when we include a cosmological constant, the properties of the homogeneous and isotropic Universe. However, the cold dark matter model faces persistent challenges on the scales of galaxies. Indeed, N-body simulations predict some galaxy properties that are at odds with the observations. These discrepancies are primarily related to the dark matter distribution in the innermost regions of the halos of galaxies and to the dynamical properties of dwarf galaxies. They may have three different origins: (1) the baryonic physics affecting galaxy formation is still poorly understood and it is thus not properly included in the model; (2) the actual properties of dark matter differs from those of the conventional cold dark matter; (3) the theory of gravity departs from General Relativity. Solving these discrepancies is a rapidly evolving research field. We illustrate some of the solutions proposed within the cold dark matter model, and solutions when including warm dark matter, self-interacting dark matter, axion-like particles, or fuzzy dark matter. We also illustrate some modifications of the theory of gravity: Modified Newtonian Dynamics (MOND), MOdified Gravity (MOG), and \( f(R) \) gravity.

PACS numbers:
Keywords: axions; cosmology; dark matter; dark matter theory; dwarf galaxies; gravity; Local Group; modified gravity; particle physics - cosmology connection; rotation curves of galaxies.

Contents

I. Introduction 2

II. Overview of the Cold Dark Matter model 4

III. Observational challenges of the Cold Dark Matter model 6
   A. The rotation curves of disk galaxies and the baryonic scaling relations 6
      1. Observational evidence 6
      2. Possible solutions within the CDM model 8
   B. The cusp/core problem 10
      1. Observational evidence 10
      2. Possible solutions within the CDM model 13
   C. The missing satellites problem 14
      1. Possible solutions within the CDM model 15
   D. The too-big-to-fail problem 17
      1. Possible solutions within the CDM model 19
   E. Planes of satellite galaxies 20
      1. Evidence of the orbital alignment of the satellites 21
      2. Comparison with the CDM predictions 21
      3. Possible solutions within the CDM model 22

IV. Possible solutions beyond the standard Cold Dark Matter 25
   A. Warm Dark Matter model 25
      1. Solutions to the observational challenges 26
   B. Self-interacting dark matter 28
      1. Solving small-scale issues with SIDM 28
      2. SIDM in particle physics 30

* Correspondence: ivan.demartino@unito.it
I. INTRODUCTION

The fundamental nature of matter and the concept of its constituent particles have always been a central topic of philosophy and of natural sciences throughout the history of human thought. The development of new tools applied to the observation of nature has often led to the discoveries of new physical phenomena. These discoveries directly or indirectly implied the existence of particles that were previously invisible [1].

In the last century, the concept of invisible or dark matter has reached the current connotation, although with some oversights, such as believing that the dark matter was made up of stars and/or interstellar medium too faint to be detected. It was J. H. Oort who, in 1932, first estimated the total amount of matter density near the Sun from dynamical data, and pointed out a discrepancy of a factor of up to 2 with the amount of the visible stellar populations [2]. Although this result is often considered to be the first evidence of the existence of dark matter, the discrepancy has now been alleviated by using more accurate observations of the stellar disk population [3–5]. For a comprehensive review, we refer to [6].

In 1933, F. Zwicky pointed out a discrepancy between the observed velocity dispersion along the line-of-sight of 8 galaxies in the Coma cluster (∼1000 km/s) and the one expected in a system of N massive galaxies in dynamical equilibrium (∼80 km/s) [7]. This discrepancy implied the presence of a large amount of invisible mass; this mass was still thought to be in the form of stars and/or gas which were not yet observable. This result of Zwicky traditionally marks the birth of the dark matter problem.

A new era began in the 1970s, when V. C. Rubin and W. K. Ford measured the rotation curve of the Andromeda galaxy (M31) out to 110 arcminutes away from the galactic centre, and estimated a mass-to-light ratio of $13 \pm 0.7 \, M_\odot / L_\odot$ at $R = 24$ kpc, corresponding to a total mass of $M = (1.85 \pm 0.1) \times 10^{11} \, M_\odot$ [8]. The measure of the 21-cm line emission of neutral hydrogen also suggested that the rotation curves of spirals fall off at large radii less rapidly than they should when most of the galaxy mass is concentrated in the optically luminous component [9, 10]. This flatness of the rotation curves led to the conclusion that galaxies are embedded in massive halos extending to large radii, as was suggested by theoretical studies of the stability of disk against the development of a bar [11].

In the last decades, the quest to unravel the intriguing puzzle of the dark matter has travelled along different paths. It was pointed out that the nature of dark matter could be either baryonic, or non-baryonic. Alternatively, it could be gravitational, namely the phenomena associated to dark matter actually is the signature of a break down of the standard theory of gravity. Each one of these hypotheses followed different paths to be validated.

The search for baryonic dark matter focused on sub-luminous compact objects, such as planets, brown dwarfs, white dwarfs, neutron stars, and black holes. These massive astrophysical compact halo objects (MACHOs) are now believed to form only a small fraction of dark matter. Microlensing surveys have been used to set an upper limit of 8% to the contribution of MACHOs in the mass fraction of the dark matter halo of the Milky Way [12, 13].

Nevertheless, there is the possibility that primordial black holes, which formed before the Big Bang nucleosynthesis and have masses below the sensitivity range of microlensing surveys, may form a substantial fraction of the total dark matter density. This idea was originally discussed by Carr and Hawking [14] in 1974. However, generating a relevant abundance of primordial black holes requires a substantial degree of non-Gaussianity in the power spectrum of the primordial density perturbations [15, 16]. Recently, the detection of gravitational waves has set a tight upper limit on the abundances of these black holes. This limit suggests that the black hole contribution to the dark matter abundance is at the level of a few per cent [17], as allowed by...
the constraints on non-Gaussianity obtained from the Planck satellite. [18–20]. Similar results were obtained by a recent microlensing survey of M31. In fact, if primordial black holes constitute the Milky Way and M31 dark matter halos, we expect that the M31 stars observed from Earth should generate \( \sim 10^5 \) microlensing events. However, in a 7 hour-long survey, only a single candidate event was identified. This result implies that the fraction of dark matter in primordial black holes in the mass range \( 10^{-11} - 10^{-6} \) M\(_\odot\) is \( \Omega_{PBH}/\Omega_{DM} < 0.001 \) [21].

One or more species of elementary particles beyond the Standard Model of particle physics are expected to make up the non-baryonic dark matter: Weakly Interacting Massive particles (WIMPs), QCD axions or ultra-light bosons are some of the suggested hypothetical particles. WIMPs are expected in the theory of supersymmetry, which dates back to the 1970s and supposes that, for any given fermion, there is a boson with the same quantum numbers, and vice versa [22]. In this case, there would be many electrically neutral and weakly interacting massive particles that could be cosmologically abundant and could play the role of dark matter. Intriguingly, a supersymmetric particle with mass in the range \( \sim 1 \) GeV/10 TeV would give rise to a relic density consistent with the observed dark matter density [23].

The searches for these particles, based on methods of direct or indirect detection, are ongoing. Despite the large effort, no direct detection has been claimed to date [24] and claimed indirect detections have been questioned [25–27]. However, supersymmetric particles are not the only particles that can play the role of dark matter. As mentioned above, there are other strongly motivated candidates, such as QCD axions and fuzzy dark matter [28, 29]. QCD axions are suggested by the solution of the Strong CP problem in the Standard Model [30, 31], while fuzzy dark matter arises from the compactification of extra dimensions in the String Theory landscape [32]. We will discuss these models in Section IV.

Alternatively, the phenomena attributed to dark matter could originate from a modification of the theory of gravity. Along this line, in 1983, the Modified Newtonian Dynamics (MOND), was proposed to explain the observational evidence attributed to dark matter by modifying the second law of dynamics with the introduction of an acceleration scale [33–35]. Later on, more theories of gravity have been proposed [36–38], all of them reporting successes and failures. Unfortunately, none of them appears to be able to explain all the relevant observations in a single general framework. In Section V we will discuss some of these models.

In the standard cosmological model, dark matter consists of massive particles that weakly interact with ordinary matter and that decoupled from the primordial plasma when they were non-relativistic. This Cold Dark Matter (CDM) scenario encounters some difficulties in describing structures at galactic scales [39–44]. These difficulties include the cuspy/core problem, the too-big-to-fail problem, and the problem of the planes of satellite galaxies.

The cusp/core problem is the discrepancy between the flat dark matter density profile observed at the centres of dwarf and ultra-faint galaxies, and the cuspy profile predicted in collisionless N-body simulations [45–48]. In particular, N-body simulations show cuspy density profiles for dark matter halos of galaxy size; the halo density increases with decreasing radius \( r \) as \( r^{-\beta} \), with \( \beta \) in the range \( \sim [1 – 1.5] \). These slopes do not match the cores favored by the observed rotation curves [39, 45, 46, 49, 50]. Nevertheless, modelling the kinematics of stars in dwarf galaxies does not lead to a clear conclusion to whether these galaxies are dominated by a core or a cusp in their innermost regions [51]. The missing satellites problem is the fact that the dark matter halos of galaxies like the Milky Way are predicted to have a number of dark matter subhalos which is an order of magnitude larger than the number of satellites observed around the Milky Way or other comparable galaxies [52, 53]. A related, but independent, issue is the too-big-to-fail problem [41, 54–56], which manifests in the central densities of the most massive dark matter subhalos formed in \( \Lambda \)CDM simulations; these densities are systematically larger than the central densities of the brightest classical Milky Way satellites, as inferred from their stellar kinematics. In principle, associating the classical satellites to dark matter subhalos with smaller central densities and smaller mass would erase the discrepancy; however, this association would clearly imply that the most massive subhalos would not host a galaxy. In other words, these subhalos “failed” to form stars even though less massive subhalos succeeded in doing so. Finally, the problem of the planes of satellites refers to the fact that, in the galaxy systems of the Milky Way, of M31, and of Centaurus A, the satellite galaxies reside in a thin plane and they generally corotate. This satellite configuration appears unlikely in the simulations of the standard CDM model [57].

In addition to these problems, the current CDM paradigm faces other challenges. The Tully-Fisher relation suggests the existence of an acceleration scale that is adopted as a fundamental constant in MOND [58]. The total angular momentum of the visible component also differs by a factor of 2 – 3 from the angular momentum of CDM halos [59, 60].

This review will focus only on some of the aforementioned problems that the standard dark matter paradigm encounters at galactic scales. We will present the state of the art of these problems in the context of the standard CDM model, and then review some alternative solutions that may reside either in a change of the dark matter paradigm or in models of modified gravity.

Section II recalls the standard cosmological model with the current constraints on the cosmological parameters. Section III discusses some of the observational challenges: the radial acceleration relation (Section § IIIA), the cuspy/core problem (Section § IIIB), the missing satellites problem (Section § III C), the too-big-to-fail problem (Section § III D), and the planes of satellite galaxies problem (Section § III E). Section IV is devoted to possible explanations beyond the Standard Model of particle physics such as warm and self-interacting dark matter (Sections § IV A and § IV B), QCD axions (Section § IV C), and fuzzy dark matter (Section § IV D). Finally, Section V explores possible explanations provided by alternative theories of gravity, specifically Modified Newtonian Dynamics (MOND) (Section § V A), Scalar-Vector-Tensor theory of gravity, also known as MODified Gravity (MOG) (Section § V B), and \( f(R) \)-gravity (Section § V C). In both Sections IV and V, we discuss whether each of the specific models may resolve the issues at galactic scales that we mention in Section III. Finally, Section VI
summarizes the main topics of the review and provides some future perspectives.

II. OVERVIEW OF THE COLD DARK MATTER MODEL

Over the last decades, the standard model for the evolution of the Universe, the ΛCDM or concordance model [61], was established by many independent observations, including the Cosmic Microwave Background (CMB) temperature fluctuations [20, 62–66], the power spectrum of the matter density perturbations [67–69], the luminosity distances to supernovae SNeIa [70–77], and the expansion rate of the Universe [78–82]. The current constraints on the cosmological parameters have reached unprecedented accuracy [83].

In this model, the present period of accelerated expansion is driven by the cosmological constant Λ that provides an energy density \( \Omega_{\Lambda,0} = 0.686 \pm 0.020 \) which is in units of the critical density \( \rho_c = 3H_0^2/8\pi G \) [63]. The cosmological constant can be assimilated to a perfect fluid with an equation of state \( p = w \rho \), between pressure \( p \) and energy density \( \rho \), with the constant parameter \( w = -1 \). The second largest contribution to the total energy-density budget of the Universe is dark matter, which is needed to explain the dynamics of galaxies and the large-scale structure. Its energy density is \( \Omega_{\text{DM},0} = 0.314 \pm 0.020 \) [63]. The present density of ordinary, or baryonic, matter is \( \Omega_{b,0}h^2 = 0.02207 \pm 0.00033 \) [63]. Summing the contributions of the cosmological constant, dark matter, and baryonic matter yields the curvature of the Universe \( \Omega_k,0h^2 = -0.037^{+0.044}_{-0.042} \) [63], where \( h \) is the Hubble constant in units of 100 km s\(^{-1}\) Mpc\(^{-1}\). This curvature makes the Universe spatially flat.

The concordance model successfully describes the homogeneous and isotropic Universe, and the dynamics of cosmic structures. However, there are some tensions with the values of its parameters \( \Omega_{\Lambda,0}, \Omega_{k,0} \), and the Hubble constant \( H_0 \).

\( \Omega_{\Lambda,0} \) is extremely small compared to the expectations from quantum field theory [84]. Quantum chromodynamics predicts a vacuum energy density of \( \rho_{\Lambda} \sim 10^{71} \text{ GeV}^4 \), whereas the cosmological upper bounds on the cosmological constant give \( \rho_{\Lambda} \sim 10^{-47} \text{ GeV}^4 \). This large discrepancy leads to a fine-tuning problem [84]. To reconcile the values of the vacuum energy density at cosmological and quantum scales, different solutions have been proposed, such as the anthropic principle or a cyclic model of the Universe [85–90]. Comprehensive reviews on the cosmological constant and dark energy are given in [91–93].

Similarly, tensions arise in the values of the Hubble constant estimated from the CMB and from measurements in the nearby Universe, and also, to a less degree, in the values of \( \sigma_8 \), the normalization of the power spectrum of the mass density perturbations on cosmic scales, inferred from the CMB and from large-scale structure surveys [63]. The Hubble constant from the CMB measurements \( H_0 = 67.27 \pm 0.60 \text{ km s}^{-1}\text{Mpc}^{-1} \) [63] is at a 4.4σ tension with the local value \( H_0 = 73.52 \pm 1.62 \text{ km s}^{-1}\text{Mpc}^{-1} \) [94, 95]. The clustering parameter \( \sigma_8 \) for the ΛCDM model, with fixed effective number of neutrino species \( N_{\nu \text{eff}} \) and fixed total mass \( \Sigma m_{\nu} \), obtained by the Planck collaboration [63], is \( \sigma_8 = 0.8149 \pm 0.0093 \), implying \( S_8 \equiv \sigma_8(\Omega_\text{m}/0.3)^{0.5} = 0.811 \pm 0.011 \). On the contrary, a tomographic weak gravitational lensing analysis of data from the Kilo Degree Survey [96] led to \( S_8 = 0.745 \pm 0.039 \), which is at a 2.3σ tension with the Planck results. Recently, it has also been shown that the last data release of the Planck experiment might suggest a spatially closed, rather than flat Universe [97]. All these tensions remain open problems to date and the subject of intense investigation.

Dark matter is expected to consist of stable massive particles beyond the Standard Model of elementary particle physics. Dark matter can thus clump in self-gravitating structures embedding galaxies and clusters of galaxies. Many pieces of evidence support the existence of dark matter based on its gravitational effects, once General Relativity is assumed to be the correct theory of gravity. For example, the mass distribution within galaxy clusters can be reconstructed with the analysis of the gravitational anisotropy fluctuations of the CMB, to grow and form the cosmic structures we observe today [e.g., 100, 101]. In fact, in the history of cosmic expansion, for models where the dark matter and dark energy are not separated from the other components of the Universe, dark matter decouples from the primordial plasma much earlier than baryons; the primordial fluctuations in the dark matter density field thus starts growing earlier and, at recombination epoch, they are larger than the baryon density perturbations that are still coupled to the perturbations of the background radiation field that generates the CMB anisotropies. After the recombination epoch, baryonic matter falls into the gravitational potential wells of the dark matter halos and form the cosmic structure (see [e.g., 1, 103–106] for details).

In these models, dark matter particles are thus in thermal equilibrium with the cosmic plasma before the decoupling epoch, when they get out of kinetic equilibrium, as they become non-relativistic, at temperature \( T = T_d \ll m_{\text{DM}} \), where \( m_{\text{DM}} \) is the mass of the dark matter particle. In the standard CDM model, the dark matter particles are so massive that \( T_d \gg T_b \), where \( T_0 \) is the plasma temperature at recombination epoch. They thus decouple from other particles and start moving freely at non-relativistic speeds in the early Universe. The comoving number density of dark matter particles freezes out when the creation and annihilation of the dark matter particles are inhibited. In general, the freeze-out cosmic temperature \( T_f \gg T_d \). The comoving
The number density of dark matter particles is thus set by their annihilation cross section at this epoch [23]

\[
\Omega_{\chi} h^2 = \frac{3 \times 10^{-27} \text{ cm}^3/\text{sec}}{\langle \sigma v \rangle_{\text{ann}}}. \tag{1}
\]

Intriguingly, for WIMPs, with mass in the range \(\sim 1\text{ GeV} - 10\text{TeV}\), the annihilation cross section \(\langle \sigma v \rangle_{\text{ann}}\) gives \(\Omega_{\chi} h^2\) comparable to the observed dark matter density \(\Omega_{\text{DM,0}} \sim 0.3\) [63]. This coincidence is usually called “the WIMP miracle”.

Moreover, WIMPs naturally arise in the supersymmetric (SUSY) extension of the Standard Model of particle physics [23], where each Standard Model particle has a supersymmetric partner, and the lightest one is a good candidate for the dark matter particle. Another possibility is that WIMPs arise from compactifications of extra dimensions [107].

Nowadays, direct and indirect searches for WIMPs as well as other dark matter particle candidates are ongoing. Direct searches look for a scattering process where a dark matter particle in the Milky Way halo interacts with an atomic nucleus of a detector whose recoil generates a detectable release of energy. Experiments of this kind are, for example, LUX and XENON 1T [24, 108].

Indirect detections of dark matter particles include searches for neutrinos arising from the annihilation of WIMPs in the centre of the Earth or the Sun, where WIMPs can concentrate at the bottom of the gravitational potential well. This approach is adopted by experiments like IceCube and Super-K [109, 110]. Similarly, WIMP annihilation in the Milky Way can enrich the cosmic rays that we detect on Earth with positrons and antiprotons [111, 112]. Similar enrichment is expected in particle colliders, like CMS at the Large Hadronic Collider (LHC), which looks for the interactions between dark matter particles and the fermions of the Standard Model. A comprehensive review on the results of the direct and indirect searches is given in [113].

Figure 1 shows the allowed region, that is below each sensitivity curves, in the plane of the dark matter-nucleon cross section and the dark matter particle mass. The case of the DAMA/LIBRA experiment is representative of the complexity of this kind of investigation. The DAMA/LIBRA experiment searches for dark matter particles in the Milky Way’s halo. The number of events in the 2-6 keV energy range has an annual modulation that can mirror the relative velocity between the dark matter particle and the Earth orbiting around the Sun. This modulation has a statistical significance of \(\sim 9\sigma\) [25, 26]. However, the signal is only consistent with values of the interaction cross-section and of the mass of the dark matter particle that are within the regions of Figure 1 that appear to be excluded by other experiments [24, 108, 111, 112]. Alternative explanations of this signal are still intensively debated [27].

Figure 1: Upper limits on the spin-independent (SI) elastic dark matter-nucleon cross-section as a function of mass of the dark matter particle. The figure is reproduced from [113].
Although the overall picture appears solid, the fundamental nature of dark matter still remains an intriguing mystery. In the next section, we focus on the main problems that the CDM model encounters at galactic scales and the possible solutions within the context of CDM. In Sections IV and V, we discuss some alternatives to the CDM paradigm.

III. OBSERVATIONAL CHALLENGES OF THE COLD DARK MATTER MODEL

Despite being successful on cosmological scales, the CDM model faces persistent challenges on the scales of galaxies, which are mostly related to the dark matter distribution in the central regions of the galactic halos and to the properties of the dwarf galaxies [41–43, 114, 115]. Here, we focus on these issues and discuss viable solutions within the CDM scenario.

A. The rotation curves of disk galaxies and the baryonic scaling relations

The rotation curve of disk galaxies is one of the most important pieces of evidence of the existence of dark matter: most of the luminous matter, namely stars, gas, and dust, is concentrated in the inner part of the galaxy and, if this luminous matter is the only galaxy component, we expect the rotation speed to fall off as \( v^2(r) \propto r^{-1} \) in the outer regions where the luminous matter density substantially decreases. However, the outermost visible objects, namely clouds of neutral hydrogen, move as fast as the inner objects. We thus observe a fairly flat rotation curve for most disk galaxies [100, 101].

The flatness of the rotation curve can be generated by a pressure-supported spheroidal halo embedding the entire galaxy [116]. This matter is dark as it remains undetected in any electromagnetic band. In the CDM model, assuming spherical symmetry, the mass density profile of the dark halo is well described by the Navarro-Frenk-White (NFW) profile [117]: it contains two free parameters, the virial mass \( M_{\text{vir}} \) and the concentration parameter \( c \) [120]. The introduction of the dark halo in the description of galaxy dynamics comes with an additional degree of freedom: reproducing the measured rotation curve of a galaxy requires a fine-tuning of the relative contribution of the galactic disk and the dark matter halo to the gravitational pull [121, 122]. This fine-tuning is known as the disk-halo conspiracy. A detailed analysis of a sample of more than a thousand disk galaxies quantifies the correlation between the luminous and dark matter and suggests the existence of a universal rotation curve [123, 124].

Additional pieces of evidence quantify the discrepancy between the amount of mass estimated by its electromagnetic emission and the mass required to describe the observed kinematics of disk galaxies: the Baryonic Tully-Fisher Relation (BTFR) [125, 126], the Mass Discrepancy Acceleration Relation (MDAR) and the Radial Acceleration Relation (RAR) [127–129]. We describe these relations in details below. Here, it suffices to say that they show a tight correlation between kinematic quantities and quantities associated to the baryonic component. The crucial feature of these correlation is their small intrinsic scatter, generally consistent with the observational uncertainties alone. This feature is surprising, because, in the CDM scenario, cosmic structures form hierarchically through stochastic mergers, and we would expect a large scatter, mirroring the merging history of each galaxy. On the contrary, the lack of a relevant scatter implies that, irrespective of the merging history of the galaxy, its dark matter halo, that represents \( \sim 90\% \) of the galaxy mass and sets its dynamical properties, adjusts its properties to those of the luminous disk, that only contains \( \sim 10\% \) of the galaxy mass. In other words, in spite of the dark matter halo contributing to the majority of the total mass of the galaxy, the kinematic properties of the disks are found to be strongly correlated with the luminous matter. This correlation appears thus challenging for the CDM paradigm.

1. Observational evidence

Rotation curves of high surface brightness (HSB) galaxies are properly described when we adopt the NFW density profile for the dark matter halo. However, some issues emerge from the analysis of the diagnostic data of the SPARC sample, that contains 171 galaxies of different luminosity and morphological type [120, 130]: (i) the \( c-M_{\text{vir}} \) relation, or mass-concentration relation, is steeper than expected from cosmological N-body simulations [131–134]; (ii) the mass-to-light ratio in the 3.6 \( \mu m \) band, \( M_3/L_{3.6} \), appears unphysically negative for 51 out of 171 galaxies [120]; and (iii) the Stellar-Population-Synthesis (SPS) models overestimate the observed correlation between \( M_3/L_{3.6} \) and the color index \( (B-V) \), independently of the chosen Initial Mass Functions (IMF) [120].

---

1 The virial radius \( r_{\text{vir}}(z) \) at redshift \( z \) is the radius of a spherical volume within which the mean mass density is \( \Delta_r(z) \) times the critical density of the Universe \( \rho_c(z) = 3H^2(z)/8\pi G = 2.775\times 10^2(z)^2(1+z)^3 \text{ M}_\odot \text{ Mpc}^{-3} \), with \( E(z) = [(\Omega_m(1+z)^3 + (1-\Omega_m - \Omega_L)(1+z)^2 + \Omega_L)] \), and \( \Omega_m \) and \( \Omega_L \) the parameters of the background Friedmann model. \( \Delta_r(z) \) is the solution to the collapse of a spherical top-hat density perturbation at the time of virialization. The virial mass is thus \( M_{\text{vir}} = 4\pi\Delta_r(z)\rho_c(z)r_{\text{vir}}^3/3 \) [118, 119].
Additional challenges for the CDM model arise from the baryonic-dark matter scaling relations mentioned above. The total baryonic mass of a galaxy, \( M_d \), and the asymptotic flat velocity of its rotation curve, \( V_c \), set by the depth of the gravitational potential well of the dark matter halo, obey the BTFR \[125, 126, 135, 136]:

\[
M_d = AV_c^4 ,
\]

(2)

where the normalization constant is \( A = 47 \pm 6 \, M_\odot \, \text{km}^{-4} \, \text{s}^4 \) \[135\]. \( A \) can also be written as \( A \sim (Ga_0)^{-1} \), the product between the gravitational constant \( G \) and an acceleration scale \( a_0 = 1.2 \times 10^{-10} \, \text{m/s}^2 \).

The observed BTFR is shown in Figure 2. Circles indicate measurements of the velocities from the asymptotically flat part of the rotation curves; the squares represent measurements of \( V_c \) as half of the HI emission line width at 20% of the peak value, whereas the stellar mass is inferred from photometric observations in the \( I \)-band \[137\], \( H \)-band \[138\], \( K' \)-band \[139\], or \( B \)-band \[140, 141\]. The total baryonic mass of the disk is estimated as \( M_d = M_* + M_{\text{gas}} \), where \( M_* \) and \( M_{\text{gas}} \) are the total mass in stars and gas, respectively. \( M_* \) is estimated as \( M_* = \Upsilon_* L \), where \( L \) is the total disk luminosity and \( \Upsilon_* \) is a constant mass-to-light ratio. \( M_{\text{gas}} \) is given by \( 1.4M_{\text{HI}} \), where \( M_{\text{HI}} \) is the total estimated mass of neutral hydrogen and the factor 1.4 is the standard correction to account for helium and metals. The black solid line shows an unweighted linear fit to the points with stellar mass estimated in the \( I \)-, \( H \)-, and \( K' \)-bands. The \( B \)-band data are not included in the fit because this band is not a robust indicator for the stellar mass. The slope of this straight line is \( b \approx 3.98 \pm 0.12 \).

**Figure 2:** Baryonic Tully Fisher Relation (BTFR): total baryonic disk mass of galaxies against the rotation velocity \( V_c \). Circles and squares represent the data derived from \[138\](red), \[139\](black), \[137, 142\](green), \[140\](light blue) and \[141\](dark blue). The black solid line shows an unweighted linear fit. The figure is reproduced from \[125\].

For mass-to-light ratios in the 3.6 \( \mu \)m band, \( \Upsilon_{[3.6]} \), larger than \( \sim 0.2 \, M_\odot/\text{L}_\odot \), the intrinsic scatter of the BTFR is smaller than the minimum intrinsic scatter of 0.15 dex predicted by semi-analytic models of galaxy formation in the \( \Lambda \)CDM model \[143, 144\], based on the \( c-M_{\text{vir}} \) relation of CDM simulations \[131\]. The intrinsic scatter of the BTFR reaches a minimum of \( \sim 0.10 \) dex for \( \Upsilon_{[3.6]} \gtrsim 0.5 \, M_\odot/\text{L}_\odot \). More importantly, the BTFR residuals show no correlations with galaxy radius and surface brightness, at odds with \( \Lambda \)CDM galaxy formation models \[145\]. In addition, \( \Lambda \)CDM simulations predict a slope \( b \sim 3 \), which implies an 8\( \sigma \) tension with the observed slope \( b = 3.98 \pm 0.12 \) \[125\].
The MDAR relates, at each radius $R$ in the galactic disk, the Newtonian acceleration $g_N(R)$, generated by the observed baryonic surface mass density, to the mass discrepancy $[V(R)/V_b(R)]^2$, where $V$ is the observed velocity and $V_b$ is the velocity that would be generated by the baryonic matter. The Newtonian acceleration and the mass discrepancy are anti-correlated [144, 146]. The mass discrepancy $[V(R)/V_b(R)]^2$ remains close to one for large baryonic accelerations $g_N(R)$, whereas it increases for accelerations smaller than the acceleration scale $a_0$ found with the normalization of the BTFR (see Figure 3).

For both the BTFR and the MDAR, the scatter depends on the value of the adopted mass-to-light ratio [146]. Specifically, the intrinsic scatter in both relations is minimized by the same mass-to-light ratio, consistent with SPS models [146]. Intriguingly, when we use MOND (Section V A) to describe the rotation curves [146, 147], the required mass-to-light ratio approximately coincides with the best mass-to-light ratio predicted by the SPS model based on a scaled Salpeter IMF [148]

$$T_* = \chi T^\text{Salpeter}_*,$$

where $T^\text{Salpeter}_*$ is the mass-to-light ratio provided by the Salpeter IMF and $\chi \sim 3/4$.

![Figure 3: Mass Discrepancy Acceleration Relation (MDAR), where the galaxy mass discrepancy, the squared ratio between the observed velocity $V$ and the velocity due to baryons $V_b$, is plotted against the Newtonian centripetal acceleration $g_N = V^2/R$ derived from the observed baryonic surface mass density. The black dots show several hundreds of individual resolved measured points from the rotation curves of almost one hundred spiral galaxies. The figure is reproduced from [149].](image)

A slightly different perspective is provided by the RAR, that correlates the observed centripetal acceleration $g_{\text{obs}}(R) = V^2(R)/R$, where $V$ is the measured rotation speed, and the Newtonian acceleration $g_{\text{bar}}(R)$, due to baryonic matter alone [127]. Figure 4 shows the RAR for the SPARC sample. The two quantities and their respective uncertainties are completely independent of each other [129], and follow the relation

$$g_{\text{obs}}(R) = \frac{g_{\text{bar}}(R)}{1 - \exp \left( - \sqrt{\frac{g_{\text{bar}}(R)}{g^\dagger}} \right)}.$$  

(4)

The only free parameter is $g^\dagger$. The fit with 153 galaxies from the SPARC sample yields $g^\dagger = (1.20 \pm 0.02 \pm 0.24) \times 10^{-10}$ m s$^{-2}$ [127, 130], which is within 1σ from the acceleration scale $a_0$.

The observed scatter is 0.13 dex, and it is comparable to the scatter 0.12 dex derived from the different mass-to-light ratios of the individual galaxies, and from the errors on the velocity measurements and the uncertainties on the inclination of the disk and the galaxy distance. When marginalizing over these uncertainties and over the whole sample of galaxies, the observed intrinsic scatter can be reduced to 0.057 dex. Therefore, the scatter appears much smaller than the scatter 0.09 dex predicted by the EAGLE simulations; the measured scatter should actually be larger than 0.09 dex, because the measurement errors are not included in the EAGLE prediction [129]. The issue remains open, however, because when the observed sample includes dwarf galaxies with slowly rising rotation curves, the scatter also appears to substantially increase at small $g_{\text{bar}}$ in the observed relation [150].

An additional unsettled controversy is the correlation between the residuals of the observed rotation curves from relation (4) and the properties of the galaxies observed in galaxy samples [151, 152] different from SPARC, where this correlation appears to be absent [130].

2. Possible solutions within the CDM model

The disk-halo conspiracy can be solved within the CDM model by adopting the maximum-disk hypothesis: the disk maximally contributes to the measured rotation curve [101, 153–155]. This hypothesis yields mass-to-light ratios for the disk stellar
Figure 4: Radial Acceleration Relation (RAR), the observed centripetal acceleration versus the Newtonian acceleration due to baryonic matter alone. The blue color-scale rectangles show 2694 individual measured points from the rotation curves of 153 SPARC galaxies. The red solid line is the fitted RAR given by equation (4), whereas the black dotted line represents the one-to-one relation, for comparison. The figure is reproduced from [129].

population in agreement with SPS models, and it is supported by several observational probes.

For example, the near infrared mass-to-light ratio constrained from the Milky Way terminal velocity curve describes this curve without the need of a dark matter halo up to $\gtrsim 5$ kpc [156]. If some dark matter is included within this distance from the Galaxy centre, the microlensing optical depth of the source stars within the Baade Window (Galactic longitude = 1°, Galactic latitude = -3.9°) would probably decrease, becoming more discrepant from the observed value of this quantity derived from the MACHO measurements in [157, 158].

Moreover, the mass-to-light ratios determined in the Milky Way from star counts, the radial force within the disk, and the vertical force from the disk are all consistent with each other and with the BTFR; they also properly reproduce the terminal velocity curve of the Galaxy by requiring a close to maximum disk [159]. At last, only maximum disks can properly reproduce the observed ratio $R \sim 1.2 \pm 0.2$ between the corotation radius of barred galaxies and the semi-major axis of the bar; this value is the evidence that the corotation radius of barred galaxies is just beyond the end of the bar [e.g., 160]. However, the maximum-disk hypothesis only works for HSB galaxies; dwarf spheroidal and LSB galaxies appear to be dark matter-dominated systems even in their innermost regions [161, 162], although there are indications of a correlation between the distributions of the luminous and dark matter also in these systems [163].

ΛCDM hydrodynamical simulations and semi-analytical models can reproduce the normalization and slope of the BTFR but not its small scatter [135, 143, 144, 164–166]: the models yield a scatter of $\sim 0.15$ dex, compared to the observed $\sim 0.10$ dex. Accounting for the BTFR, including its relatively small intrinsic scatter and the lack of correlations between their residuals and galaxy properties, requires an accurate balance between the star formation efficiency and the stellar feedback processes to regulate the relation between the properties of the baryonic matter and the properties of the dark matter halo [135, 136].
The shape and the scatter of the MDAR are properly reproduced by a semi-empirical model [144] based on the ΛCDM paradigm and on the existence of different galaxy scaling relations between (1) the concentration of the dark matter halo and its mass [167], (2) the neutral hydrogen mass and the stellar mass [168], (3) the disk mass and its size [169], and (4) the half-mass radius of the bulge and its mass [170]. The same model can also reproduce a BTFR with the correct normalization and slope but with an intrinsic scatter of 0.17 dex, larger than the observed one of ~0.10 dex.

The key assumption for this success is the dependence of the inner and outer slopes of the dark matter density profile on the stellar-to-halo mass ratio [171, 172], which accounts for the feedback of the baryonic processes. However, for this mechanism to be effective, the stellar-to-halo mass ratio must be in the small range [0.01% – 0.03%] that might be barely appropriate only for ultra-faint or LSB galaxies but it is inappropriate for HSB galaxies. In fact, when the stellar-to-halo mass ratio is within this range, stellar feedback leads to the formation of cores and the BTFR slope is reproduced. However, for smaller stellar-to-halo mass ratios stellar feedback cannot convert the cuspy NFW profile into a core and for greater stellar-to-halo mass ratios, galaxies are even more concentrated, up to a factor of 2.5, than predicted by CDM-only simulations.

In ΛCDM, it remains difficult to explain the emergence of the acceleration scale, \( a_0 = 1.2 \times 10^{-10} \text{ m s}^{-2} \), that appears in all the scaling relations mentioned above. This coincidence is assumed \textit{ab initio} in MOND, a modification of the law of gravity [33]. Alternatively, it could suggest the necessity of new physics affecting the dark sector [128].

\section*{B. The cusp/core problem}

Collisionless N-body simulations of the CDM model show that the mass density profile \( \rho(r) \) of dark matter halos is described by a steep power-law in the innermost regions: in other words, the CDM model predicts cuspy density profiles. At small radii, the NFW density profile suggested by N-body simulations in the 1990s [117, 173] is approximately \( \rho \propto r^{-\alpha} \), with \( \alpha = -1 \) [46]. Later investigations suggested that the inner slope may depend on the total mass of the halo [174] and be steeper, with \( \alpha = -1.5 \) [40]. This prediction appears to disagree with observations, because dwarf galaxies require a shallower dark matter density profile in their central regions [175, 176]. Dwarf galaxies are known to be among the darkest galaxies observed. They show a central velocity dispersion of \( \sim 10 \text{ km s}^{-1} \), compared to the expected \( \sim 1 \text{ km s}^{-1} \) for self-gravitating systems of the same luminosity and scale radius \( \sim 100 \text{ pc} \) at equilibrium [177]. Although their luminosity may vary by several orders of magnitude among different systems, from \( \sim 10^2 \text{ L}_\odot \) to \( \sim 10^{10} \text{ L}_\odot \) [178, 179], they show similar velocity dispersions, suggesting that they are dominated by a similar dark matter distribution [180].

In 1988, the best fit model of the measured HI rotation curve of the dwarf irregular galaxy DDO 154 [175] showed that its dark matter density profile is well described by a pseudo-isothermal model [e.g., 116] with a core radius of 3 kpc: the profiles of dwarf galaxies having a core rather than a cusp is a general result [175, 176]. Furthermore, the recently discovered ultra-faint galaxies orbiting the Milky Way and M31 [e.g., 181–189] have a half-light radius of a few kpc, which favors a core over a cuspy density profile; the core is hardly explicable in the context of the CDM model without resorting to baryonic feedback, and/or tidal effects [189].

This discrepancy between observations and simulations is known as the “cusp/core problem” (CCP) [43] and is still an unsolved topic of fervent debate.

\subsection*{1. Observational evidence}

The NFW density profile is too steep to fit the observations in the innermost regions of the dwarf and LSB galaxies [140, 190]: the observations favor a density profile with a core [191–196]. Figure 5 shows an example of how well a density profile with a core (black solid line) fits the measured rotation curve of the LSB galaxy UGC 5750. Neither the NFW profile (long-dashed line) nor the singular isothermal sphere (SIS, dotted line) are able to fit the measured rotation curve.

High spatial resolution rotation curves of dwarf and LSB galaxies demonstrate that these galaxies are dynamically dominated by dark matter. However, no evidence for a steep inner slope of the density profile was found on scales below \( \sim 0.15 \text{ kpc} \). Galaxies exhibit inner slopes \( \alpha = -0.2 \pm 0.2 \) [201]. Later measurements of the rotation curves of a sample of 165 low-mass galaxies yield a median inner slope \( \alpha = -0.22 \pm 0.08 \) [202]. Similarly, high resolution observations of the dwarf irregular galaxy NGC 6822 [203], and HI observations of the dwarf galaxy NGC 3741 [204] favor a core rather than the cuspy NFW profile. Similar conclusions have been recently obtained with the high-resolution rotation curves of 26 dwarfs from LITTLE THINGS (Local Irregulars That Trace Luminosity Extremes, The HI Nearby Galaxy Survey) [205]. The mean slope \( \alpha = -0.32 \pm 0.24 \) of this sample is consistent with the previous result found for the LSB galaxies [194, 202].

Dwarf spheroidals (dSphs) also show the same cusp/core controversy. Wide-field multi-object spectrographic observations of dSphs [51] provide high-quality kinematic data. A feasible analysis of those data relies on using the Jeans equations [e.g., 51, 206, 207]. By assuming that the dwarf galaxies contain one or more pressure-supported stellar populations in dynamical equilibrium tracing the underlying dark matter gravitational potential, the dynamical mass distribution \( M_{\text{dyn}}(r) \), which accounts
Figure 5: Rotation curve (left panel) of UGC 5750 (the LSB galaxy, showed in the right panel). The data of the rotation curve are obtained with the integrated field H$\alpha$ spectroscopy (squares) [197], long slit optical observations of the Balmer transition (circles) [193, 198], and radio observations of the 21 cm atomic hydrogen spin flip transition (stars) [199]. The isothermal sphere with a core (CIS) profile (solid line) fits the data. $r_c$(zero) and $r_c$(max) are the values of the core radius obtained from the fit in the no disk and maximum disk case, respectively; the solid line is the case of no disk. Neither the NFW profile (dashed line), whose parameters are fixed and given by the $\Lambda$CDM cosmology, nor the singular isothermal sphere (SIS) profile can describe the dark matter halo of this LSB galaxy. The figure is reproduced from [200].

for both the stellar and the dark matter distribution, is related to the stellar distribution through the Jeans equation (see Eq. 4-55 & 4-56 in [208]). For a spherical symmetric system, the Jeans equation reads

$$\frac{1}{\nu} \frac{d}{dr} (\nu \bar{v}_r^2) + \frac{2 \beta \bar{v}_r^2}{r} = -\frac{G M_{\text{dyn}}}{r^2},$$

(5)

where $\nu(r)$ is the three-dimensional stellar number density, $\bar{v}_r^2(r)$ is the radial velocity dispersion, and $\beta \equiv 1 - \bar{v}_\theta^2/\bar{v}_r^2$ is the orbital velocity anisotropy parameter of the stellar component. For a constant $\beta$, the solution of the previous equation is [209]

$$\nu \bar{v}_r^2 = G r^{-2\beta} \int_r^\infty s^{2\beta-2} \nu(s) M_{\text{dyn}}(s) ds,$$

(6)

which must be projected along the line of sight to be compared with observations [208].

Under the assumption that the stellar mass is negligible compared to the halo mass, $M_{\text{dyn}}$ can be identified with the dark matter mass, and Eq. (5) is a tool to probe the dark matter model. As an example, in [51, 207], the projected velocity dispersion profiles of the Milky Way dwarf satellites were fitted using a Markov Chain Monte Carlo method. Although the analysis did not place tight constraints on the dark matter parameters, it showed that, as illustrated in Figure 6, both cuspy dark matter density distributions and distributions with a core were able to reproduce the kinematic data of these dSphs, without any statistical evidence favouring one model over the other.

However, Jeans modelling is affected by a degeneracy between the velocity anisotropy $\beta$ and the total halo mass [208], which prevents unambiguous constraints on the dark matter parameters. One way around would be the addition of information from multiple stellar populations [210–212], or higher velocity moments [213]. Using the multiple stellar populations would help to lift the degeneracy but would not improve the constraints on the dark matter parameters. Using higher velocity moments would be more appropriate when proper motion measurements are available [214, 215]. For example, the Sculptor and Fornax galaxies are known to have two distinct stellar components. Using the kinematics data of both populations, the degeneracy between the mass and the anisotropy parameter is removed, and the dispersion velocity profiles is fitted by a dark matter halo with a core [210, 216]. At odds with these results, LSB [217] and late-type dwarf galaxies [59] appear to have dark matter density distributions consistent with the cuspy NFW profile.

Caution should be taken in the analysis of the data. In fact, if the asphericity of the distribution of the stellar populations is not taken into account, the methodology applied in [210] for dSphs may mistakenly favour a density profile with a core even if a cuspy profile is present [48]. This overlook can also introduce a bias in the estimate of the dwarf mass of the order of 10% for a galaxy with a viewing angle, defined as the angle between the line of sight and the major axis of the distribution of the metal-poor sub-population, of 90° [48]. Similarly, a more careful analysis, based on modelling the stellar orbits and leading to
Figure 6: Projected velocity dispersion profile of the eight brightest dwarf satellites of the Milky Way. The red lines represent the best fit of the dark matter density distribution with a core, while the blue ones derive from a cuspy profile. The long-dashed and dotted lines show the isothermal and power law models, respectively. The figure is reproduced from [51, 207].

A non-parametric estimation of the dark matter density distribution, yields a cuspy profile with $\alpha = -1.0 \pm 0.2$ for the Draco dSph [218].

Despite the increasing number of studies, and the increasing accuracy of data, whether dSphs exhibit a cuspy dark matter density profile or a profile with a core is still unclear. A recent analysis of Draco found an inner dark matter density of $\rho_{DM}(150\text{pc}) = 2.4^{+0.5}_{-0.6} \times 10^8 M_\odot \text{ kpc}^{-3}$, which is consistent with a CDM cuspy density profile [219]. Nevertheless, the discovery of ultra-diffuse galaxies such as Crater II [187], and Antlia II [189], and the discovery of galaxies that appear to be devoid of dark matter, such as NGC 1052-DF2 and NGC 1052-DF4 [220–225], can challenge the standard CDM paradigm. Specifically, galaxies such as DF2 and DF4 appear to be at $2.6\sigma$ and $4.1\sigma$ tension with the standard model, respectively: according to $\Lambda$CDM simulations, the probability of finding DF2-like galaxies at a distance 11.5 Mpc from the observer is at most $10^{-4}$; this probability drops to $4.8 \times 10^{-7}$ at a distance 20.0 Mpc [188]. However, more accurate dynamical models can substantially alleviate this tension [226–228]. In addition, a recent analysis [229] suggests that the distance to DF2 is 13 Mpc rather 20 Mpc, as previously estimated; the closer distance increases to 75% the dark matter content of the galaxy mass. Similarly, properly
taking into account the uncertainty on the velocity measurements of the small sample of globular clusters of DF2 suggests that its mass-to-light ratio is at the low end of the distribution of the mass-to-light ratios of dwarf galaxies [230], but its velocity dispersion and mass are still consistent with the universal mass profile of the Local Group dwarf galaxies [51].

2. Possible solutions within the CDM model

Possible solutions to the CCP, in the context of the CDM scenario, can reside either in neglected physical processes, mostly affecting the baryonic matter, or in systematic effects and/or observational limits.

There are processes that, in principle, can convert a cuspy density profile into a profile with a core. For example, an inner Lindblad-like resonance, which couples the rotating bar to the orbits of the star through the cusp, was suggested to cause an angular momentum transfer from the bar-pattern to the dark matter halo [231]; on turn, this transfer flattens the density profile in the central region of the halo. However, later investigations show opposite results: the generation of a bar in the disk would actually make steeper, albeit slightly, the inner dark matter density profile [232].

The most popular solutions rely on supernova feedback and dynamical friction. Winds driven by supernovae can be an effective mechanism to transform the cusp of the dark matter density profiles into a core [45, 233]. A similar effect can be generated by stellar winds [234, 235]. Supernova and stellar winds produce energy feedback that can drastically modify the shape of the dwarf galaxies by forcing the gas and the dark matter particles to move outwards, change the gravitational potential well and flatten the density profile [235].

In less massive galaxies, starbursts heat the gas that expands and inhibits the star formation. Radiative cooling makes the gas collapse again and a starburst is reignited [236]. This cyclic starburst periodically develops density waves whose resonance with the dark matter particles generates a cusp-core transition in the halo dark matter density [236].

Decreasing the steepness of the density profile might not always alleviate the tension between the data and the CDM model. The profiles with a core derived from the high-resolution HI observations of LSBs can not explain the gravitational lensing signal [237, 238]: to match this signal, the dark matter density profile should be steeper [200].

Recent high-resolution simulations of isolated dwarf galaxies show that cores of size comparable to the stellar half-mass radius may form if the star formation lasts for $\sim$4 Gyr for a dwarf mass $M_{200} = 10^8 M_\odot$ and $\sim$14 Gyr for a mass $M_{200} = 10^9 M_\odot$ [239]. However, the Auriga and EAGLE simulations suggest that gas outflows are unable to modify the dark matter profile of dwarf galaxies, and that repeated outbursts cannot explain the transition from a cusp to a core [240].

An additional process advocated for solving the CCP is dynamical friction between gas clumps with individual mass $10^5 – 10^6 M_\odot$ [241]. This friction between the clumps would transfer angular momentum from the gas to the dark matter particles that, on turn, would move away from the central region of the halo and flatten its density profile. This effect should be efficient in the early phase of the galaxy formation when the halo size is smaller [241, 242]. The main difficulty to make this solution efficient is to have gas clumps sufficiently massive [243]. The dynamical friction is found to be three times weaker [243] than previously assumed in [241], and, to increase the efficiency, the mass of the clumps should reach $\sim 7\%$ of the total halo mass. Gas naturally forms clumps through the Jeans instability and the clumps can in principle be as massive as $\sim 10^6 M_\odot$ [244]. However, in dwarf galaxies, star formation and cloud fragmentation may be inefficient [245], and gas clumps may form via thermal instability due to galaxy mergers. This process leads, for a total mass of the halo of the order of $10^7 M_\odot$, to clump masses of the order $\sim 10^4 M_\odot$ [246], which are two orders of magnitude less massive than required to solve the CCP [243].

The globular clusters of a dwarf can also play the role of the gas clumps. This process was applied to the five globular clusters of the Fornax dSph. High-resolution simulations [247] have been used to demonstrate that if globular clusters are embedded in a dark matter mini-halo, which make them more massive and fall more rapidly towards the galaxy center, and were accreted within the last 3 Gyr, early accretion favours the disruption of globular clusters due to repeated interactions with the galaxy; thus, their crossing in the central region of Fornax can generate the transition from a cusp to a core, due to momentum exchange between the globular clusters and the galaxy dark matter halos, with the frequency of the crossing which sets the size of the core.

The CCP may not originate from neglecting some physical processes, but rather from systematic effects [248] due to observational limits, including non-circular motions and pointing effects hiding the signature of a cusp in the innermost region of the galaxies.

Systematic bias towards shallower slopes of the density profile can originate from non-circular motions of gas and stars of at least $\sim 20 \text{ km s}^{-1}$, offsets of $\sim 3 − 4$ arcsec from the correct center because of inaccurate telescope pointings, or a separation between the dynamical and photometric centers of at least $\sim 0.5 − 1 \text{ kpc}$ [194].

However, the existence of this pointing effects was questioned by two independent observation campaigns of the same sample of galaxies made by independent groups with two different telescopes. One set of observations was made with the 193-cm telescope at the Observatorio de Haute Provence [193], while the other set of observations was made at the 3.6-m Telescopio Nazionale Galileo at Canary Islands [249]. Pointing effects in these two independent data sets are of the order of 0.3 arcsec, an order of magnitude below the offset required to bias a cuspy density profile towards a core [248].

An additional source of error might come from the too simplistic modelling. The gas is usually supposed to move on circular orbits. If this assumption is invalid and the gas motion is disturbed, the steepness of the slope of the density profile can be
underestimated [194, 250]. However, using a sample of galaxies from THINGS, non-circular motions have been estimated to be of the order of few kilometres per seconds for dwarf galaxies [251]. These velocity perturbations are too small to overlook the presence of a cusp [248, 252]. Nevertheless, high-resolution simulations suggest that, in cuspy and triaxial CDM halos, non-circular motions can substantially affect the observed rotation curves [150] and that correcting for these motions in observations is far from trivial. Therefore, the possible relevant role of non-circular motions in shaping the observed rotation curves cannot yet be ruled out.

C. The missing satellites problem

The stellar mass function derived for field galaxies and satellite galaxies in the Local Group is significantly less steep at low masses than the mass function expected for dark matter halos in the CDM model: $dn/dM \propto M^\alpha$, with $\alpha \simeq -1.5$ for faint-end galaxies and $\alpha \simeq -1.9$ for dark matter halos. Indeed, high-resolution CDM cosmological simulations of Milky Way-like dark matter halos predict thousands of subhalos with mass sufficiently large ($M \gtrsim 10^7 M_\odot$) to host galaxies. However, this number is $\sim 20$ times larger than the number of satellites observed around the Milky Way and M31 [e.g., 42, 52, 53, 253]. Recently, $\sim 50$ ultra-faint galaxies with stellar mass as small as 300 $M_\odot$ have been detected within the virial volume of the Milky Way [254], and many more may be discovered using the Gaia satellite [e.g., 255, 256]. However, it is very unlikely that their number will reach the thousands satellites predicted by the CDM simulations [42]. This discrepancy between the predicted and observed numbers of satellites is known as the missing satellites problem (MSP).

Figure 7 gives a visual representation of the MSP: on the left panel, thousands of subhalos predicted by the CDM simulations are shown, whereas on the right panel we can see the classical Milky Way satellites [115].

The MSP implies that either the CDM model produces too many subhalos with low mass or the formation of galaxies in these halos becomes less and less efficient as the halo mass declines. While the former view calls for a modification of the CDM paradigm, the latter suggests the need for an appropriate treatment of baryonic physics.

Figure 7: Left panel: Projected dark matter distribution (600 kpc on a side) of a $10^{12} M_\odot$ dark matter halo from the ELVIS ΛCDM simulations [56]. The number of small subhalos strongly exceeds the number of known Milky Way satellites (missing satellites problem; Section III C). The circles highlight the nine most massive subhalos. Right panel: Spatial distribution of the closest nine of the eleven most luminous (classical) satellites of the Milky Way (the diameter of the outer sphere is 300 kpc). For these satellites, the central mass inferred from stellar kinematics is a factor of $\sim 5$ lower than the mass predicted for the central regions of the subhalos highlighted in the left panel, preventing the association of the classical satellites to the most massive subhalos of the dark matter halo of Milky Way-like galaxies (too-big-to-fail problem; Section III D). The figure is reproduced from [115].
The MSP can be tackled within the CDM model by using the abundance matching (AM) technique. This technique matches the cumulative distribution of an observed property of galaxies with the predicted cumulative distribution of the mass of their dark matter halos [42]: for example, by adopting the mean star formation rate as the observed property, the MSP in the Milky Way appears to be solved for satellite masses larger than $10^9 \, M_\odot$ [257].

For smaller masses, the solution rests upon the suppression of star formation by UV reionization, when the mass of the dark matter halo is smaller than $M_{\text{vir}} \approx 10^9 \, M_\odot$ [e.g., 42, 258, 259], or by atomic cooling in the early Universe for mass smaller than $M_{\text{vir}} \approx 10^8 \, M_\odot$ [e.g., 42, 260]. Both these effects become dominant in ultra-faint galaxies with stellar mass $M_* \lesssim 10^5 \, M_\odot$ [42]. AM can be used to quantify this expectation. Figure 8 shows the correlation between the stellar mass, $M_*$, and the total mass of the galaxy, $M_{\text{halo}}$. In observations, the stellar mass $M_*$ can be obtained by fitting broad-band photometric data with a Spectral Energy Distribution model [e.g., 261–263], whereas the total galaxy mass $M_{\text{halo}}$ can be inferred from either gravitational lensing or HI rotation curves [e.g., 264–266]. Figure 8 also shows various models that adopt different galaxy data sets and the dark matter halo mass function from different $N$-body simulations to derive the expected $M_* - M_{\text{halo}}$ relation. In addition, they adopt different assumptions on the star formation rate and the star formation history of the galaxies to infer the dark matter halo mass function from different $N$-body simulations [e.g., 261–263]. Both these effects become dominant in ultra-faint galaxies with stellar mass $M_* \lesssim 10^5 \, M_\odot$ [42].

AM can be used to quantify this expectation. Figure 8 shows the correlation between the stellar mass, $M_*$, and the total mass of the galaxy, $M_{\text{halo}}$. In observations, the stellar mass $M_*$ can be obtained by fitting broad-band photometric data with a Spectral Energy Distribution model [e.g., 261–263], whereas the total galaxy mass $M_{\text{halo}}$ can be inferred from either gravitational lensing or HI rotation curves [e.g., 264–266]. Figure 8 also shows various models that adopt different galaxy data sets and the dark matter halo mass function from different $N$-body simulations to derive the expected $M_* - M_{\text{halo}}$ relation. In addition, they adopt different assumptions on the star formation rate and the star formation history of the galaxies to infer $M_*$ from the galaxy luminosities. Therefore, the models can be different, especially at low stellar and halo masses. In Figure 8, the solid lines show the predictions derived for stellar masses $M_* \gtrsim 10^8 \, M_\odot$, where observational data are available, whereas the dashed lines show the extrapolation of the models at smaller masses, where the observational information is missing. The orange shaded area shows the 1σ log-normal scatter around one of the models shown in [267]. For $M_* \gtrsim 10^8 \, M_\odot$, all the models are consistent with each other. However, at smaller stellar masses, the incompleteness of the observational surveys and the increased stochasticity of the star formation mechanism in the models make the estimation of the stellar mass uncertain and, consequently, the AM relation less constrained: for example, a stellar mass $M_* \sim 10^6 \, M_\odot$ would correspond to $M_{\text{halo}} \sim 10^9 \, M_\odot$ for the Behroozi model, but to $M_{\text{halo}} \sim 5 \times 10^9 \, M_\odot$ for the Brook model. These differences translate into an additional uncertainty on the number of low-mass satellites in the Local Group. This uncertainty is further increased by the fact that this halo mass, $M_{\text{halo}} \approx 10^9 \, M_\odot$, is close to the mass scale, $\approx 10^5 \, M_\odot$, below which star formation is expected to be suppressed.

Dooley et al. [272] applied the above-mentioned AM models, both with and without the inclusion of reionization effects, to a set of Milky Way-like dark matter halos from the dark-matter only Caterpillar simulations. For each AM model, Dooley et al. estimate, for any dark matter halo with a given mass, the expected number of satellite galaxies with a given stellar mass. The results of this exercise, for a dark matter halo with mass $M_{\text{halo}} = 1.4 \times 10^{12} \, M_\odot$, is shown in Figure 9: the Brook AM model introduced in [271] (red line in Figure 8) agrees with the complementary cumulative$^2$ stellar mass function, $N_{\text{sats}}(> M_*)$, of the 40 detected Milky Way’s dwarf satellites with $M_* > 10^3 \, M_\odot$ (black dashed line), irrespective of the inclusion of reionization (solid and dashed red lines). On the other hand, the other AM models shown in Figure 8 overestimate the observed complementary cumulative stellar mass function of the Milky Way satellites, even when reionization suppresses the formation of low-mass galaxies. A better agreement between these models and the complementary cumulative stellar mass function of the Milky Way satellites is obtained by assuming a lower halo mass for the Milky Way. In this case, the Brook model would clearly underestimate the number of satellites. The numbers of Milky Way satellites with $M_* > 10^3 \, M_\odot$ predicted by the other models in Figure 8 are consistent with other analyses [254, 273] and would solve the MSP if more than 80 satellites with $M_* > 10^3 \, M_\odot$ were detected. However, the other models agree with each other and with the Milky Way complementary cumulative stellar mass function for $M_* \gtrsim 4.5 \times 10^5 \, M_\odot$. This agreement suggests that there might be no MSP for halos with stellar mass $M_* > 4.5 \times 10^5 \, M_\odot$ or, equivalently, for halos with total mass $M_{\text{halo}} \gtrsim 10^{10} \, M_\odot$ [272].

Although the stellar mass $M_*$ has been successfully used for the AM method, it suffers from several biases. For example, the suppression of star formation due to the infall of the galaxy toward a larger galaxy [e.g., 274, 275] or tidal stripping [e.g., 54, 276, 277] introduces an intrinsic scatter in the $M_* - M_{\text{halo}}$ relation [e.g., 278–280] which must be taken into account to correctly predict the number of satellites [e.g., 281, 282]. One possible solution relies on replacing the stellar mass with the mean star formation rate, $\langle SFR \rangle$, averaged over the star-forming time of the galaxy, as anticipated at the beginning of this section [257]: for isolated galaxies, the stellar mass rises monotonically with the halo mass [265, 266, 280], and $\langle SFR \rangle$ is reasonably expected to behave similarly. By adopting $\langle SFR \rangle$ rather than $M_*$, the scatter in the AM relation at a given $M_{\text{halo}}$ decreases, and the accuracy of the AM mass estimator increases [257]. The $\langle SFR \rangle - M_{200}^3$ relation derived in [257] is based on surveys of faint galaxies [283], whose stellar mass function is

$^2$ In the relevant literature, adopting the expression cumulative distribution function for $N(> X_*)$, namely the number of objects with a physical properties $X$ larger than a threshold $X_*$, dates back to the pioneering paper by Moore et al. [253]. However, in statistics, the cumulative distribution function actually is $N(< X_*)$. $N(> X_*)$ is the complementary cumulative distribution function. Here, we prefer to adopt this more rigorous terminology.

$^3$ The radius $r_{200}(z)$ at redshift $z$ is the radius of a spherical volume within which the mean mass density is 200 times the critical density of the Universe $\rho_c(z)$. The mass enclosed within $r_{200}$ is thus $M_{200}(z) = 4\pi r_{200}^3 \rho_c(z) r_{200}/3$. $M_{200}$ is approximately equal to the virial mass $M_{\text{vir}}$, defined in the footnote.
Abundance matching (AM) relations from the models by Behroozi [268], Garrison-Kimmel GK14 [269], Garrison-Kimmel GK16 [267], Moster [270], and Brook [271]. The solid lines show the models in a region of the parameter space where observational data are available; the dashed lines show the extrapolation of the models where the observational information is missing. The orange shaded area shows the 1σ log-normal scatter around the model GK16. The figure is from [272].

The ⟨SFR⟩−M₂₀₀ relation is derived by combining estimates of the star formation rate from the data and the halo mass function from the model. The mass M₂₀₀ of the subhalos are identified with the mass of the dark matter halos just before falling into the main host halo.

Figure 10 shows the complementary cumulative subhalo mass function of a complete volume-limited sample of the bright Milky Way satellites within 280 kpc from the Galaxy center. In the left panel, the subhalo mass function is derived from the Mₚ − M₂₀₀ AM relation, whereas the right panel shows the subhalo mass function based on the ⟨SFR⟩ − M₂₀₀ AM relation [257]. The complementary cumulative subhalo mass function based on the Mₚ − M₂₀₀ relation (left panel) shows the existence of the MSP for masses M₂₀₀ ≲ 2 × 10⁹ M⊙: there are not enough quenched satellites, such as Sculptor and Leo I, in the Milky Way to be consistent with the predictions of the ΛCDM model. On the contrary, the complementary cumulative subhalo mass function based on the ⟨SFR⟩ − M₂₀₀ AM relation (right panel) suggests that no MSP exists for masses above M₂₀₀ ∼ 10⁹ M⊙.

The subhalos with pre-infall mass M₂₀₀ ∼ 5 × 10⁸ − 5 × 10⁹ M⊙ correspond to ultra-faint dwarf galaxies [e.g., 257, 286–288]: the number of Milky Way satellites with mass M₂₀₀ ≲ 5 × 10⁸ M⊙ is smaller than the number predicted by ΛCDM, and it is still unclear whether a MSP really exists below this mass scale. These satellites might be fainter than the completeness limit of the current galaxy surveys [e.g., 289–291], or they might be completely dark [e.g., 292].

1, in a Universe with Ω_m = 1 [119].
Figure 9: Mean number of satellite galaxies with mass larger than a given stellar mass threshold, $N_{\text{sats}}(>M_*)$, as a function of the threshold mass, $M_*$, around a Milky Way-sized galaxy hosted by a dark matter halo with virial mass $M_{\text{halo}} = 1.4 \times 10^{12} M_\odot$, as predicted by various AM models applied to the dark-matter only Caterpillar simulations. The colored dashed lines indicate the predictions of the AM models shown in Fig. 8; the solid lines show the predictions of the same models after the inclusion of reionization, that suppresses the formation of low-mass galaxies. The grey box represents the prediction of 37-114 satellites with stellar luminosity $L_*>10^3 L_\odot$ by [273], derived by combining a number of toy models applied to dark matter-only simulations with the sample of dwarfs corrected for completeness, observed by the Sloan Digital Sky Survey. For comparison, the complementary cumulative stellar mass function (derived from $V$-band luminosities, assuming a mass-to-light ratio equal to $1 M_\odot/L_\odot$) of the 40 known satellite galaxies of the Milky Way with $M_* > 10^3 M_\odot$ is shown with a black dashed line. The figure is reproduced from [272]. Here, we corrected the misprint that is present in the label of the ordinate axis of the original figure.

D. The too-big-to-fail problem

Dissipationless ΛCDM simulations of Milky Way-like dark matter halos predict that the most massive subhalos of the Milky Way are too dense to host any of the brightest ($L_V > 10^5 L_\odot$) Milky Way dSph satellites [41, 54, 55]. The density discrepancy can be erased by assuming that the most massive, dense subhalos are dark and the brightest dwarfs reside in subhalos that are a factor of $\sim 5$ less massive. Clearly, this scenario generates a new, serious conflict, because the largest dark matter subhalos, characterized by the deepest potential wells, are expected to be able to retain gas and form stars: in other words, they are “too big to fail” to form stars and should thus host observed dwarfs. Solving this latter conflict requires going back to the simplest scenario, where the most massive subhalos do host the brightest dSphs, and where the density conflict appears. This issue was dubbed as the too-big-to-fail (TBTF) problem by Boylan-Kolchin et al. [41].

An illustration of the TBTF issue (together with the MSP; see Section III C) is presented in Figure 7 for the Milky Way: in the left panel, the nine most massive subhalos of a dark matter distribution in a simulated Milky Way-like dark matter halo are highlighted with circles; the right panel shows the actual distribution of the nine, most closeby “classical” satellite galaxies of the Milky Way. The overdensity of the most massive subhalos with respect to the “classical satellites” prevents their reciprocal association, generating the TBTF conflict.

It is important to note that the TBTF problem is an issue distinct from the MSP discussed in Section III C. It can be seen as a problem of “missing dense satellites” related to the internal mass distribution of subhalos, rather than to their abundance [56]. Therefore, the TBTF problem is largely independent of the exact relationship between halo mass and stellar mass (see, e.g., the AM technique in Section III C). The relevance of the mismatch between estimated and predicted central densities depends upon the specific realization of the dark matter halo substructure. However, the comparison of the observational data with a series of simulations shows that, for the Milky Way, the discrepancy appears too large to be a statistical fluke [41, 55].

The density discrepancy manifests as a kinematic discrepancy between estimated and predicted circular velocities. Specif-
Figure 10: Complementary cumulative subhalo mass function of the Milky Way satellites in a sphere of radius 280 kpc centred on the Galaxy. The mass function is derived from either the $M_\star - M_{200}$ AM relation (left panel) or the (SFR) – $M_{200}$ AM relation (right panel). The peak masses before the infall are taken as the subhalos masses $M_{200}$. The names of the individual galaxies used to build the complementary cumulative subhalo mass functions are indicated in the plots. The median relations are shown as blue solid lines, whereas the $\pm 68\%$ confidence intervals are delimited by the blue dashed lines. The green lines have the same meaning as the blue lines but they also include the sample of ultra-faint dwarf galaxies of [178]. The grey shaded areas represent the variation of the complementary cumulative subhalo mass functions resulting from ten dark-matter-only “zoom-in” simulations of the Milky Way in the $\Lambda$CDM model. The red shaded areas have the same meaning as the grey shaded areas but they include a model for the stellar disk of the Milky Way. The figure is from [257]. Here, we corrected the misprint that is present in the label of the ordinate axis of the original figure.

ically, for dispersion-supported systems as the Milky Way dSph galaxies, the circular velocity at any given radius, $V_{\text{circ}}(r)$, is indicative of the dynamical mass enclosed within that radius and thus of the matter density within that radius: $V_{\text{circ}}^2(r) = G M(< r)/r$. As shown by Wolf et al. [293], an estimate of the luminosity-weighted square of the line-of-sight velocity dispersion enables to accurately determine the dynamical mass of a spherical, dispersion-supported system within a characteristic radius, approximately equal to the deprojected half-light radius, $r_{1/2}$, through the relation $M_{1/2} \equiv M(< r_{1/2}) = 3 G^{-1}(\sigma_{\text{los}}^2) r_{1/2}$: the relevant feature of this estimate is that it is nearly independent of the spatial variation of the stellar velocity anisotropy, $\beta(r)$, as long as the radial profile of the velocity dispersion is fairly flat near $r_{1/2}$, as typically observed. From the relation between circular velocity and dynamical mass given above, it follows that the circular velocity at the half-light radius can be written, in terms of the observable line-of-sight velocity dispersion, as $V_{\text{circ}}(r_{1/2}) = \sqrt{3(\sigma_{\text{los}}^2)}$.

Because dSphs are dark-matter dominated at all radii, the dynamical mass $M_{\text{dwarf}}(< r_{1/2})$ is approximately equal to the dark matter mass enclosed within the half-light radius. A necessary, although not sufficient, condition for a dark matter subhalo to host such a dispersion-supported galaxy is that, at the observed half-light radius of the galaxy, $r = r_{1/2}$, the mass $M_{\text{dwarf}}(< r_{1/2})$ agrees with the mass of the subhalo at the same radius, $M_{\text{sub}}(< r_{1/2})$ [41]. This agreement, in turn, requires that the observed and simulated values of $V_{\text{circ}}$ must agree with each other at $r = r_{1/2}$.

For the brightest Milky Way dSphs, this requirement is fulfilled only by subhalos that are not among the most massive subhalos produced in $\Lambda$CDM simulations. On the contrary, the most massive subhalos are grossly inconsistent with the kinematic properties of the dSphs: they are characterized by $M_{\text{sub}}(< r_{1/2})$ systematically larger than $M_{\text{dwarf}}(< r_{1/2})$, and consequently by $V_{\text{circ}}(r_{1/2})$ systematically higher than observed. For instance, in the Aquarius simulations [294], at least 10 massive subhalos are characterized by circular-velocity profiles that are not consistent with any of the observed dSph circular velocities, $V_{\text{circ}}(r_{1/2})$: these subhalos have their maximum circular velocity (at $z = 0$) systematically higher ($V_{\text{max}} > 25 \text{ km s}^{-1}$) than those of the halos that best fit the observational data of the Milky Way dSph galaxies ($V_{\text{max}} \sim 12 - 25 \text{ km s}^{-1}$) [55]. Similar results are obtained by Garrison-Kimmel et al. [56] with the ELVIS simulations: here, the number of dense halos unaccounted for by observations ranges from 2 to 25 within 300 kpc from the Milky Way. Figure 11 illustrates this result.

Even though it was originally identified for the Milky Way system, the TBTB issue also appears in the M31 system, where the dSph satellites have circular velocities $V_{\text{circ}}(r_{1/2})$ systematically lower than expected for the most massive subhalos of $\Lambda$CDM simulations of Milky Way-like dark matter halos [295]. Furthermore, the conflict emerges in isolated field galaxies of the Local Group, indicating that this is not a satellite-specific problem: none of these galaxies is denser than the densest satellite of the
Milky Way or M31 [296], and a comparison of the observational data with the ELVIS simulations shows that the number of halos that are too dense to be consistent with the observations increases from the above-mentioned 2–25 to a number of 12 to 40, when moving to the outskirts of the Local Group [56].

The TBTF issue also appears beyond the Local Group. By applying the AM technique to the number density of galaxies in the ALFALFA sample as a function of their rotational velocity as inferred from the width of their HI emission line, Papastergis et al. [297] show that either the dwarf galaxies are hosted by halos that are significantly more massive than indicated by their rotation velocity or the number density of galaxies at the scale of dwarfs is much larger than observed. Clearly, neither solution is consistent with the observations.

Similar tensions are suggested by photometric observations alone. The systems of galaxies surrounding both M94 and M101 appear to contain galaxies that are substantially less massive than expected in the CDM framework. M94 is a group with a central Milky Way-mass galaxy that displays only two satellites, each with mass $M_* \lesssim 10^6 M_\odot$, instead of the $\sim 10$ satellites expected for similar systems [298]. The luminosity function of the galaxies of the M101 group is similar to the luminosity function of the Milky Way system, suggesting a similar lack of intermediate-mass galaxies [299]. In the ultra-faint dwarf galaxy regime, the M101 group also appears to exacerbate the MSP (see Section III C), by containing only half of the satellites of the Milky Way system [300].

1. Possible solutions within the CDM model

A systematic overestimate by a factor of $\sim 2$ of the mass of the dark matter halo of the Milky Way can in principle remove the TBTF problem for the Galaxy [55]. If the Milky Way mass were $M \sim 5 \times 10^{11} M_\odot$, the number of extra massive subhalos that do not correspond to observed dwarfs would be $\sim 3$ and might be consistent with statistical fluctuations. On the contrary, mass estimates of $M \sim 1 \times 10^{12} M_\odot$ [301–304] or $M \sim 2 \times 10^{12} M_\odot$ [305, 306] imply larger numbers of extra subhalos, $\sim 6$ or $\sim 12$, respectively, and generate the TBTF problem. However, there are currently no indications of the existence of systematic errors that can reduce the estimate of the Milky Way mass by the required factor. In addition, the TBTF problem appears in the M31 and in other galaxies; therefore, the overestimate of the Milky Way mass would not fully remove the tension.

The most promising way out is to resort to the effects of baryonic physics. N-body/hydrodynamical simulations suggest that energetic feedback from supernovae can substantially reduce the amount of dark matter mass in the central region of the subhalos, thus decreasing the observable circular velocity [307–309]. The mass depletion in the central regions appears to be less relevant in the APOSTLE simulations [259, 310, 311] when the EAGLE models for galaxy formation [312, 313] are implemented; in these simulations, the TBTF problem is instead mostly solved by subhalo disruption and mass loss due to tidal stripping. Similar results are obtained in the FIRE simulations of Local Group-like and isolated Milky Way-like volumes [314]. However, in these simulations, Local Group-like volumes still contain $\sim 10$ more subhalos than observed with stellar mass $M_* > 10^5 M_\odot$, showing how reproducing the properties of dwarfs both in high- and low-density regions remains challenging.

The most relevant limitation of the N-body/hydrodynamical simulations is the mass resolution of the baryonic particles. This limitation makes the final results dependent on the details of the recipes implemented for the baryonic physical processes, mostly gas cooling, star formation, stellar and supernova feedback, and chemical enrichment of the interstellar medium. For example, the APOSTLE simulations typically have particle mass $\sim 10^5 M_\odot$, whereas in the simulations of Brooks and Zolotov [307] the particle mass is $\sim 2 \times 10^4 M_\odot$ and in the FIRE simulations it is $\sim 3.5 \times 7 \times 10^4 M_\odot$ [314]. Therefore, the physical processes are self-consistently resolved only on scales comparable or larger than the total stellar mass of the least massive classical dwarfs. Self-consistently resolving the baryonic physics on smaller scales is still beyond the current capability of the simulations and the full implication of the processes on these scales remain to be assessed.

The TBTF problem might be partially solved without resorting to these baryonic subgrid processes: tidal effects of the baryonic disk of the host galaxy can be sufficient to reduce the circular velocities at 300 pc by 20-30% for most subhalos, as demonstrated by dark-matter only simulations where a growing disk potential is embedded at the centre of the dark matter halo [315]. However, this solution generates an anticorrelation between the central density of the subhalos and their pericentre values that lacks in the observations; in addition, this solution cannot be applied to field dwarf galaxies.

Ostriker et al. [316] also suggest that resorting to baryonic physics is actually unnecessary. They identify the TBTF problem with the lack, in galaxy groups, of moderately luminous galaxies, 1–2 magnitudes fainter than the brightest member. According to their analysis, N-body simulations, including the EAGLE and IllustrisTNG simulations, provide magnitude gaps between the most and the second most luminous galaxies comparable to or even larger than observed. It follows that the TBTF problem would vanish. Because these large gaps are also present in dark-matter only simulations, where a mass gap replaces the magnitude gap, Ostriker and collaborators conclude that these gaps are driven by gravitational dynamics rather than baryonic physics. However, it remains to be shown that this photometric interpretation of the TBTF problem also removes the original discrepancy between the circular velocities of the real dwarfs and the circular velocity of the most massive dark matter subhalos in galaxy groups.
Figure 11: Circular velocity profiles of simulated dark matter subhalos compared to measured velocity dispersions of Milky Way dSph satellites to illustrate the too-big-to-fail problem in the Milky Way system. The circular velocity profiles, shown with lines of different colors and styles, are derived by Garrison-Kimmel et al. [56] for the resolved, massive dark matter subhalos characterized by maximum circular velocity $V_{\text{peak}} > 30 \text{ km s}^{-1}$ at the time when the halo reaches its maximal mass; these subhalos are identified in the ELVIS simulations within 300 kpc of the centre of a dark matter halo with virial mass $M_{\text{vir}} = 1.3 \times 10^{12} \text{ M}_\odot$. The circular velocities, $V_{\text{circ}}(r_{1/2})$ measured at the half-light radius of the Milky Way dSph satellites brighter than $2 \times 10^5 \text{ L}_\odot$, are compiled by Wolf et al. [293] and are plotted as filled squares whose size is proportional to the logarithm of the dSph stellar mass. The cyan solid lines represent the circular velocity profiles of the so-called “strong massive failures”, subhalos that are too dense to host any of the observed Milky Way dSphs (i.e. whose $V_{\text{circ}}(r_{1/2})$ is above the $1 \sigma$ constraints for all the dwarfs in the sample). The black solid lines indicate the “massive failures”, i.e. additional subhaloes that are not accounted for by the dense galaxies in the observational sample. The subhalos selected to host the high-density galaxies, Draco and Ursa Minor, are indicated by dotted magenta lines, with their associated galaxies plotted as magenta squares. The dotted lines plot the subhalos that are consistent with at least one of the remaining seven dwarfs in the sample. The grey dashed line indicates the sole subhalo expected to host a Magellanic Cloud ($V_{\text{max}} > 60 \text{ km s}^{-1}$). Not plotted are 40 resolved ($V_{\text{max}} > 15 \text{ km s}^{-1}$) subhalos with $V_{\text{peak}} < 30 \text{ km s}^{-1}$. Overall, there are 12 unaccounted-for massive failures, including eight strong massive failures that are too dense to host any bright Milky Way dSph. The figure is reproduced from [56].

**E. Planes of satellite galaxies**

The orbits of the satellite galaxies of the Milky Way, of M31, and of Centaurus A (CenA), whose properties are well-measured, tend to be aligned in significantly flattened configurations [317–324]. Moreover, the satellites have significant kinematic correlations. These spatial alignments and kinematic coherence are extremely rare in CDM simulations [325].
puzzling fact is referred to as the **planes of satellite galaxies** problem (PSP). Unlike some other small scale problems of the CDM model, this issue may not be tackled by incorporating the baryonic feedback alone, because the satellites are at larger distances compared to the size of the baryonic disk of the host galaxy. For an extensive review of this topic, the reader may refer to [326]. Here, we briefly mention the observational evidence of such alignments, how rarely they occur in the simulations performed so far, and some suggested solutions within the CDM framework.

1. Evidence of the orbital alignment of the satellites

For a long time, it has been known that the well measured satellite galaxies of the Milky Way lie almost on the polar great circle [327, 328]. A more recent study confirms the existence, around the Milky Way, of a **vast polar structure** (VPOS) that includes distant globular clusters and stellar streams [317]. Depending on the selection of the samples, the plane of VPOS has an r.m.s. thickness of 20 to 25 kpc, an axis ratio of 0.18 to 0.30 and an inclination of 73° to 87° with respect to the galactic disk [326] (Figure 12). New samples of fainter satellites from the Sloan Digital Sky Survey [329] and the Dark Energy Survey [330] also support this spatial correlation [326, 331]. In addition, kinematic measurements imply that at least 8 out of 11 well-measured satellites co-rotate in the plane of VPOS [326, 332, 333]. The most recent data from Gaia also confirm both these spatial and kinematic correlations of the Milky Way satellites [334, 335].

Due to the lack of measures, the plane of satellites of M31 was not apparent until the Pan-Andromeda Archaeological Survey [336] discovered new satellites, and provided distance measures. These data show that 15 out of 27 satellites lie on a plane, the **giant plane of Andromeda** (GPoA), which is aligned with the Giant Stellar Stream in the M31 halo [318, 319]. The GPoA has an r.m.s. thickness of 12.6 kpc, an axis ratio of 0.1 and an inclination of about 50° with respect to the M31 disk [326, 337]. The GPoA is viewed almost edge-on from our vantage point, and line-of-sight velocities of the satellites suggest a strong kinematic correlation (see Figure 13). Thirteen out of the fifteen in-plane satellites follow the coherent kinematic trend and suggest a co-rotating plane of satellites. Interestingly, most of the in-plane satellites are on the Milky Way side of the GPoA whereas off-plane satellites are randomly distributed [319]. The angle between the two satellite planes of the Milky Way and M31 is between 40° to 50° and they have similar spin directions.

In addition to the satellite planes of the two major galaxies mentioned above, 14 out of the sample of 15 dwarf galaxies located within the Local Group but beyond the virial volumes of the Milky Way and M31, are part of two planes called the **Local Group Plane 1 and 2** [337, 338]. The first and second planes contain 9 and 5 galaxies, respectively. Both planes are about 300 kpc from the Milky Way and M31, and are strongly flattened, with r.m.s. thicknesses of about 60 kpc and axis ratios of about 0.1. The probability of having these configurations purely by coincidence is less than 0.003 in the CDM model [338].

Observing satellite galaxies outside the Local Group is challenging. Nevertheless, several independent studies of the satellites around CenA strongly suggest preferential alignments of their orbits [320–324]. Earlier studies with smaller samples suggested the existence of two parallel planes of satellites with r.m.s. thicknesses of about 60 kpc, axis ratios of about 0.2 and a separation of 300 kpc between the planes. A recent study [323] with larger samples claims the existence of only one satellite plane, the **Centaurus A satellite plane** (CASP), with similar orientation but larger thickness. The CASP is found to be almost perpendicular to the plane of the galactic disk, similar to the satellite plane of the Milky Way. From our position, the CASP is seen almost edge-on and 14 out of 16 satellites that have line-of-sight velocity measures, are found to be co-rotating in the plane [324] (Figure 14).

2. Comparison with the CDM predictions

The planar distribution of the orbits of the 11 best-measured Milky Way satellites was initially thought to be inconsistent with the isotropic distribution predicted by the CDM model [339]. However, even in the CDM paradigm, the preferential accretion along the cosmic filaments [340] and the triaxiality of the halo of the host galaxy [341] might be responsible for the anisotropy of the spatial distribution of the satellites. In high-resolution simulations [342], only 5-10% of the simulated samples resemble the spatial flatness of the observed satellite distribution.

However, the combination of the spatial and kinematic coherence has probability \( \sim 10^{-3} \) of appearing in CDM simulations [325]. In addition, finding two host galaxies (such as the Milky Way and M31) in the simulations, each with a co-orbiting satellite plane, is extremely rare with a probability \( \sim 10^{-6} \) [325]. Increasing the resolution of the simulations does not seem to alleviate the issue [326]. In the literature, there are disagreements about the severity of the tension between the plane of satellites and the CDM predictions which mostly stem from the different selection criteria of the satellite halos adopted in the simulations [326].
Figure 12: Edge-on view of the vast polar structure (VPOS) around the Milky Way and the disk (solid black line at the center). The orientation and the width of the best fit satellite plane are indicated by the dashed and dotted lines, respectively. The colored triangles, red upward (blue downward) for receding from (approaching towards) an observer at rest with respect to the host galaxy, indicate coherent kinematics of the co-orbiting satellites. The satellites with no proper motion measurements are plotted as crosses. The grey area corresponds to the region ±12° from the Milky Way disk which is obscured by galactic foreground. The figure is reproduced from [326].

3. Possible solutions within the CDM model

Although several solutions have been suggested for the PSP within the CDM model, none of them has been unanimously accepted, because each solution either fails to reproduce all the observational aspects or it is based on poorly-investigated assumptions.

The accretion along the filaments of the cosmic web results into anisotropic spatial distributions of the satellites [340, 341, 343, 344], because the massive subhalos tend to be accreted along the spine of the filament. There is suggestive evidence that the satellite planes of M31 and CenA have the correct alignments as expected from this mechanism [345], but this same mechanism appears unable to generate the satellite plane of the Milky Way. Preferential alignments due to the cosmic filaments increase the degree of anisotropy in the subhalo distribution, but are not strong enough to reproduce the observed planar structure [326, 346].
Moreover, when both the spatial and kinematic correlations are required, only 0.6% of the simulated samples are similar to the observed ones [332, 347].

Certain amount of spatial and kinematic correlations among the satellites are expected if the satellite plane is the result of the infall of a group of satellites [348, 349]. Tight alignments of the Magellanic stream with the VPOS of the Milky Way and the alignments of the Giant Stellar stream with the GPoA of M31 provide further support to this mechanism [350, 351]. However, there are several unsolved issues. Firstly, the observed number of luminous satellites may require more than one falling group [342, 352–354]. Secondly, the falling group or groups must also be sufficiently compact and narrow to survive the tidal disruption from the host galaxy and result into the observed planar structure of the satellites [349, 355, 356]. Thirdly, support from the Magellanic or Giant Stellar stream is not quite robust because some satellites in the planes are located in regions which, according to the simulations, are found to be inaccessible by such streams even if the streams are allowed to complete more than
One of the most promising solutions is that the satellite planes consist of tidal dwarf galaxies (TDGs) that originated from the debris of interacting or colliding major galaxies [357–359]. They are common in hydrodynamical simulations [360], and can account for the strong phase-space correlation required to match the observations [317, 339, 361]. However, the low metallicity of the observed satellites disfavors their formation from a recent merger since, in that case, they would be more metal-rich as their parent galaxies [362]. This tension might disappear if the satellites formed $\sim 10$ Gyrs ago from pre-enriched large-mass galaxies [317]. However, this solution might not be viable: according to hydrodynamical CDM simulations, TDGs are expected to be baryon dominated whereas the observed satellites appear to be dark matter dominated if they are in virial equilibrium [178]. Clearly, if the satellites are out of equilibrium, their observed velocity dispersion profiles might not necessarily imply the presence of dark matter [361, 363, 364], but this hypothesis would contradict the assumption of their early formation.

One of the solutions that go beyond the CDM framework is to assume new properties of the dark matter, e.g. self-interacting dissipative dark matter (see Section IV B), which allows the existence of a dark disk in the host galaxy, so that dark matter-
dominated TDGs may form during the merger events [365, 366]. However, the presence of a dark matter disk in the Milky Way is disfavored by Gaia data [367, 368]. Alternatively, if we adopt MOND as the theory of gravity (see Section V A), the formation of TDGs is enhanced due to the additional self-gravity [369, 370]. However, it is not clear yet whether TDGs produced in MOND would show sufficient non-Newtonian behavior to account for the observed dark matter-like signatures [326].

IV. POSSIBLE SOLUTIONS BEYOND THE STANDARD COLD DARK MATTER

Several astrophysical aspects of dark matter are related to its particle nature, like the minimum mass of a dark matter halo created by the hierarchical formation of structure [371]. The two parameters that regulate the structure formation are the characteristic free-streaming wavenumber \(k_{fs}\) and the characteristic interaction or decay rate \(\Gamma\) of the dark matter particles [371]. The wavenumber \(k_{fs}\) is set by the mean velocity of the dark matter particles at the time of decoupling, and depends upon the decoupling temperature \(T_{di}\) and the mass of the dark matter particle; \(\Gamma\) is set by the dark matter-dark matter scattering rate or the lifetime of the dark matter particles.

Figure 15 shows the allowed regions in the \(\Gamma - k_{fs}\) parameter space for different dark matter candidates. Deviations from the CDM paradigm arise when the gravitational collapse of dark matter is inhibited or modified above a characteristic comoving wavenumber. This wavenumber translates into a characteristic halo mass below which the number of halos is reduced. Alternatively, the deviations can be driven by the interactions of the dark matter particles with the Standard Model particles throughout the evolution of the Universe. Both effects can erase existing structures or change the velocity distributions of the dark matter particles and the densities of their structures. For WIMPs, whose decoupling temperature varies from 15 MeV to 1500 MeV and \(k_{fs} \sim 1\text{pc}^{-1}\) for \(m_{\chi} = 100\text{ GeV}\), one gets a minimum halo mass of the order of \(\sim [10^{-8} - 10^{-2}] \text{M}_\odot\). Sterile neutrinos, which are a warm dark matter (WDM) candidate with mass ranging from \(10^3\) to \(10^5\) keV, can then also create structures in the range \(M_{\text{halo}} \sim [10^{-6} - 10^{11}] \text{M}_\odot\). Similarily, gravitinos lead to \(M_{\text{halo}} \sim [10^{-17} - 10^{-14}] \text{M}_\odot\) while their mass is set by the supersymmetry breaking scale which is in the range 100 eV to 100 TeV. Fuzzy Dark Matter models also erase structures below the de-Broglie wavelength of the particle, of the order of the kiloparsec, which leads to a minimum halo mass \(M_{\text{halo}} \sim 10^{10} \text{M}_\odot\) for a particle mass \(\sim 10^{-22} \text{eV}\).

In the following sections, we will review the state of the art of some dark matter models, and discuss both successes and failures of each model in solving the challenges faced by the standard CDM model discussed in Section III.

A. Warm Dark Matter model

In contrast to CDM, WDM particles decouple when they are still relativistic; they thus erase primordial fluctuations on sub-galactic scales, and produce a cut-off in the primordial power spectrum [372]. WDM particles can play the role of CDM in the cosmological evolution of the Universe [373] and may also alleviate some of the problems of the CDM model at galactic scales as we explain below.

One of the most powerful tools to investigate the suppression of the primordial power spectrum on small scales is the Lyman-\(\alpha\) forest. A lower limit to the mass of the WDM particles were initially set to \(m_{\chi} > 750\text{ eV}\) [374] by fitting the Lyman-\(\alpha\) forest in quasar spectra. More recent analyses of the Lyman-\(\alpha\) forest and of the Milky Way satellites have increased the bound to several keV [375–378].

One of the most stringent bounds on the mass of WDM particles comes from the high-resolution HIRES/MIKE spectrographs: \(m_{\chi} > 4.65 \text{ keV}\) [379]. The most promising candidate of WDM is sterile neutrino, with mass \(m_{\alpha}\), which is mixed with an ordinary neutrino [380, 381]. For small mixing angles such as \(\sin^2 2\theta \sim 10^{-7}\), the total amount of sterile neutrinos is only a small fraction of the ordinary neutrinos. In Figure 16, we give a visual representation of the allowed parameter space for sterile neutrinos in the plane \(\sin^2 2\theta - m_{\chi}\). The claimed detections are based on the so-called 3.55 keV line emission, which is attributed to the decay of dark matter particles, and are obtained by using Chandra X-ray observations of galaxies in the Local Group [382], studies on dwarf galaxies [383], and Suzaku observations of the Perseus galaxy cluster [384]. These investigations imply a sterile neutrino of mass \(\sim 7\) keV and \(\sin^2 (2\theta) = [2, 20] \times 10^{-11}\). In contrast to these results, the full-sky Fermi Gamma-ray Burst Monitor data [385] do not reveal any significant signal for sterile neutrino decay lines in the energy spectrum, and improve previous upper limits by an order of magnitude (for a comprehensive review on sterile neutrino we refer the reader to [386]).

Additional constraints have been obtained by using gravitationally lensed quasars: under the assumption of a thermal relic dark matter particle, modelling the image positions and the flux-ratios of several gravitationally lensed quasars implies a lower limit of \(m_{\chi} > 5.58 \text{ keV}\) [388] or \(m_{\chi} > 5.2 \text{ keV}\) [388], consistent with the lower limits \(m_{\chi} > 5.3 \text{ keV}\) or \(m_{\chi} > 3.5 \text{ keV}\), depending on the assumed thermal history of the intergalactic medium, derived from the analysis of the Lyman-\(\alpha\) forest [389–391].
Figure 15: Allowed parameter space of the currently most popular dark matter candidates. The bottom axis shows the characteristic free-streaming wavenumber of the model, which is set in the primordial Universe. The top axis shows the corresponding halo mass, and the y-axis shows the characteristic interaction or decay rate which quantifies the evolutionary effects of dark matter. ADM, BEC, and FDM stand for Asymmetric Dark Matter, Bose-Einstein Condensate, and Fuzzy Dark Matter, respectively. The figure is from [371].

1. Solutions to the observational challenges

WDM particles are moving freely when they are relativistic; they can thus travel distances of the order of the horizon size. It follows that density fluctuations are suppressed on scales below the inverse of a characteristic comoving wavenumber [392]:

\[ k_{\text{WDM}} \sim 15.6 \frac{h}{\text{Mpc}} \left( \frac{m_X}{\text{keV}} \right)^{4/3} \left( \frac{0.12}{\Omega_{DM} h^2} \right)^{1/3}, \]

with obvious meaning of the symbols. This suppression of the density perturbations may lead to a solution of some of the problems that the CDM model encounters at galactic scales (see Section III). As shown in the left panel of Figure 17, the matter power spectrum drops below a certain length scale \( k^{-1} \) depending upon the mass of the WDM particles. For example, the power spectrum is suppressed below \( k^{-1} \sim 100 \) kpc for a particle mass of \( \sim 1 \) keV [393]. Therefore, the subhalo mass function can be
Figure 16: The allowed parameter space for sterile neutrinos obtained from the X-ray emission. The colored regions are 1, 2 and 3σ confidence levels. Numbers from 1 to 9 mark the centre of the 2σ region detections obtained using nine galaxy clusters [387]. The lines show constraints at the 90% confidence level from Chandra [382], Suzaku [384] and dwarf galaxies dataset [383]. The stars represent several models of sterile neutrino. The figure is reproduced from [386].
B. Self-interacting dark matter

A model of warm and self-interacting dark matter was firstly introduced in 1992 [407], and subsequently constrained few years later [408, 409]. In 1999, Spergel and Steinhardt [410] proposed the idea of cold and self-interacting dark matter in the concordance cosmology to solve two small scale issues of CDM: the core-cusp and the missing satellites problems (see Sections III B and III C). The newly proposed self-interacting dark matter (SIDM) particles behave like collisionless CDM at larger length scales, where the rate of collisions becomes negligible due to the smaller density. The typical collision rate of SIDM particles in the central region of a dwarf galaxy is

$$R_{\text{coll}} = \rho_{DM} v_{\text{rel}} \left( \frac{\sigma}{m} \right) \approx 1.0 \text{ Gyr}^{-1} \left( \frac{\sigma/m}{1 \text{ cm}^2/\text{g}} \right) \left( \frac{\rho_{DM}}{0.1 M_\odot/\text{pc}^3} \right) \left( \frac{v_{\text{rel}}}{50 \text{ km/s}} \right),$$

where $m$ and $\rho_{DM}$ are the mass and mass density of the dark matter particles, $v_{\text{rel}}$ and $\sigma$ are the relative velocity and the scattering cross section, respectively. SIDM models are commonly parametrized by the cross-section per unit mass $\sigma/m$, which, in general, is a function either of the relative velocity $v_{\text{rel}}$ of the dark matter particles, or of the total mass of the virialized halo $M_{\text{halo}}$. In addition, $v_{\text{rel}}$ and $M_{\text{halo}}$ are related by the fact that the velocity dispersion of the dark matter particles is larger in more massive halos. Here, we will briefly discuss the small-scale challenges addressed by SIDM and where SIDM stands in particle physics. For a detailed review, the reader may consult Tulin & Yu [411].

1. Solving small-scale issues with SIDM

Cosmological simulations of SIDM halos without the baryonic feedback predict constant density cores [412–415] (see Figure 18) contrary to the cuspy profiles from the simulations of collisionless CDM, as discussed in Section III B. The SIDM simulations solve the cusp-core problem if $\sigma/m \gtrsim 0.5 \text{ cm}^2/\text{g}$ for galactic halos with $M_{\text{halo}} \sim 10^{11} M_\odot$ [415, 416]. However, galaxy clusters with $M_{\text{halo}} \sim 10^{14} M_\odot$ require $\sigma/m \sim 0.1 \text{ cm}^2/\text{g}$ [411, 416–418]. The difference in $\sigma/m$ at galaxy and cluster scales implies velocity-dependent cross-sections which is a crucial aspect of the SIDM models.

---

Figure 17: **Left panel:** comparison of the linear matter power spectrum as a function of wavenumber in the WDM and CDM scenarios. The solid black line shows the CDM model. The dot-dashed lines represent the WDM power spectra for different values of the particle mass. **Right panel:** density profiles of CDM and WDM halos with mass $M = 10^9 M_\odot$. The black line shows the CDM profile; the red, green, blue, cyan, and magenta lines show the WDM profiles for particle mass $0.25, 0.5, 0.75, 1.0, \text{and } 1.25 \text{ keV}$, respectively. The solid colored lines show the case where there are no thermal relic velocities; the dot-dash lines denote the case where it is assumed the existence of a relic thermal velocity distribution for the WDM particles. In the panel, it is indicated the scale radius of the halo for the various WDM particle mass considered. The figure is reproduced from [372].

---

*In [411], the numerical value was mistakenly reported as 0.1 Gyr$^{-1}$.**
Figure 18: Inner density profile of a dark matter halo with $M_{\text{halo}} = 0.9 \times 10^{10} M_\odot$ in collisionless CDM and in SIDM for different values of $\sigma/m$. The figure is reproduced from [415].

Although SIDM with large scattering rates tends to reduce the number of subhalos, the MSP (see Section III C) is not easily resolved by SIDM [414, 419–421], unless non-minimal SIDM interactions are assumed [411]. Halos found in SIDM-only simulations are more spherical in the inner regions compared to their CDM counterparts; moreover, observed ellipticities of dark matter halos inferred from gravitational lensing have put several stringent constraints on the SIDM parameters at both galaxy and cluster scales [422]. However, there are disagreements among various constraints [422, 423] which may be resolved by properly accounting for the effects from baryons [411]. From SIDM-only simulations, the values of $\sigma/m$ required to solve the MSP are most likely ruled out by constraints from the measured ellipticities of dark matter halos (see [411] and references therein). However, as we discuss below, the baryonic contributions may significantly change this scenario on galaxy scales. On the scale of galaxy clusters, values of $\sigma/m$ in the range $0.3$ to $10^4$ cm$^2$/g are excluded based on the fact that, unlike the number of satellite galaxies, the number of galaxies in a cluster is comparable to the number found in the simulations of the formation of the large-scale structure [414, 424].

Simulations of SIDM and baryons [425, 426], with sufficient resolution at small scales, are computationally expensive. To overcome this drawback, a semi-analytical approach, known as the Jeans method, has been pursued to study the baryonic
contributions in the SIDM paradigm [414, 427, 428]. The Jeans method determines the inner density profile of the halo by assuming the inner halo to be isothermal and in hydrostatic equilibrium with the baryons. However, the outer part of the halo matches the NFW profile since SIDM particles are effectively collisionless there. For a galaxy like the Milky Way, the results using the Jeans method suggest that the inner halo shape is oblate rather than spherical as found in SIDM-only simulations, and the radius of the central core is smaller by an order of magnitude compared to the core size in SIDM-only simulations [427]. These modifications upon including the effects of baryons demand corrections of the constraints on SIDM models based on the observed halo shapes.

In addition, if baryonic effects are taken into account by the Jeans method, SIDM may solve the issues related to the disk-halo conspiracy [429] (see Section III A), i.e. it can reproduce diverse inner rotation curves from a single value of $\sigma/m$ for galaxies with similar maximum rotation speed $V_{\text{max}}$ [430, 431] (see Figure 19). Collisionless CDM, even with baryonic feedback, does not solve this issue because of the large central dark matter density in the central region. Thanks to the smaller dark matter density provided by the SIDM scenario, the baryons are more effective at setting $V_{\text{max}}$.

The self-interactions between the dark matter particles also alleviates the TBTF problem (see Section III D) for satellite and field dwarf galaxies. Even though self-interaction has little effect on the abundance or total mass of the subhalos, it effectively decreases the central density of the most massive subhalos by removing mass from these regions, characterized by cuspy density profiles. Velocity-independent SIDM models are consistent with the kinematics of the Milky Way dSphs for values of the cross-section per unit mass $\sigma/m \approx 1$ cm$^2$/g [421, 432]; lower values of $\sigma/m$ generate a population of subhalos too dense to be consistent with the observations. On the other hand, velocity-dependent SIDM models successfully solve the TBTF [420, 432, 433]. However, the class of velocity-dependent SIDM models remains largely unconstrained.

In principle, SIDM with dissipative or inelastic scattering may also alleviate the problem of the planes of satellite galaxies (see Section III E) by allowing the observed satellites to be dark matter dominated tidal dwarfs [365, 366]. In CDM simulations, dwarf galaxies formed due to tidal disruption during the merging of two galaxies are baryon dominated. On the contrary, a dissipative dark matter scenario allows a thin dark matter disk in a halo, and merging galaxies with such dark disks may produce dark matter dominated dwarf galaxies with strong phase-space correlations. However, Gaia data do not currently support the presence of a dark matter disk in the Milky Way [367, 368].

The collisional nature of SIDM at high relative velocities ($\sim 1000$ km/s) can be probed by merging clusters, such as the Bullet Cluster, in different ways [434]. Clusters, that nearly survive the merger, have scattering depth $\tau = \sigma \Sigma/m < 1$, where $\Sigma$ is the surface density of dark matter [435]. Additional constraints come from the amount of dark matter that is stripped off during the merger: we can infer this mass by estimating the mass-to-light ratios of the two merging clusters [435]. Another probe of the self-interaction is to look for any possible offset between the centroid of the luminous part of the merging cluster and that of its dark matter content inferred from weak-lensing [435]: self-interactions between the dark matter contents of two colliding clusters would generate a drag force that may exceed the gravitational binding between the dark matter and the luminous components. Combining the constraints from all the three methods for the Bullet Cluster shows that $\sigma/m$ is less than 0.7 cm$^2$/g [435]. For a list of such constraints for other mergers, see [411] and references therein.

2. SIDM in particle physics

SIDM may solve several small-scale issues of the CDM paradigm, provided $\sigma/m \sim 1$ cm$^2$/g on the scale of galaxies and $\sim 0.1$ cm$^2$/g on the scale of galaxy clusters. We now briefly describe the role of SIDM in the context of particle physics.

The simplest model of SIDM could be a real scalar field $\phi$ with quartic self-interactions ($\lambda/4!)\phi^4$. There are several ways to build this kind of models [436–438]. However, the contact-type self-interactions have a major drawback, namely the cross section is independent of the particle velocity and this feature generates dark matter cores both in galaxies and in galaxy clusters. Unfortunately, real galaxy clusters do not show any presence of a core.

The issue of the velocity independence can be tackled by considering a model where the dark matter particles ($\chi, \bar{\chi}$) are fermions and are self-interacting via a light mediator $\phi$, which can be a scalar or a vector field [439–441]. The self-coupling strength in the dark sector is characterized by $\alpha$, which is similar to the fine structure constant in electrodynamics. In the limit of $\alpha \chi m_\chi \ll m_\phi$, the differential cross-section for elastic self-scattering is given by [439, 441]

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 m_\chi^2}{\frac{1}{2}(1 - \cos \theta)m_\phi^2 v_{\text{rel}}^2 + m_\phi^2},$$

where $m_\chi$ and $m_\phi$ are the mass of the dark matter particle and the dark mediator, and $\theta$ is the scattering angle. In the limit of $m_\chi v_{\text{rel}} \ll m_\phi$, the cross-section is constant, whereas, for $m_\chi v_{\text{rel}} \gg m_\phi$, the cross-section varies as $1/v_{\text{rel}}^4$. The model is consistent with both galaxy and cluster scales, if the transition between the two regimes occur around $v_{\text{tran}} \sim 300$ km/s or equivalently, $m_\phi/m_\chi \sim v_{\text{tran}}/c \sim 10^{-3}$ [428]. More accurate cross-sections have been calculated to study the full parameter space of this model [423, 439–442]. The dependence of the cross section on the velocity varies with the limits of the parameters [441]. Figure 20 shows the preferred regions of the $(m_\chi, m_\phi)$ space for a fixed coupling $\alpha = 1/137$ [428].
Other models of SIDM include composite states of strongly interacting dark sector particles \( (m \sim 1\ TeV) \), e.g. dark hadrons in a QCD-like dark sector, and bound states of different dark sector particles, e.g. dark atoms and dark matter with excited states (see [411] and references therein). The SIDM paradigm is rich in phenomenology and each of the particle candidates of SIDM can be probed by both terrestrial searches for dark matter as well as astrophysical observations.

C. QCD axions

QCD axions have the double virtue of solving the Strong CP problem in the Standard Model (SM) of particle physics [30, 31, 443, 444] and of being a potential candidate for dark matter [445–447]. Starting with a brief description of their origin in particle physics, we will discuss the cosmology of axions, the constraints on their mass and when they behave differently from WIMPs.

1. Emergence of QCD axions

Quantum chromodynamics (QCD) describes the strong interactions between quarks and gluons. The CP violating parameter \( \bar{\theta} \) in QCD results into a non-zero electric dipole moment of the neutron [448] which has an experimental upper bound [449]:

\[
d_n \approx 5 \times 10^{-16} \bar{\theta} \ \text{e-cm} < 2.9 \times 10^{-26} \ \text{e-cm} ,
\]

i.e. the parameter \( \bar{\theta} \) is constrained to be unusually small, \( \bar{\theta} \lesssim 10^{-10} \). Given that the CP violating term is required to solve the \( U_A(1) \) problem [450] and that the CP violation occurs in the weak sector of the SM [113], \( \bar{\theta} \) is expected to have a value of \( O(1) \)
Combined fit
95% CL
99% CL
Dwarfs
LSBs
Clusters
Bullet
Cluster excl.
Cluster mergers excl.

0.01 0.1 1 10 100 1000
0.1
1
10
100

Dark matter mass (GeV)

Mediator mass (MeV)

Figure 20: Parameter space spanning the mass \( m_\chi \) of the SIDM particle and the vector mediator mass \( m_\phi \) with self-coupling \( \alpha_\chi = \alpha = 1/137 \). Regions preferred by dwarfs (red), LSB spiral galaxies (blue) and galaxy clusters (green), each at 95% confidence level, are indicated. Combined 95% (99%) region is shown by the solid (dashed) contour. The region estimated to be excluded by the Bullet Cluster (other observed merging clusters) lies below the dot-dashed (long-dashed) curve. The figure is reproduced from [428].

between 0 and \( 2\pi \). Why the value of \( \bar{\theta} \) is so small, i.e. why CP is so weakly violated in the strong sector of the SM, is known as the Strong CP problem. Several solutions to the Strong CP problem have been proposed (see [451–453] and references therein). Among them, the most attractive and popular solution is the Peccei-Quinn (PQ) solution [443, 444] which implies the existence of a new pseudo-scalar boson axion [30, 31]. In the PQ solution, the CP-violating parameter is promoted to be a dynamical variable \( a(x)/f_a \) where \( a(x) \) is the axion field and \( f_a \) is the axion decay constant with a dimension of energy. As the axion field \( a(x) \) relaxes to its vacuum expectation value \( \langle a \rangle \) to minimize the energy, the value of \( \bar{\theta} \) becomes zero and that solves the Strong CP problem. The constant \( f_a \) is given by the energy scale at which the PQ symmetry is broken. For QCD axion, the mass is related to the decay constant by [454]:

\[
m_a \approx 6 \times 10^{-6} \text{eV} \left( \frac{10^{12}\text{GeV}}{f_a} \right). \tag{11}
\]

However, for axion-like particles (ALPs), both \( m_a \) and \( f_a \) are considered to be independent parameters. ALPs not obeying Eq. (11) do not solve the Strong CP problem. The existence of ALPs or ultra-light ALPs (ULALPs) is motivated from string theory [32]. In the literature, the term axions may refer to QCD axions or ALPs or ULALPs depending upon the context.

After QCD axions were proposed to solve the Strong CP problem, they were quickly recognized to be a dark matter candidate [445–447], i.e. it was shown that they could be abundantly produced in the early Universe and their energy density behaves as that of CDM. Couplings of QCD axions with the SM particles can also be made very small since \( f_a \) can be allowed to be much larger than the QCD energy scale [455–458]. Here, we will briefly mention the key features of axion cosmology. The reader may consult [454, 459] for extensive reviews. In the early Universe, after the PQ symmetry is broken at temperature \( T_{PQ} \sim f_a \), axions could be produced by both thermal and non-thermal mechanisms. However, the present number density of non-thermal or cold axions is much larger than that of the thermal axions. Cold axions are produced by the non-thermal mechanisms such as vacuum realignments, decays of axion strings, and decays of axion domain walls [454]. At any given cosmic time \( t \), the
velocity dispersion of the cold decoupled axions is, in units of the speed of light $c$ [454],

$$\beta_a(t) \sim 3 \times 10^{-17} \left( \frac{10^{-5} \text{ eV}}{m_a} \right) \hat{a}(t_0) \hat{a}(t),$$  \hspace{1cm} (12)

where $a(t)$ is the scale factor of the expanding Universe and $t_0$ is the age of the Universe today. This velocity dispersion corresponds to an effective temperature of $0.5 \times 10^{-34} \text{ K} \left(10^{-5} \text{ eV}/m_a\right)^{2/3}$ today.

Axion mass is constrained by cosmology, astrophysics, as well as terrestrial experiments. If $f_a$ is larger than about $10^{12} \text{ GeV}$, there would have been a large amount of axions to overclose the Universe [454]. This implies a lower bound on the mass of QCD axions: $m_a \gtrsim 6 \times 10^{-6} \text{ eV}$. This bound is rather soft due to a number of uncertainties in the axion energy density from decays of strings and domain walls. Astrophysical phenomena, such as stellar evolution and the supernova SN1987A, also put constraints on the axion mass (see, [e.g., 460, 461]). Axions were thought to be invisible until it was shown by Sikivie [462] that, in a cavity with a strong inhomogeneous magnetic field, an axion can be resonantly converted into two photons when the tuning frequency of the cavity matches the axion mass. The cavity technique is used to design both axion haloscopes and helioscopes. Axion haloscopes, such as ADMX [463], are designed to search for axions in the Milky Way halo, while axion helioscopes, such as CAST [464] and IAXO [461], are designed to probe axions emitted from the Sun via Primakoff conversions. Heavier QCD axions with $m_a \gtrsim 50 \text{ keV}$ have been ruled out by beam dump experiments and rare meson decays (see, [e.g., 465]). Searches for QCD axions exploit their coupling $G_{a\gamma\gamma}$ to two photons which is related to the mass by [113]

$$G_{a\gamma\gamma} = C \times 10^{-15} \text{ GeV}^{-1} \left( \frac{m_a}{6 \times 10^{-6} \text{ eV}} \right),$$ \hspace{1cm} (13)

where $C$ is a model dependent number of order one. The current constraints on axion mass and their coupling to photons are shown in Figure 21 where the yellow band corresponds to the QCD axions in different models. There has been a significant boost in the interest to search for axions or ALPs in the last decade. The reader may consult [113, 466, 467] for an overview of many novel ideas and experiments that have been proposed.

2. Distinctive features of QCD axions

In the context of cosmology, QCD axions may have two distinctive features: isocurvature perturbations and axion miniclusters (see [454, 459] and references therein). Quantum fluctuations of the axion field may cause isocurvature perturbations either on the length scale of the CMB (for the PQ phase transition before inflation), or on the scale of QCD horizons i.e. the horizon size at the time of the QCD phase transition (for the PQ phase transition after inflation). The CMB observations put tight constraints on how much isocurvature perturbations are allowed (see [e.g., 468–470]). For the PQ phase transition after inflation, the misalignment angle may be different in different QCD horizons and result into large inhomogeneities in the density of axions. Such inhomogeneities grow in size due to gravitational instabilities and lead to the formation of axion miniclusters with densities many order of magnitude larger than the average density (see, [e.g., 471, 472]). Searches for axions would be more challenging if a substantial fraction of them are present in the form of miniclusters which have typical mass of many order of magnitude larger than the average density (see, [e.g., 476]). Nevertheless, if the QCD axions undergo thermalization and form a Bose-Einstein condensate (BEC), they may have distinctive features like caustic rings and axion stars.

It has been shown by Sikivie and Yang [477] that, due to their huge phase-space degeneracy and very small velocity dispersion, the axions may thermalize via gravitational self-interactions and form a BEC when the photon temperature reaches about $500 \text{ eV} \left(f_a/10^{12} \text{ GeV}\right)^{1/2}$. It has been argued that, as the halo grows in size by accreting more dark matter, the axions continuously rethermalize to track the lowest available energy state [478]. This fact supports the necessary initial conditions for the formation of caustic rings which are ring-like substructures in the galactic plane [479, 480]. Caustic rings should have astrophysical signatures in the forms of bumps in the rotation curves, stellar overdensities and characteristic imprints on the...
Figure 21: The current constraints on QCD axions and ALPs. The yellow band represents the various models of QCD axions. The constraints from the Sun and the Horizontal Branch stars are the regions above the corresponding lines. The regions ruled out by various experiments are indicated by the colored areas. The figure is reproduced from [113].

density of the interstellar medium [481–483]. Several pieces of evidence of the existence of caustic rings have been reported in [481]. The existence of vorticities in the axion BEC causes a depletion of dark matter density near the galactic centre [478] which may explain why the inner rotation curves can be reconstructed only from baryonic contributions. The rethermalizing axion BEC can also solve [478] the galactic angular momentum problem which is the discrepancy between the observed angular momentum distributions in the dwarf galaxies and the predicted ones with ordinary CDM [484].

When the quantum pressure balances the gravitational attraction at small length scales, the axions may form gravitationally bound and stable objects like axion stars [485]. Dilute stars from QCD axions are expected to have mass of order $10^{-13} M_\odot$ and radius of order $10^{-4} R_\odot$. The mass of dense axion stars may range from $10^{-20} M_\odot$ to $M_\odot$. If an axion star collides with a neutron star, the axions may be converted into photons in the strong magnetic field of the neutron star and emit a strong flux of radiation. Interested readers may refer to [486] and references therein, for a detailed descriptions and observable consequences of such objects.

3. QCD axions and small-scale problems of CDM

As mentioned above, rethermalizing QCD axions with vortices may solve the galactic angular momentum problem and may also explain the depletion of dark matter in the central regions of the halos [478]. The caustic ring model, which is an outcome of rethermalizing axial BEC, is found to reproduce the inner and outer rotation curves of the Milky Way [487]. Despite these successes, QCD axions ($m_a \sim 10^{-6} \text{ eV}$) are hardly explored in the context of the small-scale problems of CDM. In fact, ULALPs of mass $\sim 10^{-22} \text{ eV}$ (see Section IV D) are attractive candidates to solve the small-scale challenges such as the CCP and the MSP. An attempt to solve the CCP with QCD axions by simply extrapolating the ULALPs scenario does not work, because, while ULALPs allow a core of size $\sim 1 \text{ kpc}$, the size of the core from QCD axions would be of $O(1) \text{ km}$ [488].

Whether QCD axions can address the small-scale problems requires further investigations. From the numerical perspective, resolving small scale structures through N-body simulations of QCD axions with mass $m_a \sim 10^{-6} \text{ eV}$ is formidable. From
the theoretical point of view, QCD axions are usually treated as classical fields (see, for example, [489, 490]). However, it has been shown that thermalization of axions [477] is a quantum phenomenon and cannot be explained by the classical field theory [475, 491] unless an artificial cut-off wavevector is introduced [489]. It has been argued [492] that the wave-mechanics assuming axions as classical fields may not unveil all the physical aspects and the quantum field effects play important roles for self-interacting axions.

D. Fuzzy Dark Matter

Another excellent alternative to CDM is Fuzzy Dark Matter (FDM) which consists of ultra-light bosons with mass in the range $10^{-23} - 10^{-20}$ eV [28, 29]. These light bosons are naturally generated from symmetry breaking due to the misalignment mechanism [445, 446, 493], and are very common in string theory (for more details see [32]). FDM is considered to be a real scalar field $\phi$ with mass $m_\phi$ which is minimally coupled to the metric [494]. The field $\phi$ is initially massless until the Universe cools down to some critical temperature [495]. It acquires the mass by rolling down and oscillating about the minimum of a potential generated non-perturbatively [495]. The typical de Broglie wavelength of a FDM particle is a few kpc:

$$\lambda_{DB} = \frac{\hbar}{m_\phi v} = 1.92 \text{kpc} \left(\frac{10^{-22} \text{eV}}{m_\phi}\right) \left(\frac{10 \text{km/s}}{v}\right).$$

Therefore, the physics of FDM on length scales below $\lambda_{DB}$ differs from that of the standard CDM. In particular, small density fluctuations unstable for masses larger than the Jeans mass (see Eq. 42 in [494]) lead to a minimum halo mass of $\sim 10^7 M_\odot$ for a boson mass of $\sim 10^{-22}$ eV. On the contrary, on a scale above $\lambda_{DB}$, the large scale structure of FDM is indistinguishable from CDM [476].

I. Solutions to the observational challenges

In 2014, novel N-body simulations with unprecedented high resolution showed the rich small scale structures of FDM halos [476]. The uniqueness of these N-body simulations was their ability to capture the quantum nature of the dark matter particles by combining the Schrödinger’s and the Poisson’s equations [496]. Each virialized halo has a core of dark matter in the innermost part, which represents the ground state solution of the Schrödinger-Poisson equations. This core is surrounded by an interference pattern represented by fluctuations in the velocity and density fields of the particles. The cores, more often called solitons, exhibit flat density profiles that can naturally explain the wide cores in dwarf galaxies, and match the NFW density profiles [117] in the outer regions of the halos (see Figure 22). As it was shown in [476], the solitonic density distribution of the dark matter is well fitted by

$$r_c \sim 1.9 \left(\frac{m_\phi}{10^{-23} \text{eV}}\right)^{-2} \left(\frac{r_c}{\text{kpc}}\right)^{-4} \left[1 + 9.1 \times 10^{-2}(r/r_c)^2\right]^6 M_\odot \text{pc}^{-3},$$

where $r_c$ is the core radius of the soliton; $r_c$ also is related to the boson mass and the halo mass $M_h$ by the scaling relation [476, 497]:

$$r_c = 1.6 \left(\frac{10^{-22} \text{eV}}{m_\phi}\right)^{1/2} \left(\frac{\zeta(z)}{\zeta(0)}\right)^{-1/6} \left(\frac{M_h}{10^9 M_\odot}\right)^{-1/3},$$

where $a$ is the cosmic scale factor, $\zeta(z) \equiv (18\pi^2 + 82(\Omega_m(z) - 1) - 39(\Omega_m(z) - 1)^2)/\Omega_m(z)$ is the critical overdensity at redshift $z$. These results have been validated by N-body simulations [498, 499].

Since solitons have a constant central density thanks to pressure support, they may potentially solve the CCP. To shed light on this issue, two independent analyses used stellar kinematics data of dwarf spheroidal galaxies by carrying out a Markov Chain Monte Carlo fitting procedure of the projected velocity dispersion profiles [488, 501–503]. Both analyses found that the data prefer soliton-generated cores (with boson mass $\sim 10^{-22}$ eV) over cuspy NFW profile. Figure 23 shows the effectiveness of the FDM halo profile at describing the kinematics of stars in the Milky Way dwarf satellites.

In addition, a detailed study of the recently discovered ghostly galaxy Antlia II has shown that the solitonic structure of the FDM halo may help to explain the presence of the wide core ($\sim 2.8$ kpc) of this dwarf galaxy [504], and of other ultra-faint galaxies [503]. However, in contrast to previous results, numerical solutions [506] of the evolution of a scalar field in a spherically symmetric space-time fail to reproduce the scaling relation between the core density and the core radius [507]. Nevertheless, it is worth noting that this scaling relation originates from a fitting procedure of the the so-called Burkert profile to the measured rotation curves of disk galaxies. This approach might not be appropriate, because in FDM, the core density and the core radius depend upon the boson mass and the total halo mass as in the scaling relation of Eq. (17). In addition to the
cores of ultra-faint galaxies, the solitonic structure can provide a solution to the excessive central velocity dispersion of the stars in the bulge of the Milky Way [508]. Finally, N-body simulations show the suppression of the halo number density for mass $\leq 10^{10} \, M_\odot$, and how this cut-off depends upon the mass of the FDM particle (Figure 24). Such suppression may provide a solution to the MSP [509].

The debate on whether or not FDM may be a viable dark matter candidate is still ongoing. The presence of the soliton in every virialized halo can affect the dynamics of the disk by enhancing the circular velocity in the inner part of the rotation curve; this feature can provide a way to probe the model [510]. One should note that the inner dynamics of a disk galaxy is affected by the baryonic contributions to the gravitational potential, and the breakdown of the spherical symmetry may affect the geometry of the solitons [510]. Additionally, the effect of quantum pressure in structure formation, which is suppressed on small scales, is still under investigation due to the poor resolution of the N-body simulations. A box size of at least $\sim 500 \, \text{Mpc/h}$ on a side is needed to predict the matter power spectrum, and evaluate how much suppression is introduced [511]. Furthermore, the clustering properties of hot gas at high redshift have been used to constrain the dark matter properties. The analysis of the Lyman-\(\alpha\) forest data constrains the boson mass to be larger than $7 \times 10^{-21} \, \text{eV}$ [512], which is almost two order of magnitudes larger than the boson mass required to describe dwarf galaxies ($\sim 10^{-22} \, \text{eV}$). Since the power spectrum of WDM and FDM is the same on large scales [513], and is suppressed below a certain characteristic scale [513, 514], the constraints based on Lyman-\(\alpha\) forest data are obtained by translating the observational bounds on WDM into the corresponding bounds on FDM by matching the suppression scale [512]. However, this analogy between WDM and FDM is still under debate [494, 515]. Hydrodynamical simulations are required for FDM to quantify the impact on the matter power spectrum of the small scale density fluctuations on the de Broglie scale which are absent in WDM [476, 498, 499]. These effects may play a fundamental role in distinguishing between the models [494], although the uncertainties on the different thermal histories and underlying reionisation models of the WDM and FDM particles may weaken these constraints [494, 516].
V. POSSIBLE SOLUTIONS BEYOND NEWTONIAN DYNAMICS

The observational challenges of the CDM model on galactic scales (see Section III) may also be interpreted as a breakdown of the law of gravity. Modifications of the law of gravity conceived to explain the observed kinematics of visible matter started to be systematically investigated back in the 1980s [147, 155, 517, 518], although some suggestions were put forward much earlier [e.g., 519]. On cosmological scales, observational data from the Planck mission do not seem to provide statistical evidence in favor of any particular theory of gravity [63], whereas at galactic scales, where the physics is complicated by the relevant role of baryons, the issue remains open. In the following, we briefly touch upon three modified gravity models, proposed in the literature, that focus on the scales of galaxies.

A. MOND

In 1983, Milgrom suggested to explain the mass discrepancy in cosmic structures with a modification of the law of gravity rather than with the presence of dark matter [33–35]. His phenomenological proposal rests upon the hypothesis that there is an acceleration scale \( a_0 \simeq 1.2 \times 10^{-10} \text{ m s}^{-2} \) above which Newtonian gravity holds and below which Newtonian gravity breaks down. This idea goes beyond the naive idea that gravity should be modified simply beyond a length scale [140, 519].

According to Milgrom’s suggestion, the magnitude \( a \) of the acceleration experienced by a test particle in a gravitational field is

\[
a = \nu \left( \frac{a_N}{a_0} \right) a_N, \tag{18}
\]

where \( a_N \) is the magnitude of the gravitational acceleration, estimated in Newtonian gravity, originated by the distribution of the baryonic matter alone, as dark matter is assumed to be nonexistent; \( \nu \) is an interpolation function whose asymptotic behaviors are \( \nu \to 1 \) when \( a_N \gg a_0 \) and \( \nu \to (a_N/a_0)^{-1/2} \) when \( a_N \ll a_0 \). A number of interpolation functions has been proposed in...
Rather than as a modification of the law of gravity, the introduction of an acceleration scale can be alternatively interpreted as a modification of the law of inertia $F = ma$, where the inertial mass differs from the gravitational mass when $a_N \ll a_0$ [521–523].

In both cases, Milgrom’s suggestion yields a MOdified Newtonian Dynamics (MOND) (see [149] for an extensive review). The introduction of an acceleration scale makes the MOND formulation manifestly purely phenomenological: in General Relativity, the acceleration is linked to the affine connection $\Gamma^{\mu}_{\lambda\nu}$ which is not a tensor; therefore, MOND cannot be easily formulated in a covariant form.

This drawback has two important consequences: in the MOND framework, we can neither build a cosmological model, which is the most relevant success of the $\Lambda$CDM model, nor quantify the phenomenology of gravitational lensing, which is an important probe of the mass distribution on large scale. An additional shortcoming is that MOND is unable to explain the observed mass discrepancy on the scale of galaxy clusters and on larger scales, although the amount of required dark matter is substantially reduced [35].

Attempts to provide MOND with a covariant formulation include, for example, AQUAL [517], TeVeS [524], and bimetric MOND [525]. These theories introduce additional scalar, vector or tensor fields and reduce to MOND in the non-relativistic limit [526]. Their number shows that a covariant theory that reduces to MOND is not uniquely determined. Therefore, invalidating
one of these theories does not necessarily invalidate MOND.

For example, the detection of gravitational waves originating from the merging of two neutron stars [527] combined with the observation of a gamma-ray burst within a few seconds [528, 529] implies that the speed of light and the speed of gravitational waves coincide within one part in $10^{-15}$. In the original formulation of TeVeS, the speed of gravitational waves in general is different from the speed of light and therefore TeVeS appears to be ruled out [530]. Nevertheless, there is a family of tensor-vector-scalar theories, that still reduce to MOND in the non-relativistic limit, where the speed of gravitational waves equals the speed of light [531].

We can also build a hybrid model that merges the success of MOND on small scales with the properties of the large-scale structure provided by the presence of dark matter. This idea was suggested by Angus in 2009 [532]. He assumed MOND as the theory of gravity and added a hot dark matter component made of sterile neutrinos of mass $\sim 11$ eV. The existence of a sterile neutrino still appears to be a solution to the detection of the excess of electron-like events in short-baseline neutrino experiments [533]. Hot dark matter has the advantage of clumping on scales larger than the scale of galaxies. MOND phenomenology would thus be preserved on small scales, whereas dark matter starts becoming relevant on larger scales, where MOND apparently disagrees with observations (see [534] for a review). Unfortunately, this hybrid model is unable to reproduce the mass function of galaxy clusters [535–537] and currently appears unviable.

1. Disk galaxies

Many of the observational challenges described in the previous sections were predicted by MOND many years before they were actually observed and posed unexpected challenges to the traditional dark matter framework. This feature is specific to MOND and makes it fundamentally different from the other suggested theories of modified gravity: these latter theories attempt to describe these observations only after they become available and never anticipate them.

Indeed, the predictions of MOND on the scales of galaxies are so distinctive that it has become customary to collect them in the so-called MOND phenomenology. These observational facts are clearly independent of the theory of gravity; therefore, other theories, alternative to the standard model, must mimic the MOND phenomenology on these scales. Recent models that explicitly attempt to reproduce the MOND phenomenology include superfluid dark matter [538], dipole dark matter [539], refracted gravity [152, 540], and emergent gravity [38].

MOND was motivated by the observations of the flat rotation curves of disk galaxies [33]. The normalization of the Tully-Fisher relation [541] sets the magnitude of the acceleration scale $a_0$. As reminded in Section III A, the Tully-Fisher relation links the luminosity $L$ of a disk galaxy to the width $W$ of the profile of the global neutral hydrogen line in emission; if $L$ and $W$ are proxies of the galaxy mass and the asymptotic rotation velocity respectively, the Tully-Fisher relation becomes a relation between mass and centripetal acceleration.

At large $r$, where we can consider that the cumulative baryonic mass $M(< r) = M_d$ is constant, we have, in Newtonian gravity, the acceleration $a_N = GM_d/r^2$, and, from Eq. (18), in the limit $a_N \ll a_0$, we have $a^2 = a_N a_0 = a_0 GM_d/r^2$. At large $r$, we have $a = v_{\text{rot}}^2/r$, with $v_{\text{rot}} = \text{const}$, and we obtain $v_{\text{rot}}^4 = a_0 GM_d$. We thus see that the normalization of the Tully-Fisher relation is $a_0 G$, and its slope is 4. In the 1980s, observations provided a considerably uncertain slope, in the range 2.5 – 5, as reviewed in [34]; nowadays, this slope is confirmed to converge to 4 [159].

In this perspective, the Tully-Fisher relation simply is a law of gravity, similar to the Kepler’s laws, and it is not linked to any property of galaxies other than their baryonic mass. Consequently, the Tully-Fisher relation should hold for any self-gravitating system; this prediction is relevant because we should expect the Tully-Fisher relation to hold irrespective of the surface brightness of the galaxy. In other words, MOND predicts the existence of a Baryonic Tully-Fisher relation between the baryonic mass, including stars and gas, and the asymptotic flat rotation velocity. The BTFR is now neatly supported by the data and holds from LSB to HSB galaxies over six orders of magnitude in baryonic mass [542] with a slope confirmed to be in the range 3.5 – 4 [543] (see Figure 2).

The BTFR is intimately connected to another MOND prediction: the correlation between the disk surface brightness $\mu$ and the centripetal acceleration $a = v_{\text{rot}}^2/r$, that directly derives from the relation $a \sim 2\pi G\mu$. Accurate spectroscopic measures of the disk galaxies in SPARC [130] and other data sets [e.g., 544] support this relation.

MOND also implies the Faber-Jackson relation $M \propto \sigma^4$ between the mass $M$, if proportional to luminosity, and the stellar velocity dispersion $\sigma$ of elliptical galaxies [545]. Unlike the Tully-Fisher relation however, the Faber-Jackson relation would be exact in MOND only if ellipticals were isothermal and their velocity fields were isotropic. In fact, matching the observed Fundamental Plane of ellipticals requires that the velocity anisotropy parameter varies with radius [546, 547].

The fit of the rotation curve in the standard model has two free parameters that describe the dark matter halo density profile when assumed to be spherical, in addition to the disk-to-halo mass ratio. In MOND, there is only one free parameter: the mass-to-light ratio $M/L$. For MOND to be viable, the best fit $M/L$ has to be consistent with stellar population synthesis models. Indeed, MOND fits yield mass-to-light ratios $M/L$ that agree with the expectations and also recover the expected relations between color and $M/L$ [159]. Claimed failures of MOND fits often derive from inaccurate measures of the galaxy distance and/or disk inclination or inappropriate bulge-disk decompositions [548].
By requiring that the velocity of the baryonic matter is directly set by its distribution, MOND makes a clear prediction on the shape of the rotation curves of disk galaxies: HSB galaxies are expected to have steeply rising rotation curves that flatten at small radii, whereas LSB galaxies are expected to have slowly rising rotation curves that converge to the asymptotic constant velocity at large radii. This prediction by Milgrom in 1983 [34] appeared well before LSB galaxies were known to be a substantial fraction of the galaxy population and well before their systematic observations confirmed the MOND prediction [549].

An additional consequence of the absence of dark matter is the expected one-to-one correspondence between the irregularities in the surface brightness distribution and the features of the rotation curves. This correspondence is widely observed in disk galaxies [550] and is known as the Renzo’s rule [551]. These features in HSB disk galaxies are also naturally explained in the standard framework, because the stars dominate the gravitational potential at small radii, according to the maximum-disk hypothesis. However, in LSB disk galaxies this hypothesis does not hold and the observed correspondence remains unexplained. MOND naturally explains these observations in both HSB and LSB disk galaxies [552].

The existence of LSB disk galaxies and their properties have, in MOND, a relevant role that needs to be emphasized. In the 1970s, the idea of disk galaxies embedded in a halo of dark matter was introduced to describe the observations of flat rotation curves inferred from the spectroscopy of neutral hydrogen [10]. The community easily accepted this idea because it had recently become clear that bare cold self-gravitating disks, as spirals appear to be, are violently unstable and rapidly generate a bar [553, 554], unless the disk is embedded in a massive halo [11]: a single massive halo, that was initially thought to be composed of too faint stars to be observed, made the disk dynamically stable and explained the flat rotation curves.

This consideration has an important consequence on the dynamics and morphology of LSB disk galaxies: the rotation curves of LSB disks indicate that these galaxies are dynamically dominated by dark matter in the standard model even at small radii, where the baryonic matter dominates in HSB disk galaxies [549]. Consequently, the disk of LSB galaxies is over-stabilized [555] and its mass should be smaller than the limit required to generate spiral modes, unlike what happens in HSB disks [556]; nevertheless, spiral arms and bars are observed in LSB disk galaxies [557–560]. As a consequence, explaining the amplitude of the rotation curves and the presence of spiral arms and bars in the standard model requires mass-to-light ratios that are inconsistently larger than the expectations from stellar population synthesis models [561–563]. This tension does not appear in MOND, where there is no dark matter and the spiral modes are naturally generated as in the HSB disk galaxies [549, 564, 565].

At the same time when the introduction of the massive halos was thought to be necessary, Freeman noted that the mean central surface brightness of disk galaxies is $\mu_F = 21.65\ 	ext{mag arcsec}^{-2}$ in the $B$-band with a little scatter of 0.30 mag arcsec$^{-2}$: this relation became known as the Freeman law [9]. However, the Freeman law was the result of an observational bias, as it became clear a few years later [566]. In fact, disks appear in a wide range of central surface brightness [567]. In addition, LSB galaxies are actually more than half of the total galaxy population and the number of galaxies with central surface brightness brighter than $\mu_F = 21.65\ 	ext{mag arcsec}^{-2}$ drops substantially faster than for a normal distribution [568]: the Freeman law actually is a Freeman limit.

This observational result obviously makes us wonder what sets the Freeman limit. The Freeman limit is a property of the baryonic matter, whereas dark matter dominates the dynamics of disk galaxies in the standard model. Therefore, the Freeman limit must originate by the interplay between dark matter and baryonic matter. The CDM model has not yet an obvious solution for deriving this limit [569].

In MOND, the origin of the Freeman limit simply derives from gravitational dynamics. Without dark matter, disks are unstable in Newtonian gravity. However, if MOND is the theory of gravity, the disk becomes stable against the development of bars [564, 565, 570–574]. It follows that disks without any dark matter are gravitationally stable only if they are in the MOND regime $a_N < a_0$. If mass is proportional to luminosity, we have the acceleration $a_N \sim 2\pi G\Sigma$, in Newtonian gravity, where $\Sigma$ is the surface mass density. The limit $a_N < a_0$ becomes $a_N \sim 2\pi G\Sigma < a_0$; in other words, gravitationally stable disks must have $\Sigma < a_0/(2\pi G)$; this limit $a_0/(2\pi G) \approx 143\ M_\odot\ 	ext{pc}^{-2}$ neatly returns the Freeman limit $\mu_F = 21.65\ 	ext{mag arcsec}^{-2}$ [34, 564]. This result demonstrates that the acceleration scale $a_0$ enters in another context which is completely different from the Tully-Fisher relation shown above.

In MOND, from the distribution of the baryonic matter, we can also predict the profile of the stellar vertical velocity dispersion, namely the velocity dispersion perpendicular to plane of the disk, as a function of the radial coordinate in disk plane [34]. Testing this prediction is particularly challenging, because we have to measure both the vertical velocity dispersion and the rotation curve and, ideally, we would need face-on galaxies for the former and edge-on galaxies for the latter. The task was performed by the DiskMass Survey collaboration [575–577], who found that the kinematic properties, when interpreted with Newtonian dynamics, require a stellar population with a mass-to-light ratio $\sim 0.3\ M_\odot/L_\odot$ in the $K$-band, a factor two smaller than expected from stellar population synthesis models [578, 579]. In other words, the vertical velocity dispersion is too small compared to the rotation curve and the disks are too cold for their stellar mass.

When interpreted in MOND, the mass-to-light ratio in the $K$-band is $0.55 \pm 0.15\ M_\odot/L_\odot$, in agreement with stellar population synthesis models, but the disks are a factor of two thinner than expected from the observations of edge-on galaxies [580], or, alternatively, the vertical velocity dispersion is overpredicted by $\sim 30\%$ [581]. However, the shape of the vertical velocity dispersion profile in the model is consistent with the observed shape over the entire radial range; this concordance suggests that the measurements might suffer from a systematic bias [582]. Indeed, the estimate of the velocity dispersion could be dominated by the younger, and dynamically colder, stellar population, whereas the estimate of the disk thickness in edge-on galaxies could
be dominated by the older, and dynamically hotter, stellar population [583].

To date, the conundrum remains unsolved, although a number of additional observations might support the MOND framework: superthin disk galaxies, that appear largely self-gravitating [584–586], would be naturally stabilized by the enhanced MOND acceleration [549]; similarly, the large vertical velocity dispersion of the gas in the outer region of some disk galaxies, when interpreted in the standard model, would require an embedding disk of dark matter or highly flattened halos that are difficult to reconcile with the conventional framework [587]. A similar dark matter disk, or alternatively an extended dark matter core with a large core radius of 10 kpc [588], might be required for the Milky Way, according to the kinematics of red clump stars from the RAVE survey [589].

Overall, the solidity of the existence of the acceleration scale in the data has been increasing for the last decades (see however [590]). Its existence requires a natural explanation, irrespective of the correctness of MOND. In addition to the Tully-Fisher relation and the Freeman limit, appears in the mass-discrepancy relation and in the radial acceleration relation.

In the standard model, we estimate the dynamical mass with \( M_d \sim \frac{rv^2}{G} \) from the luminosity we can estimate the baryonic mass \( M_b \). In MOND, Newtonian gravity holds when the Newtonian gravitational acceleration generated by the baryonic mass is larger than \( a_0 \), therefore we expect \( M_d/M_b \sim 1 \) in this regime and \( M_d/M_b \) increasingly larger than 1 at increasingly smaller accelerations (Figure 3). This mass-discrepancy relation [146, 155], predicted in 1983 [34], implies increasingly large mass-to-light ratios for galaxies with increasingly fainter surface brightness, as it was confirmed years later with the observations of dwarf spheroidals and LSB disks [549, 591].

Similarly, we can see the appearance of \( a_0 \) in the RAR, between the observed centripetal acceleration derived from the rotation curve of disk galaxies and the Newtonian gravitational acceleration generated by the baryonic mass distribution (Figure 4): the deviation from the one-to-one relation appears at Newtonian accelerations smaller than \( a_0 \) and the relation is described by the interpolation function of Eq. (19) [127]. Although the agreement shows no scatter and no dependence of the residuals on the galaxy properties [128, 129], as expected for a relation driven by gravity alone, additional investigations are required to confirm these results, because some dependence of the residuals on the galaxy properties appears to be present in other galaxy samples different from SPARC [151, 152].

2. Dwarf galaxies

In MOND, the cusp/core problem is absent by definition, because there is no dark matter. MOND can explain the velocity dispersion profiles of dwarf spheroidals with mass-to-light ratios \( M/L \) consistent with stellar population synthesis models for the classical dwarfs, except Carina [592, 593]. Carina is the closest dwarf spheroidal to the Milky Way and detailed N-body simulations in MOND show that tidal forces and the external field effect, an effect that lacks in Newtonian gravity, can only partly alleviate the tension [594]: the best-fit \( M/L \) in the V-band required to match the observed velocity dispersion profile is \( M/L \sim 5.3 - 5.7 \, M_\odot/L_\odot \), a value \( \sim 10\% \) greater than the upper limit for the old stellar population of Carina \( \sim 5 \, M_\odot/L_\odot \) [595]. However, there might still be the possibility to alleviate this tension both observationally and theoretically: more accurate measurements of the proper motion of the dwarf spheroidals and larger samples of stars with accurate photometry can provide a better understanding of the physical properties of the dwarf spheroidals; similarly, the modelling of these systems can be improved by considering a triaxial three-dimensional stellar distribution, more sophisticated treatments of stellar binaries and even more accurate stellar population synthesis models.

If MOND is the correct theory of gravity, interpreting the dynamics of dwarfs with Newtonian gravity implies dark matter distributions that might differ from the standard model expectations. The external field effect in MOND mostly generates (1) an offset of the dark matter distribution from the stellar distribution of the order of the half-mass radius of the dwarf, and (2) possible large concentrations of dark matter along the direction to the Milky Way, especially in Fornax and Sculptor, which have no reason to appear in the standard model [596]. In principle, accurate measurements of the surface brightness distribution and measurements of the proper motions of the dwarf stars with future astrometric missions [597, 598] could test these predictions and eventually falsify MOND.

In the standard model, the problem of the plane of satellites would be most easily solved if these satellites were collapsed tidal debris formed during galaxy interactions [357–359]. Unfortunately, the dwarf galaxies appear to be dark matter dominated, whereas these Tidal Dwarf Galaxies (TDGs) are expected to be dark-matter free. In fact, the dark matter halo of the parent galaxy is supported by the velocity dispersion of the dark matter particles and it is dynamically hot, unlike the dynamically cold baryonic galactic disk supported by rotation. Therefore the dark matter halo does not participate in the formation of the TDGs orbiting in a dynamically cold plane: this tidal tail can only originate from the galactic disk. In MOND, this mechanism would work without the complication of the existence of the dark matter, as shown by N-body simulations [370, 599]; the mechanism is actually favoured by the enhanced self-gravity of the baryons.

Observations of recently formed TDGs are still unable to confirm whether these systems have mass discrepancies consistent with MOND or rather no mass discrepancies and are thus consistent with dark-matter free TDGs in Newtonian gravity [600–602]. However, this latter solution would suggest the existence of two dwarf populations in the standard model: dark-matter dominated dwarfs, presumably formed at early time, and more recently formed dark-matter free TDGs. However, the existence
of these two populations appears to be inconsistent with the samples currently available [603].

In conclusion, MOND predicted, rather than solved, most of the dynamical properties of disk and dwarf galaxies and might present a viable mechanism for the formation of the plane of satellites. MOND appears mostly consistent with the observed kinematics of dwarf spheroidals and makes additional predictions on their density and velocity fields that can further falsify the theory. Additional theoretical and observational investigations are required to clarify these open issues. In addition, it still remains to be seen how MOND can be embedded in a more extended theory capable of providing a cosmological model and describing the formation and evolution of the large-scale structure.

### B. MOdified Gravity (MOG)

Scalar-Vector-Tensor theory of gravity, also renamed MOdified Gravity (MOG), is built to describe the observational effects related to dark matter. In this model, scalar, tensor and massive vector fields are added to the standard Hilbert-Einstein action [36]. Thus, the total action reads:

\[
A = A_G + A_\phi + A_S + A_M, \tag{20}
\]

where the first term is

\[
A_G = \frac{1}{16\pi} \int \frac{1}{G} (R + 2\Lambda) \sqrt{-g} \, d^4x, \tag{21}
\]

with \( \Lambda \) the cosmological constant. The second term is the action related to the massive vector field:

\[
A_\phi = -\int \left[ \frac{1}{4} B^\sigma \tau B_\sigma \tau - \frac{1}{2} \mu^2 \phi_\beta \phi^\beta \right] \sqrt{-g} \, d^4x, \tag{22}
\]

where \( \phi^\beta \) is the vector field, and \( B_\sigma \tau = \partial_\sigma \phi_\tau - \partial_\tau \phi_\sigma \). The third term represents the action for the scalar fields \( G \) and \( \mu \):

\[
A_S = -\int \frac{1}{G} \left[ \frac{1}{2} g^{\tau \beta} \left( \nabla_\tau G \nabla_\beta G \right) + \nabla_\tau \mu \nabla_\beta \mu \right] + \frac{V_G(G)}{G^2} + \frac{V_\mu(\mu)}{\mu^2} \sqrt{-g} \, d^4x, \tag{23}
\]

where \( \nabla_\tau \) is the covariant derivative with respect to the metric \( g_{\tau \beta} \), and \( V_G(G) \) and \( V_\mu(\mu) \) are self-interaction potentials of the \( \mu \) and \( G \) fields. Finally, the fourth term describes the matter action.

In the weak field limit, the gravitational potential shows a Yukawa-like correction to the Newtonian term. Such a correction generates a repulsive gravitational force that cancels out gravity at short range (galactic, sub-galactic, or smaller scales) and stabilizes the system. On the other hand, at larger scales, the repulsive term becomes weaker and one obtains the Newtonian gravity with a larger gravitational constant [604]. The gravitational potential in MOG is given by [604]:

\[
\Phi_{\text{eff}}(\vec{x}) = -G_N \int \frac{\rho(\vec{x'})}{|\vec{x} - \vec{x'}|} \left[ 1 + \alpha - \alpha e^{-\mu |\vec{x} - \vec{x'}|} \right] d^3\vec{x'}, \tag{24}
\]

where \( \mu \) is the inverse of the characteristic length of the gravitational force that acts at a certain scale, and it depends on the total mass of the self-gravitating system [604], and \( \alpha = (G_\infty - G_N)/G_N \) accounts for the modification of the gravitational constant, where \( G_\infty \) is the effective gravitational constant at infinite distance from the self-gravitating system and \( G_N \) is the standard gravitational constant [605]. The theory has been tested on a wide range of observational scales. At extragalactic and cosmological scales, MOG describes a wide variety of phenomena ranging from the gravitational bending of light [606, 607], to the X-ray and Sunyaev-Zel’dovich emissions of galaxy clusters [606, 608–611], and the accelerated expansion of space-time [612, 613].

#### 1. Solutions to the observational challenges

In the weak-field limit, the gravitational potential in Eq. (24) has been used to fit the observed rotation curves of both LSB and HSB galaxies from the THINGS catalogue, achieving an excellent agreement between the model and the observations, as shown in Figure 25. This comparison implies a single set of the parameters \( \alpha \) and \( \mu \) of the gravitational potential: the gravitational strength is \( \alpha = 8.89 \pm 0.34 \), and the characteristic scale is \( \mu = 0.042 \pm 0.004 \text{ kpc}^{-1} \) [604]. MOG also yields a good fit to the rotation curve of the Milky Way with total mass \( M \approx 5 \times 10^{10} \text{ M}_\odot \) [614]. Finally, no significant difference with Newtonian dynamics is found by studying the stability of the disk in MOG [615]; this result implies that the formation and evolution of self-gravitating systems is not altered.
Rotation curves of less massive systems also are well accommodated in the MOG theory. The rotation curves of the dwarf galaxies in the LITTLE THINGS catalogue constrain the parameters \((\alpha, \mu)\) with a relative error of 10\% [616]. Furthermore, MOG is consistent with the RAR observed in the SPARC galaxies [617]. Finally, the observed velocity dispersion of the ultra-diffuse galaxy NGC1052-DF2 \(\sigma = 3.2^{+5.0}_{-3.2}\) km/s [222] is consistent with the value \(\sigma = 3.9\) km/s estimated in MOG [618].

However, other analyses encounter some difficulties. A recent analysis shows that the theory fails to reproduce the observed rotation curve of the Milky Way at radii < 20 kpc [619], at odds with previous results showing the capability of MOG to fit the rotation curve of the Milky Way [614]. This inconsistency remains unexplained. Furthermore, a detailed analysis of the velocity dispersion profile of the Milky Way’s dwarf spheroidals leads to mass-to-light ratios in the V-band in the range \(M_*/L = 5.2^{+1.0}_{-0.9}\) M\(_{\odot}/L\odot\) for Fornax to \(M_*/L = 152.3^{+78.5}_{-62.2}\) M\(_{\odot}/L\odot\) for Draco. This wide interval of values is inconsistent with the stellar population synthesis models. Additionally, a single set of the parameters \(\alpha\) and \(\mu\) for all the galaxies in the sample cannot be found. These results are summarized in Figure 26; this figure also shows a difference of a few order of magnitudes between the parameters constrained by the dSphs galaxies (colored data points) and by the more massive galaxies in the THINGS catalogue (black lines) [620].

These difficulties lead to argue that MOG may fail at reproducing the dynamics of pressure-supported systems and, thus, at solving the CCP problem. Along the same lines, the results obtained by comparing the predicted velocity dispersion profile of
Figure 26: The top and bottom panels show the best-fit $\alpha$ and $\mu$ parameters of the MOG gravitational potential in Eq. (24) with their uncertainties as obtained from the analysis of dSphs. The solid lines show the values obtained by fitting the rotation curves of more massive galaxies [604]. For each galaxy, three assumptions on the mass-to-light ratio are used, i.e. $M/L_V = [0.5, 2, 5]$, corresponding to blue triangles, red squares, and green circles respectively. The figure is reproduced from [620].

Antlia II with the observed one [189] show that MOG cannot explain the existence of such a large core ($\sim 2.8$ kpc) without introducing a strongly negative anisotropy parameter $\beta \approx -18$, which means that stars have a tangential velocity much larger than the radial counterpart [621].
In addition, high-resolution N-body simulations show substantial differences between MOG and Newtonian gravity with dark matter. In MOG, the growth rate of the bar in disk galaxies is slower, and the final size of the bar is almost an order of magnitude smaller [622]. Moreover, at cosmological scales, where MOG is supposed to behave like the ΛCDM model, the growth rate of the perturbations is reduced, and the value of the normalization of the power spectrum σ8 is increased to 1.44 [623], which is inconsistent with σ8 = 0.802 ± 0.018 obtained by the Planck satellite [63]. Finally, due to the lack of cosmological N-body simulations in the context of the MOG theory, the MSP and PSP can not be currently addressed.

C. f(R)-gravity

Extensions of General Relativity are obtained by including higher-order curvature invariants in the Lagrangian, such as \( R^2 \), \( R_{\mu\nu} R^{\mu\nu} \), \( R^{\mu\nu\alpha\beta} R_{\mu\nu\alpha\beta} \), \( R \square R \), or \( R \square^2 R \), and minimally or non-minimally coupled terms between scalar fields and geometry, such as \( \phi^2 R \) [624–626]. These theories can be roughly classified as Scalar-Tensor Theories if the geometry is non-minimally coupled to some scalar field, and Higher-Order Theories if the action contains derivatives of the metric components of order higher than two. Combinations of both types of theories give rise to higher-order/scalar-tensor theories of gravity.

Among them, \( f(R) \)-theories extend General Relativity by replacing the Einstein-Hilbert Lagrangian, which is linear in the Ricci scalar, with a more general function \( f(R) \) of the curvature \( R \):

\[
S = \int d^4x \sqrt{-g} \left[ f(R) + \frac{8\pi G}{c^4} L_m \right],
\]

where \( L_m \) is the matter Lagrangian. The field equations thus read

\[
f'(R) R_{\mu\nu} - \frac{1}{2} f(R) g_{\mu\nu} - f'(R)_{,\mu\nu} + g_{\mu\nu} \square f'(R) = \frac{8\pi G}{c^4} T_{\mu\nu},
\]

where \( T_{\mu\nu} \) is the energy momentum tensor of matter. Back in the 1980s, A. Starobinsky made the first attempt to describe the acceleration of the Universe by extending General Relativity along this line [627]. Afterwards, many other \( f(R) \) models have been considered and tested to give an alternative explanation to the cosmological constant [628–637]. However, there is no statistical evidence favouring any of these models over ΛCDM [638, 639].

Usually, \( f(R) \) gravity is invoked to replace either the Einstein-Hilbert Lagrangian with a cosmological constant or dark energy models. However, \( f(R) \) gravity can also partly remove the need for dark matter on small scales. In 1977, Stelle already noted that, in the weak-field limit, \( R^2 \)-gravity gives rise to Yukawa-like corrections to the Newtonian gravitational potential [640]. Under the hypothesis that the \( f(R) \) Lagrangian is expandable in Taylor’s series [37]:

\[
f(R) = \sum_n \frac{f_n(R_0)}{n!} (R - R_0)^n \approx f_0 + f_1' R + f_2' R^2 + \ldots,
\]

the gravitational potential can be recast as follows:

\[
\Phi(r) = -\frac{GM(r)}{(1 + \delta)^r} \left( 1 + \delta e^{-\frac{r}{\Lambda}} \right),
\]

where the parameter \( L \) represents the effective scale length above which the corrections to the gravitational potential are relevant, and the parameter \( \delta \) is related to the strength of the gravitational force [37]. These parameters are related to the coefficients of Taylor’s series as: \( \delta = f_1' - 1 \) and \( L = -6f_2'' / f_1'^{3/2} \). These corrections may affect the astrophysical scales of galaxies and galaxy clusters while being negligible at the Solar systems scale [37, 641–653]. Therefore, geometric modifications of this type may serve also to fully or partially replace dark matter in the energy-density content of the Universe.

1. Solutions to the observational challenges

Back in the 1980s, R.H. Sanders found the Yukawa-correction term to be able to describe the rotation curves of several spiral galaxies [518]. More recently, it was shown that \( R^2 \)-gravity models are able to describe the kinematics of the stars in spiral galaxies without resorting to dark matter [641–644, 654]. Hybrid models combining \( f(R) \)-gravity and dark matter have also been proposed: studies of galactic rotation curves find an 8σ Bayesian evidence favouring \( f(R) \)-gravity plus a NFW dark matter halo over the CDM model [655]. In addition, the rotation curves expected in \( f(R) \)-gravity combined with cuspy NFW density profiles are favored over the rotation curves expected in CDM profiles with a core, offering a potential solution to the CCP [645].

A recent study of the chameleon-\( f(R) \) gravity [e.g. 656], where higher order curvature terms can be recast as a scalar field strongly coupled to matter, tightly constrains the parameters of the model with HI/He rotation curves of the disk galaxies in
the SPARC sample. Specifically, the chameleon-$f(R)$ gravity with a cuspy NFW dark matter halo fits well the galaxy rotation curves; the left panel of Figure 27 shows the example of NGC 3741. Nevertheless, in contrast with previous results [655], there is no statistical evidence favouring this model over Newtonian gravity with a NFW dark matter halo with a core [657], as illustrated by the Bayesian Information Criteria estimator shown in the right panel of Figure 27.

Whereas rotation curves are well accommodated in $f(R)$-gravity, there is a lack of studies of pressure supported systems such as dwarf galaxies. Nevertheless, due to the similarities between the gravitational potential in $f(R)$-gravity and in MOG (see Eq. (24) in Section V B), one may expect to encounter difficulties similar to MOG to explain the dynamics of dwarf galaxies, and to provide a solution to the CCP. It is worth mentioning that the possibility that these systems might not be in dynamical equilibrium in the context of these theories may invalidate the constraints [658–660].

Finally, N-body simulations of $f(R)$-gravity alone (see, for example, [661–663]), and of $f(R)$-gravity with dark matter, such as massive neutrinos or WDM particles, indicate the existence of a degeneracy between these models and the standard ΛCDM paradigm [664–666]. For instance, N-body simulations of $f(R)$ gravity with $f_{R,0} = -1 \times 10^{-4}$ plus massive neutrinos with a total neutrino mass of $\Sigma m_\nu = 0.4 \text{ eV}$ show a matter power spectrum and a halo mass function consistent with the predictions of ΛCDM at 10% and 20% accuracy levels, respectively [664].

VI. SUMMARY AND DISCUSSION

The cosmological parameters of the ΛCDM model are measured at an accuracy of $\sim 1\%$ or smaller [62, 63], effectively confirming the capability of the model to describe the cosmological evolution of the Universe. In this model, dark matter is $\sim 85\%$ of the total matter density of the Universe [63], but its fundamental nature is still unknown [1]. WIMPs are the most promising candidate for CDM [22, 23]. WIMPs are massive and weakly interacting particles, expected in supersymmetric theories, which decoupled from the primordial plasma when they were non-relativistic. Collider, direct, and indirect searches for such particles are ongoing [24, 108–112]. Nevertheless, LHC has not provided any evidence of supersymmetry so far [113], and the claimed DAMA/LIBRA/CoGeNT annual modulation is still under debate [25–27] (for further details see Section II).

The lack of a direct detection of the dark matter particles is not the only issue encountered by the CDM model: the issues on galactic scales discussed in Section III are difficult to accommodate in the context of the CDM model [39–43]. Supernovae feedback and dynamical friction from baryonic clumps may help to explain the transition from the cuspy dark matter profiles, predicted by CDM N-body simulations, to the cores suggested by the observed kinematic properties of LSB, dwarf and ultra-faint
galaxies [239]. In principle, these baryonic mechanisms may also help to reduce the overabundance of subhalos and describe the observed radial acceleration relation of disk galaxies. However, the effectiveness of baryonic feedback is still under debate, as discussed in Section III.

At this point, a question arises: are the lack of detection of a CDM particle and the observational issues at galactic scales indicating a breakdown of the model? The current state of the debate is inconclusive and puzzling. A change of the dark matter paradigm or modified gravity models can help to solve some of the issues but they do not offer a definitive answer to the question. A visual overview of whether the models discussed in Sections IV and V either solve or do not display the challenges of the CDM model is given in Table I; for the sake of completeness, Table I also displays the ability of each model to explain the large-scale structure of the Universe, successfully described by the \( \Lambda \)CDM model.

| Model | Rotation curves and scaling relations | Cusp/core problem | Missing satellites problem | Too-big-to-fail problem | Planes of satellites problem | Large scale structure and cosmological scales |
|-------|-------------------------------------|-------------------|--------------------------|------------------------|-----------------------------|----------------------------------|
| Warm DM | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ |
| Self-interacting DM | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ |
| QCD axions | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fuzzy DM | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| MOND | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| MOG | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |
| \( f(R) \)-gravity | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

✓ = Solved or Not present ✗ = Not solved = Under investigation

WDM, SIDM, ALPs, and FDM, discussed in Section IV, allow the existence of a dark matter core in dwarf galaxies, hence, providing, in principle, a solution to the CCP. Moreover, WDM and FDM show a cut-off in the matter power spectrum that suppresses the formation of halos below a given mass threshold [396, 476]. However, WDM does not alleviate any of the CDM issues at galactic scale when the constraints from the Lyman-\( \alpha \) forest or the gravitational lensed quasars are taken into account. The cores in the WDM halos are indeed too small to solve the CCP, although they might contribute to fix the TBTF problem. It thus remains unclear whether WDM models may represent a viable solution [373]. SIDM solves the CCP, but its ability to solve the MSP and PSP requires further investigations that take into account the baryonic feedback. QCD axions are highly motivated dark matter candidates from the perspectives of both particle physics and cosmology. However, whether they can solve the small-scale issues of CDM has not been properly investigated yet, because the role of the quantum nature of thermalizing QCD axions on cosmic small scales demands a more rigorous theoretical framework. Finally, FDM also encounters its own difficulties. The boson mass, \( \sim 10^{-22} \) eV, required to explain the dynamics of the dwarf galaxies and to solve the CCP and the MSP is almost two orders of magnitude lower than the boson mass, \( \sim 7 \times 10^{-21} \) eV, needed to account for the Lyman-\( \alpha \) data [512]. Nevertheless, the debate is far from being settled [494]. Uncertainties on the thermal histories and the underlying reionisation model may invalidate these constraints [516, 667].

While the CCP and MSP may usually be solved in paradigms beyond the standard model of particle physics, both are hardly addressed in modified theories of gravity. Both MOG and \( f(R) \)-gravity are able to reproduce the rotation curve of the giant spiral galaxies [604, 614, 641–644, 654], whereas a deeper investigation of dwarf and ultra-faint galaxies, which appear to be dark-matter dominated and show the presence of wide cores, is still missing. Recent results obtained in the context of MOG have shown that, although this modified gravity model has been built to specifically replace dark matter, it is unable to self-consistently reproduce the dynamics of the dwarf galaxies orbiting the Milky Way [616, 621] (Section V B). While similar studies for \( f(R) \)-gravity are missing, one may expect similar results on the basis of the theoretical analogies of the weak field limits of MOG and \( f(R) \). Nevertheless, one must be careful when comparing predictions from modified gravity models to the observational data. In fact, the assumption that all real systems are in dynamical equilibrium may not hold and, if so, their amount of dark matter would be substantially overestimated [658–660].

MOND is fundamentally different from the other suggested solutions that attempt to remove the requirement of dark matter (Section V A). MOND predicted many of the challenges of the conventional dark matter paradigm related to disk and dwarf galaxies many years before these challenges were actually observed. MOND might also have a natural solution for the formation of the plane of satellites, although it is yet to be demonstrated that this is indeed the case. In addition, MOND remains a phenomenological model whose extension providing a cosmological model and describing the formation and evolution of the
large-scale structure is not yet available.

To sum up, no unambiguous signature of dark matter has been found yet. The standard CDM paradigm encounters several issues at galactic scales, and its capability to solve them is still unclear, as discussed in Section III. Ideally, the next generation facilities, such as the Maunakea Spectroscopic Explorer (MSE, [668]) or the Large Synoptic Survey Telescope (LSST, [669]), will achieve the sensitivity needed to discover other ultra-faint and ultra-diffuse galaxies, whose observed properties can strongly constrain the theoretical models. In fact, understanding the inner structure of these galaxies has a fundamental role in our comprehension of the nature of dark matter. Additional observational constraints may come from the next generation astrometric missions, such as the proposed Theia satellite [597, 598]. Theia is expected to measure the proper motion of stars in the Milky Way dwarf satellites that, together with high precision measurements of the position of their stars, would shed light on the dynamics of these galaxies. We can thus either constrain the theory of gravity or the nature of the dark matter, if the presence of either cores or cusps in their dark matter density profile is definitely confirmed. This evidence will complement the constraints from other astrophysical and cosmological probes and from indirect or direct searches of dark matter.

Author Contributions: All authors contributed equally to this review article.

Funding: This research received no external funding.

Acknowledgments: We sincerely thank three referees whose suggestions and corrections improved our review. Possible residual inaccuracies remain our full responsibility. IDM and SSC are supported by the grant “The Milky Way and Dwarf Weights with Space Scales” funded by University of Torino and Compagnia di S. Paolo (UniTO-CSP). We also thank Compagnia di San Paolo (CSP) for funding the fellowship of VC and the Italian Ministry of Education, University and Research (MIUR) under the Departments of Excellence grant L.232/2016 for funding the fellowship of AG. We also acknowledge partial support from the INFN grant InDark. This research has made use of NASA’s Astrophysics Data System Bibliographic Services.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:
[1] Bertone, G.; Hooper, D. History of dark matter. Rev. Mod. Phys. 2018, 90, 045002, doi:10.1103/RevModPhys.90.045002.
[2] Oort, J.H. The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems. Bull. Astron. Inst. Neth. 1932, 6, 249.
[3] Kuijken, K.; Gilmore, G. The mass distribution in the galactic disc—II. Determination of the surface mass density of the related disc near the Sun. Mon. Not. R. Astron. Soc. 1989, 239, 605–649, doi:10.1093/mnras/239.2.605.
[4] Kuijken, K.; Gilmore, G. The mass distribution in the galactic disc—III. The local volume mass density. Mon. Not. R. Astron. Soc. 1989, 239, 651–664, doi:10.1093/mnras/239.2.651.
[5] Holmberg, J.; Flynn, C. The local surface density of disc matter mapped by Hipparcos. Mon. Not. R. Astron. Soc. 2004, 352, 440-446, doi:10.1111/j.1365-2966.2004.07931.x.
[6] Read, J.I. The local dark matter density. J. Phys. G Nucl. Phys. 2014, 41, 063101, doi:10.1088/0954-3899/41/6/063101.
[7] Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. Helv. Phys. Acta 1933, 6, 110–127.
[8] Freeman, K.C. On the Disks of Spiral and S0 Galaxies. Astrophys. J. 1970, 160, 811, doi:10.1086/150474.
[9] Roberts, M.S.; Rots, A.H. Comparison of Rotation Curves of Different Galaxy Types. Astron. Astrophys. 1973, 26, 483–485.
[10] Lasserre, T.; Afonso, C.; Albert, J.N.; Andersen, J.; Ansari, R.; Aubourg, É.; Bareyre, P.; Bauer, F.; Beaulieu, J.P.; Blanc, G.; et al. Not enough stellar mass Machos in the Galactic halo. Astron. Astrophys. 2000, 355, L39–L42.
[11] Carr, B.J.; Hawking, S.W. Black Holes in the Early Universe. Mon. Not. R. Astron. Soc. 1974, 168, 399–415, doi:10.1093/mnras/168.2.399.
[12] Motohashi, H.; Hu, W. Primordial black holes and slow-roll violation. Phys. Rev. D 2017, 96, doi:10.1103/physrevd.96.063503.
[13] Passaglia, S.; Hu, W.; Motohashi, H. Primordial black holes and local non-Gaussianity in canonical inflation. Phys. Rev. D 2019, 99, doi:10.1103/physrevd.99.043536.
[14] Wang, S.; Wang, Y.F.; Huang, Q.G.; Li, T.G.F. Constraints on the Primordial Black Hole Abundance from the First Advanced LIGO Observation Run Using the Stochastic Gravitational-Wave Background. Phys. Rev. Lett. 2018, 120, 191102,
[520] McGaugh, S.S. Milky Way Mass Models and MOND. Astrophys. J. 2008, 683, 137–148, doi:10.1086/589148.

[521] Milgrom, M. Dynamics with a Nonstandard Inertia-Acceleration Relation: An Alternative to Dark Matter in Galactic Systems. Ann. Phys. 1994, 229, 384–415, doi:10.1006/aphy.1994.1012.

[522] Milgrom, M. The modified dynamics as a vacuum effect. Phys. Lett. A 1999, 253, 273–279, doi:10.1016/S0375-9601(99)00077-8.

[523] Milgrom, M. MOND as modified inertia. EAS Publ. Ser. 2006, 20, 217–224, doi:10.1051/eas:2006074.

[524] Bekenstein, J.D. Relativistic gravitation theory for the modified Newtonian dynamics paradigm. Phys. Rev. D 2004, 70, 083509, doi:10.1103/PhysRevD.70.083509.

[525] Milgrom, M. Bimetric MOND gravity. Phys. Rev. D 2009, 80, 123536, doi:10.1103/PhysRevD.80.123536.

[526] Cardone, V.F.; Radicella, N. Can MONDian vector theories explain the cosmic speed up? Phys. Rev. D 2009, 80, 063515, doi:10.1103/PhysRevD.80.063515.

[527] Abbott, B.P.; Abbott, R.; Abbott, T.D.; Acernese, F.; Ackley, K.; Adams, C.; Adams, T.; Addesso, P.; Adhikari, R.X.; Adya, V.B.; et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. Phys. Rev. Lett. 2017, 119, 161101.

[528] Goldstein, A.; Veres, P.; Burns, E.; Briggs, M.S.; Hamburg, R.; Kocevski, D.; Wilson-Hodge, C.A.; Feree, R.D.; Poolakkil, S.; Roberts, O.J.; et al. An Ordinary Short Gamma-Ray Burst with Extraordinary Implications: Fermi -GBM Detection of GRB 170817A. Astrophys. J. 2017, 848, L14, doi:10.3847/2041-8213/aaa8f4.

[529] Savchenko, V.; Ferrigno, C.; Kuulkers, E.; Bazzano, A.; Bozzo, E.; Brandt, S.; Chenevez, J.; Courvoisier, T.J.L.; Diehl, R.; Domingo, A.; et al. INTEGRAL Detection of the First Prompt Gamma-Ray Signal Coincident with the Gravitational-wave Event GW170817. Astrophys. J. 2017, 848, L15, doi:10.3847/2041-8213/aa8f94.

[530] Gong, Y.; Hou, S.; Liang, D.; Papantonopoulos, E. Gravitational waves in Einstein-æther and generalized TeVeS theory after GW170817. Phys. Rev. D 2018, 97, 084040, doi:10.1103/PhysRevD.97.084040.

[531] Skordis, C.; Zło´snik, T. Gravitational alternatives to dark matter with tensor mode speed equaling the speed of light. Phys. Rev. D 2019, 100, 104013, doi:10.1103/PhysRevD.100.104013.

[532] Angus, G.W. Is an 11eV sterile neutrino consistent with clusters, the cosmic microwave background and modified Newtonian dynamics? Mon. Not. R. Astron. Soc. 2009, 394, 527–532, doi:10.1111/j.1365-2966.2008.14341.x.

[533] Aguilar-Arevalo, A.; Brown, B.; Bugel, L.; Cheng, G.; Conrad, J.; Cooper, R.; Dharmapalan, R.; Diaz, A.; Djuricic, Z.; Finley, D.; et al. Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment. Phys. Rev. Lett. 2018, 121, doi:10.1103/PhysRevLett.121.221801.

[534] Diaffiero, A.; Angus, G.W. The Acceleration Scale. Modified Newtonian Dynamics, and Sterile Neutrinos. arXiv 2012, arXiv:1206.6231.

[535] Angus, G.W.; Diaffiero, A. The abundance of galaxy clusters in modified Newtonian dynamics: cosmological simulations with massive neutrinos. Mon. Not. R. Astron. Soc. 2011, 417, 941–949, doi:10.1111/j.1365-2966.2011.19321.x.

[536] Angus, G.W.; Diaffiero, A.; Famaey, B.; van der Heyden, K.J. Cosmological simulations in MOND: The cluster scale halo mass function with light sterile neutrinos. Mon. Not. R. Astron. Soc. 2013, 436, 202–211, doi:10.1093/mnras/stt1564.

[537] Angus, G.W.; Diaffiero, A.; Famaey, B.; Gentile, G.; van der Heyden, K.J. Modified Baryonic Dynamics: Two-component cosmological simulations with light sterile neutrinos. J. Cosmol. Astropart. Phys. 2014, 2014, 079, doi:10.1088/1475-7516/2014/10/079.

[538] Berezhiani, L.; Khoury, J. Theory of dark matter superfluidity. Phys. Rev. D 2015, 92, 103510, doi:10.1103/PhysRevD.92.103510.

[539] Blanchet, L. Gravitational polarization and the phenomenology of MOND. Class. Quantum Gravity 2007, 24, 3529–3539, doi:10.1088/0264-9381/24/14/001.

[540] Matsakos, T.; Diaffiero, A. Dynamics of galaxies and clusters in [refracted gravity]. arXiv 2016, arXiv:1603.0494.

[541] Tully, R.B.; Fisher, J.R. A new method of determining distance to galaxies. Astron. Astrophys. 1977, 500, 105–117.

[542] McGaugh, S.; Lelli, F.; Li, P.; Schombert, J. Dynamical Regularities in Galaxies. arXiv 2019, arXiv:1909.02011.

[543] McGaugh, S.S.; de Blok, W.J.G. Testing the Hypothesis of Modified Dynamics with Low Surface Brightness Galaxies and Other Evidence. Astrophys. J. 1998, 499, 66–81, doi:10.1086/305629.

[544] Swaters, R.A.; Sancisi, R.; van der Hulst, J.M.; van Albada, T.S. The link between the baryonic mass distribution and the rotation curve shape. Mon. Not. R. Astron. Soc. 2012, 425, 2299–2308, doi:10.1111/j.1365-2966.2012.21599.x.

[545] Sancisi, R. The visible matter—dark matter coupling. In Dark Matter in Galaxies; Ryder, S., Pisano, D., Walker, M., Freeman, K., Eds., Cambridge University Press: Cambridge, UK, 2004, Volume 220, p. 233.

[546] Sanders, R.H. The prediction of rotation curves in gas-dominated dwarf galaxies with modified dynamics. Mon. Not. R. Astron. Soc. 2019, 485, 513–521, doi:10.1093/mnras/stz2353.

[547] Miller, R.H.; Prendergast, K.H.; Quirk, W.J. Numerical Experiments on Spiral Structure. Astrophys. J. 1970, 161, 903, doi:10.1086/150593.

[548] Hoh, F. Numerical Experiments with a Disk of Stars. Astrophys. J. 1971, 168, 343, doi:10.1086/151091.
[660] McGaugh, S.S. MOND Prediction for the Velocity Dispersion of the “Feeble Giant” Crater II. Astrophys. J. Lett. 2016, 832, L8, doi:10.3847/2041-8205/832/1/L8.

[661] Li, B.; Zhao, G.B.; Teyssier, R.; Koyama, K. ECOSMOG: An Efficient COde for Simulating MOdified Gravity. J. Cosmol. Astropart. Phys. 2012, 2012, 051, doi:10.1088/1475-7516/2012/01/051.

[662] Puchwein, E.; Baldi, M.; Springel, V. Modified-Gravity-GADGET: A new code for cosmological hydrodynamical simulations of modified gravity models. Mon. Not. R. Astron. Soc. 2013, 436, 348–360, doi:10.1093/mnras/stt1575.

[663] Llinares, C.; Mota, D.F.; Winther, H.A. ISIS: A new N-body cosmological code with scalar fields based on RAMSES. Code presentation and application to the shapes of clusters. Astron. Astrophys. 2014, 562, A78, doi:10.1051/0004-6361/201322412.

[664] Baldi, M.; Villaescusa-Navarro, F.; Viel, M.; Puchwein, E.; Springel, V.; Moscardini, L. Cosmic degeneracies—I. Joint N-body simulations of modified gravity and massive neutrinos. Mon. Not. R. Astron. Soc. 2014, 440, 75–88, doi:10.1093/mnras/stu259.

[665] Baldi, M.; Villaescusa-Navarro, F. Cosmic degeneracies - II. Structure formation in joint simulations of warm dark matter and f(R) gravity. Mon. Not. R. Astron. Soc. 2017, 473, 3226–3240, doi:10.1093/mnras/stx2594.

[666] Arnold, C.; Li, B. Simulating galaxy formation in f(R) modified gravity: Matter, halo, and galaxy statistics. Mon. Not. R. Astron. Soc. 2019, 490, 2507-2520, doi:10.1093/mnras/stz2690.

[667] Kobayashi, T.; Murgia, R.; De Simone, A.; Irič, V.; Viel, M. Lyman-α constraints on ultralight scalar dark matter: Implications for the early and late universe. Phys. Rev. D. 2017, 96, 123514, doi:10.1103/PhysRevD.96.123514.

[668] Li, T.S.; Kaplinghat, M.; Bechtol, K.; Bolton, A.S.; Bovy, J.; Carleton, T.; Chang, C.; Drlica-Wagner, A.; Erkal, D.; Geha, M.; et al. Astrophysical Tests of Dark Matter with Maunakea Spectroscopic Explorer. arXiv 2019, arXiv:1903.03155.

[669] Drlica-Wagner, A.; Mao, Y.Y.; Adhikari, S.; Armstrong, R.; Banerjee, A.; Banik, N.; Bechtol, K.; Bird, S.; Boddy, K.K.; Bonaca, A.; et al. Probing the Fundamental Nature of Dark Matter with the Large Synoptic Survey Telescope. arXiv 2019, arXiv:1902.01055.