Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Dynamics of Dark Matter Halos in Compact Groups of Galaxies

Tesis presentada al

Posgrado en Física Aplicada

como requisito parcial para la obtención del grado de

Doctor en Ciencias

por

Erick Munive Villa

Asesorado por

Dra. Ana Aurelia Avilez López

Dr. Oscar Mario Martínez Bravo

Puebla Pue.
Julio de 2023

Título: Dynamics of Dark Matter Halos in Compact Groups of Galaxies
Estudiante: Erick Munive Villa

COMITÉ

Dr. Cupatitzio Ramírez Romero Presidente	Dr.Carlos Enrique Varela Secretario
Dr. Luis Alberto Martínez Medina Vocal	Dra. Tula Bernal Marín Vocal
Dr. Francisco Shidartha Guzmán Murillo Vocal	Dra. Olga Guadalupe Félix Beltrán Suplente
Dra. Ana Aurelia Avilez López Asesor	Dr. Oscar Mario Martínez Bravo Asesor

A mis padres, familiares y amigos.

Contents

Acknowledgments

I would like to express my deep gratitude to all those who have been instrumental in my academic and personal journey during my doctoral studies.

First and foremost, I would like to thank my thesis advisors, Dr. Ana Aurelia Aviléz López and Dr. Oscar Mario Martínez Bravo, for their unwavering guidance, support, and dedication throughout my research. Their expertise and guidance have been invaluable in the development of this work.

I would also like to thank the members of my thesis committee, Dr. Cupatitzio Ramírez Romero, Dr. Carlos Enrique Varela, Dr. Luis Alberto Martínez Medina, Dr. Tula Bernal Marín, Dr. Francisco Shidartha Guzmán Murillo, and Dr. Olga Guadalupe Félix Beltrán, for their time and for their valuable comments and suggestions that have enriched my research.

My gratitude also extends to my colleagues and collaborators, who have provided technical support, intellectual discussions, and significant contributions throughout this project.

I cannot fail to express my appreciation to my university, the Benemérita Universidad Autónoma de Puebla, for providing me with the opportunity to pursue my doctoral studies and for providing the necessary resources and facilities to carry out my research successfully.

Furthermore, I would like to acknowledge the financial support provided by the Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONACYT) through the CONACYT Scholarship for Graduate Studies in Mexico during my doctoral studies.

My research would not have been possible without the participation and contribution of the participants and research subjects. I sincerely thank each and every one of them for their time and willingness to be part of this study.

I would also like to acknowledge the technical and support staff of the following computing centers: the Laboratorio Nacional de Supercómputo del Sureste de México CU-BUAP, the Big Mamma cluster at the Laboratory of Artificial Intelligence and Supercomputing, IFM-UMSNH, and the Phoebe cluster at CEICO, FZU. They have provided assistance and support in carrying out the simulations reported in this thesis.

Finally, I want to express my gratitude to my friends and loved ones for their unwavering support, understanding, and encouragement throughout my doctoral journey. Their presence has been a fundamental pillar on this path.

Resumen

Este trabajo tiene como objetivo principal estudiar la dinámica de los halos de materia oscura y la influencia del término de presión cuántica que surge en el modelo de materia oscura escalar ultra-ligero en la formación de estructuras a pequeña escala, específicamente a escalas galácticas. Además, buscamos desarrollar una rutina que permita obtener soluciones para la evolución temporal del sistema Gross-Pitaevskii-Poisson en el límite Newtoniano a escalas astronómicas usando el método de diferencias finitas, sistema que determina la evolución de la materia oscura en este modelo. Para lograrlo, abordamos dos modelos diferentes de halos de materia oscura: el modelo de materia oscura fría propuesto en el marco del modelo estándar ΛCDM, y el modelo de materia oscura de campo escalar ultra-ligero (SFDM, por sus siglas en inglés).

El trabajo se divide en tres partes fundamentales. En primer lugar, se realiza un estudio detallado del modelo estándar ΛCDM, que se utiliza como modelo fiducial para establecer una referencia sobre los resultados a gran escala. Luego, se lleva a cabo una revisión exhaustiva del modelo SFDM, que servirá como nuestro modelo de prueba a pequeña escala para investigar el impacto de la dinámica de este campo en contraste con ΛCDM.

En la última parte del trabajo, nos enfocamos en la generación de condiciones iniciales, dividiéndose en dos enfoques principales. Por un lado, creamos condiciones iniciales basadas en simulaciones cosmológicas, explorando también el modelo de materia oscura generalizada (GDM) para obtener diferentes configuraciones de halos de materia oscura, como se discutirá más adelante en el texto. Por otro lado, utilizamos observaciones directas de galaxias a través del catálogo MaNGA para delimitar nuestros resultados y contrastarlos con las observaciones disponibles, sin embargo este es un trabajo en desarrollo.

Además de la exhaustiva revisión de la literatura relacionada con el modelo SFDM y de desarrollar un código capaz de resolver las ecuaciones asociadas a este modelo. En el código se implementó el método de malla móvil adaptativa (AMM, por sus siglas en inglés). También se ejecutó con éxito en unidades de procesamiento gráfico (GPUs) utilizando la biblioteca CuPy, y se realizó una ejecución en paralelo en múltiples núcleos de una sola unidad central de procesamiento (CPU).

Se estudiaron los efectos del modelo GDM en la formación de estructuras, especialmente en grupos compactos, y se encontró que pequeñas variaciones en parámetros como la presión, la viscosidad o la ecuación de estado conducen a estructuras diversas, lo que se refleja en las características de los grupos compactos. Además, se observó un aumento considerable en la cantidad de grupos compactos en comparación con ΛCDM, lo cual demuestra una mejora significativa en las predicciones del modelo de materia oscura en comparación con los datos observacionales.

Para estudiar la dinámica de la materia oscura desde una perspectiva observacional, se llevó a cabo una revisión exhaustiva de la literatura relacionada con las curvas de rotación y se obtuvo información sobre el perfil de materia oscura de algunas galaxias del catálogo MaNGA. Esto implicó el uso de técnicas de análisis de datos para eliminar señales de ruido, así como el Método de Monte Carlo basado en cadenas de Markov (MCMC). El catálogo MaNGA es especialmente relevante para nuestros fines, ya que contiene información de cerca de diez mil galaxias que pueden proporcionar datos de gran importancia sobre la dinámica de la materia oscura de manera observacional y permitir la comparación con los resultados de las simulaciones cosmológicas. El análisis

se encuentra en progreso y hay perspectivas prometedoras para la aplicación masiva del algoritmo en desarrollo a miles de galaxias en el catálogo.

En el caso de las galaxias espirales, es necesario realizar una descomposición de las componentes disco-bulbo para extraer adecuadamente la contribución de la materia oscura. Dado que el catálogo de Mendel examina un gran número de galaxias en el catálogo SDSS, se presentan las condiciones adecuadas para realizar un estudio utilizando redes neuronales profundas. Para ello, utilizamos información de catálogos simulados (catálogos ficticios) creados a partir de simulaciones cosmológicas del modelo ΛCDM, con el fin de recrear los valores observados por Mendel y probar el poder predictivo de las redes neuronales. Hasta la fecha de redacción de esta tesis, el trabajo se encuentra en etapas finales de desarrollo y aproximándose a su finalización.

Abstract

The main objective of this thesis is to study the dynamics of dark matter halos and the influence of the quantum pressure term on the formation of small-scale structures, specifically at galactic scales. Additionally, we aim to develop a routine that allows us to obtain solutions for the temporal evolution of the Gross-Pitaevskii-Poisson system in the Newtonian limit at astronomical scales using finite difference methods, which determines the evolution of dark matter in this model. To achieve this, we address two different models of dark matter halos: the cold dark matter model proposed in the framework of the ΛCDM standard model, and the ultra-light scalar field dark matter (SFDM) model.

The thesis is divided into three main parts. Firstly, a detailed study of the ΛCDM standard model is conducted, which serves as a fiducial model to establish a reference for large-scale results. Subsequently, an exhaustive review of the SFDM model is carried out, which will serve as our small-scale test model to investigate the impact of the dynamics of this field in contrast to ΛCDM.

In the final part of the thesis, we focus on the generation of initial conditions, which is divided into two main approaches. On one hand, we create initial conditions based on cosmological simulations, exploring the generalized dark matter (GDM) model to obtain different configurations of dark matter halos, as will be discussed later in the text. On the other hand, we use direct observations of galaxies through the MaNGA catalog to delimit our results and compare them with the available observations, although this is an ongoing work.

In addition to the exhaustive literature review related to the SFDM model and the development of a code capable of solving the equations associated with this model, the adaptive mesh refinement (AMR) method was implemented in the code. It was also successfully executed on graphic processing units (GPUs) using the CuPy library, and parallel execution was performed on multiple cores of a single central processing unit (CPU).

The effects of the GDM model on structure formation, especially in compact groups, were studied, and it was found that small variations in parameters such as pressure, viscosity, or equation of state lead to diverse structures, which is reflected in the characteristics of the compact groups. Furthermore, a considerable increase in the number of compact groups compared to ΛCDM was observed, demonstrating a significant improvement in the predictions of the dark matter model compared to observational data.

To study the dynamics of dark matter from an observational perspective, an exhaustive review of the literature related to rotation curves was conducted, and information about the dark matter profile of some galaxies in the MaNGA catalog was obtained. This involved the use of data analysis techniques to remove noise signals, as well as the Markov chain Monte Carlo (MCMC) method. The MaNGA catalog is particularly relevant for our purposes, as it contains information from nearly ten thousand galaxies that can provide highly important observational data on the dynamics of dark matter and allow for comparison with the results of cosmological simulations. The analysis is still in progress, and there are promising prospects for the massive application of the developing algorithm to thousands of galaxies in the catalog.

In the case of spiral galaxies, a decomposition of the disk-bulge components is necessary to accurately extract the contribution of dark matter. Given that the Mendel catalog examines a large number of galaxies in the SDSS catalog, the conditions are suitable for conducting a study using

deep neural networks. For this purpose, we used information from simulated catalogs (fictitious catalogs) created from cosmological simulations of the ΛCDM model to recreate the values observed by Mendel and test the predictive power of the neural networks. As of the writing of this thesis, the work is in its final stages of development and approaching completion.

Chapter 1

Introduction

1.1 Observational Foundations of Modern Cosmology

In this section, we will explore some fundamental results that have laid the foundations of modern cosmology. We will begin by analyzing the discovery of the universe's expansion, a crucial milestone that gave rise to the Big Bang model. Additionally, we will examine the prediction of the abundance of light elements, based on the notion of a primordial universe with extreme conditions of high density and pressure, in accordance with the Big Bang model. We will also delve into the discovery of the cosmic microwave background. This cosmic signal represents the remnant of the decoupling of photons from baryonic matter and provides valuable clues about the origins and evolution of the universe. We will also explore the implications of an isotropic and homogeneous universe in relation to the expansion and how gravitational sources can modify this expansion. Lastly, we will investigate observations of cosmic structures, such as galaxy clusters and super-clusters, which allow us to study the formation and evolution of the large-scale structure of the universe.

1.1.1 Hubble's Law and the Accelerated Expansion of the Universe

The radial velocity of galaxies, measured through the Doppler effect with a redshift of spectral lines (Fig. 1.1), is positive for almost all galaxies, indicating that they appear to be moving away from us. In 1928, Edwin Hubble Hubble (1929) discovered that this velocity v increases with the distance of the galaxy. He identified a linear relationship between the radial velocity v and the distance D of galaxies, which is known as Hubble's Law:

$$v = H_0 D, \tag{1.1}$$

where H_0 is a constant, known as the Hubble constant. If we plot the radial velocity of galaxies as a function of their distance, as shown in the Hubble diagram in Fig. 1.2, the resulting points approximate a straight line, whose slope is determined by H_0. The fact that all galaxies appear to be moving away from us with a velocity that increases linearly with their distance is interpreted as the expansion of the Universe.

For a long time, the value of H_0 was uncertain by a factor of about two. However, in recent years, the uncertainty has been significantly reduced Planck Collaboration (2020), obtaining

$$H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}.$$

The main problem in determining H_0 lies in measuring the absolute distance of galaxies, while the Doppler redshifts are easily measurable. Assuming that (1.1) is valid, the radial velocity of a galaxy is a measure of its distance. The redshift z of an object is defined as the change in

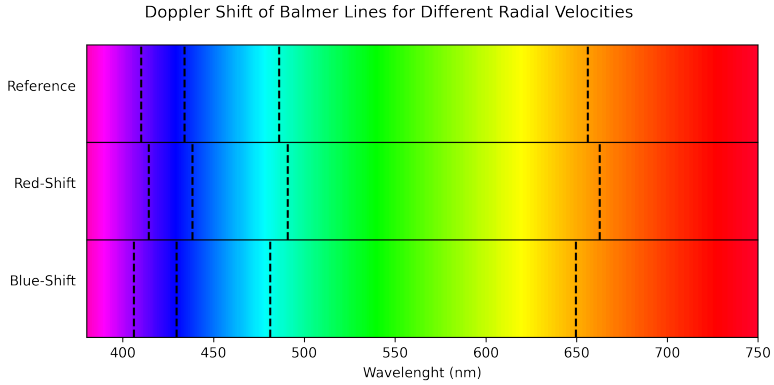

Figure 1.1: Doppler effect of Balmer lines for different radial velocities. The top panel represents the reference spectrum, while the middle and bottom panels show the spectra shifted to red and blue, respectively. Balmer lines are indicated by dashed lines. In this example, we use redshift $z = \pm 0.01 \sim \pm 3000$ km/s.

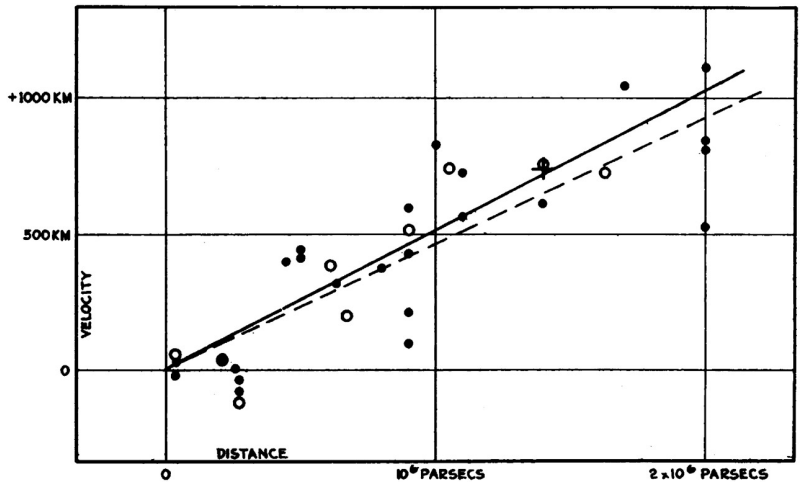

Figure 1.2: Radial velocities, corrected for solar motion, are plotted as a function of estimated distances derived from the involved stars and mean luminosities of nebulae in a cluster. The black disks and solid line represent the solution for solar motion using the nebulae individually; the circles and dashed line represent the solution combining the nebulae in groups; the cross represents the average velocity corresponding to the mean distance of 22 nebulae whose distances could not be individually estimated. Image and caption reproduced from Hubble (1929).

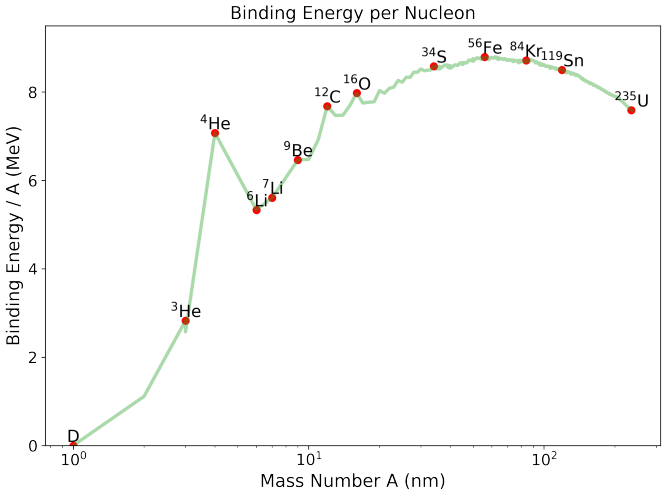

Figure 1.3: Binding energy per nucleon as a function of mass number (A) for atomic nuclei. The data points represent various elements, including deuterium (D), helium-3 (^{3}He), helium-4 (^{4}He), lithium-6 (^{6}Li), lithium-7 (^{7}Li), beryllium-9 (^{9}Be), carbon-12 (^{12}C), oxygen-16 (^{16}O), sulfur-34 (^{34}S), iron-56 (^{56}Fe), krypton-84 (^{84}Kr), tin-119 (^{119}Sn), and uranium-235 (^{235}U). The data was obtained from Ghahramany et al. (2012).

wavelength of spectral lines:

$$z := \frac{\lambda_{\text{obs}} - \lambda_0}{\lambda_0}, \quad \lambda_{\text{obs}} = (1+z)\lambda_0,$$

where λ_0 denotes the wavelength of a spectral transition in the rest frame of the emitter and λ_{obs} is the observed wavelength. For small redshifts $z \ll 1$, it is valid to consider

$$v \approx zc.$$

Hubble's law indicates that the universe is undergoing accelerated expansion. Therefore, in the past, the universe was more compact, i.e., the components of the universe were closer to each other. This gives rise to the possibility of a highly dense early universe.

1.1.2 Big Bang Nucleosynthesis

During the early stages of the universe, when the temperature was very high, neutral atoms did not exist, and the formed nuclei would quickly decay. This means that the primordial plasma was in thermodynamic equilibrium. As the universe cooled to around 30 MeV (equivalent to approximately 3.5×10^{11} K), nuclei started to form, and eventually, the first atoms appeared. This process, known as primordial nucleosynthesis, was responsible for the production of light elements. We know that protons and neutrons can fuse to form atomic nuclei when the temperature and density of the plasma are sufficiently high. However, if the temperature exceeds the binding energy threshold, the nuclei will decay, and no light elements would be stable. Helium-4 (^{4}He) has the highest binding energy among the light elements, which is 28 MeV (Fig. 1.3).

The allowed interactions during this stage of the universe are shown in Table 1.1, with the Neutron-Proton rate being the dominant one. Higher-order interactions (greater than two bodies) are unlikely and therefore do not contribute significantly. Based on these interactions, heavier elements can be formed through a series of nuclear reactions. For example, deuterium can be formed

through the interaction $p+n \rightarrow D+\gamma$, helium-3 through $D+D \rightarrow n+{}^3\text{He}$, and ${}^3\text{He}$ through ${}^3\text{He}$ + D $\rightarrow p+{}^4\text{He}$. These reactions depended on the abundance of neutrons and protons in the primordial universe. These nuclear fusion processes are driven by the strong and electromagnetic interactions acting on subatomic particles present in the primordial plasma. The relative abundance of the elements produced in primordial nucleosynthesis is determined by the combination of temperature, density, and reaction time in the early universe.

Table 1.1: Reactions $1 + 2 \longleftrightarrow 3 + 4$

	1	2	3	4
Neutron-Proton Ratio	n	ν_e or e^+	p	e^- or $\bar{\nu}_e$
Recombination	e	p	H	γ
Dark Matter Production	X	X	l	l

After approximately one second of the Big Bang Sarkar (1996), the temperature of the universe had decreased enough to allow the formation of the first atomic nuclei. During this brief period, known as primordial nucleosynthesis, a series of nuclear reactions took place resulting in the creation of light elements such as deuterium, helium-3, and helium-4. The specific abundance of these light elements generated during primordial nucleosynthesis has left its imprint on the observed abundances in the present-day universe (Fig 1.4).

The detailed observation of the relative quantities of deuterium, helium-3, and helium-4 provides valuable information about the conditions that prevailed in the early universe and supports the predictions of the Big Bang model. These observations closely match the theoretical predictions based on primordial nucleosynthesis. The existence and relative abundances of these light elements are consistent with the idea of a universe that has undergone expansion from an extremely hot and dense initial state, as postulated by the Big Bang model.

1.1.3 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is a microwave radiation that currently has an average temperature of around 2.7 K. It is extremely uniform in all directions and exhibits small temperature fluctuations that are crucial for understanding the formation of structures in the universe. This radiation originated when the universe cooled down enough for electrons and protons to combine and form neutral hydrogen atoms, approximately 300,000 years after the Big Bang. Prior to this moment, the universe was filled with a hot and dense plasma in which photons continuously interacted with electrons and protons, preventing them from freely propagating and forming atomic nuclei. However, when neutral atoms formed, photons became free to travel through space unhindered, thus creating the CMB. The released photons have a redshift of approximately $z = 1100$.

The cosmic background radiation contains a wealth of information about the initial conditions of the universe. The small temperature fluctuations in the CMB provide clues about the irregularities in the density of the early universe, which later, through the expansion of the universe, led to the formation of structures such as galaxies and galaxy clusters. Detailed studies of the CMB, conducted by space missions such as the COBE satellite Smoot (1999), the WMAP Bennett et al. (2013), and the Planck Planck Collaboration (2020), have provided precise measurements of the cosmic microwave background radiation. These measurements have confirmed many predictions of the Big Bang model and have allowed for the determination of key parameters of the universe, such as its age, composition, and rate of expansion. Observations from the Planck satellite Planck Collaboration (2020) (see Figure 1.5) show that 27% of the mass contained in the Universe is composed of Dark Matter, while 4% corresponds to ordinary matter, and the remaining 69% remains completely unknown and corresponds to Dark Energy. The cosmic microwave background is a remnant of the decoupling of photons from matter: it is known to be isotropic up to the scale of 10^{-5} K and is the best example of blackbody radiation that exists in the Universe. It is also

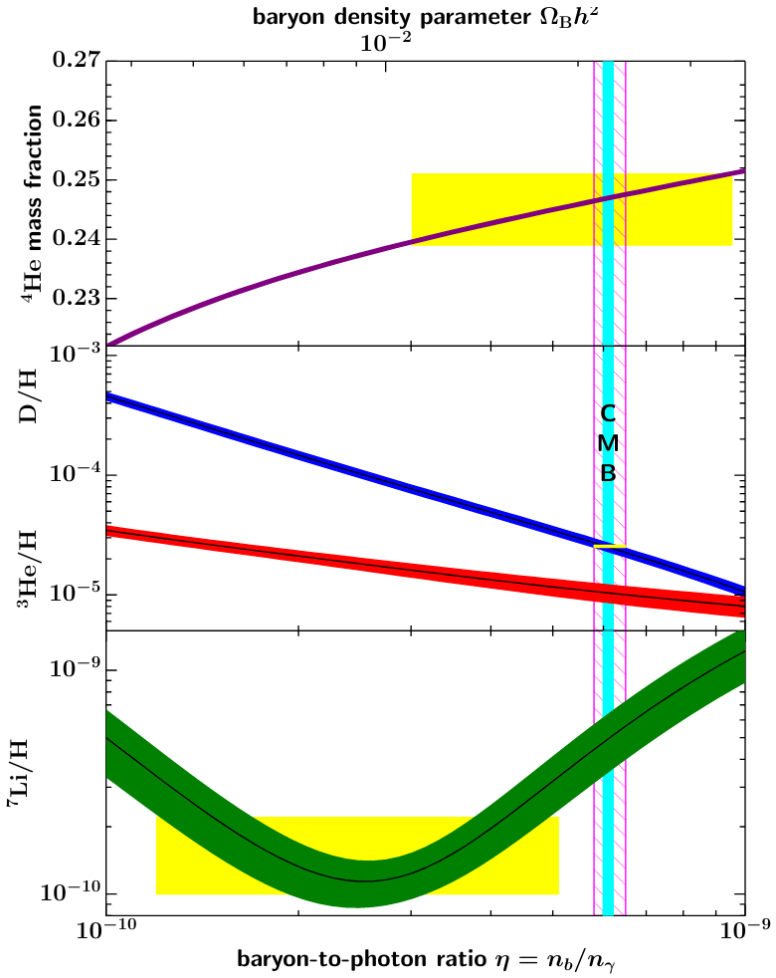

Figure 1.4: The primordial abundances of ^{4}He, D, ^{3}He, and ^{7}Li, as predicted by the standard Big Bang nucleosynthesis model, are shown in the shaded bands representing the 95% confidence range Fields et al. (2020). The boxes indicate the observed abundances of the light elements. The narrow vertical band indicates the measurement of the cosmic baryon density obtained through the cosmic microwave background (CMB), while the wider band indicates the D+^{4}He concordance range of primordial nucleosynthesis (both at the 95% confidence level). Image extracted from the Particle Data Group Workman et al. (2022).

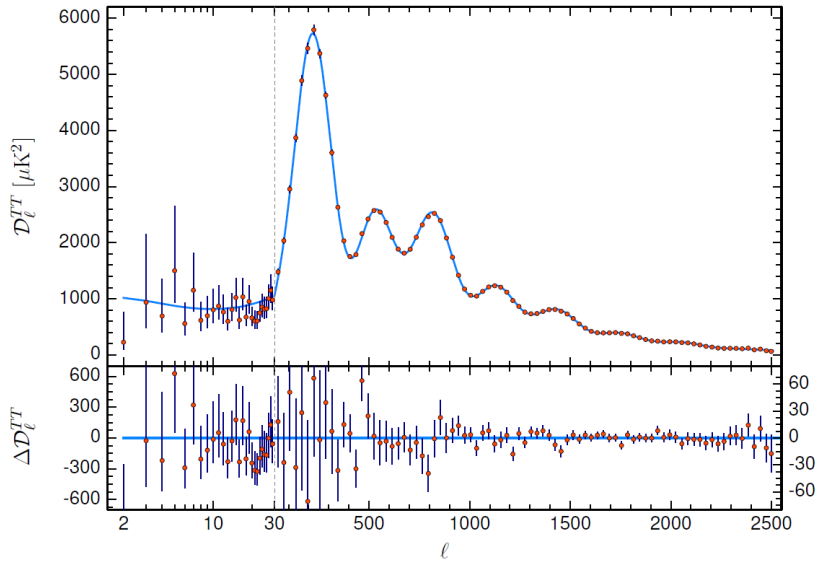

Figure 1.5: Results from the Planck satellite Planck Collaboration (2020) where the CMB is fitted, obtaining $\Omega_c h^2 = 0.120 \pm 0.001$, $\Omega_b h^2 = 0.0224 \pm 0.0001$, *corresponding to the abundances of dark matter and baryonic matter in the Universe.*

known to have a temperature of $T = 2.726$ K. Below 10^{-5} K, variations from the average value of the CMB temperature are observed. These anisotropies are expressed using a spherical harmonic expansion

$$T(\theta, \phi) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\theta, \phi),$$

where $Y_{\ell m}(\theta, \phi)$ are the spherical harmonics. The angular power spectrum is defined as

$$C_\ell = \left\langle |a_{\ell m}|^2 \right\rangle,$$

and in Figure 1.5, $D_\ell^{TT} = \ell(\ell+1)C_\ell/2\pi$ is shown.

1.1.4 Large-scale structures

In Figure 1.6, the results from different surveys are shown. For instance, the *CfA2 Great Wall* Geller and Huchra (1989) is a large-scale structure in the Universe identified in the CfA2 galaxy catalog (Center for Astrophysics Redshift Survey). It extends over approximately 500 million light-years and takes the form of a thin and elongated filamentary structure. Discovered in the 1980s, its name derives from its size and shape, resembling a cosmic wall. At the center of the CfA2 Great Wall lies the Coma cluster, one of the most massive and closest galaxy clusters to Earth. This discovery revealed the presence of large-scale organization in the distribution of galaxies in the Universe and provided valuable insights into the formation and evolution of structures on a large scale.

Adittionally, the *Sloan Great Wall* III et al. (2005) is an immense cosmic structure discovered within the framework of the Sloan Digital Sky Survey (SDSS). With over 10,000 galaxies and an extent of more than 1,37 billion light-years, this structure represents a fascinating example of the large-scale organization in the Universe. Its discovery has provided valuable information about the formation and evolution of galaxies, as well as the distribution of matter on a large scale in the Universe. Additionally, the *Millennium* simulation Springel et al. (2005) is a powerful tool in cosmology that simulates the formation and evolution of the large-scale structure in the Universe.

Based on current models and theories, the simulation utilizes sophisticated algorithms to study the distribution of dark matter and the formation of galaxies over billions of years. From this simulation, simulated catalogs called *mock catalogs* are generated (see for instance Behroozi et al. (2012a)), which mimic real observations and allow for the comparison and validation of theoretical results with observational data. These simulated catalogs are essential for understanding galaxy formation and the evolution of the Universe. The observations of the Great Wall and other large-scale structures in the Universe have led to several important conclusions.

Firstly, they reveal a cosmic organization in the distribution of galaxies, with filamentary structures that suggest flows and cosmic webs of matter. Furthermore, they highlight the connection between large-scale structures and the formation of massive galaxy clusters, such as the Coma cluster in the *CfA2 Great Wall* which is located at a node of this cosmic web. These observations of galaxy distribution provides clues about the processes of formation and evolution of these structures. Lastly, they have validated theoretical models of the formation and evolution of large-scale structures by comparing observational data with the results of numerical simulations Springel et al. (2006).

1.2 The ΛCDM Cosmological Model

In this section, we will present the standard ΛCDM cosmological model, which describes the structure and evolution of the universe based on accelerated expansion, cold dark matter, and dark energy. We will explain its main features and how it fits with the available observations.

1.2.1 Background Cosmology

This section will focus on background cosmology, which encompasses fundamental concepts such as the Cosmological Principle, the Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime, and the Friedmann equations. We will also analyze the equations of a perfect fluid and the different species of particles present in the universe. We will explore the cosmological solutions that describe the expansion and evolution of the universe based on its parameters.

1.2.2 Cosmological Principle

It is a fundamental concept in cosmology that states that the universe is homogeneous and isotropic on large scales. In other words, it asserts that the properties of the universe are the same at all points and in all directions on sufficiently large scales. This principle is a basic assumption in the construction of cosmological models and provides the foundation for understanding the structure and evolution of the universe.

The apparent isotropic distribution of galaxies, when averaged on large scales, and the isotropy of the cosmic microwave background suggest that the universe around us is isotropic on large angular scales. Therefore, we first consider a model that describes the universe around us as isotropic. If we further assume that our location in the cosmos is not privileged over any other location, then the assumption of isotropy around us implies that the universe appears isotropic when viewed from any other location. The homogeneity of the universe immediately follows from isotropy around every location. The combination of the assumptions of homogeneity and isotropy of the universe is also known as the Cosmological Principle. We will see that a model based on the Cosmological Principle provides an excellent description of numerous observational facts.

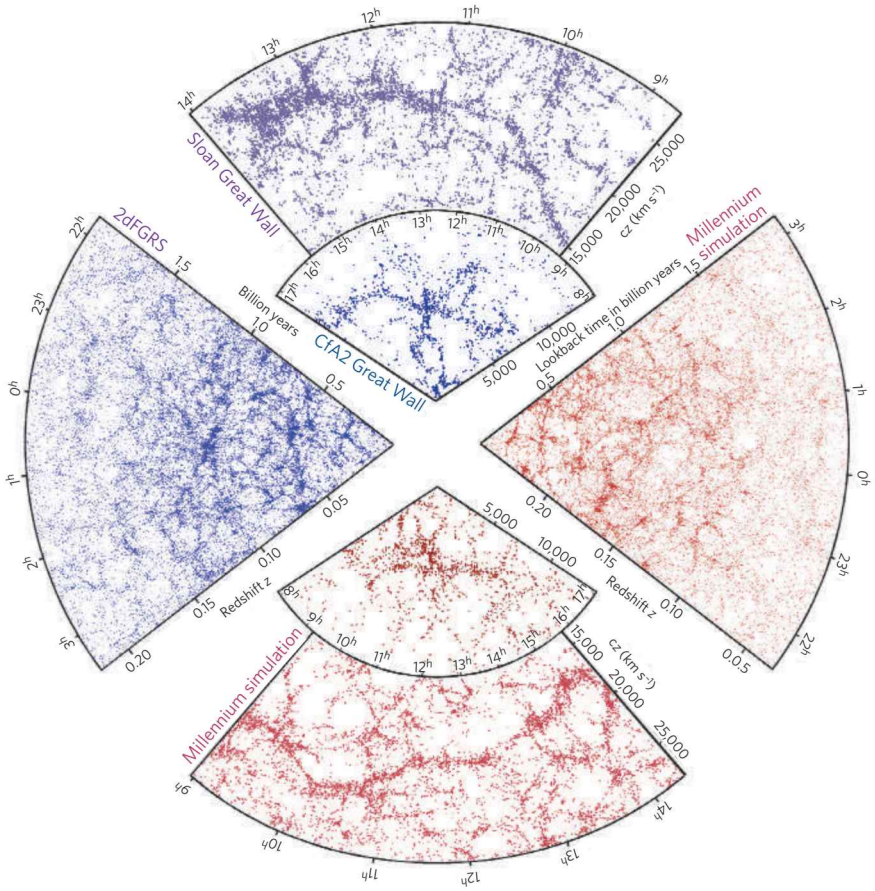

Figure 1.6: The galaxy distribution obtained from spectroscopic redshift surveys and from mock catalogs constructed from cosmological simulations. The small slice at the top shows the CfA2 Great Wall Geller and Huchra (1989), with the Coma cluster at the center. Drawn to the same scale is a small section of the SDSS, in which an even larger Sloan Great Wall has been identified III et al. (2005). This is one of the largest observed structures in the Universe, containing over 10,000 galaxies and stretching over more than 1,37 billion light years. The cone on the left shows one-half of the 2dFGRS, which determined distances to more than 220,000 galaxies in the southern sky out to a depth of 2 billion light years. The SDSS has a similar depth but a larger solid angle and currently includes over 650,000 observed redshifts in the northern sky. At the bottom and on the right, mock galaxy surveys constructed using semi-analytic techniques to simulate the formation and evolution of galaxies within the evolving dark matter distribution of the Millennium simulation Springel et al. (2005) are shown, selected with matching survey geometries and magnitude limits.

1.3 Friedmann-Lemaître-Robertson-Walker (FLRW) Spacetime

This is the metric that describes the universe on large scales within the framework of cosmology. This metric is a solution to Einstein's field equations and is based on three fundamental assumptions: homogeneity, isotropy, and expansion.

The FRW metric can be expressed as follows in comoving coordinates:

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right), \tag{1.2}$$

where ds^2 is the infinitesimal line element, dt is the time coordinate (proper time), $a(t)$ is the scale factor that describes the expansion of the universe as a function of time, dr, $d\theta$, $d\phi$ are the radial, angular, and azimuthal spatial coordinates, respectively, r is the radial coordinate, and k is the curvature parameter that determines the geometry of space: $k = -1$ for hyperbolic space (negative curvature), $k = 0$ for flat space (zero curvature), $k = 1$ for spherical space (positive curvature).

This metric serves as the basis for describing the expansion of the universe in the Friedmann-Lemaître-Robertson-Walker (FLRW) cosmological model. From this metric and Einstein's field equations, the Friedmann equations can be derived, which relate the expansion of the universe to the distribution of energy and matter within it.

The study of the FLRW spacetime is crucial for understanding the evolution and structure of the universe on large scales, as well as to research phenomena such as cosmic microwave background radiation, the formation of cosmic structures, and the nature of dark energy.

1.3.1 Friedmann Equations

These are a set of equations that describe the evolution of the universe within the framework of the Friedmann-Lemaître-Robertson-Walker (FLRW) cosmological model. These equations are derived from Einstein's field equations and relate the expansion of the universe to the distribution of energy and matter within it. The Friedmann equations are expressed as follows:

$$\begin{aligned} \text{Friedmann equation} H^2 &= \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}, \\ \text{Acceleration equation} \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}, \end{aligned} \tag{1.3}$$

where $H = \dot{a}/a$ is the Hubble parameter, which, along with the continuity equation derived from the temporal component of the conservation equation of the perfect fluid energy-momentum tensor,

$$\dot{\rho} + 3H(\rho + P) = 0, \tag{1.4}$$

describes the rate of expansion of the universe as a function of time (in the equations, G is the gravitational constant, ρ is the energy density, p is the pressure, a is the scale factor that describes the expansion of the universe, k is the curvature parameter that determines the geometry of space, and Λ is the cosmological constant). The continuity equation (1.4) can be easily solved if we consider an equation of state of the form $P = w\rho$.

The first equation in the system (1.3), known as the Friedmann equation, relates the rate of expansion of the universe to the energy density, the curvature parameter, and the cosmological constant. The second equation in the system (1.3), known as the acceleration equation, relates the acceleration of the expansion of the universe to the energy density, the pressure, and the cosmological constant. These equations allow us to study the evolution of the universe in terms of the different energy and matter components present within it. Moreover, they provide information about the geometry of space and the possible presence of dark energy. The analysis of the Friedmann equations is essential for understanding the dynamics and evolution of the universe as

a whole, as well as investigating cosmological phenomena such as inflation, accelerated expansion, and the large-scale structure of the universe.

1.3.2 Cosmological Solutions

Cosmological solutions are mathematical descriptions that model the expansion and evolution of the universe based on its parameters and initial conditions. These solutions are obtained from the Friedmann equations and are based on the Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime. Let's study some solutions to these equations. To study the solutions of these equations, it is necessary to consider the definition of the critical density and relative density for radiation, non-relativistic matter, curvature, and cosmological constant:

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}, \tag{1.5}$$

$$\begin{aligned} \Omega_\gamma &= \frac{\rho_\gamma}{\rho_{\text{crit}}}, \quad \Omega_m = \frac{\rho_m}{\rho_{\text{crit}}}, \\ \Omega_k &= \frac{-k}{a^2H^2} \quad \Omega_\Lambda = \frac{\Lambda}{3H^2}. \end{aligned} \tag{1.6}$$

Let's first consider the case of dust-like matter (pressureless), for which the continuity equation implies:

$$\rho = \rho_0 a^{-3}, \tag{1.7}$$

assuming that $a_0 = 1$. For the case of radiation-like matter, with $P = \frac{1}{3}\rho$, the continuity equation implies:

$$\rho = \rho_0 a^{-4}. \tag{1.8}$$

Now, using equation (1.6), we can rewrite the Friedmann equation (1.3) as:

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2(\Omega_{0m}a^{-3} + \Omega_{0\gamma}a^{-4} + \Omega_{0k}a^{-2} + \Omega_{0\Lambda}), \tag{1.9}$$

where we have considered the relative density for non-relativistic and relativistic species. Let's now study some solutions of the Friedmann equations and solve the simplified version (1.9).

i) de Sitter Universe:

This solution describes a universe in which the energy density is dominated solely by the cosmological constant. The expansion of the universe is exponentially accelerated. In this universe, both the pressure and density are zero, so the Friedmann equation (1.3) can be written as:

$$\left(\frac{\dot{a}}{a}\right)^2 = \Omega_\Lambda.$$

The solution to this equation can be easily obtained as:

$$a(t) = \mathrm{e}^{-\sqrt{H_0^2\Omega_{0\Lambda}}(t-t_0)},$$

where t_0 corresponds to the present time, with $a(t_0) = 1$.

ii) Einstein-de Sitter Universe

This solution describes a universe in which the energy density is solely dominated by non-relativistic matter. The expansion of the universe follows a power law, and the long-term evolution is determined by the gravitational interaction of matter. In this universe, the pressure is zero, P = 0, so the Friedmann acceleration equation (1.3) can be written as

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2(\Omega_{0m}a^{-3} - \Omega_{0k}a^{-2}),$$

The solution to this equation can be easily obtained for a flat universe, k = 0, as

$$a(t) = \left(H_0^2\Omega_{0m}\right)^{1/3}\left(\frac{3}{2}(t+t_0)\right)^{2/3}.$$

iii) Radiation-Dominated Universe

This solution describes a universe in which the energy density is solely dominated by radiation and relativistic matter. The expansion of the universe follows a power law, and the long-term evolution is determined by the gravitational interaction of matter. In this universe, the pressure is $P = \rho/3$, so the acceleration equation (1.9) can be written as:

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2(\Omega_{0\gamma}a^{-4} - \Omega_{0k}a^{-2}).$$

If we consider a flat universe with $k = 0$, the solution is expressed as:

$$a(t) = \left(H_0^2\Omega_{0\gamma}\right)^{1/4}\sqrt{2(t+t_0)}.$$

In the general case where $k \neq 0$, this equation can be easily solved by multiplying both sides by a^2 and separating variables. The solution is expressed as:

$$a(t) = H_0\sqrt{\Omega_{0\gamma} - \Omega_{0k}(t+t_0)^2}.$$

Matter-Dominated Universe with Cosmological Constant

This solution describes a universe in which the energy density is dominated by both the cosmological constant and non-relativistic matter. We consider the density as in (1.7). In this universe, the first Friedmann equation (1.9) can be written as:

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2(\Omega_{0m}a^{-3} + \Omega_{0\Lambda}).$$

If we consider a flat universe with $k = 0$, the solution is expressed as:

$$a(t) = \left[\sqrt{\frac{\Omega_{m0}}{\Omega_{\Lambda 0}}}\sinh\left(\frac{3H_0}{2}\sqrt{\Omega_{\Lambda 0}}(t+t_0)\right)\right]^{2/3}.$$

These are some of the cosmological solutions that can be easily found analytically. The solutions are shown in Figure 1.7. As we will see later, these are the solutions most relevant during the evolution of the universe.

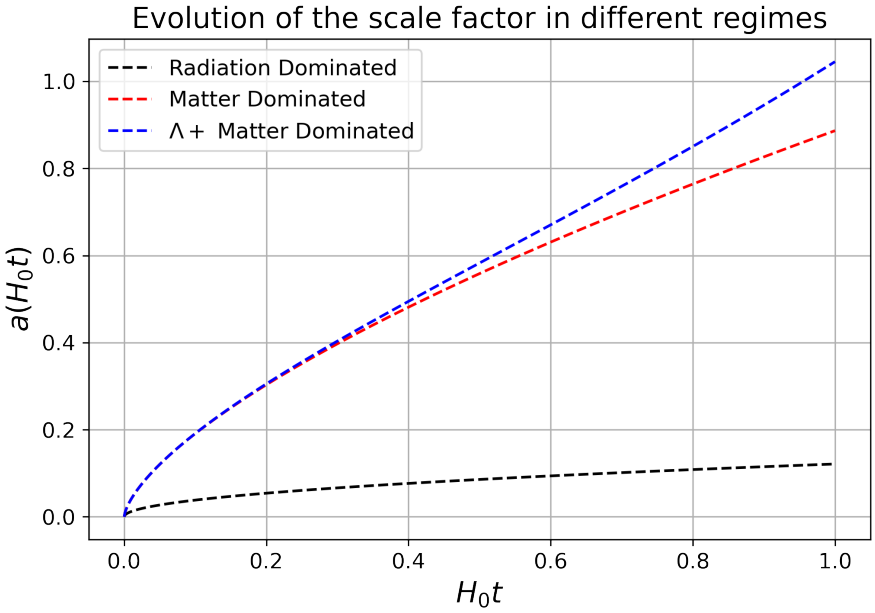

Figure 1.7: Evolution of the scale factor for different scenarios: a radiation-dominated universe, a non-relativistic matter-dominated universe, and a universe considering the cosmological constant Λ and non-relativistic matter. The expansion of the universe is more accelerated when non-relativistic matter is considered. In each case, a flat curvature $k = 0$ is assumed according to observations, and Planck results Planck Collaboration (2020) for the relative densities and the Hubble constant H_0 are used.

1.4 Perfect Fluids and Particle Species

As seen previously in Section 1.1, the Hubble expansion law implies that the universe had a singularity in density and temperature at a finite past time, which is commonly referred to as the Big Bang. In this context, we study the thermodynamics of a universe dominated by relativistic species assuming that the entropy per comoving volume is constant:

$$s(T)a^3 = \text{constant}.$$

This hypothesis implies a conservation law for energy:

$$T\frac{dP}{dT} = \rho + P.$$

1.4.1 Recombination

Recombination refers to the stage in the universe's evolution when the first hydrogen atoms are formed. In this process, we assume that after the Big Bang, the reaction rates for particle interactions were much faster than the expansion rate, allowing the cosmic fluid to be in thermal equilibrium. During this stage, the dominant reaction is:

$$e^- + p \leftrightarrow H + \gamma.$$

In equilibrium, the Gibbs potential is constant:

$$dG = \mu_p dN_p + \mu_p dN_p + \mu_e dN_e + \mu_H dN_H + \mu_\gamma dN_\gamma = 0,$$

where μ_i represents the chemical potential of each species i. Additionally, the number of particles is conserved: $dN_p + dN_H = dN_e + dN_H$. Since the chemical potential of photons is zero, we have $\mu_e + \mu_p - \mu_H = 0$ from this reaction. The particles involved obey the Maxwell-Boltzmann statistics:

$$n_i = \frac{2}{(2\pi)^3}(2\pi m_i T)^{3/2} \mathrm{e}^{-\frac{m_i - \mu_e}{T}},$$

for $i = e,\ p,\ H$. Let's define $\Delta = m_e + m_p - m_H$. Moreover, considering that the universe is neutral, $n_e \sim n_p$, we have:

$$\frac{n_e^2}{n_H} = \frac{n_e n_p}{n_H} = \left(\frac{m_e T}{2\pi}\right)^{3/2} \mathrm{e}^{-\Delta/T}.$$

Defining the ionization fraction as $x_e = n_e/(n_e + n_H)$ and the baryon entropy as $\sigma = (4\pi^2/45)T^3/n_B$, we obtain the Saha equation:

$$\frac{x_e^2}{1 - x_e} = \frac{45\sigma}{4\pi^2}\left(\frac{m_e}{2\pi T}\right)^{3/2} \mathrm{e}^{-\Delta/T}. \tag{1.10}$$

From this equation, we can see that for $T \gg \Delta$, $x_e \sim 1$, implying $n_H = 0$. We define the recombination epoch as the temperature when $x_e = 1/2$. It can be deduced that $T_{\mathrm{rec}} \sim 3757$ K ~ 0.32 eV.

1.4.2 Cosmic Microwave Background

We will now explore the evolution of electrons, photons, and neutrinos in the early universe. At these scales, several processes dominate the dynamics and play a crucial role in shaping the cosmic landscape. The key processes can be summarized as follows:

$$e^- + e^+ \leftrightarrow 2\gamma, \quad e^- + \bar{\nu} \leftrightarrow e^- + \bar{\nu},$$
$$\nu + \bar{e} \leftrightarrow \nu + \bar{e}, \quad e + \bar{e} \leftrightarrow \nu + \bar{\nu}.$$

These reactions involve particle interactions of photons and neutrinos. During the early stages of the universe, these processes are the most efficient. The interaction rate of neutrinos compared to the expansion rate of the universe is expressed as:

$$\frac{\Gamma_F}{H} \sim \left(\frac{T}{1,4\ \mathrm{MeV}}\right)^3.$$

As the temperature of the universe drops below approximately 1,4 MeV, the interaction rate for neutrinos becomes significantly smaller than the expansion rate. Consequently, neutrinos gradually decouple from other particles, resulting in the universe becoming *transparent* to neutrinos. Moreover, the interaction between photons and electrons remains highly efficient until the temperature decreases substantially below the electron mass, m_e. At this critical point, photons also decouple from electrons, leading to a distinct thermal evolution. The decoupling of photons occurs at a temperature represented by T_γ, which can be related to the decoupling temperature of neutrinos as:

$$T_\gamma = \left(\frac{11}{4}\right)^{1/3} T_\nu,$$

where T_ν denotes the temperature of neutrinos at the time of their decoupling. This relationship highlights the interdependence of photons and neutrinos and reveals the existence of a cosmic background composed of both photons and neutrinos.

1.4.3 Nucleosynthesis

The observed abundance of helium, denoted as Y, is an important parameter in understanding the nucleosynthesis process in the early universe. It represents the ratio of the number density of helium nuclei (n_{He}) to the number density of hydrogen nuclei (n_H), normalized by their respective masses (m_{He} and m_H). The value of Y is approximately 0,24, indicating that helium is a significant component in the early universe. The abundance of helium is a result of various nuclear reactions that occurred during the early stages of the universe. These reactions involve the formation and interconversion of protons, neutrons, and helium nuclei. The detailed analysis of these processes is crucial in understanding the composition and evolution of the universe. The abundance of helium is observed to be:

$$Y \equiv \frac{n_{He} m_{He}}{n_H m_H} \sim 0,24.$$

Let's now examine the reactions and conditions that govern the nucleosynthesis process in the early universe. These reactions are crucial in determining the abundance of light elements such as helium. Several reactions are in equilibrium until a certain temperature, known as the decoupling temperature (T_D), which is approximately 0,7 MeV. These reactions involve neutrinos, photons, protons, electrons, antineutrinos, and neutrons. The reactions can be represented as follows:

$$\nu + \gamma \leftrightarrow p + e, \quad \bar{e} + n \leftrightarrow p + \bar{\nu}, \quad n \leftrightarrow p + e + \bar{\nu}.$$

At temperatures above T_D, these reactions are efficient and keep the involved species in thermal equilibrium. However, as the temperature drops below T_D, the expansion of the universe causes the reaction rates to decrease, leading to the decoupling of neutrinos, photons, and the other particles involved. This decoupling process allows the universe to transition into the next stage of its evolution, with the formation and abundance of light elements becoming more prominent. To quantify the rates of these reactions, we can introduce the baryon-to-photon ratio, denoted as $\eta_B = n_B/n_\gamma$. The baryon-to-photon ratio represents the ratio of the number density of baryons (protons and neutrons) to the number density of photons in the universe. The rate for the process $n + p \leftrightarrow^2 H + \gamma$ is given by:

$$\Gamma_{np} \sim 10^{12} \eta_B \left(\frac{T}{\text{MeV}}\right)^3 \text{s}^{-1}.$$

As the universe expands and cools down, the conditions become favorable for the formation of the first nuclei. Specifically, for the neutron and proton species, their abundances can be expressed as $Y_n = 1/8$ and $Y_p = 7/8$, respectively. This distribution arises from the equilibrium reactions and processes that occur until a certain temperature threshold is reached.

Following the formation of neutrons and protons, the next significant step is the synthesis of helium-4 nuclei (^{4}He). The abundance of helium-4 is approximately $Y_{^4He} \sim 1/4$. The successful formation of helium-4 is a crucial outcome of the nucleosynthesis process and contributes to the enrichment of the universe with this element.

To further illustrate the predictions and observations of the abundances of these species, we can refer to Figure 1.4. This figure presents data and theoretical predictions that showcase the relative abundances of hydrogen, helium, and other light elements in the universe. The agreement between observations and predictions is an important validation of our understanding of the nucleosynthesis processes and the evolution of element abundances in the early universe.

1.5 The Problem of Dark Matter

Fritz Zwicky was the first astronomer to report a mass deficit while studying the Coma galaxy cluster in 1933 Zwicky (1937)[1]. Assuming that the system is in equilibrium (i.e., satisfies the Virial

[1]For a translated version in English, please refer to Andernach and Zwicky (2017)

theorem), Zwicky reported a mass-to-light ratio of $M/L \sim 300$, indicating that the mass is 300 times greater than that of luminous matter. Even greater density is expected than observed if the system is not assumed to be stationary.

Subsequently, in 1970, Vera Rubin reported her findings from detailed observations of the rotation velocities of stars in spiral galaxies Rubin et al. (1980). She found that the stars in the outer regions of galaxies rotated faster than expected according to the laws of Newtonian gravity, indicating the presence of invisible additional mass or dark matter. These findings led to the conclusion that there was a significant amount of additional mass in galaxies beyond visible matter. This additional mass was associated with the gravitational force necessary to keep the stars in the outer regions together and explained the high observed rotation speeds. This missing mass, which does not emit electromagnetic radiation, became known as dark matter.

1.5.1 Observational Evidence

Rotation Curves

The orbital velocities of stars at different radial distances from the center of a galaxy give rise to the rotation curve of a galaxy. Thus, the rotation curve represents the orbital velocity of stars as a function of their radial distance from the galactic center. In the case of a spiral galaxy, we find a central nucleus where most of the mass is concentrated, while spiral arms extend through a disk. For a star in this galaxy, located at a distance r from the galactic center and moving with an orbital velocity $v(r)$, the gravitational force is balanced by the centrifugal force through the following equation:

$$\frac{mv(r)^2}{r} = \frac{GmM(r)}{r^2},$$

where $M(r)$ represents the mass enclosed within the radius r. If the star is within the dense central region or the nucleus of the galaxy, the following relation holds:

$$M(r) = \frac{4}{3}\pi r^3 \rho. \tag{1.11}$$

Here, ρ represents the average density of the central nucleus. In this region, the orbital velocity $v(r)$ is expected to increase proportionally with distance r, i.e., $v(r) \sim r$. However, for stars outside the galaxy, it can be considered that the enclosed mass $M(r)$ is constant, and therefore, the orbital velocity $v(r)$ varies inversely with the square root of r, i.e., $v(r) \sim 1/\sqrt{r}$. Thus, the variation of the orbital velocity $v(r)$ with distance r in a spiral galaxy should show an initial increase, followed by a decline (known as Keplerian falloff) that behaves as $1/\sqrt{r}$. However, observational measurements of rotation curves for several spiral galaxies indicate that $v(r)$ remains constant for large r. From equation (1.11), it is deduced that $M(r)$ grows proportionally with r, suggesting the presence of an enormous invisible mass in the galaxy. This invisible mass, known as dark matter, is believed to form a dark halo in which the galaxy is immersed. A couple of examples of this phenomenon are shown in Figure 1.8.

Gravitational Lensing

The gravitational lensing effect is another piece of evidence for the existence of invisible matter in the Universe. It has been estimated that the mass required to produce this effect in some galaxy clusters exceeds the observed mass by a large margin ($\sim 4 \times 10^2$ times). This phenomenon allows us to search for the effects of dark matter on large scales. The event known as the Bullet Cluster provides evidence of the behavior of dark matter, as it is found to not interact with itself or with ordinary matter. Gravitational lenses are predicted in Albert Einstein's General Theory of Relativity, which states that gravity determines the geometry of space. Light follows a curved path around a massive body, behaving as if it were passing through a lens. This is why this

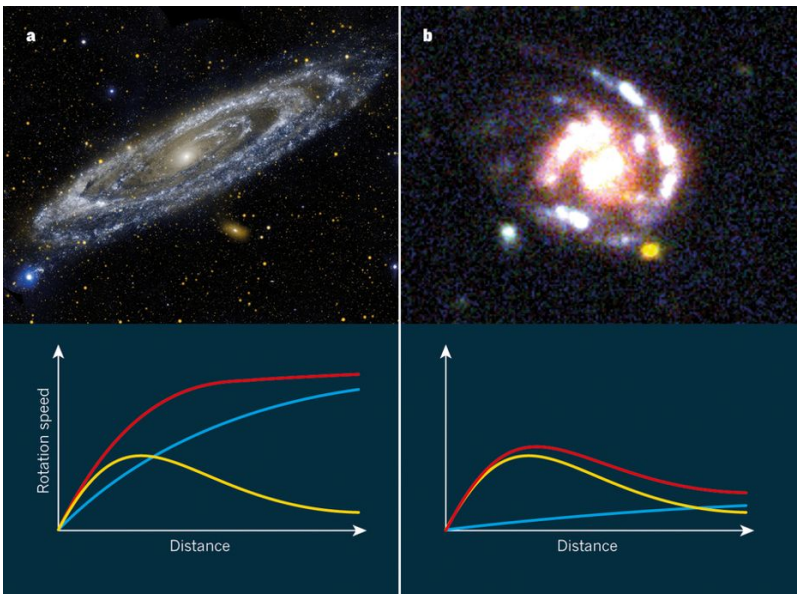

Figure 1.8: a) Nearby spiral galaxies, such as Andromeda (shown here), exhibit flattened rotation curves: as we move away from the galactic center, the rotational velocities of stars remain approximately constant with distance from the center (red). This contrasts with the expected rotation curves based on the distribution of visible matter (yellow). To explain this discrepancy, astronomers proposed that the mass of galaxies is dominated by the so-called dark matter (blue). b) In Genzel et al. (2017), an investigation of six massive star-forming galaxies in the distant Universe was conducted (shown here is an example of one of these galaxies, UDFJ033237-274751). The authors discovered that these galaxies have rotation curves that, after reaching a peak, decline with distance, suggesting that they contain relatively little dark matter. Image and description taken from Swinbank (2017).

Figure 1.9: This is a composite image of the Bullet Cluster. The Bullet Cluster is a widely studied pair of colliding galaxy clusters. One has passed through the other, like a bullet passing through an apple. This happens in our line of sight, so we can clearly see the two clusters. The optical image from the Magellan and Hubble Space Telescopes shows the galaxies in orange and white in the background. The hot gas, which contains most of the normal matter in the cluster, is shown in the Chandra X-ray image, which reveals the hot cluster gas (pink). The gravitational lensing effect, the distortion of background images caused by the mass in the cluster, reveals that the cluster's mass is dominated by dark matter (blue), an exotic form of matter that is abundant in the Universe, with very different properties compared to normal matter. This was the first clear separation observed between normal matter and dark matter. Image released by The European Space Agency (ESA). Source: ***https://www.esa.int/ESA_Multimedia/Images/2007/07/The_Bullet_Cluster2***

phenomenon is known as gravitational lensing. In cases like the one shown in Figure 1.9, this method can be used to estimate the mass of the gravitating object. However, in most cases, it is not straightforward because galaxy clusters are generally composed of many gravitating bodies. It is very interesting that in Figure 1.9, the blue marks represent the distribution of dark matter and appear to be spherical, which facilitates the task of measuring the portion of dark matter present in that region. Figure 1.10 illustrates the method used to perform such measurements. When a massive object, referred to as the lens object (since it behaves like one), is located between a point source of light and an observer, the light bends, and the image is ultimately displaced from the source. This process can be approximated by a pair of straight lines separated by an angle α_D (deflection angle), which is given by

$$\alpha_D \approx \frac{2Gm}{c^2 R},$$

and can be used to estimate the mass of dark matter by observing this phenomenon with background objects as sources.

Cosmic Microwave Background (CMB) Anisotropies

As reviewed in section 1.1.3 and 1.4.2, the CMB is the oldest observable information about the Universe. Figure 1.11 shows the dependence of the CMB temperature spectrum on different components of the fundamental parameters. Observations from the Planck satellite Planck Collaboration (2020) (see Figure 1.5) show that 27% of the mass in the Universe is composed of dark matter, while 4% corresponds to ordinary matter. Additionally, the remaining 69% is completely unknown and corresponds to Dark Energy.

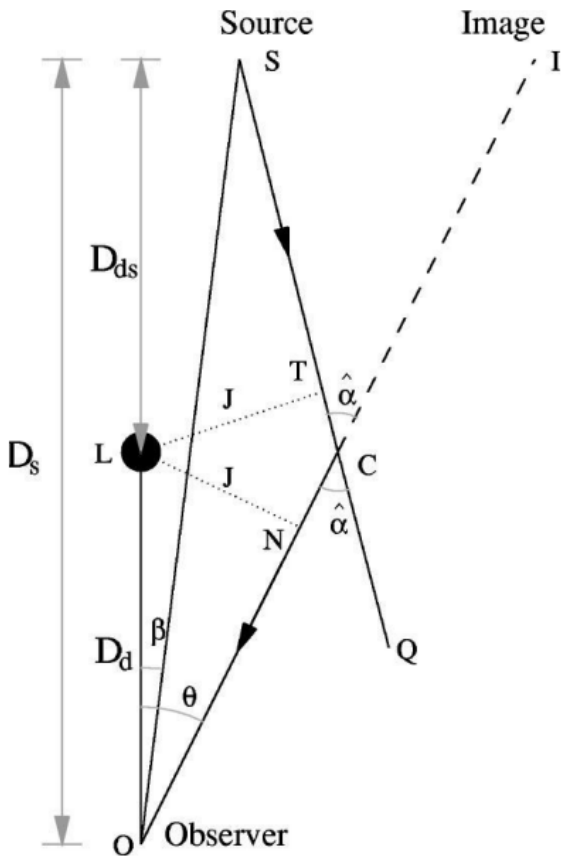

Figure 1.10: The lens diagram: O, L, and S represent the positions of the observer, the deflector (lens), and the source, respectively. OL is the reference axis (optic axis). ∠LOS and ∠LOI are the angular separations between the source and the image with respect to the optic axis. SQ and OI are tangents to the null geodesic at the positions of the source and the observer; LN and LT, the perpendiculars to these tangents from L, represent the impact parameter J. ∠OCQ is the Einstein deflection angle. D_s indicates the observer-source distance, D_{ds} the lens-source distance, and D_d the observer-lens distance. Image and description taken from Virbhadra and Ellis (2000)

1.5.2 Candidates for Dark Matter

Although the exact nature of dark matter is still unknown, within cosmology, it is possible to classify dark matter based on its possible production (thermal or non-thermal) or the nature of the particles that compose it or the mass of the dark matter particles. In general, dark matter can be classified based on whether it was thermally or non-thermally produced in the early Universe. In the case of thermal production, dark matter is produced through the collision of cosmic plasma in the radiation-dominated era.

Non-thermal dark matter particles can be produced by other mechanisms, such as the decay of massive particles or certain symmetry conditions that do not involve thermal production, etc. Dark matter can also be explained by considering that gravity behaves differently depending on the scale. In these models, gravity can behave differently from the predictions of general relativity, especially at galactic and cosmological scales, and there are other possible origins as well. In this work, we will focus on the scalar dark matter called ultra-light scalar dark matter[2]. We will discuss

[2]It is also found in the literature under names such as Wave Dark Matter, Fuzzy Dark Matter, Ultralight Dark Matter, Scalar Field Dark Matter, to name a few.

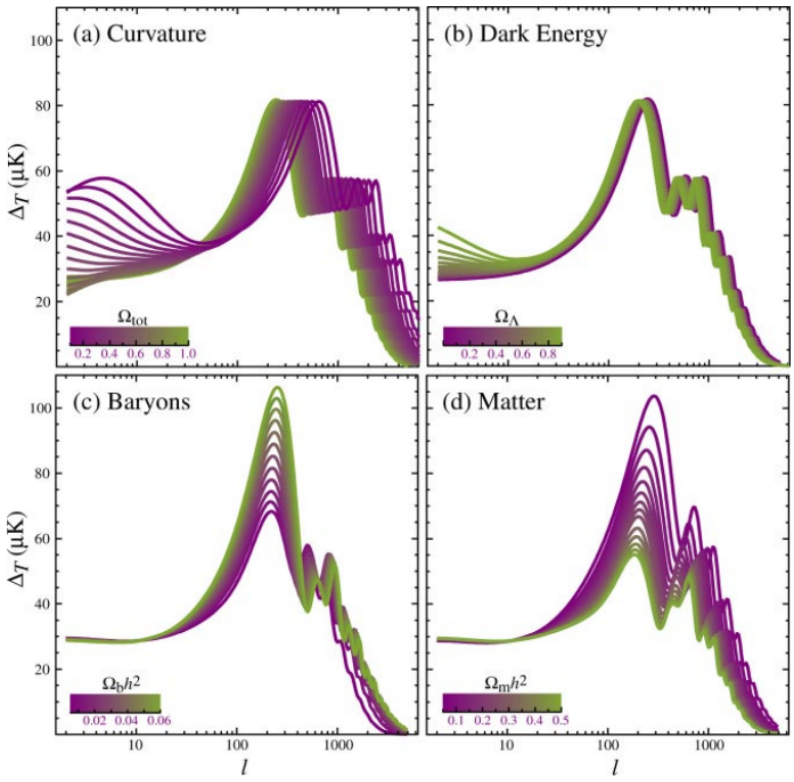

Figure 1.11: Sensitivity of the acoustic temperature spectrum to four fundamental cosmological parameters. (a) Curvature quantified by Ω_{tot}. (b) Dark energy quantified by the cosmological constant Ω_Λ. (c) Physical density of baryons $\Omega_b h^2$. (d) Physical density of matter $\Omega_m h^2$. All of them are varied around a fiducial model with $\Omega_{tot} = 1$, $\Omega_\Lambda = 0.65$, $\Omega_b h^2 = 0.02$, $\Omega_m h^2 = 0.147$, $n = 1$, $z_{ri} = 0$, $E_i = 0$. In the image, n corresponds to the spectral index, z_{ri} to the reionization redshift, and E_i to the inflation energy scale. Image taken from Hu and Dodelson (2002).

this topic later on.

Thermal Dark Matter:

This type of dark matter is thought to have been in thermal and chemical equilibrium with cosmic plasma. Due to the expansion of the Universe, this plasma cools down, and dark matter decouples, leaving the comoving density of dark matter constant. The relic density of this type of dark matter must be consistent with experimental constraints on the annihilation cross-section of dark matter. If we denote the dark matter particles as χ and their density as n_χ ($\bar{n}_\chi$ for antiparticles), a simple study can show that

$$n_\chi - \bar{n}_\chi \sim \left(\frac{m_\chi T}{2\pi}\right)^{3/2} \exp\left\{-\frac{m_\chi}{T}\right\},$$

where m_χ is the mass of χ, and T is the temperature of cosmic plasma. We can observe that if dark matter decouples at $T_f \neq 0$, there will be a relic of dark matter that cannot annihilate. Theoretical calculations of the relic density require the annihilation cross section of dark matter with relative velocity v to be of the order of

$$\langle \sigma_{\text{ann}} v \rangle \sim 10^{-26} \frac{\text{cm}^3}{\text{s}}.$$

For this reason, this type of dark matter is known as WIMPs (Weakly Interactive Massive Particles). Additionally, this type of matter must be stable, as otherwise it could be observed through its decay. The relic density is given in terms of the dark matter annihilation cross-section and is estimated to be

$$\Omega_\chi h^2 \approx \frac{3 \times 10^{-27} \text{cm}^3 \text{s}^{-1}}{\langle \sigma_{\text{ann}} v \rangle}.$$

For neutrinos, there is a constraint on the order of

$$\Omega_\nu h^2 \sim 0.03,$$

which rules out the possibility of neutrinos being the total contriburion dark matter.

Non-Thermal Dark Matter:

There is a possibility that the dark matter candidate was not in thermodynamic equilibrium with the rest of the matter in the Universe. For example, Axions must be of non-thermal nature, as otherwise, they would have already decayed. Axions arise from a symmetry called Peccei-Quinn, which is introduced to solve the *strong* CP problem that arises in Quantum Chromodynamics (QCD). This symmetry gives rise to pseudo-Goldstone bosons with a small mass ($\sim$ eV) for which the Standard Model decouples in the present Universe. There is a wide variety of dark matter candidates that arise when extending the Standard Model.

1.5.3 Cold Dark Matter Candidates

The Standard Model of particle physics is a fundamental theory that describes the interactions and properties of elementary particles and has been very successful in explaining a wide range of observed phenomena in the universe. The Lagrangian of the Standard Model, $\mathcal{L}_{SM}$, is composed of terms that describe electromagnetic interactions, weak interactions, strong interactions, and the dynamics of the Higgs field. The Lagrangian of the Standard Model can be written as:

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{W}} + \mathcal{L}_{\text{S}} + \mathcal{L}_{\text{Higgs}},$$

where $\mathcal{L}_{\text{EM}}$ represents the electromagnetic part, $\mathcal{L}_{\text{W}}$ represents the weak interactions, $\mathcal{L}_{\text{S}}$ represents the strong interactions, and $\mathcal{L}_{\text{Higgs}}$ describes the dynamics of the Higgs field. However, the Standard

Model does not provide a satisfactory explanation for the existence of dark matter, which is a form of matter that does not interact with light or electromagnetic and strong forces. To address this problem, various extensions to the Standard Model have been proposed that incorporate additional particles that could be candidates for dark matter. Throughout this discussion, we will explore two of these extensions: models with a scalar singlet and axions. These models modify the Lagrangian of the Standard Model by introducing new terms and symmetries, which potentially can explain the presence of dark matter in the universe. Any candidate particle for dark matter certainly does not belong to the Standard Model. Currently, there is a plethora of dark matter candidates. We will now discuss some examples of extended Standard Model models and possible candidates for dark matter as shown in Majumdar (2014).

Kaluza-Klein Dark Matter

Theories with extra dimensions provide candidates for dark matter. The Universe is 4-dimensional, so any extra dimension must be considered as compactified. The motivation behind these theories is to try to understand the unification of gravity with gauge interactions, the cosmological constant problem, the hierarchy problem, etc. In the context of quantum field theory and general relativity, the coordinates are denoted as x_μ, where $\mu = 0, 1, 2, 3$ for the temporal and spatial coordinates. If an extra spatial dimension y is considered, the Lagrangian density of a non-massive scalar field can be written as

$$\mathcal{L} = -\frac{1}{2}\partial_A \Phi \partial^A \Phi, \quad A = 0, 1, 2, 3, 4,$$

where the extra dimension is compactified on a circle of radius R. The coordinates can be decomposed and the scalar field can be written as $\Phi = \Phi(x_\mu, y)$, where x_μ are the usual spacetime components and y is the extra spatial coordinate. Since the extra dimension is compactified on a circle, when $y \to y + 2\pi$, it satisfies $\Phi(x, y) = \Phi(x, y + 2\pi)$. Therefore, we can expand it in a Fourier series and extract the extra dimension y:

$$\Phi = \sum_{n=-\infty}^{\infty} \phi_n(x) e^{\frac{iny}{R}}. \tag{1.12}$$

With these considerations, we can rewrite the Lagrangian as follows:

$$\begin{aligned} \mathcal{L} &= -\frac{1}{2}\partial_A \Phi \partial^A \Phi \\ &= -\frac{1}{2}\sum_{n,m}\left(\partial_\mu \phi_n \partial^\mu \phi_m - \frac{nm}{R^2}\phi_n \phi_m\right) \exp\left\{i(n+m)y/R\right\}. \end{aligned} \tag{1.13}$$

Now we can extract the zero mode term from the series described in Equation (1.12), and by substituting this into the Lagrangian (1.13), we have:

$$\mathcal{L} = -\frac{1}{2}\left[\partial_\mu \phi_0 \partial^\mu \phi_0 + \sum_{n,m}\left(\partial_\mu \phi_n \partial^\mu \phi_m - \frac{nm}{R^2}\phi_n \phi_m\right)\right] \exp\left\{i(n+m)y/R\right\}, \tag{1.14}$$

where, of course, $n, m \neq 0$. If we also consider the relation $\phi_n^* = \phi_{-n}$, it is possible to manipulate the action of this Lagrangian. By substituting (1.14) to integrate over the fifth coordinate, we can easily obtain that

$$\begin{aligned} S &= \int d^4x \int_0^{2\pi R} dy \left[-\frac{1}{2}\partial_\mu \phi_0 \partial^\mu \phi_0 - \frac{1}{2}\sum_{n,m}\left(\partial_\mu \phi_n \partial^\mu \phi_m - \frac{nm}{R^2}\phi_n \phi_m\right)\right] \exp\left\{i(n+m)y/R\right\} \\ &= \int d^4x \left[-\frac{1}{2}\partial_\mu \psi_0 \partial^\mu \psi_0 - \frac{1}{2}\sum_{n=1}^{\infty}\left(\partial_\mu \psi_n \partial^\mu \psi_n^* - \frac{n^2}{R^2}\psi_n \psi_n^*\right)\right], \end{aligned}$$

where it has been considered that $\psi_n(x) = \sqrt{2\pi R}\phi_n$, as well as the property

$$\int_0^{2\pi R} dy \exp\left\{\frac{i(n+m)y}{R}\right\} = 2\pi R\delta_{nm}.$$

The result is a Lagrangian density that contains an infinite sum of Klein-Gordon Lagrangian densities for massive scalar fields (Kaluza-Klein (KK) tower of excitations) and a Klein-Gordon Lagrangian density for a non-massive scalar field in four dimensions. According to this model, each field of the Standard Model can propagate in an extra dimension, and therefore, each one would have a KK tower. The proposed candidate for dark matter in this model is the particle B^1, which is the first KK companion of the hypercharge gauge boson B of the Standard Model and is known as the Lightest KK Particle (LKP). Since the fermions of the Standard Model are chiral, chiral fermions must be obtained in this type of theory. To satisfy this, the compactification of the extra dimension is considered on an S^1/Z_2 orbifold, where S^1 is the circular compactification and Z_2 is the reflection symmetry about the fifth coordinate $y \to -y$. With the reflection symmetry, the orbifold is simply reduced to the segment $0 \leq y \leq \pi R$. With this compactification, we can rewrite the field by considering two of its components, one even and one odd under Z_2:

$$\phi_+(x,y) = \sqrt{\frac{1}{\pi R}}\phi^0 + \sqrt{\frac{2}{\pi R}}\sum_{n=1}^{\infty}\cos\frac{ny}{R}\phi_+^n(x),$$
$$\phi_-(x,y) = \sqrt{\frac{2}{\pi R}}\sum_{n=1}^{\infty}\sin\frac{ny}{R}\phi_-^n(x).$$

These fields remain invariant under a transformation $y \to y + \frac{n\pi}{R}$ for even n (n-KK number), and for odd n, they have the factor $(-1)^n$, which is known as the KK parity and is conserved. The particle B^1 is stable as long as this parity is conserved.

Dark Matter in Models with a Scalar Singlet

By adding a real scalar singlet to the Lagrangian of the Standard Model with $SU(2)_L \times U_Y(1)$ symmetry, the Lagrangian density is written as

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{SM} + \partial_\mu S\partial^\mu S - \frac{\delta_1}{2}H^\dagger HS - \frac{\delta_2}{2}H^\dagger HS^2 \\ - \left(\frac{\delta_1 m^2}{2\lambda}\right)S - \frac{\kappa_2}{2}S^2 - \frac{\kappa_3}{3}S^3 - \frac{\kappa_4}{4}S^4.\end{aligned} \tag{1.15}$$

To avoid couplings with fermions, a Z_2 symmetry is imposed on the Lagrangian such that $\mathcal{L} \to \mathcal{L}$ when $S \to -S$. This implies that $\delta_1 = \kappa_3 = 0$. In the unitary gauge, we have

$$H = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix},$$

so that the term mixing the Standard Model Higgs with the new singlet is written as

$$\frac{\delta_2}{2}H^\dagger HS^2 = \frac{\delta_2}{2}\frac{v^2S^2 + h^2S^2 + 2vhS^2}{2},$$

from which we obtain a mass term for S given by $-\delta_2 v^2/4\, S^2$. On the other hand, from (1.15), we have the term $-\kappa_2/2S^2$, so that

$$m_S^2 = \kappa^2 + \frac{\delta_2 v^2}{2},$$
$$m_h^2 = \frac{\lambda v^2}{2}.$$

In the Lagrangian density (1.15), one can see the interaction vertex between two scalars S and a Higgs boson from the Standard Model, which serves as a gateway between dark matter and the Standard Model.

Axions

The problem faced in the 1970s was that in QCD there is an operator that violates CP symmetry

$$\mathcal{L} = -\bar{\Theta}\frac{\alpha_s}{8\pi}G^{\mu\nu a}\tilde{G}^a_{\mu\nu}.$$

It is known that CP symmetry is conserved at our energy scale, so $\bar{\Theta}$ must be extremely small, $\left\langle\bar{\Theta}\right\rangle < 10^{-10}$, which is problematic to explain. For N flavors of quarks, in the limit where $m_f \to 0$, we have a global symmetry $U(N)_V \times U(N)_A$, and since $m_u, m_d \ll \Lambda_{QCD}$, we expect at least QCD interactions to be invariant under $U(2)_V \times U(2)_A$. However, the breaking of $U(2)_A \simeq SU_A(2) \times U_A(1)$ symmetry would provide 4 Nambu-Goldstone bosons, but we only have three light mesons: the three pions, while the η meson is heavier ($m_\eta \gg m_\pi$). To remedy this problem, Roberto Peccei and Helen Quinn introduce a new symmetry $U_{PQ}(1)$, from which the aforementioned term can now be written as

$$\mathcal{L} = \left(\frac{\phi_A}{f_A} - \bar{\Theta}\right)\frac{\alpha_s}{8\pi}G^{\mu\nu a}\tilde{G}^a_{\mu\nu},$$

from which it follows that for $\phi_A = \bar{\Theta}f_A$, the potential has a minimum and ϕ_A acquires a very small mass, inversely proportional to $10^9\text{GeV} < f_A < 10^{12}\text{GeV}$. The study of this type of dark matter is done effectively, considering the first correction of electrodynamics for photon frequencies of the order $\omega \ll m_e$, from which

$$\mathcal{L}_{\phi_A\gamma\gamma} = -\frac{1}{4}G_{A\gamma\gamma}F_{\mu\nu}\tilde{F}^{\mu\nu}\phi_A.$$

For example, if there is an external electromagnetic field, the axion has the opportunity to oscillate and convert into a photon, only then can it couple to particles of the Standard Model, which is why it is a viable candidate for dark matter. However, as shown in Figure 1.12, this candidate has been ruled out for wide ranges of masses by different experiments, as seen in Bertone (2010).

Ultra-Light Scalar Dark Matter

Considering a homogeneous and isotropic universe, the spacetime is described by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. Furthermore, there is evidence that the universe is nearly flat, and we will take that into account in this analysis. In this case, the field ψ is also isotropic and homogeneous, so we have $\psi = \psi(t)$, and the Klein-Gordon equation becomes

$$\ddot{\psi} + 3H\dot{\psi} + m^2\psi = 0.$$

As we recall, this is the equation of motion for this scalar field. For the regime $H < m$, the field oscillates with a timescale smaller than the expansion of the universe (H^{-1}). In this case, the average observable energy density is $\langle\rho_\psi\rangle \sim a^{-3}$, which means it behaves like cold dark matter (CDM). Therefore, in such a regime, there is no way to distinguish between a conventional CDM model and this SFDM model, which is a consistency test at large scales. Moreover, in the standard cosmological model (ΛCDM), there must be an era dominated by matter to allow structure formation, which implies that at least $m \geq H_{\text{eq}}$, where H_{eq} denotes the Hubble constant at the radiation-matter equality epoch and serves as a lower bound for the mass according to the ΛCDM model.

Different analyses have been conducted at a cosmological level to study the compatibility of the SFDM model with observational data. In Figure 1.13, it is shown that this model is capable

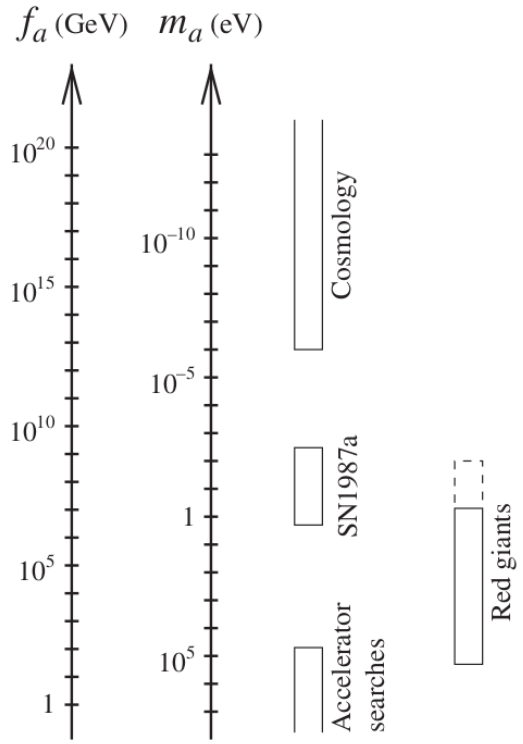

Figure 1.12: Mass ranges of the axion m_a, or equivalently the axion decay constant f_a, that have been ruled out by searches in accelerators, the evolution of red giant stars, the supernova SN1987a, and finally, the cosmological energy density of the axion. Image taken from Bertone (2010).

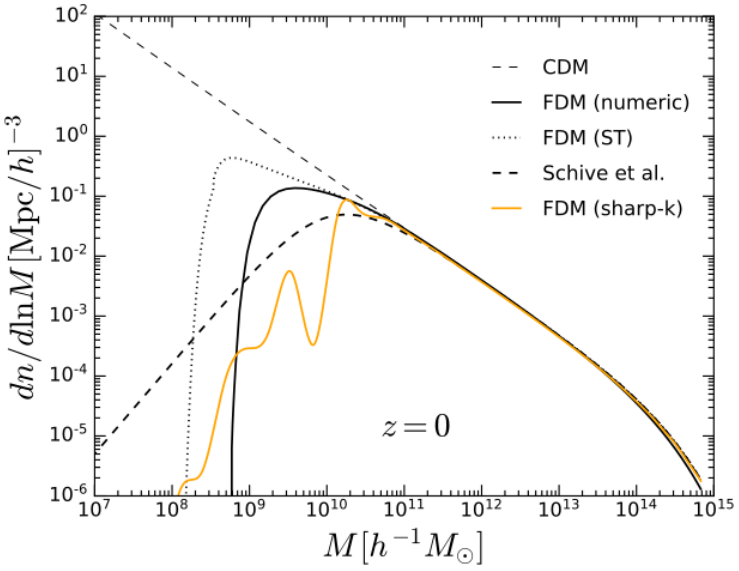

Figure 1.13: Figure taken from Niemeyer (2020) where more references can be found: Here, the mass function is shown at present time ($z = 0$) obtained for different SFDM models with $m_\psi \sim 10^{-22}$, and it is compared to CDM. It can be observed that at small scales ($m < 10^{10} M_\odot$), there is a natural suppression in contrast to the CDM model. This demonstrates how SFDM solves the missing satellites problem.

of predicting the suppression of small scales and resolving the missing satellites problem. On the other hand, in Figure 1.14, the CMB and the power spectrum $P(k)$ obtained for the SFDM model are shown and compared to CDM. It can be observed that in both cases (at large scales), these models are consistent.

1.6 Objectives, Goals, and Structure of the Thesis

The main objective of this thesis is to find solutions that provide us with halos of dark matter which will serve to evolve system of galaxies known as compact groups. In particular, we are interested in studying how to determine the initial conditions to study compact Hickson groups.z To achieve this, the model of ultralight scalar field dark matter will be used, solving the Schrödinger-Poisson system of equations in 3D. At the same time, numerical simulations of the same systems will be performed using the ΛCDM model with N-body simulations, in order to analyze and compare the evolution in both models.

Cosmologically, both models, ultralight scalar dark matter (SFDM) and the standard ΛCDM model, are compatible. However, there are discrepancies at galactic scales and in the non-linear regime, which give rise to different possibilities of evolution. Also, we considered the GDM model wich provide us with cosmological initial conditions slightly different than ΛCDM that will help us to model small-scale structures using different starting points.

Since dark matter is only detected through its gravitational effects, understanding its dynamics in galactic systems is based on speculation until the nature of this phenomenon is determined. However, rotation curves in spiral galaxies are observable properties that allow for estimating the distribution of dark matter in the system, understanding these structures is a fundamental part of this project.

To carry out this project, it is necessary to develop and study numerical tools for simulations and data processing. Numerical simulations involve studying methods to solve partial differential equations (PDEs) and using programming languages for massive calculations. Additionally, one

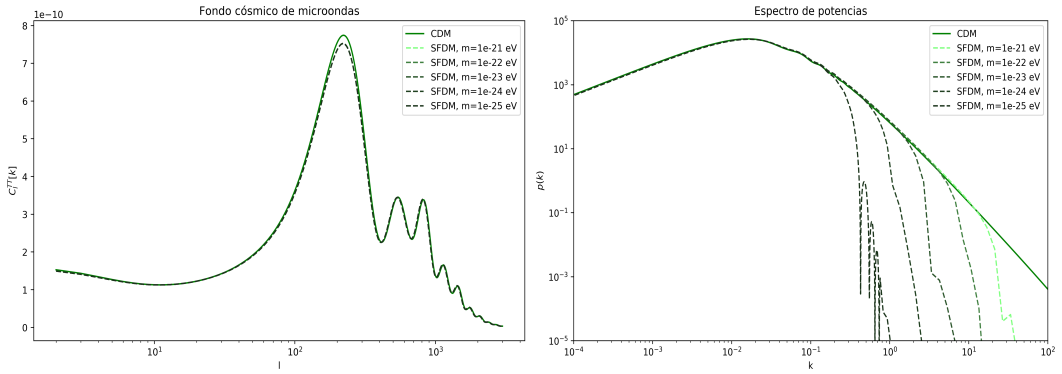

Figure 1.14: Left: CMB (cosmic microwave background) plot obtained for the SFDM model. Right: Power spectrum $P(k)$ obtained for SFDM compared to CDM considering different values of mass for the ϕ candidate.

must learn to compile, execute, and adapt existing codes for the desired purpose. Regarding data processing, it is necessary to understand the algorithms used and have the ability to modify them according to specific needs.

The specific objectives are, therefore, to numerically solve the Schrödinger-Poisson system in 3D, which is achieved using the finite difference method. The possibility of improving the solutions using the Adaptive Moving Mesh (AMM) method has also been explored, resulting in a product of this research in an article titled *Solving the Schrödinger-Poisson System using the coordinate Adaptive Moving Mesh method* published in the scientific journal Physical Review D. Additionally, two different projects are being prepared: one based on the resimulation technique and the other on the Coarse Graining method to study the dynamics of dark matter in the context of SFDM and CDM.

Perform cosmological and galactic-scale N-body simulations in the ΛCDM model, Generalized Dark Matter (GDM), and SFDM. N-body simulations have been carried out in the context of ΛCDM and GDM, and the results of this research have been published in the article titled *Compact Groups in GDM Cosmological Simulations*, published in the scientific journal Monthly Notices of the Royal Astronomical Society. The simulations for SFDM are presented as perspectives since an algorithm to create the initial conditions is yet to be created.

Study observational rotation curves using numerical techniques such as the Markov Chain Monte Carlo (MCMC) method. Currently, this project is being prepared, but the results obtained so far are reported in this text.

Study the ability of artificial intelligence algorithms to predict galaxy mass components such as bulge and disk. This research is in the final phase to be submitted for publication under the title *Machine learning algorithms to estimate the mass components of galaxies.*

In general, this text describes these objectives while supporting the reader with appendices describing the most important methods used in this project. First, the observational foundations of modern cosmology are presented, including the Hubble Law and the accelerated expansion of the universe, Big Bang nucleosynthesis, the Cosmic Microwave Background, and large-scale structures. Then, the ΛCDM cosmological model is described, focusing on background cosmology and the cosmological principle. The Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime is studied, and the Friedmann equations describing the evolution of the universe are analyzed. In addition, perfect fluids and particle species are examined, with a focus on recombination, the Cosmic Microwave

Background, and nucleosynthesis. An important aspect addressed in the thesis is the dark matter problem. Observational evidence of its existence is presented, possible candidates are explored, and the candidates for cold dark matter are further investigated. Furthermore, the

theory of dark matter as an imperfect fluid is explored, and the formation of structures in this context is analyzed.

This work is divided into four main parts. Firstly, in Chapter 2, we study how primordial perturbations give rise to the observed large-scale structure. We review why the temporal evolution of the universe cannot be studied from the perspective of linear perturbations alone and why numerical simulations are important. In Section 2.3, we present results obtained when studying compact groups of galaxies in a dark matter model treated as an imperfect fluid.

Next, in Chapter 3, we examine dark matter halos through cosmological simulations and observational data. In Section 3.1, a detailed description of the halos obtained from simulations is presented. Additionally, in Section 3.2, the methods used to estimate the contribution of dark matter from observational data are discussed. Furthermore, in Section 3.3, a disk-bulge decomposition is performed to gain a better understanding of halo characteristics using artificial intelligence algorithms.

Then, in Chapter 4, we address the study of dark matter on galactic scales through simulations of ultralight scalar field dark matter (SFDM). Section 4.1 provides a detailed description of the SFDM model, Section 4.2 presents numerical solutions of the Schrödinger-Poisson system, and Section 4.3 conducts stability tests of the solutions. Furthermore, in Section 4.4, a comparison with N-body simulations of some test initial conditions is made. Finally, in Section 4.5, the results obtained from simulations constructed with the halos obtained in Chapter 3 from cosmological simulations are shown.

Lastly, in Chapter 5, the conclusions of this work are presented, along with future prospects.

Chapter 2

Matter Perturbations in the Late Universe and Large-Scale Structure Formation

In this chapter, we will explore the field of matter perturbations in the late universe and the formation of large-scale structures. It is crucial to understand how small initial fluctuations evolve into the vast cosmic structures we observe today. We will examine both the linear and nonlinear regimes of perturbations, as well as different models of dark matter, to gain a deeper understanding of these cosmic phenomena.

2.1 Perturbation Theory

The theory of galaxy formation based on gravitational instability aims to describe how primordial fluctuations in radiation and matter grow into galaxies and galaxy clusters due to gravity. The field equations are derived from the Principle of Least Action applied to the so-called Einstein-Hilbert action, by varying the metric

$$\delta S = \delta \left(\frac{1}{2k}(R - 2\Lambda) + \mathcal{L}_{\text{Matter}} \right) \int d^4x \sqrt{-g},$$

here, $k = 8\pi G/c^4$, R is the Ricci scalar, Λ is the cosmological constant and $\mathcal{L}_{\text{Matter}}$ is the contribution of any matter extra field. The Einstein equations become

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu}, \tag{2.1}$$

where $T_{\mu\nu}$ is the energy-momentum tensor, which encapsulates the information of the energy sources. When we consider the FLRW metric (1.2), from the temporal part $\mu, \nu = 0$ and from the spatial part $\mu = i, \nu = j$, the Einstein equations become

$$\begin{aligned} G^0_{\ 0} &= 3\left(\frac{\dot{a}}{a}\right)^2 + \frac{3k}{a^2} = 8\pi G\rho, \\ G^i_{\ j} &= \delta_{ij}\left(-2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right) = -8\pi G \delta_{ij} P. \end{aligned} \tag{2.2}$$

Then, the FLRW equations (1.3) are reproduced. Perturbation theory can be formulated when the amplitude of the perturbations is small. The perturbations are characterized by two scalar

potentials ψ and ϕ which appear in the line element as

$$ds^2 = a^2(-(1+2\psi)d\tau^2 + (1-2\phi)dx^i dx_i).$$

For these perturbations, the Einstein tensor can be expressed as

$$\begin{aligned}
\delta G^0_{\ 0} &= \frac{2}{a^2}\nabla^2\phi - \frac{6}{a^2}H(\dot{\phi} + H\psi), \\
\delta G^0_{\ i} &= \frac{2}{a^2}(\dot{\phi} + H\psi)_{,i}, \\
\delta G^i_{\ j} &= -\frac{2}{a^2}\left[\ddot{\phi} + H(2\dot{\phi} + \dot{\psi}) + (2\dot{H} + H^2)\psi + \frac{1}{2}\nabla^2(\psi - \phi)\right]\delta^i_{\ j} + \frac{1}{a^2}\delta^{ik}(\psi - \phi)_{,kj}.
\end{aligned}$$

The stress tensor for a perfect fluid is defined by the following equation

$$T^\alpha_{\ \beta} = (\rho_m + p)u^\alpha u_\beta - p\delta^\alpha_{\ \beta},$$

where ρ_m is the mass density and p the fluid pressure. Perturbations of these components are purely diagonal and henceforth, using the equation (2.1) we can conclude that $\psi = \phi$ and the equations for the scalar perturbations are

$$\begin{aligned}
\nabla^2\phi - \frac{6}{a^2}H(\dot{\phi} + H\phi) &= 4\pi G a^2 \delta\rho_m, \\
(\dot{\phi} + H\phi)_{,i} &= 4\pi G a^2(\rho_{m0} + p_0)\delta u_{\parallel i}, \\
\ddot{\phi} + 3H\dot{\phi} + (2\dot{H} + H^2)\phi &= 4\pi G a^2 \delta p,
\end{aligned} \tag{2.3}$$

for a non-expanding universe ($H = 0$, $a = 1$), the first equation in (2.3) is the Poisson equation for a gravitational potential ϕ. Also, if we consider a perfect fluid $\delta S = 0$, so $\delta p = c_s\delta\rho$, with c_s the speed of sound. Using the last equation in (2.3) and the generalized Poisson equation we get a equation of motion for the gravitational potential ϕ

$$\ddot{\phi} + 3\dot{\phi}H(1 + c_s^2) + (2\dot{H} + H^2(1 + 3c_s^2) - c_s^2\nabla^2)\phi = 0.$$

2.1.1 Linear Regime

Currently (at $z \sim 0$), fluctuations cannot be considered small, as supermassive clusters typically have densities at least two orders of magnitude higher than the average density of the universe. Therefore, perturbations have been growing over time. In the linear regime, perturbations are small enough to be treated as linear perturbations in matter and the metric of the universe. We define the relative density contrast as

$$\delta(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t) - \bar{\rho}(t)}{\bar{\rho}(t)},$$

where $\bar{\rho}(t)$ is the average density of the universe and $\rho(\mathbf{r}, t)$ is the density at position $\mathbf{r}$ and time t. The anisotropies in the CMB suggest that for $z \sim 1000$, $|\delta| \ll 1$. Density fluctuations increase with time due to their self-gravity; overdense regions increase their density contrast over time, while underdense regions decrease their density contrast. In both cases, $|\delta|$ increases. Therefore, this gravitational instability effect leads to an increase in density fluctuations over time. The evolution of structure in the Universe is described by this gravitational instability effect. For simplicity, let us consider that matter in the universe can be treated as dust, i.e., pressureless matter with density $\rho(\mathbf{r}, t)$ and velocity $\mathbf{v}(\mathbf{r}, t)$. The equations of motion for this fluid are given by:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0, \\ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} &= -\frac{\nabla P}{\rho} - \nabla \Phi, \\ \nabla^2 \Phi &= 4\pi G\rho - \Lambda. \end{aligned} \tag{2.4}$$

The three equations represent the continuity equation, which states that the sum of the temporal rate of change of density and the divergence of the fluid flow is zero; the Euler equation, which describes how the velocity of an incompressible fluid changes over time and in the presence of forces such as pressure and gravity; and the Poisson equation, which states that the Laplacian of the gravitational potential Φ is proportional to the mass density ρ and the gravitational constant G, with a possible additional contribution given by the cosmological constant Λ. If we consider comoving coordinates and assume $\rho = \bar{\rho}(1+\delta)$, in the case of dust-like matter as in Equation (1.7), the fluid equations transform as:

$$\begin{aligned} \dot{\delta} + \frac{1}{a}\nabla \cdot [(1+\delta)\mathbf{u}] &= 0, \\ \dot{\mathbf{u}} + \frac{\mathbf{u} \cdot \nabla}{a}\mathbf{u} + \frac{\dot{a}}{a}\mathbf{u} &= -\frac{1}{a\bar{\rho}}\nabla P - \frac{1}{a}\nabla \Phi, \\ \nabla^2 \phi &= \frac{3H_0^2 \Omega_m}{2a}\delta. \end{aligned} \tag{2.5}$$

For the Poisson equation, we consider the potential to be the sum of the cosmological constant solution and the term ϕ, which is the solution to the equation for the inhomogeneities:

$$\Phi(\mathbf{x}, t) = \frac{\ddot{a}a}{2}|\mathbf{x}|^2 + \phi(\mathbf{x}, t).$$

When we approximate these equations in the limit where $|\delta| \ll 1$ and neglect quadratic terms (linear regime), it can be shown that the density contrast has solutions of the form:

$$\delta(\mathbf{x}, t) = D(t)\bar{\delta}(\mathbf{x}), \tag{2.6}$$

meaning that the spatial dependence can be separated from the temporal dependence. Generally, there are two solutions for $D(t)$, one that increases with time and one that decreases. However, due to observational constraints and the specific problem studied in this section, the decreasing part is irrelevant compared to the increasing part, as we mentioned before that the density contrast tends to increase in absolute value. This solution is denoted as D_+ and is known as the growth factor. In general, perturbations can be considered to qualitatively grow in the same way as the scale factor, as shown in Figure 2.1.

In general, various perturbations to the FLRW metric can be considered. The metric with these perturbations is known as the perturbed FLRW metric and is defined as follows:

$$ds^2 = a^2 \left\{ -(1+\Psi)d\tau^2 - \vec{\nabla}_i \xi d\tau dx^i + \left[\left(1 + \frac{1}{3}h \right) \gamma_{ij} + D_{ij}\nu \right] dxdy \right\},$$

where $a(\tau)$ is the conformal time scale factor, γ_{ij} is the metric for a three-dimensional space with constant curvature k, $\vec{\nabla}_i$ is the covariant derivative of γ_{ij}, and $D_{ij} = \vec{\nabla}_i \vec{\nabla}_j - \frac{1}{3}\gamma_{ij}\vec{\nabla}^2$.

For a general fluid with viscosity $\omega \neq 0$ and pressure $P \neq 0$, the fluid equations in the linear regime and in Fourier space can be written as Kopp et al. (2016):

$$\begin{aligned} \dot{\delta} &= 3H(\omega\delta - \Pi) - (1+\omega)\left[k^2(\boldsymbol{u} - \xi) + \frac{1}{2}\dot{h} \right], \\ \dot{\boldsymbol{u}} &= -(1 - 3c_{\mathrm{a}}^2)H\boldsymbol{u} + \frac{\Pi}{1+\omega} - \frac{2}{3}(k^2 - 3k)\Sigma + \Psi. \end{aligned} \tag{2.7}$$

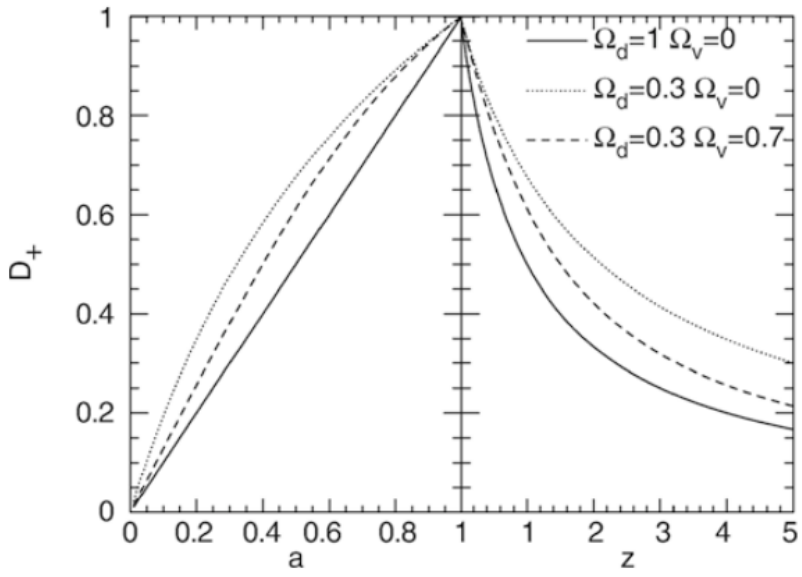

Figure 2.1: Growth factor D_+ for three different cosmological models, as a function of the scale factor a (left panel) and redshift (right panel). It is clearly visible how D_+ rapidly decreases with increasing redshift in the Einstein-de Sitter model, compared to the lower-density models. Additionally, it can be observed that the scale factor and the growth factor exhibit a strong similarity, with $D_+ \sim a$. Image extracted from Schneider (2015).

2.1.2 Power Spectrum

First, let's study the probability of finding a galaxy at a point $\boldsymbol{x}$, given that galaxies are not uniformly distributed in the universe, so the probability is not position-independent. For example, it is more likely to find a galaxy next to another galaxy than at an arbitrary point, since galaxies form galaxy clusters. Let's consider the probability of finding two galaxies at two points $\boldsymbol{x}$, $\boldsymbol{y}$ within the volume dV. First, the probability of finding one galaxy is given by

$$P_1 = \bar{n}dV,$$

where $\bar{n}$ is the average number density of galaxies. However, the probability of finding a second galaxy P_2 given the first galaxy is not simply P_1^2 because, as mentioned before, the probabilities are not independent. In fact,

$$P_2 = (\bar{n}dV)^2 \left[1 + \xi_g(\boldsymbol{x}, \boldsymbol{y})\right],$$

where $\xi_g(\boldsymbol{x}, \boldsymbol{y})$ is known as the two-point correlation function. Now, in reality, the dependence of ξ_g on the points $\boldsymbol{x}$, $\boldsymbol{y}$ is actually on the distance $r = |\boldsymbol{x} - \boldsymbol{y}|$, since the universe is considered homogeneous. Observations of this function are shown in Figure 2.2. In that image, an interesting phenomenon called baryonic acoustic oscillations (BAO) is shown. This is a signature left by the Big Bang through baryons via expansion. When the universe was very dense and hot, the cosmic soup was in thermodynamic equilibrium. However, as it cooled down, particles decoupled from this cosmic soup, leaving behind a wavefront that redistributed matter, giving rise to this observed correlation function. If we consider density perturbations as a sum of plane waves $\delta(\boldsymbol{x}) = \sum a_k \cos \boldsymbol{a} \cdot \boldsymbol{x}$, the power spectrum $P(k)$ is related to the correlation function through a Fourier transform:

$$P(k) = 2\pi \int_0^\infty dx x^2 \frac{\sin kx}{kx} \xi(x).$$

On the other hand, when considering the growth function (2.6), the power spectrum can be

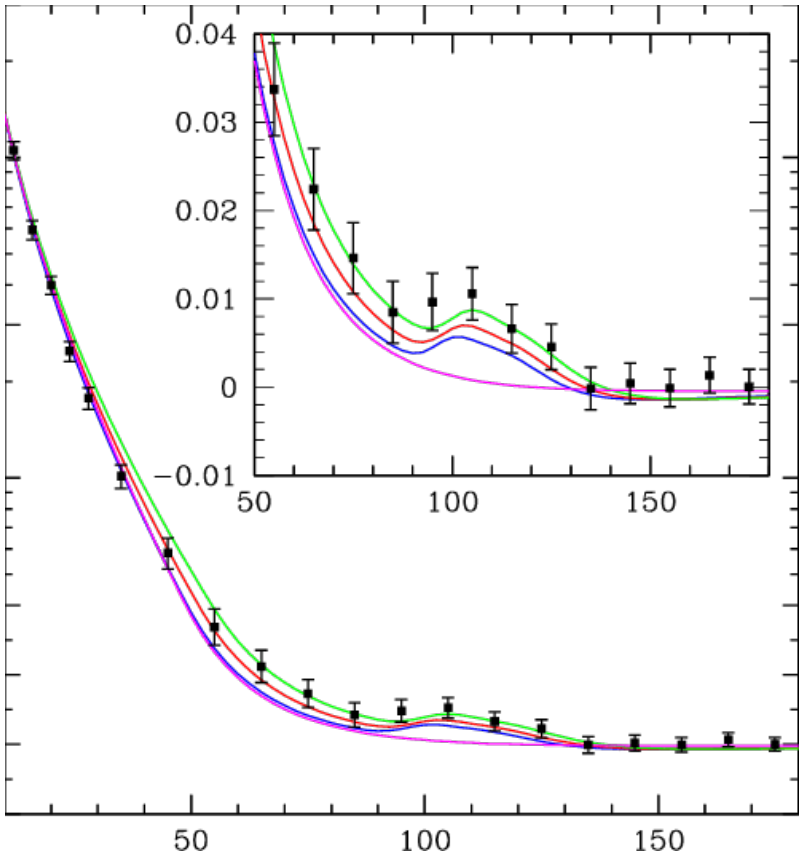

Figure 2.2: The spherically-averaged galaxy correlation function measured from the SDSS data, clearly showing a peak corresponding to the BAO scale at $\sim 100h^{-1}$ *Mpc. The lines shown correspond from top to bottom to models with* $\Omega_m h^2 = 0.12$*, 0.13, 0.14 respectively, with* $\Omega_b h^2 = 0.24$*, the last line is a model with* $Omega_b h^2 = 0$ *and* $\Omega_m h^2 = 0.105$*. Image obtained from*

rewritten as

$$P(k,t) = D_+^2(t)P_0(k),$$

defined in terms of $P_0(k)$, which is known as the primordial power spectrum. The primordial power spectrum is approximated by the Harrison-Zeldovich function

$$P(k,t) = D_+^2(t_i)Ak^{n_s}.$$

This is valid in the context of a perfect fluid. However, when studying other models such as warm dark matter, it is necessary to take into account additional effects, which are encapsulated in the transfer function $T^2(k)$.

2.2 Nonlinear Regime

As perturbations grow and become more pronounced, we enter the nonlinear regime. In particular, the density contrast δ increases, giving rise to massive structures. These structures cannot be described in the linear fluid limit, as it requires capturing collisional information that a fluid cannot efficiently capture. Gravitational interactions between density and velocity perturbations become dominant and give rise to processes such as gravitational collapse, formation of dark matter halos, and the evolution of complex structures. To alleviate these difficulties, different numerical methods are employed, such as N-body simulations. I

n this case, the flow equations cannot be linearized and must be numerically treated using advanced approximation techniques, such as higher-order approximations Rafatullah (2020) or the Smoothed Particle Hydrodynamics (SPH) method Gingold and Monaghan (1977). Analytical considerations, such as linear perturbation theory or the spherical collapse model, can only describe limiting cases of structure formation. In general, gravitational dynamics are too complicated to be examined analytically in detail. For this reason, experiments have been conducted to simulate structure formation using numerical methods for some time. Comparing the results of these simulations with observations has significantly contributed to establishing the standard model of cosmology, as only through them it was possible to quantitatively distinguish the predictions of this model from those of other models. Since the Universe is dominated by dark matter, for many purposes, it is sufficient to calculate the behavior of this matter component and consider only gravitational interactions.

Currently, there are codes that allow for these N-body simulations, such as Springel et al. (2021), which is a code that simulates dust-like dark matter. However, efforts have also been made to simulate other dark matter candidates, such as the case of Fuzzy Dark Matter Zhang et al. (2018), which considers a quantum pressure term in the equations. On the other hand, there is currently a surge of interest in an ultra-light Cold Dark Matter (CDM) candidate. This candidate has its origins in string theory Liddle and Ureña López (2006); Liddle et al. (2008); Pérez-Lorenzana et al. (2008); Hui et al. (2017) and is attractive due to its quantum nature and the potential influence it could have on the formation of large-scale structures Zhang et al. (2019).

2.2.1 CDM as an N-body System

In this section, we will focus on the standard case of cold dark matter (CDM) of dust-like nature. We will examine the methodology used to study the nonlinear evolution of the universe as an N-body system of dark matter. Additionally, we will review the challenges and issues that arise in this model at small scales, such as the lack of observed satellites, the "too-big-to-fail" phenomenon, and the discrepancy in density profiles ("core-cusp" problem).

The cold dark matter (CDM) model is based on several assumptions to study structure formation.

Representative Particles: First, individual dark matter particles are not tracked in the simulations due to their high numerical density. Instead, they are represented by mass bodies M that behave as macroscopic particles in a volume $V = M/\rho$. A representative section of the Universe is chosen in the form of a comoving cube with a side length L, which should be larger than the largest observed structures in the Universe to capture large-scale effects. However, this section cannot encompass all large structures in the Universe due to its low spatial density. The numerical effort required to compute gravitational forces in the simulation scales with the number of grid points N_{grid} and is limited by computational capacity.

The choice of L determines the numerical resolution scale and the minimum mass that can be resolved in the simulation since the total mass within the numerical volume is proportional to $\Omega_m L^3$. Additionally, particles near the boundaries of the cube also experience gravitational forces from matter outside the cube, so it cannot be assumed that the region outside the cube is empty. Assumptions must be made about the distribution of matter outside the numerical volume. To address this, it is assumed that the Universe is homogeneous on scales larger than L, and thus, the cube extends periodically.

The mass distribution and force field are periodic in these simulations, with a period of L. However, this periodicity assumption affects the mass distribution results on scales comparable to L, so quantitative analysis is limited to scales smaller or equal to $L/2$. With all these assumptions, the equation of motion is established for all particles, where the force on the i-th particle F_i is the sum of forces exerted by all other particles j using the law of universal gravitation:

$$F_i = \sum_{i \neq j} \frac{M^2 (r_j - r_i)}{|r_j - r_i|^3}, \tag{2.8}$$

where r_i and r_j are the position vectors of the i-th and j-th particles, respectively.

P^{3}M Method: The precise calculation of the force acting on each individual particle using the summation in equation (2.8) is not feasible in practice due to the large number of terms that need to be computed. For example, if the simulation tracks 10^{10} particles, it would be necessary to compute 10^{20} terms at each time step, which is not feasible even on the most powerful computers available today.

To solve this problem, an approximation is used for force calculation. It is observed that the force experienced by a particle, exerted by another particle, is not very sensitive to small variations in the separation vector $r_i - r_j$ as long as these variations are much smaller than the separation itself. Except for the closest particles, the force on a particle can be computed by introducing a grid into the cube and shifting the particles in the simulation to the nearest grid point.

This way, a discrete mass distribution on a regular grid is obtained. The force field of this mass distribution can be computed using a Fast Fourier Transform (FFT), which is an efficient and fast algorithm. However, the introduction of the grid imposes a lower limit on the spatial resolution of the force. The size of the grid cells defines the spatial resolution of the force field, so it is chosen to be approximately equal to the average separation between two particles, such that the number of points in the grid is typically of the same order as the number of particles. This method is known as the Particle-Mesh (PM) method.

Softening length: To achieve better spatial resolution, the interaction of nearby particles can be considered separately. This is accomplished by dividing the gravitational potential $\phi(r) = -\frac{GM}{r}$ of a particle into a short-range part and a long-range part, $\phi = \phi_s + \phi_l$. For example, one can choose $\phi_s = \phi(r) f(r/r_s)$, where the function f smoothly decreases from $f(0) = 1$ to $f(1) = 0$, and $f(x) = 0$ for $x > 1$. In this way, the short-range gravitational potential $\phi_s(r)$ becomes zero for $r > r_s$. On the other hand, the long-range potential is $\phi_l(r) = \phi(r)[1 - f(r/r_s)]$, which means it vanishes at $r = 0$, while for $r > r_s$, $\phi_l = \phi$. The force on a particle is calculated as the sum of the gradients of the short-range and long-range potentials. For the short-range potential, only

particles with a separation $\leq r_s$ contribute, and this can be computed as a sum of pair forces. On the other hand, the force field corresponding to ϕ_l is smooth and is calculated using the P^3M method.

The force model (2.8) also accounts for strong collisions between particles, for example, when a particle changes the direction of its velocity by approximately 90° in a collision if it gets close enough to another particle. Of course, this effect is a consequence of replacing the constituents of dark matter with macroscopic particles of mass M. As we have seen before, the typical relaxation time for a system is $\propto N/\ln N$, and since the mass in the numerical volume is defined by L, we have $N \propto 1/M$. By reducing the mass of the particles and increasing N accordingly, the abundance of strong collisions would decrease, but computational power and computer memory become limiting factors. Therefore, to correct the effect of strong collisions, the force law is modified for small separations so that strong collisions no longer occur.

The length scale below which the force equation is modified (smoothed) and deviates from $\propto 1/r^2$ is called the *softening length*. Its choice depends on the method used to evaluate the force field. If the force is computed on a mesh, as in the PM method, where the force resolution is limited by the size of the mesh cells, the smoothing length is usually chosen to be of similar size. On the other hand, for methods similar to P^3M, the smoothing length can be chosen considerably smaller, although at the expense of requiring smaller time steps. Of course, the smoothing length defines the smallest length scale at which the simulation results can be considered reliable: scales below or comparable to the smoothing length are unresolved, and the behavior at these small scales is affected by numerical artifacts.

Initial conditions: The initial conditions for the simulation are set at a very high redshift. Then, the particles are distributed in such a way that the power spectrum (see Figure 2.3) of the resulting mass distribution resembles a Gaussian random field with the theoretical (linear) power spectrum $P(k, z)$ of the cosmological model. For example, Figure 2.3 shows the power spectrum extrapolated to $z = 0$ with the best-fitting ΛCDM model based on observations. The equations of motion for the particles with the force field described above are integrated over time. The choice of the time step is a critical matter in this integration, as the force on particles with relatively close neighbors will change more rapidly than on quite isolated particles. Therefore, the time step is chosen to be short enough for the first particles, which requires substantial computation time, or the time step varies for different particles individually, which is clearly the most efficient strategy. At different moments in the evolution, the positions and velocities of the particles are stored; these results are available for further analysis.

2.2.2 Problems at Small Scales

Evidence of the viability of the ΛCDM model has been shown, for example, in sections 1.5.1 and 1.1 or in figures 1.4 and 2.3, where it can be observed how the ΛCDM model fits well with observations from different approaches. However, when examining in detail the consequences of the model at galactic scales, some discrepancies are found, which are listed below:

The Missing Satellites Problem: When conducting numerical simulations, it is found that halos form hierarchically, starting with the smallest halos. Temporal evolution causes these halos to form larger halos through collisions, and the abundance of these halos is quantified by the mass function, which is a histogram of the halo masses. Figure 2.4 shows the results obtained for the mass function in Bose et al. (2015). In that work, results from CDM and WDM simulations are compared. The initial conditions are set using the power spectrum generated for each model. It is interesting how a small variation in the model can alleviate such a challenging theoretical problem. These modifications can be as simple as modifying the dark matter model or as complex as considering baryonic dynamics, or a combination of both approaches Lovell et al. (2017); Tomozeiu et al. (2016).

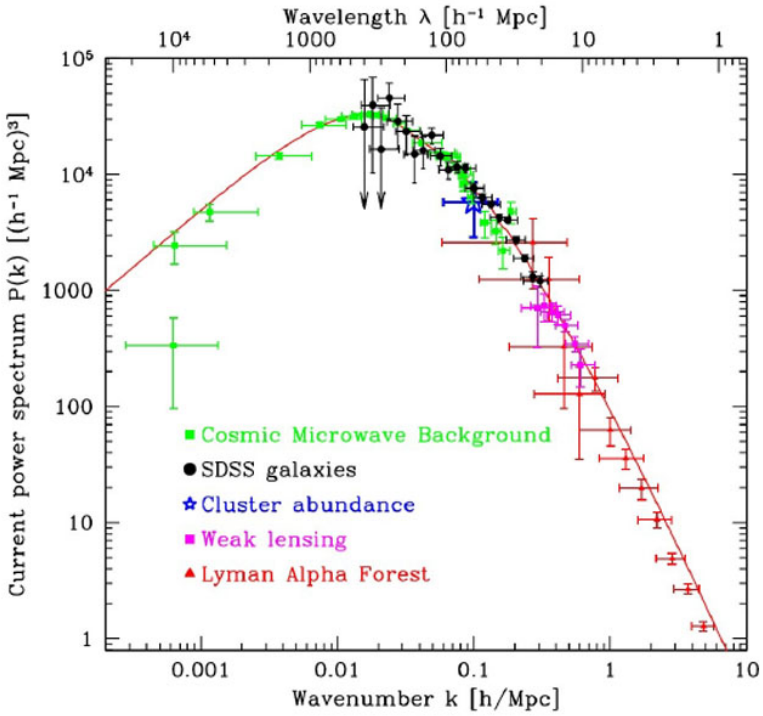

Figure 2.3: Linear matter power spectrum, $P(k)$, as a function of wave number extrapolated to $z = 0$, from various measurements of the cosmological structure. The best-fitting ΛCDM model is shown as a solid line (taken from Tegmark et al. (2004)). Image extracted from `https://ned.ipac.caltech.edu/level5/Sept11/Norman/Norman2.html`

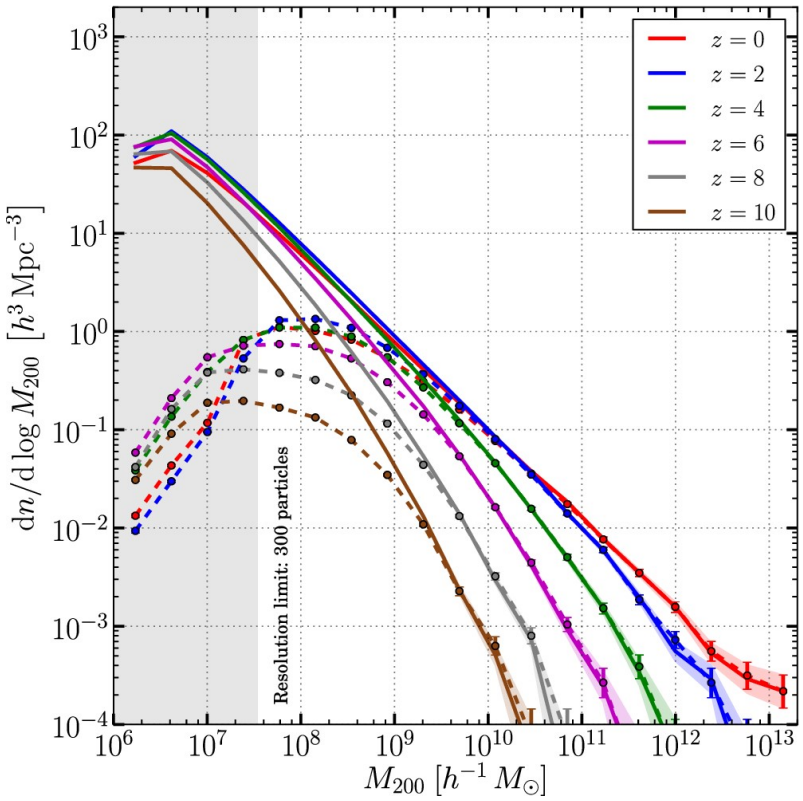

Figure 2.4: Halo Mass Function (HMF), showing the redshift evolution for a cosmological simulation in Gadget2 Springel (2005). The mass function for the CDM model exhibits a divergent behavior for $M \ll 1$, known as the Missing Satellites Problem, as the model predicts an increasing number of small-mass halos. The mass function is compared to the warm dark matter (WDM) model, where a modification to the ΛCDM model can alleviate this problem. Image extracted from Bose et al. (2015).

The Too Big to Fail Problem: The Too Big to Fail problem refers to financial institutions that, due to their size and complexity, could cause catastrophic consequences if they were to collapse. It is used similarly in cosmological N-body simulations. The problem arises when simulations predict more massive dark matter halos in small galaxies than what is observed. These galaxies should have very little dark matter compared to the predictions obtained from N-body simulations, so according to the numerical predictions, these galaxies should not be able to form stars. In other words, in the case of some observed dwarf galaxies, they seem to have significantly less dark matter than what would be expected according to the simulations Boylan-Kolchin et al. (2011). It is speculated that some unknown mechanism counteracts gravity, preventing stellar collapse. Scenarios such as the introduction of baryons Tomozeiu et al. (2016) or alternative models to ΛCDM Lovell et al. (2017) have been proposed, but there is no consensus on a definitive solution.

The Core-Cusp Problem: The core-cusp problem refers to the discrepancy between theoretical predictions and observations of the dark matter density distribution in galaxies. Cosmological N-body simulations predict a divergent distribution at the center, while observations indicate a smoother or core-like distribution in low-luminosity galaxies. Several explanations have been proposed, such as thermal or hydrodynamic interactions, interactions with undiscovered subatomic particles, and modifications to the theory of gravity. However, a definitive consensus on its solution in cosmological N-body simulations has not yet been reached.

The empirical profile that best fits N-body cosmological simulations is the Zhao profile Zhao (1996), which is a generalization of the Navarro-Frenk-White (NFW) profile Navarro et al. (1996). The NFW profile is recovered by considering $\alpha = 1$, $\beta = 3$, $\gamma = 1$ in the Zhao profile. On the other hand, the soliton proposed in Schive et al. (2014a,b); Schive and Chiueh (2018) is also considered. For Figure 2.5, only a soliton in the SFDM model is considered. On the other hand, Figure 2.6 shows the reported fits in Oh et al. (2011) using the isothermal profile considering $\alpha = 2$, $\beta = 2$, $\gamma = 0$ in the Zhao profile (2.9). In general, Zhao profiles with $\gamma = 0$ correspond to flattened profiles at the center that are free of divergences.

$$\begin{aligned} \rho_{\rm Zhao}(r|r_s,\rho_s,\alpha,\beta,\gamma) &= \frac{\rho_s}{\left(\frac{r}{r_s}\right)\left[1+\left(\frac{r}{r_s}\right)^{\alpha}\right]^{\frac{\beta-\gamma}{\alpha}}}, \\ \rho_{\rm Sol}(r) &= \frac{\rho_c}{[1+0.091(r/r_c)^2]^8}. \end{aligned} \tag{2.9}$$

There are studies that report that ΛCDM can reproduce non cuspy profiles when intriducing baryonic effects. For example, the results reported in Veltmaat et al. (2020) shows how the baryonic component affects the dark matter halo formed after some evolution (see Fig. 2.7). For ΛCDM the gas in SFDM prevents the halo to show the NFW tail the core shape is more pronounced with time. These sort of effects are very interesting and we wanted to see what happends with groups of galaxies.

2.3 Dark Matter as a non-perfect fluid

In this section, we study the evolution of matter in a universe where there are contributions from viscosity and non-zero sound speed in the system. Our main objective is to explore how perturbations in this scenario evolve in the non-linear regime. To carry out this study, we rely on the scientific article titled "Compact groups in GDM cosmological simulations" López-Sánchez et al. (2022), which is presented as one of the results obtained within the framework of this research project. The article provides a detailed and rigorous analysis of the implications of these additional conditions on the dynamics of matter in the universe.

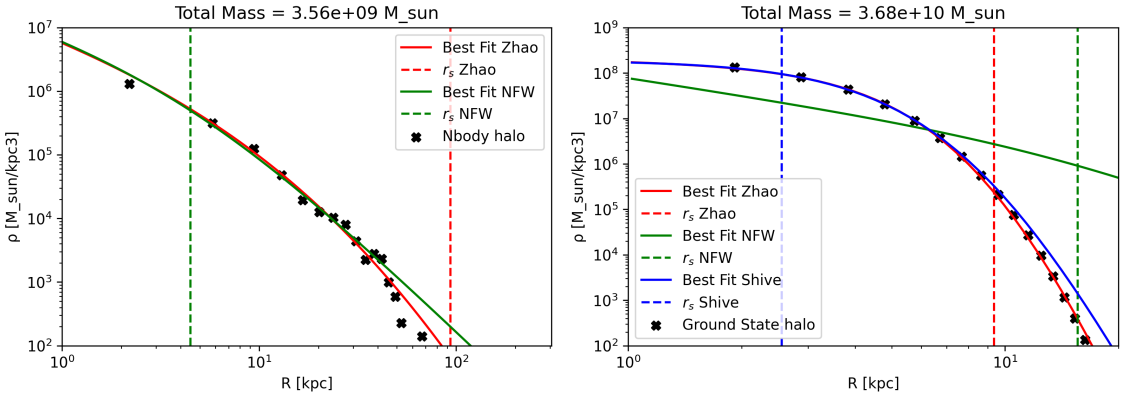

Figure 2.5: Fits for densities obtained from different models. For a halo from a cosmological simulation, the best fit for two empirical profiles, Zhao and NFW ($\alpha = 1$, $\beta = 3$, $\gamma = 1$) (2.9) (left), and for a simulation in the SFDM model, the best fits for the Zhao, NFW ($\alpha = 1$, $\beta = 3$, $\gamma = 1$), and Schive (2.9) profiles are shown (right). It can be observed that for CDM-like simulations, halos form with a density that diverges as $r \to 0$, while in models like SFDM, the profile found is flattened and does not diverge at the center. The characteristic radii of each profile r_s are shown as vertical lines.

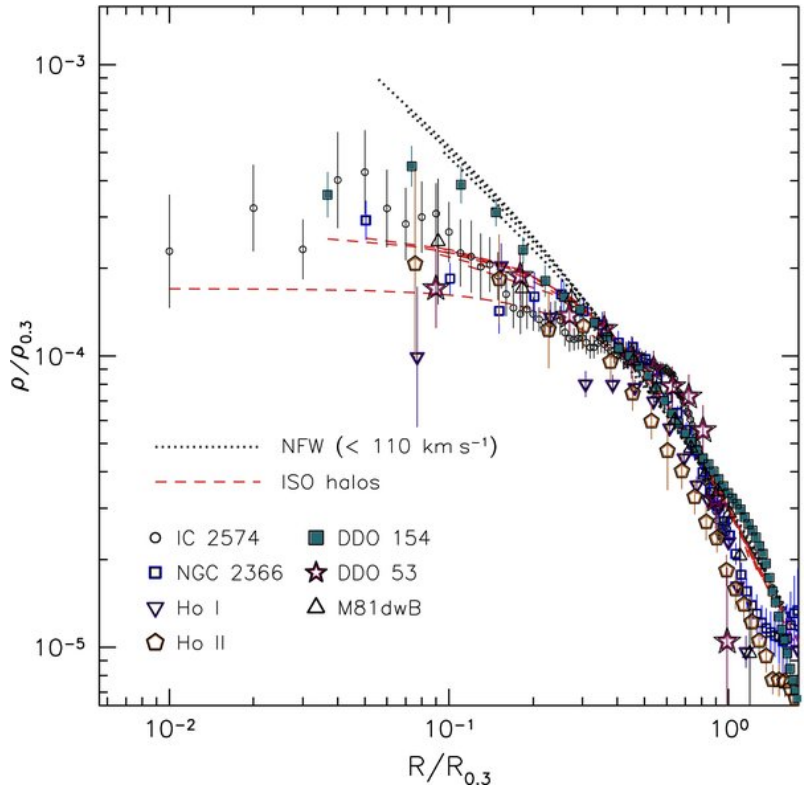

Figure 2.6: Dark matter density profiles of the seven dwarf galaxies from THINGS (The H_I Nearby Galaxy Survey). The profiles are obtained using normalized rotation curves. The dashed lines represent the mass density profiles of NFW models with V_{200} ranging from 10 to 110 km s^{-1}. The dotted lines indicate the mass density profiles of the best-fit pseudo-isothermal halo models ($\alpha = 2$, $\beta = 2$, $\gamma = 0$) in (2.9). This image shows how the NFW-like profiles deviate too much from observational data. Image extracted from Oh et al. (2011).

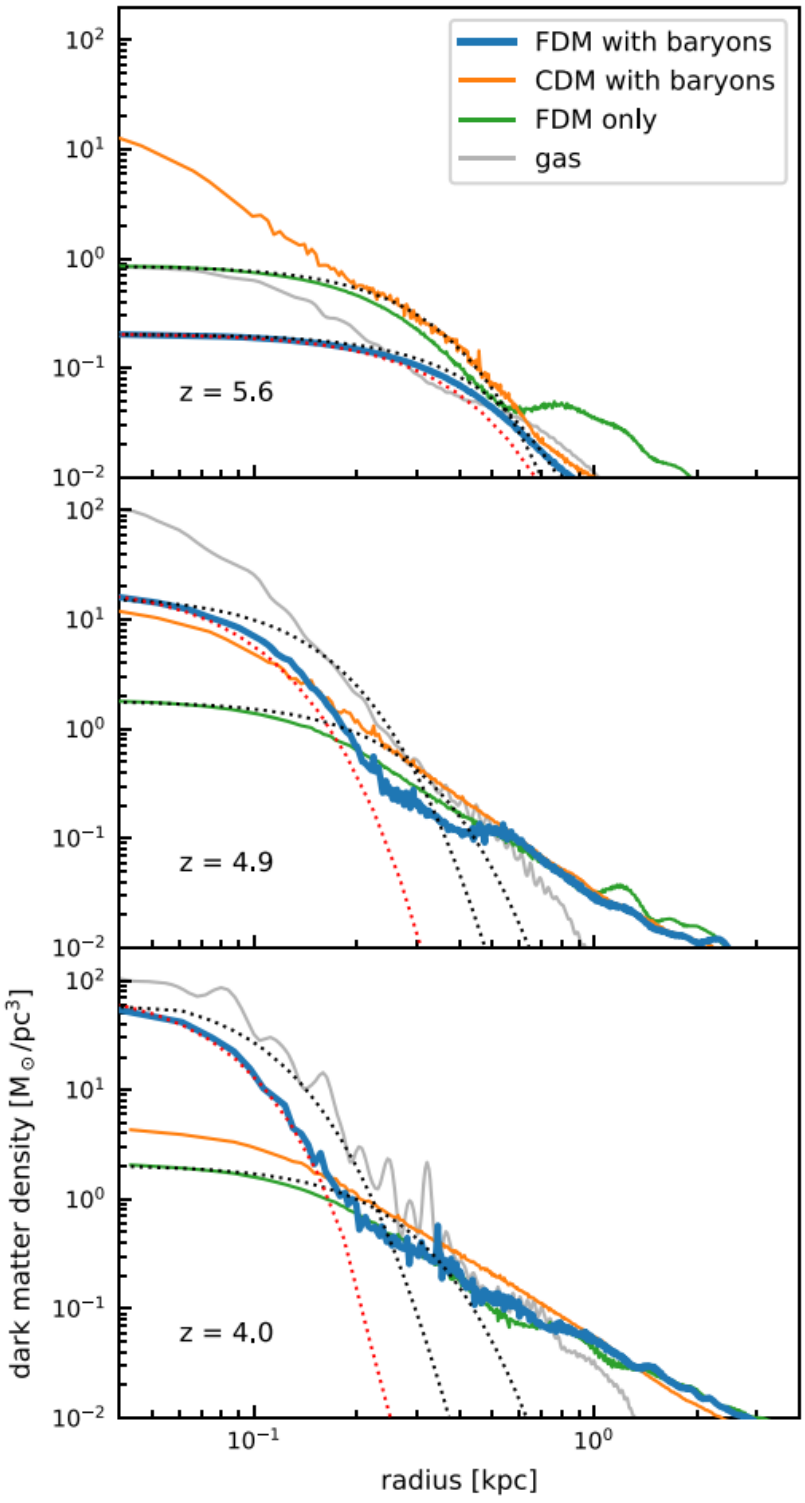

Figure 2.7: Radial dark matter density profiles of halo 1 in all three runs at three different redshifts. The inner profile of the FDM run with baryons matches the modified FDM ground-state solution (red dotted line) instead of the dark-matter-only ground-state solution (black dotted line). Also shown is the gas density profile of the FDM run with baryons. Image and description extracted from Veltmaat et al. (2020)

2.3.1 GDM Theory

The GDM (Generalized Dark Matter) model describes a fluid that allows for a non-zero pressure term and shear compared to a perfect fluid, where these quantities are zero. The energy-momentum tensor for GDM includes additional parameters:

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu} + \Sigma_{\mu\nu},$$

where $\Sigma_{\mu\nu}$ represents the additional perturbations controlled by the GDM parameters. The equation of state relates the background pressure and energy densities, while $c_{\rm s}^2$ and $c_{\rm vis}^2$ control the pressure perturbation Π and the scalar anisotropic stress Σ, respectively. These quantities are expressed through closure equations:

$$\Pi_g = c_a^2 \delta + (c_{\rm s}^2 + c_a^2)\hat{\Delta}_g,$$
$$\dot{\Sigma}_g = -3H\Sigma_g + \frac{4}{1+w} c_{\rm vis}^2 \hat{\Theta}_g,$$

where $\hat{\Delta}$ and $\hat{\Theta}$ are gauge-invariant density and velocity perturbations for GDM. The label g refers to the GDM fluid. The GDM-free functions are defined in physical terms and their values determine the structure formation in a GDM universe. The decay of the gravitational potential at small scales is determined by the values of $c_{\rm vis}^2$ and $c_{\rm s}^2$, given by

$$k_{\rm dec}^{-1}(\eta) \sim \eta \sqrt{c_{\rm s}^2 + \frac{8}{15} c_{\rm vis}^2}, \tag{2.10}$$

where η is the conformal time. If $c_{\rm s}^2 \gg c_{\rm vis}^2$, the gravitational potential oscillates below the Jeans scale. Conversely, $c_{\rm vis}^2$ damps the density perturbation without causing oscillations. The values of these free parameters determine the dynamics of the perturbations and the shape of the Matter Power Spectrum. In Kopp et al. (2016), the authors provide a detailed analysis of the implications of these additional conditions on the dynamics of matter in the universe.

Matter Power Spectrum

In this work, we consider three different scenarios for GDM, denoted as GDM I, GDM II, and GDM III. The values for the GDM parameters $c_{\rm vis}^2$, $c_{\rm s}^2$, and w in each scenario are based on the constraints reported in Thomas et al. (2019):

- **GDM I**: $w = -1 \times 10^{-6}$, $c_{\rm s}^2 = 1 \times 10^{-7}$, and $c_{\rm vis}^2 = 1 \times 10^{-7}$,
- **GDM II**: $w = -1 \times 10^{-6}$, $c_{\rm s}^2 = 1 \times 10^{-7}$, and $c_{\rm vis}^2 = 1 \times 10^{-6}$,
- **GDM III**: $w = 6 \times 10^{-4}$, $c_{\rm s}^2 = 1.92 \times 10^{-6}$, and $c_{\rm vis}^2 = 1.1 \times 10^{-7}$.

The Matter Power Spectrum was calculated using the Boltzmann code GDM-Class[1] (Ilić et al., 2021), which is a modified version of the code CLASS[2] (Blas et al., 2011). The resulting Matter Power Spectra for the different GDM models are shown in Fig. 2.8. We observe that GDM I is less viscous than GDM II, which is in turn colder (with smaller $c_{\rm s}$) and more viscous than GDM III.

In GDM III, the potential decays earlier, leading to a sharper cut-off. Although the differences between GDM I and GDM II parameters may seem minimal, GDM I exhibits a less pronounced cut-off than the subsequent model for wavenumbers corresponding to $k = 1$. For wavenumber values $k \geq 1$, the differences among the GDM models become visible, while at larger scales, the Power Spectra for all three models are identical.

[1] https://github.com/s-ilic/gdm
[2] https://lesgourg.github.io/class_public/class.html

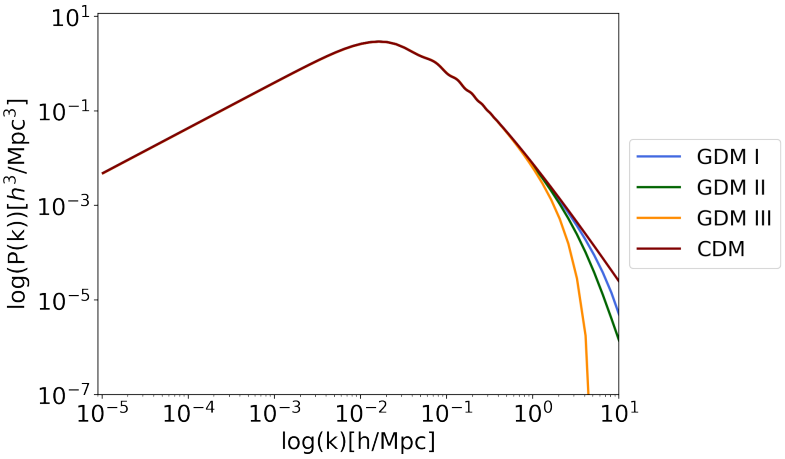

Figure 2.8: Matter Power Spectra for CDM and three different scenarios for GDM at $z = 127$. The results are identical until $k = 1$, after which the GDM models exhibit a cut-off determined by the c^2_{vis}, c^2_s, and w parameters, approaching the CDM scenario.

2.3.2 Structure formation in GDM

To create initial conditions for our simulations we used the N-GenIC code [3] (Angulo et al., 2012) which constructs an initial distribution of N-body particles from the distribution of density perturbations prescribed by a given matter power spectrum. The cosmological dm-only simulations were performed by using Gadget 2[4] (Springel, 2005) which is a Smoothed Particle Hydrodynamics (SPH) code that uses a hierarchical tree algorithm to compute the gravitational forces. Given that the sound speed and viscosity in our GDM fluid models are very small, we carry out simulations within an approximation where the effects of collisions between the N-body particles at small scales are neglected.

Basically, we are interested in studying the effects produced by the cut-off of the initial matter power spectrum (mpk) (due to the free streaming of linear perturbations) over the non-linear structure formation. In all our simulations, the boxsize is $L = 100$ Mpc$/h$ and the particle number is $N_{\rm tot} = 512^3$. The initial redshift is $z = 127$. The cosmological parameters are $\Omega_{\rm dm} = 0.25$, $\Omega_b = 0.05$, $\Omega_\Lambda = 0.7$ and $H_0 = 73$km/s Mpc. The σ_8 value varies in each model, in the case of CDM is 0.84, for GDM I is equal to 0.81, for GDM II is 0.75 and GDM III is only 0.65. The softening length corresponds to 2 per cent of $(V/N_{\rm tot})^{1/3}$, where V is the volume of the box, $V = L^3$. Snapshots for our simulations at $z = 0$ projected onto the xy plane are shown Fig. 2.9 where we used the library Pylians3[5] for making these plots. As expected, for the GDM instances, an evident lack of structure at small scales can be appreciated as an effect of the power spectrum cut-off.

We need information about the dynamics and distribution of the galactic baryonic components. Since our cosmological simulations are dark matter-only, in order to introduce information about baryons in galaxies we used a semi-empirical approach as is usual on other simulations as in Springel et al. (2005); Klypin et al. (2011); Teyssier, R. et al. (2009); Prada et al. (2012).

By using this information about structure formation provided by the simulations, it is possible to trace back the history of each halo, their progenitors and descendants halos and to identify the successive mergers occurring between them. This reconstruction is known as a *merger tree* and it can be linked to the galaxy formation by means of a semi-analytical or semi-empirical model, where the baryonic components within halos are modeled through the physical or observational

[3] https://www.h-its.org/2014/11/05/ngenic-code/
[4] https://wwwmpa.mpa-garching.mpg.de/gadget/
[5] https://pylians3.readthedocs.io/en/master/index.html

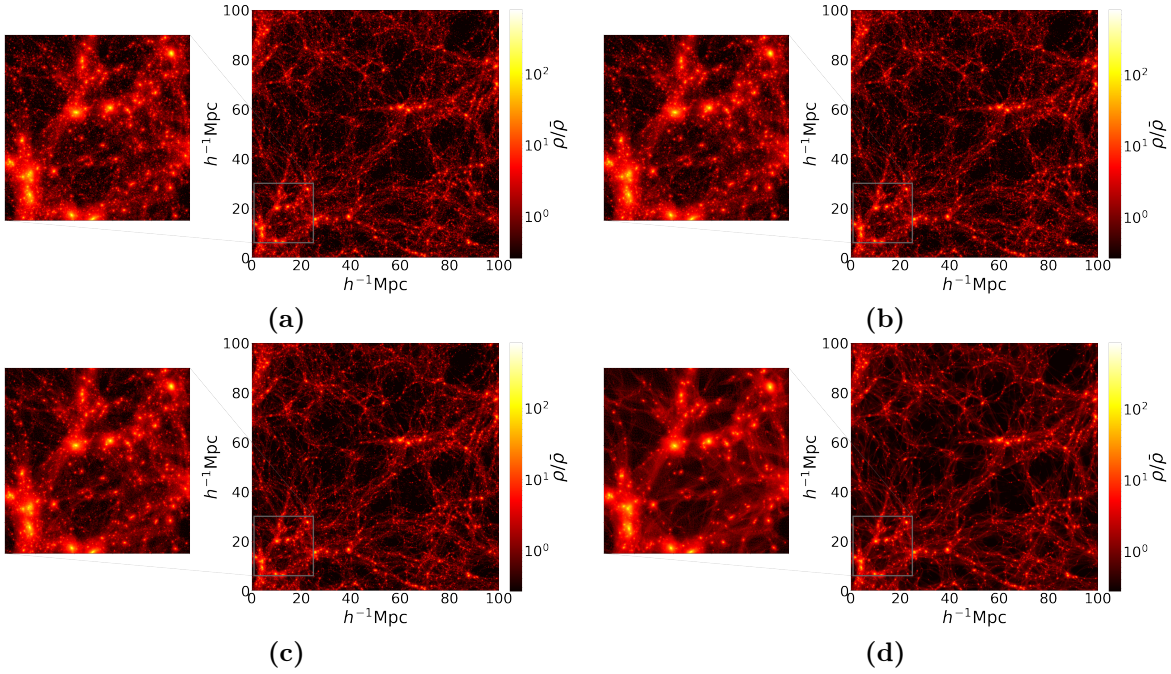

Figure 2.9: Projections in the XY plane of the simulations for CDM (a), GDM I (b), GDM II (c) and GDM III (d) for $Z \in [0, 50] Mpc/h$. In the left panels, we show an amplified region to observe the differences between models at small scales. For large scales, the structure formation seems very similar. In the high-density regions is clear how CDM has more substructure than the other models. GDM III is the model with the least structure formation and also the voids (i.e. the regions where the contrast has a negative value) are more evident.

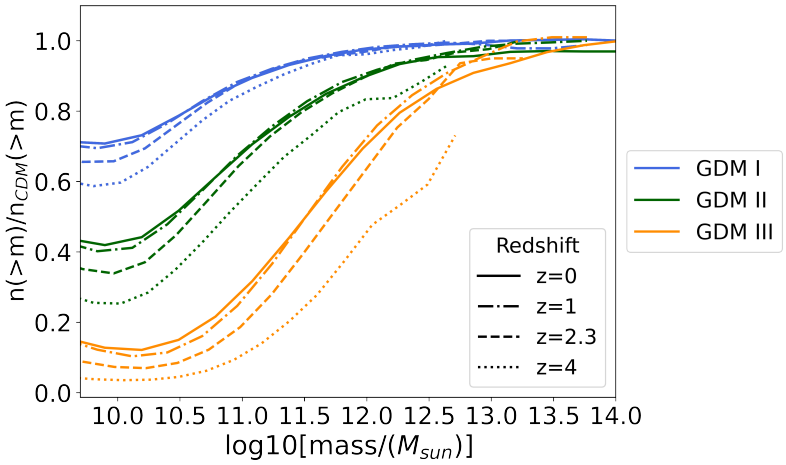

Figure 2.10: The smoothed relative change for the HMF function associated with different GDM models with respect to CDM at different redshifts. The range of the plot lies between the low limit resolution in the simulations, close to 10^9 $M_\odot$ *and the upper limit* 10^{14} $M_\odot$*. The HMF for GDM models shows a cut-off in agreement with Fig. 2.8. For higher z, the large structure is suppressed identically in all the models. Additionally, intermediate structures are less suppressed for models with smaller* c_s*, compared to those at* $z = 0$*.*

prescriptions (De Lucia and Blaizot, 2007). In particular, semi-empirical models trace the co-evolution of galaxies and their host dark matter halos over time, constraining the galaxy–halo connection at each epoch through the mass accretion and star formation rates.

In order to identify halos in the simulations, we used the Rockstar finder [6] (Behroozi et al., 2012a) which locates dark matter overdensities. Also, we used the Consistent-trees code[7] (Behroozi et al., 2012b) to generate the merger trees and the Universemachine[8] code (Behroozi et al., 2019) to create the corresponding semi-empirical model to describe baryons.

2.3.3 Halo and Stellar Mass Function

In order to quantify the amount of dark matter structures at different redshifts and compare it against the prediction of CDM, we computed the halo mass functions (HMF) in different models for different times relative to that for CDM. The results are shown in Fig. 2.10. Clearly, the mpk cut-off for GDM I, II and III is in agreement with that in2.10. Counts of large structures ($M > 10^{12.5} M_\odot$) within GDM models are similar to those for CDM at $z = 0$. When considering higher redshift values, we observe that the number of halos from small to medium scales decreases as z increases. However, counts of the biggest structures show the largest discrepancy given that the most massive halos are not created until redshifts nearby zero.

Additionally, by following the prescription described above, we reconstructed the stellar mass functions (SMF) of galaxies for simulations corresponding to different models at different redshifts. They are shown in Fig. 2.11 along with the observational stellar mass functions reported in Table 1 by Pérez-González et al. (2008). In that work, the authors analyze measurements of the stellar mass of a sample of galaxies observed with the Spitzer telescope with redshifts ranging between $z = 0$ and $z = 4$. The stellar mass functions for the models described in this work have been plotted at the nearest redshift available in our simulations snapshot to the mean one reported in the observations. Besides, the functions are shown in the ranges of masses reported for the

[6] https://bitbucket.org/gfcstanford/rockstar
[7] https://bitbucket.org/pbehroozi/consistent-trees
[8] https://bitbucket.org/pbehroozi/universemachine

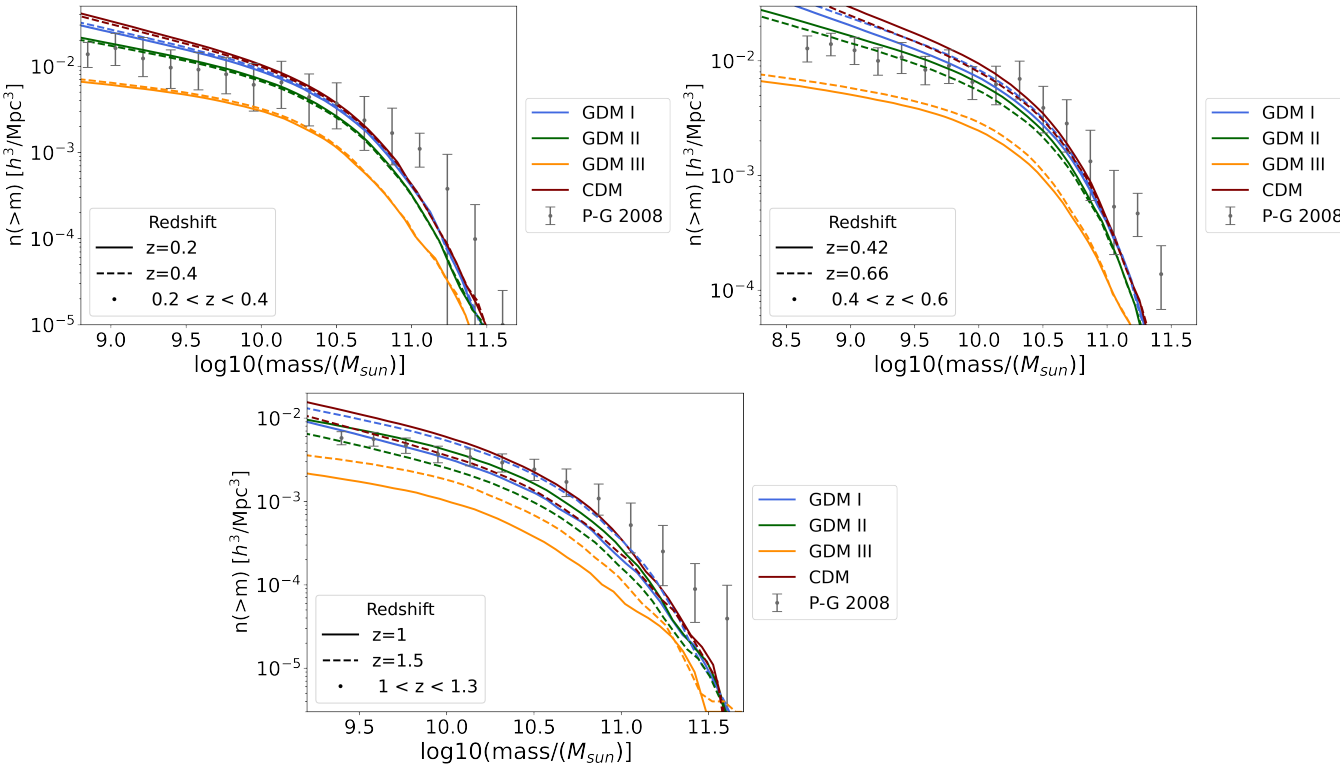

Figure 2.11: SMF for the three GDM scenarios and the standard CDM at different redshifts in the range $(0, 2)$*. It is also shown the SMF for an assembly of galaxies observed by Spitzer which was reported in Pérez-González et al. (2008).*

observed galaxies. As we can see in Figure 2.11, the GDM III prediction of SMF disagrees the most from observations for all the redshift values. In addition, the CDM prediction for the density number of galaxies at scales below $10^{10} M_{\odot}$ is overestimated for all redshifts.

In contrast, the GDM III prediction of the density number is underestimated for the entire range of masses in all redshifts. In some cases, the number density of the largest observed stellar structures is largely underestimated. This is also the range where the error bars from the observations increase significantly.

2.3.4 A probe case: Hickson Compact Groups

Since the main focus of this work is the study of properties of HCG in simulations associated with different GDM fluid models. We dedicate this section to mentioning the basics about them and how they have been usually classified over time. In 1982, Paul Hickson classified a sample of 100 HCGs following the list of criteria:

- Population: $N_{\rm pop} \geq 4$,
- Isolation: $\theta_N \leq 3\theta_G$,
- Compactness: $\bar{\mu}_G < 26$,

where N is the total number of galaxies within 3 mags below that of the brightest member, this criterion allows selecting galaxies with similar masses instead of very massive members and their satellites. Also, $\bar{\mu}_G$ is the average total magnitude of these galaxies per arcsecond^2 inside the

smallest circle (with angular diameter given by θ_G) that contains their geometric centers. Finally, θ_N is the angular diameter of the largest concentric circle containing no other (external) galaxies within this magnitude range (Hickson, 1982).

In some works, such as Barton et al. (1996) or in Tzanavaris et al. (2010), triplets with $N_{\rm pop} \geq 3$ are taken into account; for instance, the Redshift Survey (RSCG's). In this work, we follow the standard selection criterion considering $N = 4$ as the minimum number of members, given that we consider observations of HCGs from the original catalog which was constructed using this criterion. However, in the first part of the analysis we also consider triplets in order to study some effects arising from modifying the classification criteria. Also, in other works additional considerations are taken into account as in Díaz-Giménez and Mamon (2010), aiming to avoid selection biases, the following condition is imposed besides the standard ones:

$$R_{\rm brightest} \leq 14.44(\text{flux limit}),$$

where $R_{\rm brightest}$ is the R-band magnitude of the brightest galaxy. Additionally, some restrictions to the members' velocity are placed such as:

$$|v_i - \tilde{v}| \leq 1000 \text{ km s}^{-1},$$

where v_i is the radial velocity of the i-th member and $\overline{v}$ is the median radial velocity of the set (Díaz-Giménez and Mamon, 2010; Wiens et al., 2019; Díaz-Giménez, Eugenia and Zandivarez, Ariel, 2015). This last works as a filter of *fly-by* galaxies that are not actual members, that is, they are not bound by the group.

It is important to mention that this filter was applied in subsequent works by Hickson for the observational HCGs (Hickson et al., 1992). On the other hand, although in literature different algorithms using 2D projection of data to identify HCG in N-body simulation have been proposed (McConnachie et al., 2008; Díaz-Giménez and Mamon, 2010; Díaz-Giménez, Eugenia and Zandivarez, Ariel, 2015; Farhang et al., 2017) (aiming to mimic the selection procedure applied to observations), it is possible to establish a relation between the projected separation and the 3D distance between members (Wiens et al., 2019) and therefore, to modify the original compactness criteria according to the spatial information (which is available in the mock catalogs). For that reason, it is useful to take as a reference the reported median projected separation between members 39kpc/h inferred from observations, which interestingly is comparable to the sizes of the galaxies, and the median redshift of the sample of HCG $z = 0.0297$.

It is important to mention, that the last value strongly depends on the observational capability achieved up to date. In this work, these observational reference values are used to determine the relation between 3D and projected distances.

2.3.5 Group finder algorithm for mock catalogs

We adapted the selection criteria described above to the 3D available information in the mock catalogs as follows:

1) **Compactness** Instead of defining the compactness of an HCG by using the surface brightness of the group as it is done in the original selection rule, we consider the observational reference value of the median projected separation between members. This criterion will be applied once the relation between the median projected separation and the median 3D separation for the simulated HCGs candidates is established.

2) **Isolation**: We constructed a shell with a size three times the radius of the HCGs candidates ($r_{\rm shell} = 3r_{\rm group}$). Then, the ratio of the density inside that shell to the density of the group is computed. A candidate is isolated as long as this ratio is sufficiently close to zero, i.e.

$$\frac{\rho_{\rm shell}}{\rho_{\rm group}} < 10^{-4}. \tag{2.11}$$

2a) **Boundary:** Given that simulations are done within a finite box, clusters located close enough to the boundary are dropped; namely, clusters obeying $x_i - r_{\rm shell} < 0$ or $x_i + r_{\rm shell} >$L, where L is the box size and x_i are the Cartesian coordinates of the HCG candidate. These clusters close to the boundary are ruled out since they would fulfill the isolation criterion because a portion of the shell is outside the boundary.

3) **Dwarf mass limit** We establish a lower bound for the galaxy mass given by $M_{\rm dwarf} = 2 \times 10^9 M_\odot$, in order to get rid of dwarf galaxies (Wiens et al., 2019).

4) **Membership condition** We applied the filter over the velocity of members $|v - \bar{v}| < 1000{\rm km}/s$, to avoid the *fly-by* galaxies.

5) **Galaxy mass ratio** By assuming that the dominant member galaxy in luminosity corresponds to the dominant galaxy in mass (Wiens et al., 2019), we are able to select groups of galaxies with similar large masses instead of satellite galaxies. Therefore, HCGs candidates must satisfy the following condition

$$\frac{M_2 + M_3}{M_1} > 0.1,$$

where M_1, M_2, and M_3 stand for the first, the second, and the third most massive galaxies, respectively.

6) **Minimum number of members** After applying the previous filters, we double-check the total number of galaxies in each group. We are considering two instances, the one adopted in the original Hickson selection criteria, $N \geq 4$, and the one including galaxy triplets, $N \geq 3$.

2.3.6 Comparison with observations

In this section, we present a comparison of some metrics describing the geometric and physical properties of the HCG populations either in different theoretical models and observations. In the first place, we compare distributions for HCG in simulations against those in observations. Afterward, we compare different levels of agreements between models to give a qualitative analysis of the viability of GDM scenarios. In Fig. 2.12 we present the comparison between observed HCGs properties and those for the simulated data in general. Due to the fact we are considering the original catalog data, only the case for $N = 4$ will be taken into account in this section and the subsequent. Fig. 2.12 (a) shows 2D distributions in the HCG mass-density ($M_{\rm vir} - \rho_g$) plane for different models and observations. In there, we observe overlapping regions in all cases.

In addition, the mean HGC density shown in Fig 2.12, in all cases is close to $1 \times 10^{16} M_\odot /$ Mpc^3. Likewise, in (b) and (c) we show the distributions for different HGC samples in the density-crossing ($\rho_g - t_c$) time and density-velocity dispersion ($\rho_g - \sigma_v$) planes, respectively. It is well known in the literature that the crossing time of a group corresponds to a rough measure of the dynamical stability of such a system and can be defined as follows (Hickson et al., 1992)

$$t_c = \frac{4}{\pi} \frac{R}{\sigma_v},$$

where R is the median projected separation between members and σ_v is the mean total velocity dispersion of the group. In (b) we observe that the distributions for HCG in simulations partially match those of observed HCGs. Values of the cross-time below $3 \times 10^{-1} H_0^{-1}$ are not consistent with those in all simulations, but only values above that threshold. In other words, these observed groups with small t_c cannot be reproduced in any simulation considered here. A similar situation is evident in (c) where the velocity dispersion distribution for observed HCGs is much higher than those for HGC in simulations. More precisely, the mean values for the velocity dispersion in CDM, GDM I, GDM II, and GDM III distributions are 69.9, 73.2, 68.6, and 89.7 km/s respectively. On the other side, the same quantity for the observational HCGs is 344 km/s.

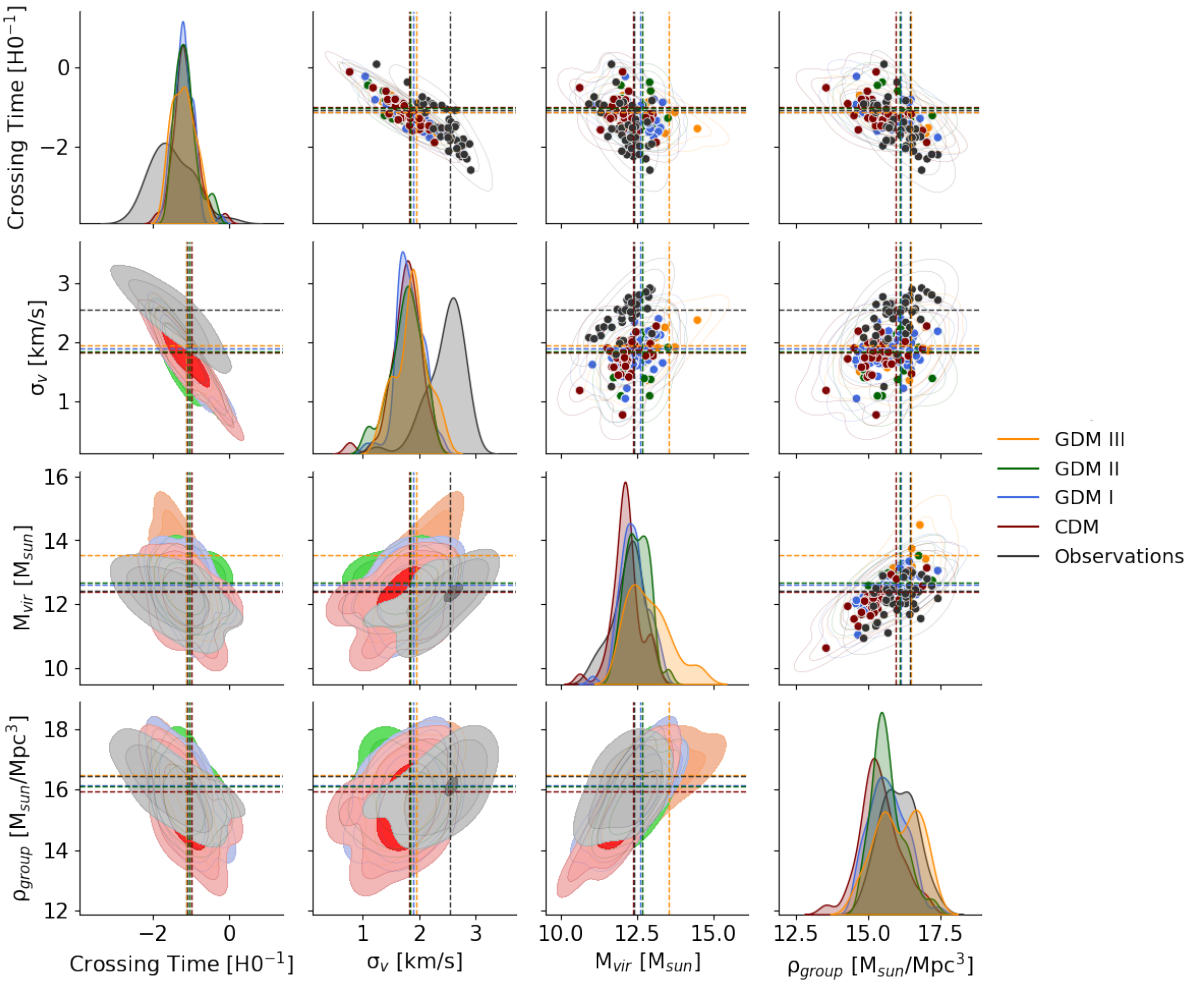

Figure 2.12: Comparison between observational and simulated data. We present the parameter space between virial mass, density, crossing time, and velocity dispersion in a matrix plot, with their respective histograms in the diagonal. We used kernel density estimation in the triangular-lower component with 1, 2, and 3-σ confidence levels in the contours, whereas in the triangular-upper component, we show the raw data.

In conclusion, such a large discrepancy cannot be solved either within CDM or GDM and it is linked with one of the most interesting open problems regarding HCG Wiens et al. (2019); Díaz-Giménez and Mamon (2010). It is important to mention that the best-fit for $\rho_{\rm g}$ inferred for GDM III is the closest to the central observational value in comparison to the best-fit for other models A similar situation happens for the virial mass, where GDM scenarios are closer to the observation predictions.

Aditionally, the mean values for the crossing time for CDM, GDM I, GDM II, and GDM III are $8.57 \times 10^{-2} H_0^{-1}$, $7.43 \times 10^{-2} H_0^{-1}$, $9.97 \times 10^{-2} H_0^{-1}$ and $7.60 \times 10^{-2} H_0^{-1}$ respectively and for the observations the value is $8.38 \times 10^{-2} H_0^{-1}$. On the other hand, the parameter space velocity dispersion-crossing time is shown in 2.12 (d). According to the previous comments we can see the discrepancies between the velocity dispersion.

Since t_c closely depends on the dynamics in HCGs, it is linked to the kinematics in such systems. Therefore, the existing tension for the velocity dispersion between observations and simulations is intimately related to that for the crossing time. The HCG distribution in the mass-crossing time and mass-velocity dispersion planes shown in panels (e) and (f) of Figs. 2.12 respectively. For the mass, there are regions where the distributions for observations and simulations match.

However, in the last panel (f), distributions for observations and models are disjoint but we can identify some points lying inside the simulated regions. Another interesting fact is the correlations between different observables can be noticed in every panel. Some of them are more prominent than others, for example, panel (d) clearly shows that there exists a linear scaling relation between the crossing time and the velocity dispersion either in models or observations. Evidently, the slope corresponding to observations differs from that for simulations. This fact might bring some hints about the differences between the dynamics of real and simulated groups.

In contrast in panel (f), a similar correlation can be noticed between the virial mass and the velocity dispersion, however, in this case, the slopes corresponding to models and observations are pretty similar. Despite the discrepancy between estimates of the velocity dispersion in observations and simulations, it is interesting that the $M_{\rm vir} - \sigma_v$ scaling relation is similar in both cases.

Chapter 3

Dark Matter Halos from Cosmological simulations and observational data

In this chapter, we will review how to construct initial conditions for dark matter halos and perform simulations at galactic scales. One well-known approach is the resimulation technique Ruszkowski and Springel (2009), which involves taking cosmological simulations, identifying target halos, and tracing their history back in time to resimulate the final halo with higher resolution Di Matteo et al. (2005). In this work, we propose cosmological simulations to carry out this resimulation and identify halos with specific characteristics that will be detailed in Section 3.1. The progenitor halos will be traced back to $z = 1$, where each progenitor will be treated as an isolated and virialized halo. We will assume that they have a density profile independent of time, specifically using the Zhao profile Zhao (1996). We will fit this profile to the density profile of the simulation and populate the new simulation using the Schwarzschild technique Schwarzschild (1979), which involves calculating the orbits followed by each particle in a stationary N-body system. For this step, we will use the GalIC (GALaxy Initial Conditions) code[1] Yurin and Springel (2014).

From a more rigorous perspective, we aim to utilize observational data, and for this purpose, we turn to the MaNGA catalog (Mapping Nearby Galaxies at Apache Point Observatory). This research project is conducted as part of the fourth Exploration Experiment of the Sloan Digital Sky Survey (SDSS-IV) Weijmans (2015). Within this catalog, there are maps of stellar velocity and various estimations of stellar velocity based on emission lines, specifically Hα emission. Hα is an emission line in the electromagnetic spectrum, specifically in the visible region. This line is formed when hydrogen atoms are excited and return to their ground state, releasing energy in the form of electromagnetic radiation at the Hα wavelength. By observing the light emitted by stars in the Hα line, precise measurements of the radial velocities of stars in galaxies can be obtained.

This information allows for the estimation of the rotation curve of spiral galaxies, which, in turn, is a measure of the total mass of the system. As mentioned before, galaxies have a dominant dark matter component. Precisely estimating this component would require knowledge of all the non-dark matter in the galaxy, which generally includes gas and stars. However, to estimate the dark matter component of a galaxy, it is necessary to resort to decomposition methods (see, for example, Kent (1985); Dimauro et al. (2018); Andredakis et al. (1995); Tabor et al. (2019)). These methods involve estimating the contribution of each stellar component in a galaxy. For spiral galaxies, there are two dominant regions, the bulge and the disk. Estimating these components

[1]GalIC (GALaxy Initial Conditions) is an implementation of an iterative method to construct steady-state composite halo-disk-bulge galaxy models with a prescribed density distribution and velocity anisotropy that can be used as initial conditions for N-body simulations.

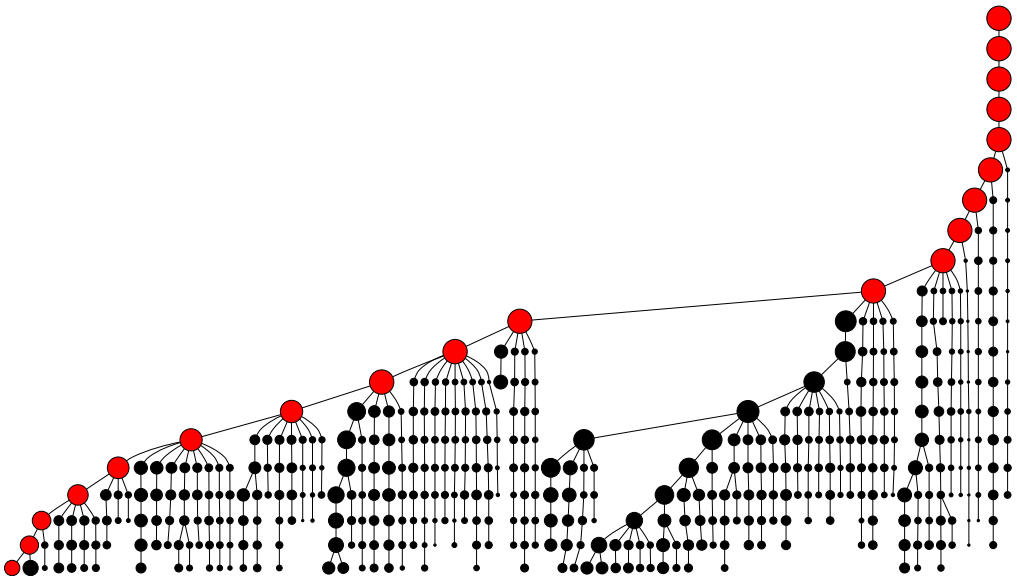

Figure 3.1: Merger Tree of a halo from a cosmological simulation performed with Gadget4Springel et al. (2021). Halos are identified using the Rockstar code Behroozi et al. (2012a), and the merger history is obtained using the Consistent Trees code Behroozi et al. (2012b).

along with the rotation curve allows for the estimation of the missing mass that cannot be directly observed, which is referred to as dark matter.

The estimation of these components is usually not an easy task. However, there are catalogs such as Mendel's Mendel et al. (2014), where this decomposition of galaxies is done for the SDSS galaxy catalog, which actually includes the MaNGA galaxies. As part of the research carried out during this period, a study is being conducted on the benefits of using deep neural networks to address the decomposition problem discussed in Section 3.3[2].

3.1 Halos from Cosmological Simulations

In sections 3.1 and 2.3, we discussed how different cosmological simulations can be performed. Here, we will briefly review how these cosmological simulations enable the construction of initial conditions for simulating galactic-scale structures. In particular, we are interested in the assembly process of a dark matter halo with a mass of $M \gtrsim 10^{11} M_{\odot}$. Later, we will see that this problem is intriguing because, when considering a higher resolution, the halos appear to still be dynamically active, implying that they could continue evolving. We are particularly interested in investigating the future of the dark matter halos that we observe today.

For cosmological simulations, we employ the Merger Trees algorithm De Lucia (2019) using the Consistent Trees code Behroozi et al. (2012b)[3]. This powerful code enables us to identify halos in cosmological simulations, as illustrated in Figure 3.1, and track their evolution over time. The cosmological simulations are discussed in Section 2, and Figure 2.9 shows some simulations of CDM-GDM models. Furthermore, Figure 2.10 displays the mass functions of different formed halos.

To extract information about the density of dark matter halos, we first estimate it using the

[2]This project is being prepared for publication.

[3]For more detailed information, please visit the official page at `https://bitbucket.org/pbehroozi/consistent-trees/src`.

KernelDensity library from scikit-learn Pedregosa et al. (2011)[4]. Once the density is estimated, we further refine the results using the Markov Chain Monte Carlo (MCMC) method Metropolis et al. (1953).

In Figure 3.2, we present some fitting results for representative halos from a cosmological simulation at $z = 1$. These halos were identified using the Consistent Trees algorithm Behroozi et al. (2012b) at $z = 0$. That is, we identify halos in the present and track their evolution back to $z = 1$. The fittings show the results for two different profiles. The most common profile in the context of cold dark matter (CDM) is the Navarro-Frenk-White (NFW) profile Navarro et al. (1996). However, a generalized version, known as the Zhao profile Zhao (1996), is proposed. It is defined by Equation (3.1), where the NFW prescription is recovered when $\alpha = 1$, $\beta = 3$, and $\gamma = 1$.

$$\rho(r; r_s\rho_s, \alpha, \beta, \gamma) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)^{\gamma}\left(1 + \left(\frac{r}{r_s}\right)^{\alpha}\right)^{\frac{\beta-\gamma}{\alpha}}}. \tag{3.1}$$

From Figure (3.2), we can observe that by allowing more freedom in the NFW profile, a better fit can be achieved. Additionally, the profiles obtained from this simulation appear to be flattened in the center, as the best fits show a value of $\gamma < 0.2$. When $\gamma = 0$, it corresponds to flattened profiles. For further reference, we can refer to Figure 3.3, where the parameter γ is varied around the Zhao profile (3.1).

3.2 Halos From Observations

MaNGA (Mapping Nearby Galaxies at Apache Point Observatory) is a research project carried out as part of the fourth Exploration Experiment of the Sloan Digital Sky Survey (SDSS-IV) Weijmans (2015). The main objective of MaNGA is to map and understand the formation and evolution of nearby galaxies using the Integral Field Unit (IFU) instrument of the Apache Point Observatory telescope in New Mexico, United States [5] Bundy et al. (2014).

MaNGA utilizes integral field spectroscopy to obtain light spectra from each pixel within a specific region of a target galaxy. This allows the analysis of the physical and chemical properties of stars and gas in different parts of the galaxy. With the data obtained by MaNGA, aspects such as stellar kinematics, chemical composition, mass distribution, and stellar formation history in nearby galaxies can be investigated. All the information in this catalog is obtained along the line of sight, as shown in Figure 3.4, which implies that geometric properties of the galaxy must be assumed to perform a deprojection and obtain an estimation of the actual data.

3.2.1 Velocity Map Fitting

As mentioned in Section 1.5.1, the rotation curve (RC) of a galaxy represents the rotational velocity of stars or gas in the galaxy as a function of the radial distance from the galaxy's center. We can extract information from the rotation curves of different galaxies using the MaNGA catalog. However, it's important to consider that the data is obtained as projections along the line of sight. Therefore, a process called *deprojection* is performed to obtain the actual rotation curve. This process involves fitting the rotation curve assuming that the stars in the galaxy move along circular orbits with a velocity Yoon et al. (2021):

$$V(r) = V_c \tanh \frac{r}{R_t} + s_{\text{out}}. \tag{3.2}$$

Figure 3.5 illustrates the functional form of the rotation curve. Initially, a constant behavior in the outer parts of galaxies, like in M31 Babcock (1939), was observed. However, as reported in

[4]For more information, refer to the official documentation at `https://scikit-learn.org/stable/index.html`
[5]For more detailed information, visit the official website at `https://www.sdss4.org/surveys/MaNGA/`

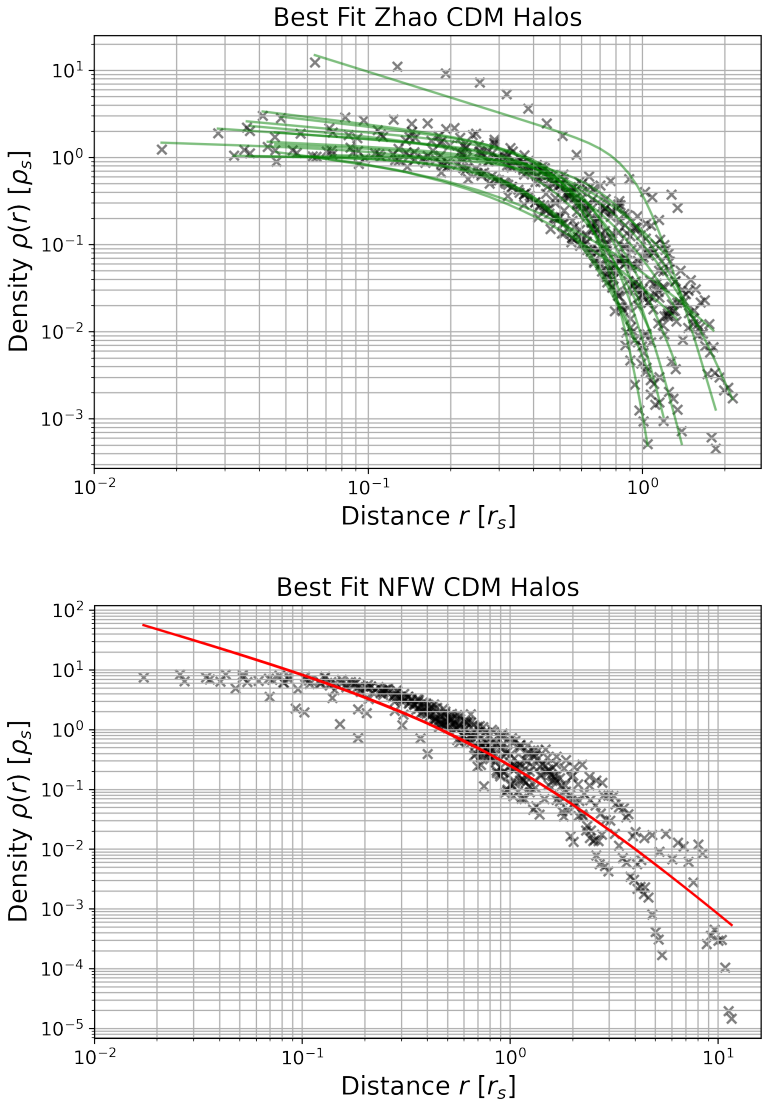

*Figure 3.2: Fitting of halos using the Zhao function (**Top**) and NFW function (**Bottom**) from a CDM cosmological simulation. The Densities $\rho(r)$ are normalized to the scale density ρ_s, and the distance r is given in units of the scale radius r_s. The black dots represent the observed data, and the red line shows the fitting performed using the Markov Chain Monte Carlo (MCMC) method to obtain the best parameters.*

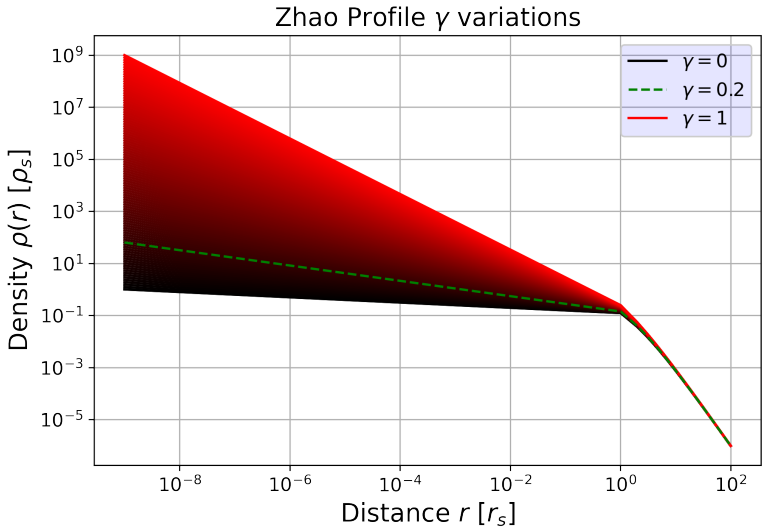

Figure 3.3: Variations of the Zhao profile with respect to the parameter γ in relation to the NFW profile, while keeping the other parameters constant at $\alpha = 1$ and $\beta = 3$. Different density curves $\rho(r)$ are shown as a function of the distance r, normalized to r_s, for specific values of γ. The solid red curve corresponds to $\gamma = 0$, the dashed green curve represents $\gamma = 0.2$, and the steeper solid red curve represents $\gamma = 1$. Additionally, multiple intermediate curves are included to explore the continuous variation of γ.

Yoon et al. (2021), rotation curves exhibit a wide variety of shapes, which is why the term s_{out} is introduced.

In some cases, such as the one shown in Figure 3.6, where two galaxies are interacting, it is not possible to estimate the rotation curve of each individual object. Additionally, these objects are being disturbed by the collision, and the velocity map does not provide reliable information about the enclosed mass of each object.

To derive Equation (3.2), we first need to account for the arbitrary positions of galaxies, as depicted in Figure 3.7. To estimate the rotation curve, we need to align them in an edge-on position, so that the approaching velocities V_+ lie in quadrants I and VI, while the receding velocities V_- lie in quadrants II and III. If we consider the galaxies as obtained from MaNGA, we need to define two coordinate systems: the celestial coordinate system x, y and the galaxy's coordinate system x', y'. The relationship between these two coordinate systems is given by:

$$\begin{aligned} x' &= (x - x_c)\cos(-\phi_0) - (y - y_c)\sin(-\phi_0), \\ y' &= \frac{(x - x_c)\sin(-\phi_0) + (y - y_c)\cos(-\phi_0)}{\cos(i)}, \end{aligned} \tag{3.3}$$

where (x_c, y_c) is the center in the celestial plane, ϕ_0 is the angle between the x and x' axes, and i is the inclination angle. This transformation assumes that the center of the galaxy is located at (x_c, y_c) in the celestial plane. Additionally, it involves a rotation by ϕ_0 and an inclination by angle i. The inclination affects only y'. Once we have obtained x' and y', we can calculate $r = \sqrt{x'^2 + y'^2}$ and consequently derive the rotation curve (3.2). Furthermore, we can estimate the polar angle $\theta = \arccos(x'/r)$, defining the polar coordinate system (r, θ) on the galaxy.

This analysis serves to estimate the rotation curve in the galactic plane. However, in reality, the galaxy is observed inclined, which is why it is necessary to project this hypothetical curve onto the celestial plane. To achieve this, we can use the polar angle θ and the inclination i, resulting in $V(r)\cos(\theta)\sin(i)$. Additionally, we must take into account the fact that the galaxy is receding

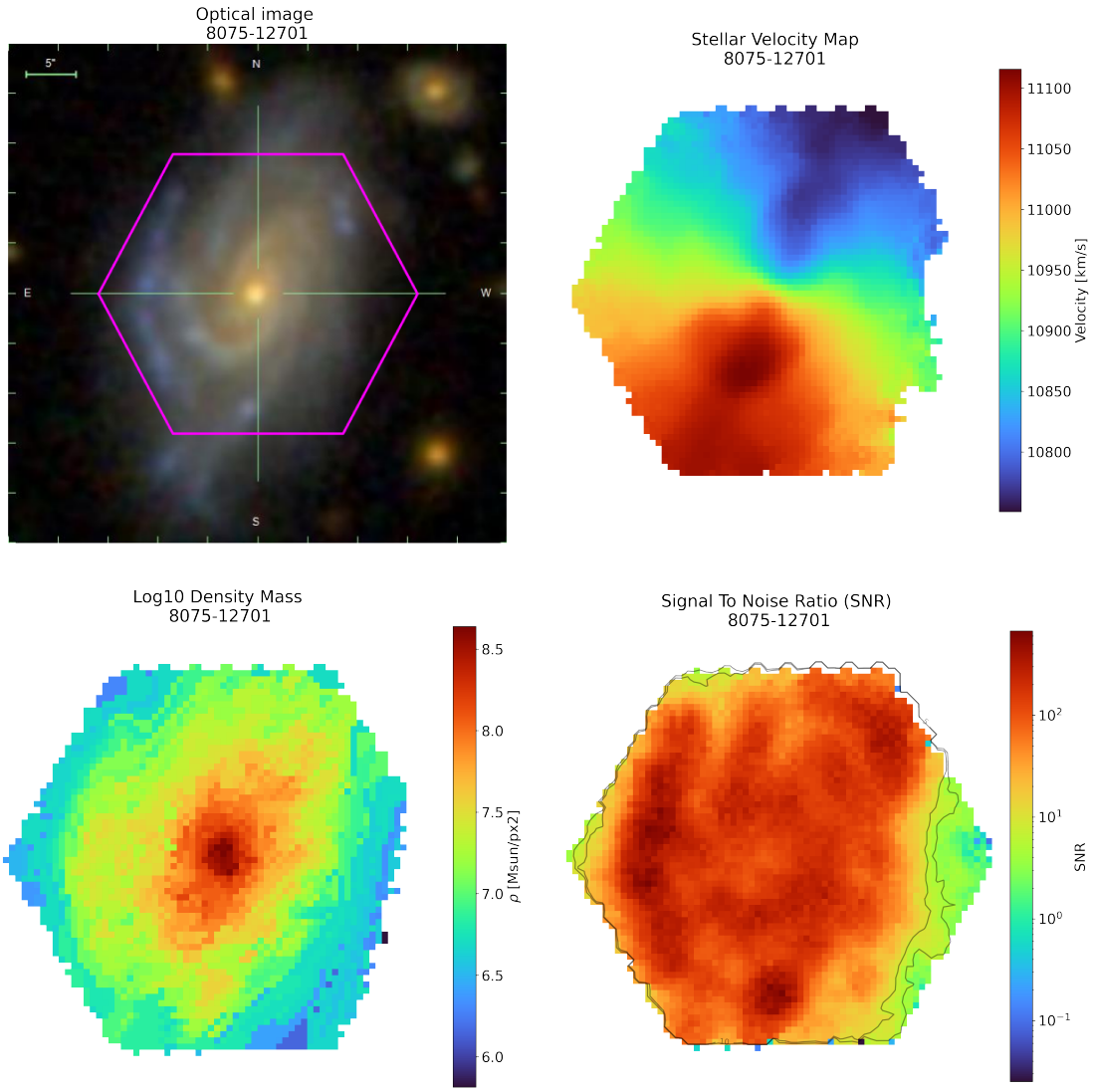

*Figure 3.4: A spiral galaxy observed by MaNGA: Hexagon showing the coverage of MaNGA's IFU instrument (**Upper Left**). Stellar velocity map (**Upper Right**), Stellar mass density map (**Lower Left**). Signal-to-noise ratio (**Lower Right**). In the stellar velocity map, points with signal-to-noise ratio values below 5 were omitted. However, it is sometimes convenient to omit signal-to-noise ratios below 10, which are indicated in the signal-to-noise ratio contours with black color curves.*

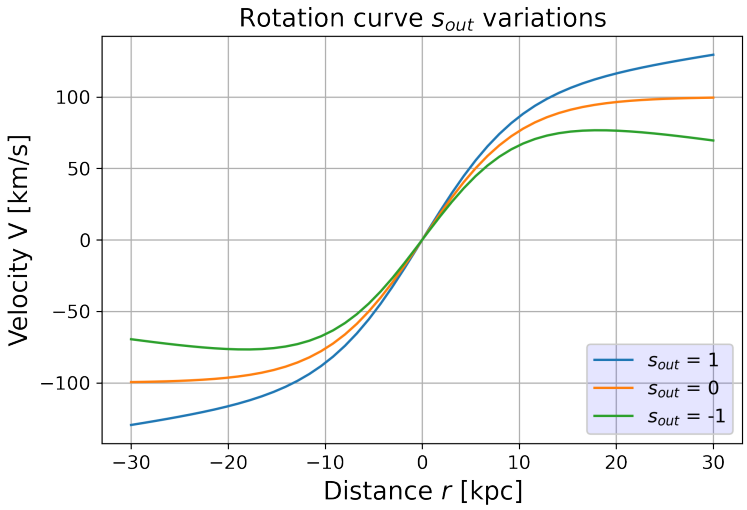

Figure 3.5: Hypothetical rotation curve with $V_c = 100$, $R_t = 10$, and $s_{out} = 1$, 0, and -1.

from us (redshift) at a velocity of V_{sys}. Finally, the velocity in the celestial plane is estimated using the following formula:

$$V_{\text{Obs}} = V_{\text{Sys}} + V(r)\sin(i)\cos(\theta). \tag{3.4}$$

Finally, with Equation (3.4), it is possible to compare with observations and estimate the set of parameters x_c, y_c, R_t, V_{sys}, V_c, ϕ_0, i, s_{out} that best fit the observational data using the Monte Carlo Markov Chain (MCMC) method Metropolis et al. (1953). By comparing the observations with the theoretical data, we obtain the best values by maximizing the likelihood function (3.5).

$$LL = -\frac{(x_d - x_{\text{theory}})^2}{2\sigma^2}. \tag{3.5}$$

An example of these fits is shown in Figure 3.8. For this example, the galaxy 8075-12701 was used, and the best fit is achieved when the following parameters are used:

Parameter	**Value**	**Parameter**	**Value**
V_c	255.17 km/s	V_{sys}	10856.92 km/s
R_t	4.53 arcsec	i	50.88 deg
θ	251.84 deg	s_{out}	-3.95 km/s/arcsec

3.2.2 Estimation of Dark Matter Halo Mass

Once the rotation curve for a galaxy is obtained, it is necessary to estimate the baryonic mass. Typically, the stellar mass and gas mass are estimated, but for this work, we will assume that the contribution of gas is negligible and focus on the stellar component. Since we have the polar coordinate r, it is possible to describe the stellar mass profile of this galaxy. Figure 3.10 shows the mass density for galaxy 8075-12701.

The mass density is obtained by measuring the mass enclosed in a shell of size Δr (see Figure 3.9). Δr is defined as a subdivision of the spatial domain of the MaNGA map into N intervals, in this case, we use $N = 40$. Spiral galaxies typically exhibit two dominant families of stars: the bulge, which is generally composed of red and old stars, and the disk, which usually consists of

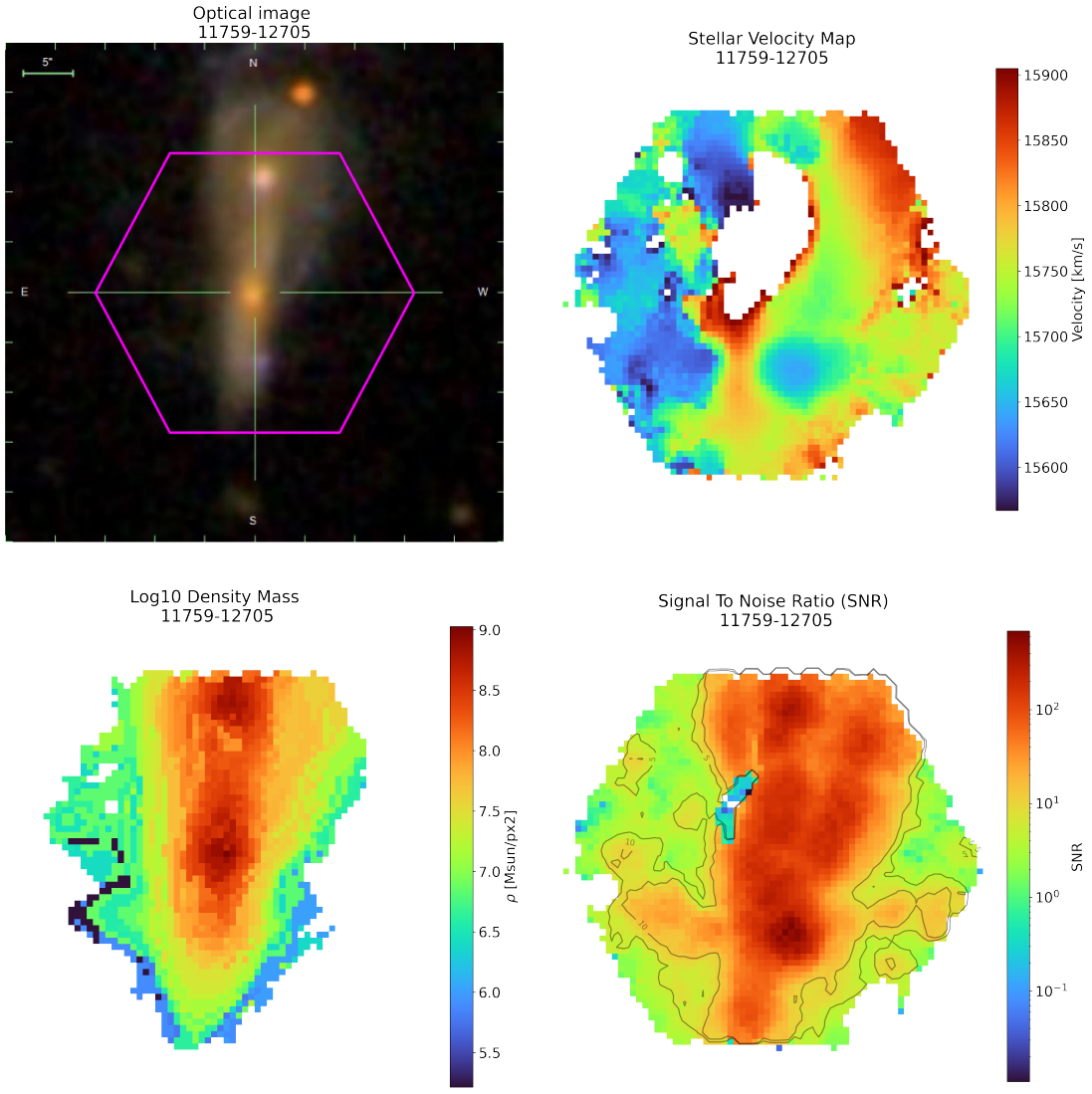

*Figure 3.6: A spiral galaxy observed by MaNGA: Hexagon showing the coverage of the MaNGA IFU instrument (**Upper Left**). Stellar velocity map (**Upper Right**). Stellar mass density map (**Lower Left**). Signal-to-noise ratio (**Lower Right**). This observed object exhibits two galaxies along the line of sight, as seen in the velocity map, and this galaxy pair cannot be easily resolved.*

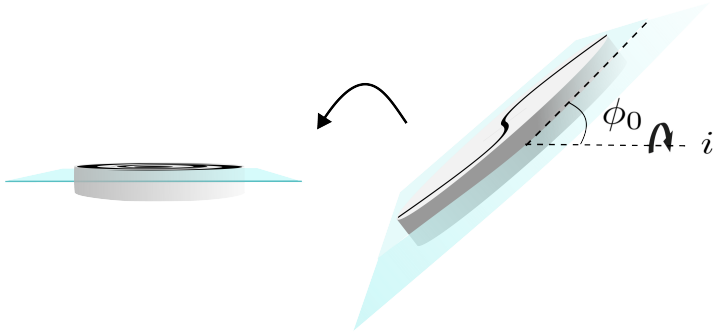

Figure 3.7: Example of a deprojection of a randomly oriented galaxy. The galaxy is initially rotated by an angle ϕ_0 and inclined by an angle i.

young stars as star formation occurs in the spiral arms. The dynamics in the disk are usually more violent than in the bulge, as it is estimated that for the galaxies in the MaNGA catalog, the galaxy's black hole has already consumed all the surrounding gas, leaving behind a relaxed system. It is common to assume that the density of the disk follows an exponential decay Yurin and Springel (2014), while the bulge is a denser structure. We can define the density in the bulge using the Plummer profile Plummer (1911). Both profiles are defined in Equation (3.6).

$$\begin{aligned} \rho_d(R,z) &= \frac{m_d}{4\pi z_0 h}\mathrm{sech}^2\left(\frac{z}{z_0}\right)\exp\left(-\frac{R}{h}\right), \\ \rho_b(r) &= \frac{3m_b}{4\pi a^3}\left(1+\frac{r^2}{a^2}\right)^{-5/2}. \end{aligned} \tag{3.6}$$

These two profiles (3.6) are used to fit the mass density obtained from the MaNGA mass map. Again, we use the Markov Chain Monte Carlo (MCMC) method, and the results are shown in Figure 3.10, where it can be seen that the disk is dominant in this region. However, with better resolution in the central part, a better estimation of the bulge contribution could be obtained, as it is usually dominant.

Finally, we can estimate the dark matter component using the prescription of Zhao Zhao (1996) defined in Equation (3.1). For this, we assume that the total rotation curve is given by Equation (3.7):

$$V_{\mathrm{Tot}} = \sqrt{\gamma_{\mathrm{gas}}\left|V_{\mathrm{gas}}\right|V_{\mathrm{gas}} + \gamma_{\mathrm{disk}}\left|V_{\mathrm{disk}}\right|V_{\mathrm{disk}} + \gamma_{\mathrm{bul}}\left|V_{\mathrm{bul}}\right|V_{\mathrm{bul}} + V_{\mathrm{DM}}^2}, \tag{3.7}$$

where γ_c is the Mass-to-Light ratio of each component. By fitting this function to the mass density shown in Figure 3.10, we can extract the dark matter component, as shown in Figure 3.11.

3.3 Disk-Bulge Decomposition

The bulge-halo decomposition of galactic systems has provided valuable insights into the evolutionary processes of galaxies and their components. By analyzing various measurements that describe the distribution of visible matter in galaxies, it has become possible to distinguish different components, specifically the disk and bulge masses, based on their distinct stellar populations with different dynamics and chemical features.

There are numerous schemes for classifying galaxies, with one of the most well-known being the morphological classification proposed by Edwin Hubble Sandage (1961). This classification categorizes galaxies into four types: elliptical, spiral, barred spiral, and irregular. Another scheme

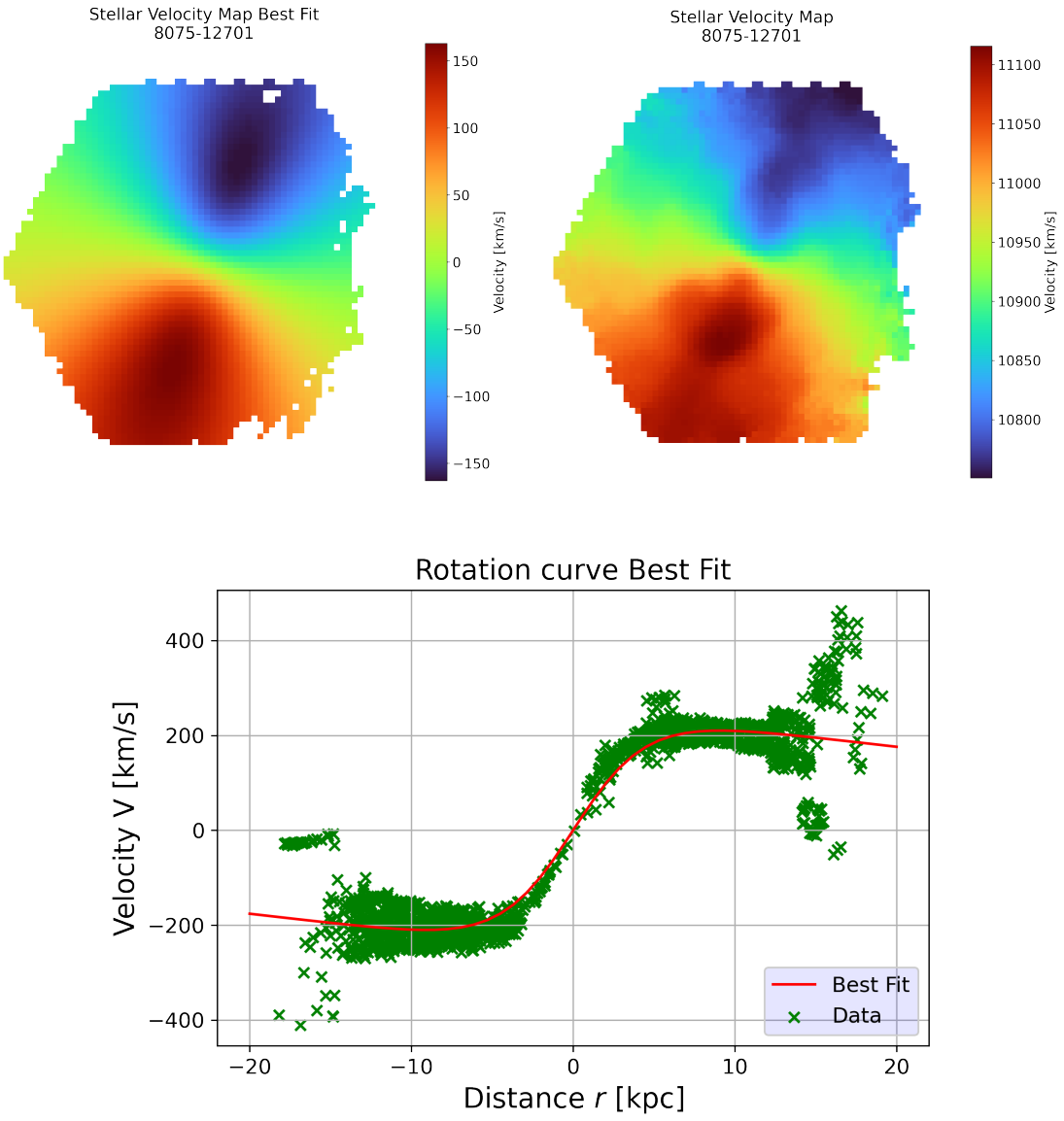

*Figure 3.8: The best fit obtained for the object 8075-12701 using the method described in this section. (**Top**) The observed velocity map compared to the fit obtained using Equation (3.4) by varying the parameters x_c, y_c, R_t, V_{sys}, V_c, ϕ_0, i, s_{out} to maximize the likelihood value (3.5). (**Bottom**) The values of the actual deprojected velocity will be used to estimate the mass contained within these galaxies.*

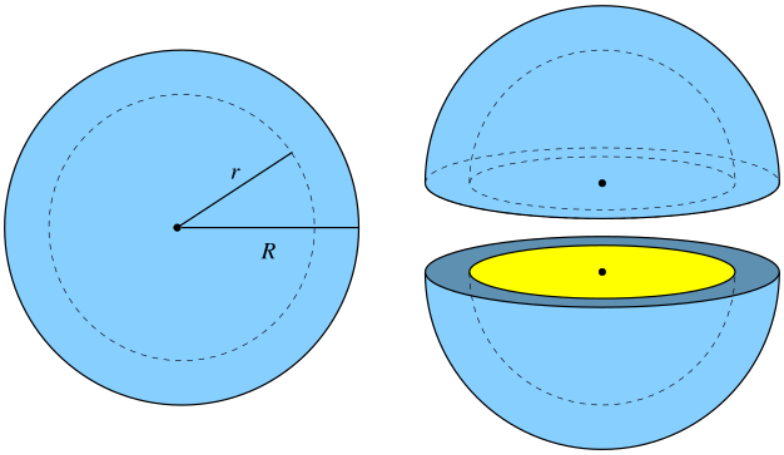

Figure 3.9: Representation of a shell created with two spheres, one with a radius of R and another with a radius of r.

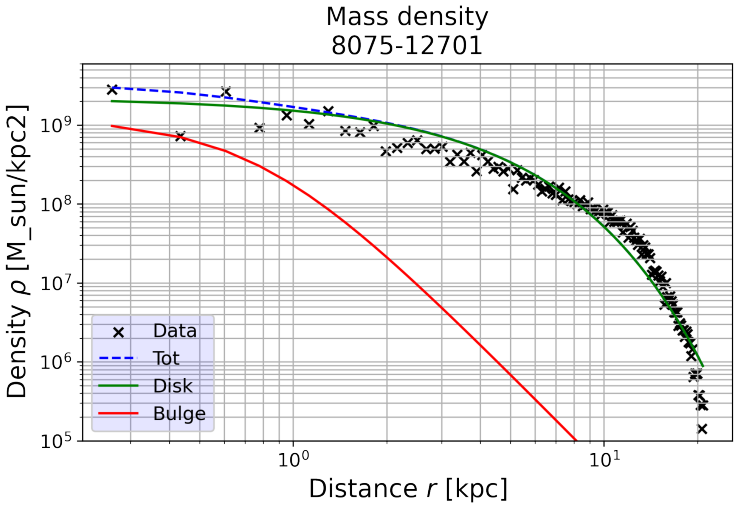

Figure 3.10: Stellar mass density obtained after fitting the rotation curve. The distance r corresponds to the distance from the center of the galaxy. The mass density is obtained from the mass map of the MaNGA catalog. The best fit obtained with the exponential and Plummer profiles (3.6) is shown in blue. The individual profiles are shown in green and red.

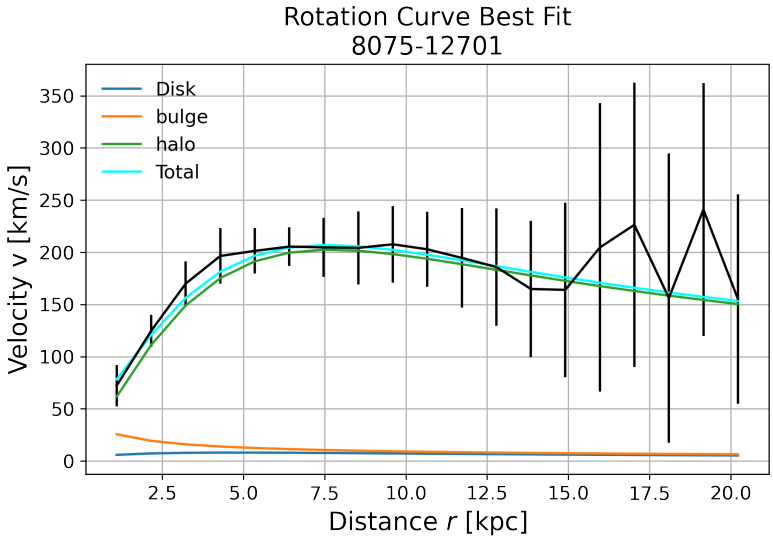

Figure 3.11: Best fit obtained for the total contribution of the disk, bulge, and dark matter in galaxy 8075-12701. From this fit, we can estimate the dark matter contribution in the system.

involves measuring the isophotal radius Holmberg (1958), which represents the size attributed to a specific level of surface brightness within a galaxy component. Additionally, concentration is used to characterize the light distribution independently of the light profile. It is defined as the ratio of two radii, each containing a fixed fraction of the total luminosity of the galaxy Kent (1985).

Another technique for reconstructing visible mass components in galaxies is the use of standardized fitting functions. Ideally, these functions would be derived from the physical principles governing galactic formation and evolution. However, due to the complexity of the physics involved in these processes, these models often become intricate and involve a large number of parameters. As a result, commonly used functions are derived empirically or based on phenomenological considerations. For example, exponential laws are well-suited for fitting disk components. Additionally, for elliptical galaxies and bulges in spiral galaxies, commonly considered models include the Hubble law Sandage (1961), the King's model King (1966), and the de Vaucouleurs's law de Vaucouleurs (1948). However, in some cases, the bulges associated with late-type galaxies are best fitted by exponential laws Andredakis et al. (1995); Freeman (1970). It is important to note that these methods are complex to implement and execute, and they require high-quality data of the objects under study in order to ensure reliability.

Moreover, spectroscopic methods have been utilized to distinguish individual galactic components, as demonstrated in Tabor et al. (2019). Essentially, by considering the light profile of the galaxy at each wavelength in the spectrum, it becomes possible to fit the bulge and disk light profiles to the distribution, similar to the photometric bulge-disk decomposition, which corresponds to the one-dimensional case Johnston et al. (2012).

Alternatively, numerical simulations have played a significant role in exploring detailed predictions of galaxy formation and dynamics within the standard ΛCDM framework Croton et al. (2006); Guo et al. (2010); De Lucia et al. (2004). Particularly, for identifying structural components of galactic systems, the most popular approach is through the use of semi-analytical models, which combine a highly simplified description of baryonic physics with Monte Carlo methods to reconstruct merging history trees De Lucia and Blaizot (2007).

A common assumption in simulations is that baryonic matter follows the same distribution as dark matter. Both components co-evolve, and material exchange occurs between the baryonic components through various processes. For example, the formation of a galactic bulge may arise

from a major or minor merger Hopkins et al. (2010). In these processes, both pre-existing and newly formed stars play crucial roles, as all the stars from the progenitors are added to the bulge component of the remnant galaxy after the merger. Additionally, it is assumed that in a galaxy merger, the gas contained within the progenitors becomes part of the resulting galaxy's disk, and the specific angular momentum of this component is equal to that of the halo in which it is embedded Guo et al. (2011); Bower et al. (2006); De Lucia and Blaizot (2007).

Estimating the virial mass of a galaxy is a challenging task. Typically, this quantity can be inferred by measuring the orbital velocities of stars and interstellar matter Lelli et al. (2016); Bottema and Pestaña (2015); Cautun et al. (2020). Precise spectroscopy techniques are required for such measurements. In the case of elliptical galaxies, the standard method follows the virial theorem Binney and Tremaine (2008), along with alternative techniques derived from it, such as projected mass estimators Heisler et al. (1985).

Furthermore, mock catalogs derived from cosmological simulations often employ a technique known as halo abundance matching. This involves establishing a monotonic relation between galaxy abundance as a function of stellar mass and the abundance of halos as a function of the original mass of their host halos Guo et al. (2010); Kravtsov et al. (2004); Vale and Ostriker (2004).

As we can observe, there are various approaches to computing the mass and properties of different galactic components. Bulge-disk decomposition can be performed based on observational morphology or using light information such as photometric and/or spectroscopic measurements, either through synthetic means using mock catalogs. On the other hand, obtaining information about the total mass involves making strong assumptions about at least one specific dark matter model or the overall kinematics of the system.

In this work, we propose an artificial intelligence (AI)-based method designed to identify the bulge and disk components of baryonic mass, as well as the total mass of a galaxy, using information derived from luminosity and features inferred from kinematics.

Recently, machine learning algorithms have gained popularity in astrophysics Rodríguez et al. (2022); BALL and BRUNNER (2010); VanderPlas et al. (2012); Baron (2019) and cosmology de Dios Rojas Olvera et al. (2022); Kamdar et al. (2016); Lucie-Smith et al. (2018) because they provide a powerful approach for inferring various astrophysical quantities and identifying correlations between them. Specifically, machine learning algorithms have been utilized to explore the features of galactic systems derived from their kinematics and luminous properties. Many studies have focused on designing methods capable of classifying galaxies, often trained using mock catalogs or extensive observational catalogs such as the Sloan Digital Sky Survey (SDSS) Adams and Woolley (1994); Vavilova, I. B. et al. (2021); Banerji et al. (2010).

The objective of this study is to implement and evaluate various machine learning algorithms for predicting the stellar mass ($M_\star$), total mass ($M_{\rm tot}$), and disk mass ($M_{\rm disk}$) of galactic systems. These quantities represent the major components in a galaxy. The predictions are made using data associated with the luminosity of visible components and the kinematics extracted from the Guo's galaxy catalog Guo et al. (2011), which is a synthetic catalog derived from the Millennium simulation Springel et al. (2006).

It should be noted that the bulge mass ($M_{\rm bulge}$) is not included in the prediction set since it can be easily derived in terms of the disk mass using the following expression Guo et al. (2011); De Lucia and Blaizot (2007):

$$M_\star = M_{\rm bulge} + M_{\rm disk}. \tag{3.8}$$

By training machine learning models on the available data, we aim to develop accurate predictors for the stellar mass, total mass, and disk mass of galaxies, thereby advancing our understanding of galactic systems and their components.

3.3.1 The Data

In order to train our machine learning algorithms, we have used a Mock catalog derived from the Millennium Simulation Guo's Galaxy Guo et al. (2011). The latter is a dark matter-only simulation carried out under the ΛCDM prescription Angulo et al. (2012), by using a customized version of the Gadget 2 code Springel et al. (2005). The Millennium simulation contains 2160^3 particles within a box of $L = 500$ Mpc/h. The corresponding mock catalog contains information on the merger tree history of each halo, redshift, position, and velocity of the corresponding galaxy. Also, data regarding the galactic components are included such as the amount of mass in metals for each component, star formation rate, and the rest frame absolute magnitudes at different bands. In this case, we are considering the information at $z = 0$. Additionally, due to the fact we are interested in studying the mass of the bulge and the disk, we only selected those galaxies whose disk and bulge components are different from zero in the mock catalog.

Features importance: It is well known that the physical and photometric properties of the stellar population of a galaxy are closely related to its dynamics and the spatial mass distribution of different components within the system. Specifically, this relation is reflected in the color-magnitude relation. For instance, it has been shown that bulge-dominant galaxies have a color-magnitude diagram mainly described by red galaxies Hogg et al. (2004). Besides, in Barsanti et al. (2021) it has been shown that the bulge is redder than the disk in galaxies within a cluster. A similar conclusion was reported in Dimauro et al. (2018).

In this work, we consider different sets of features, the first one corresponding to the u, g, r, i, and z absolute magnitudes that we dub hereafter Set I. Such magnitudes are also available within Sload Digital Sky Survey (SDSS) dataset Abdurro'uf et al. (2022), so they have an observational counterpart. Within a second set (Set II) the same features as Set I are considered in addition to $V_{\rm max}$, in order to include information about the kinematics. In both cases, the algorithm predictions turn out to be the $M_{\rm disk}$, $M_\star$, and $M_{\rm tot}$.

In fig. 3.12 the correlation matrix is shown for our set of features and it illustrates the extent that each feature contributes to the algorithm prediction. An exploration of the data is performed by using Pearson's correlation ratio which is defined for two sets of data X and Y as

$$r_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)s_X s_Y},$$

where barred symbols stand for the mean values and $S_{X,Y}$ for the standard deviation. When this quotient is $r_{X,Y} = \pm 1$ we have a perfect positive (negative) correlation, whereas for $r_{X,Y} = 0$ parameters are not correlated at all. In figure 3.12 we show the absolute value of this ratio since we are only interested in values that can contribute to a regression.

As expected, we can see that $M_\star$ is highly correlated to the magnitudes, especially with the z and i bands which correspond to the infrared and near-infrared regions of the spectrum, respectively. Actually, observational inferences of the luminous mass are strongly affected by the presence of dust given that emissions within the optical band are reddened while in the near-infrared this effect is negligible Tully et al. (1998).

The $M_{\rm disk}$ is less correlated to the magnitudes in comparison to $M_\star$. We can notice that it also shows a weak relation with the remaining quantities.

Additionally, $M_{\rm tot}$ shows a strong link with $M_\star$, this is due to the way mock catalogs are created. Then, the correlation of it with $V_{\rm max}$, which encodes information about the dynamics of all components of the system, is the highest in comparison to the masses of the rest of the components. Finally, the importance of the correlations with $M_{\rm disk}$ is weak. This suggests that the dark matter component affects the stellar one as a whole instead of each component separately.

Implementation of the machine learning algorithms: We have decided to apply a set of the most representative supervised algorithms that lead to an adequate prediction. Namely, the

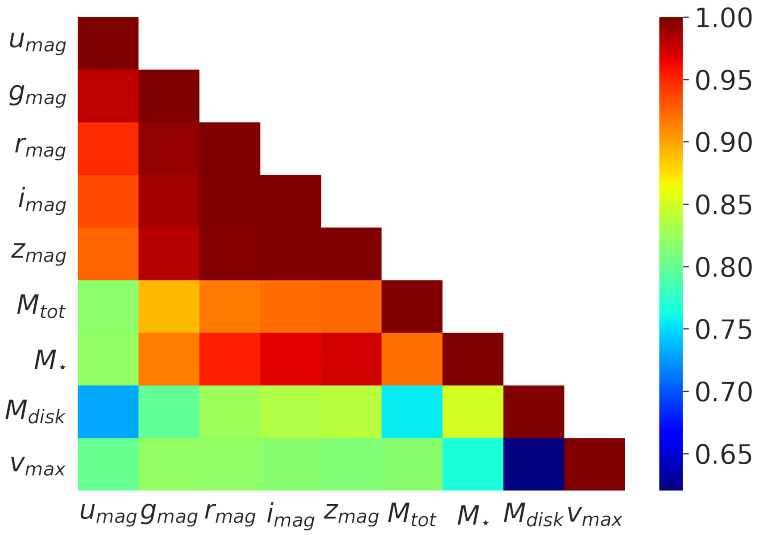

Figure 3.12: Heat map of the absolute value of the Pearson correlation coefficient between the galaxy parameters considered in this work. Number 1 represents a strong correlation. The stronger correlations occur between the stellar mass and the magnitudes in different bands, as expected. However, there exists a relation between the total mass and the magnitudes as well, although to a lesser extent.

KN-Neighbors (KNN), the traditional Linear Regression model (LR), RandomForest (RF), and Neural Network (NN). We implemented these IA methods by means of the scikit-learn library Pedregosa et al. (2011); Buitinck et al. (2013) and the Keras API Chollet et al. (2015).

For the KNN method, the distance between neighbors is defined in the hyperspace of features using the Euclidian metric, and the output value corresponds to the average of the outputs for each neighbor. Indeed, we found that the highest accuracy is achieved when this parameter is close to 18, and beyond that value, the error starts to increase. The RF algorithm is subject to the number of trees as well as their depth. Each tree contains decision nodes that split the data into smaller subsets until the branch find a homogeneous group according to the set of hyperparameters. Because of the way the trees are built, it is very easy to overfit. For that reason, it is strongly recommended to use a set of trees, instead. In this case, we used nearly 150 trees for the training. Finally, the NN is an interconnected group of nodes. Each node is stored in a layer and connected to other nodes in the network by unidirectional connections of different weights. Patterns learned in a layer are transferred to the next activated nodes. In this case, the NN was created by considering four hidden layers with 128, 64, 32, and 16 neurons, respectively.

3.3.2 Testing the algorithms performance

Relative percentage difference

In fig. 3.13 we present the relative percentage difference between the logarithm of $M_{\rm actual}$ within the mock catalog and the logarithm of the predicted value $M_{\rm pred}$ by each algorithm, computed as follows

$$\Delta = 100 \left[\frac{\log M_{\rm pred}}{\log M_{\rm actual}} - 1 \right]. \tag{3.9}$$

The left side panel shows the result when the training was carried out using Set I while on the right side that corresponds to Set II. The dashed lines correspond to the mean values for the whole set whereas the width of shaded regions corresponds to the standard deviation around the mean value $\mu \pm \sigma$. Also, the histogram for the quantities in the Mock catalog is shown.

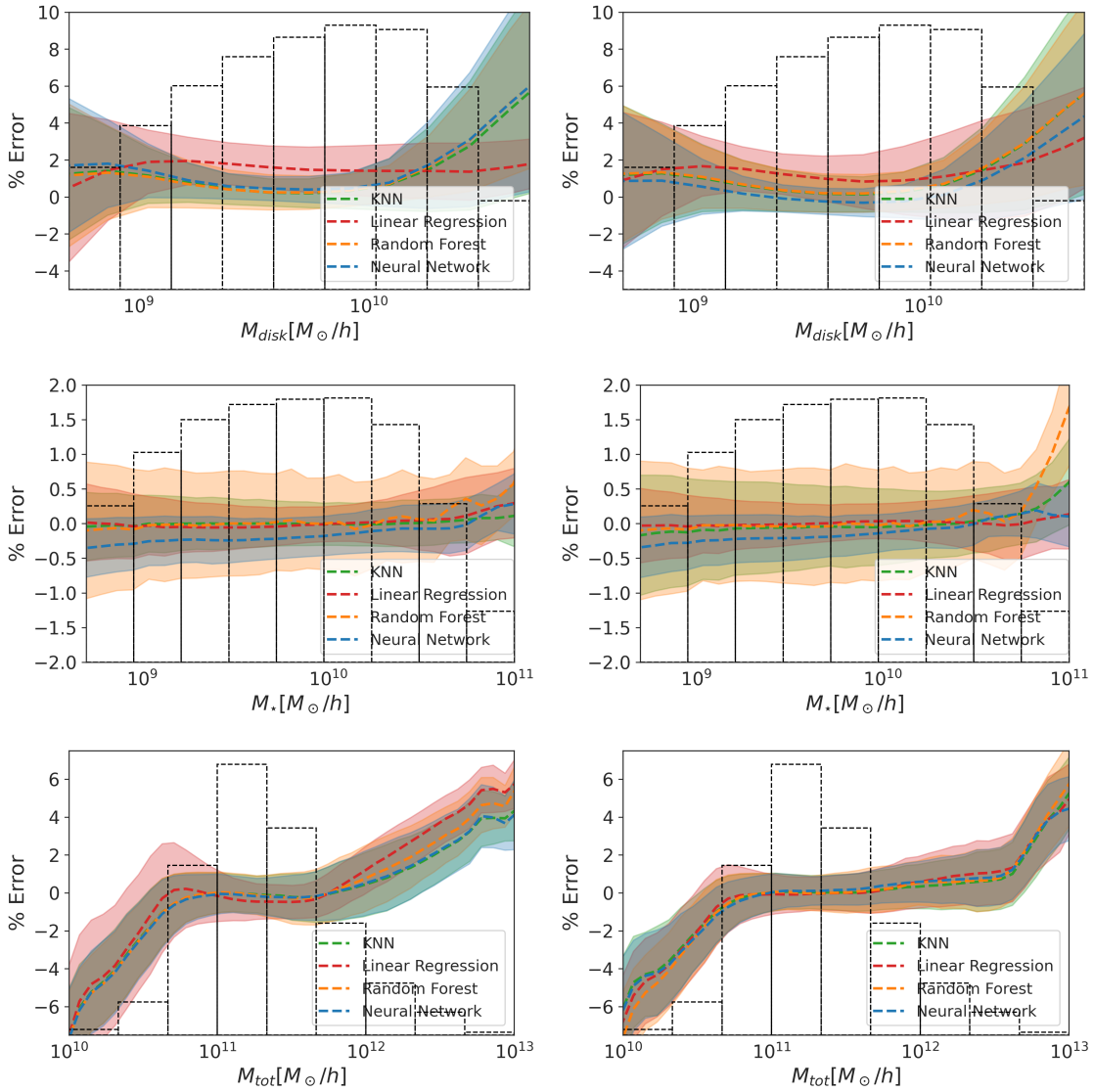

Figure 3.13: Relative percentage difference for the predictions of different Machine Learning algorithms with respect to the actual values in the mock catalogs. Set I is displayed on the left, the Set II on the right. The histograms in the figures represent the distribution of the data. As expected, the predictions are better where the density of data is higher. The lines represent the mean value μ and the bands are one standard deviation away from the mean value $\mu \pm \sigma$.

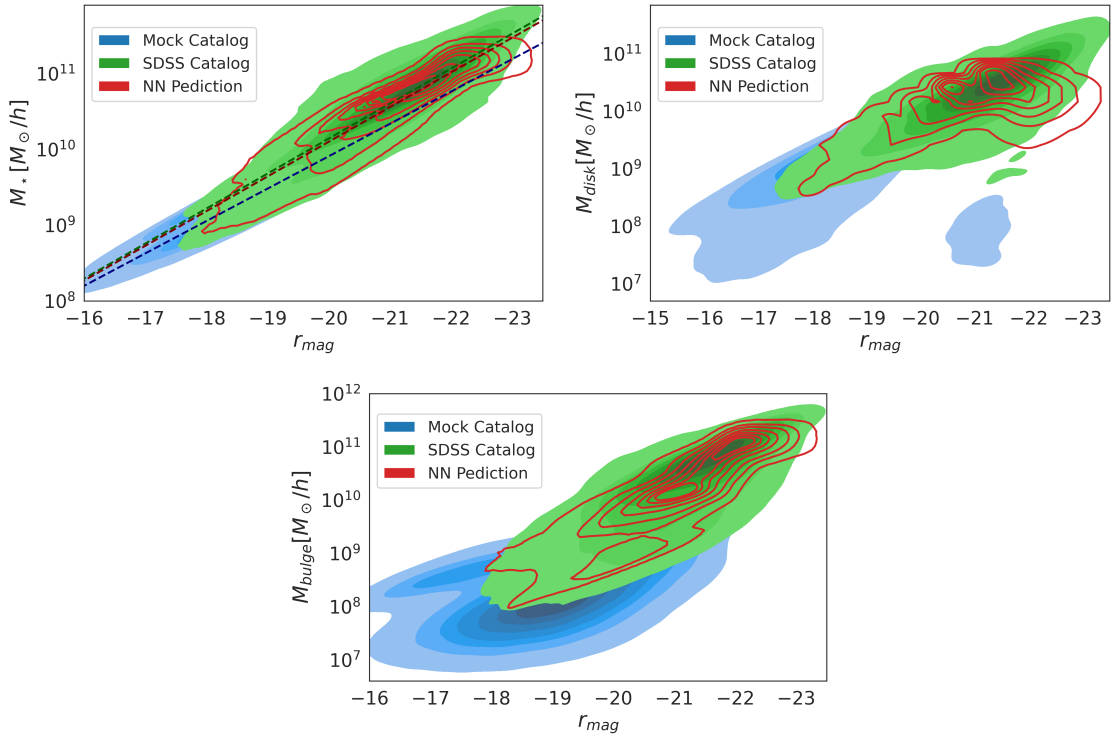

Figure 3.14: Kernel density estimation (KDE) plots of the stellar (above), disk (center) and bulge (bottom) masses components versus the r-magnitude for the simulated data in blue and the observational one in green. In the first case, we have a linear relation and the best fit for each dataset is represented by the dashed lines. The red lines are the level curves for the NN prediction.

It can be noticed in the uncertainty bands for the histogram become thinner as data counts are larger and hence the prediction is more accurate. Therefore, the highest error for the $M_{\rm disk}$ and $M_\star$ predictions (Figs. 3.13 (upper-left) and 3.13 (center-left)) lay below $10^9 M_\odot/h$. Besides, for $M_{\rm tot}$ in Figs. 3.13 (lower-left) and 3.13 (lower-right), the error increases for larger values of the mass, which means the prediction is reliable in the central region of the range around $10^{11}\,M_\odot$. In fact, it is worth mentioning that the distribution of $M_{\rm tot}$ is narrower compared to the distribution of $M_{\rm disk}$, as we can see in Fig. 3.13 (upper-left) and 3.13 (lower-left). This can be explained given that galaxies within the sample chosen from the mock catalog passed the strong condition of holding a bulge and this requirement is only fulfilled by sufficiently massive galaxies.

In general, all predictions have statistical errors around a central value equal to zero. However, $M_\star$ in fig. 3.13 (center-left) and 3.13 panel (center-right) show the smallest percentage difference in both cases, since there is a linear correlation between the magnitudes and the luminous mass Reiprich and Boehringer (2002); Kuiper (1938); Liebert and Probst (1987). For that reason, the LR model shows the best score since it was trained by directly fitting a scaling relation. In some algorithms like NN and RF, the error increases around 1% for masses $10^{10} M_\odot/h$, when Set II is considered.

Now, for the mass of the disk, figs. 3.13 (upper-left) and 3.13 (upper-right), we can notice that the percentage difference is higher than for the $M_\star$ case, but still, it is within a good prediction range for medium and high masses. In this case, the predictions for both sets of features are similar. Also, the LR method is no longer the best, since it is not a linear fit. Here, the NN and RF show the best training for Set I, and the first one improves then Set II is taken into account.

Finally, the $M_{\rm tot}$ is predicted similarly in both cases, fig. 3.13 (lower-left) and fig. 3.13 (lower-

right). Here, the correlation between Set I, which corresponds to the magnitudes and $M_{\rm tot}$, is not straightforward. However, given that the mock catalogs were created following a halo abundance matching relation, there exists a correlation to $M_\star$ and therefore, to the magnitudes. This allows having good training for $M_{\rm tot}$. Something interesting is that the prediction for Set II does not show an improvement, even though this set includes information about the kinematics of the system. Here, the best score corresponds to NN for Set I, given the lack of an explicit scale relation, while for Set II, all predictions are very similar.

By analyzing the behavior of the predictions for Set I and Set II, we can conclude the latter does not show a significant improvement in the results. Actually, the major contribution is in the $M_{\rm tot}$ as we can see in figure 3.13 (lower-right) where the curve is closer to zero above $10^{12}M_\odot$. Besides, having information about the $V_{\rm max}$ for real galaxies can be complicated because of the complexity of the dynamics of the system. For that reason, in order to keep the set of features as simple as possible, we decided to use Set I only from now on.

3.3.3 Generalized predictions from observational data

Up to this point, we have assessed the training performance by using data from the mock catalog. Now, such trained algorithms are ready to predict masses of different components in real galaxies by using observational data as input features. In other words, now our main goal is to analyze the predictions of these trained algorithms for features beyond the training values. It is important to stress that predictions in this case correspond to generalizations provided by the algorithm.

With that purpose in mind, we took data from the SDSS galaxy catalog Mendel et al. (2014), which contains data from around 660 000 galaxies. Nevertheless, we only used a portion of the total data $\sim 70\%$, which corresponds to the galaxies whose u, g, r, i, z magnitudes are known. The bulge-disk brightness profiles were reconstructed by using the Sersic profile photometric decomposition method. In order to compare the predictions of our algorithms with this observational data, we took into consideration the information on the bulge, disk, and stellar masses.

Additionally, we considered the SPARC database which contains information on irregular galaxies observed with the Spitzer telescope whose photometry lies at 3.6 μm. The catalog contains information about the observed rotation curves either for the dark matter and for each baryonic component and they were derived by using photometric and spectrometric methods. Particularly, we used the Mass Models Supplement (MMS) table to obtain the mass of the disk and the bulge.

The apparent magnitudes used in the predictions stage were obtained from the SDSS DR7 Abdurro'uf et al. (2022), we converted the apparent to absolute magnitudes following the equation Schneider (2015)

$$M = m - 5\Big(\log_{10} d - 1\Big), \tag{3.10}$$

where M and m are the absolute and apparent magnitudes, respectively, and d is the distance to the source. Distances were computed using the library Astropy The Astropy Collaboration et al. (2013) with the redshift reported in NED[6] and assuming the concordance cosmological model with the Planck 2018 cosmological parameters Planck Collaboration (2020) $H_0 = 67.66$ km/Mpc/s, and $\Omega_{m0} = 0.26$. A valuable piece of information for describing the evolution and structure of galaxies is the scaling relations between different physical quantities of a galaxy sample. In this section, we are interested in obtaining some of those relations from the input and output parameters of our algorithms. Specifically, we focus on analyzing the relation between mass components and the r-magnitude. We choose the magnitude in this color since is a common practice in different works Mahajan et al. (2017); Venhola, Aku et al. (2019); Côté et al. (2015). Furthermore, the relations for magnitudes in other colors are similar to the former. In addition, we study the $M_{\rm bulge} - M_{\rm disk}$ relation as well as the $M_\star - M_{\rm tot}$.

[6]The NASA/IPAC Extragalactic Database (NED) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

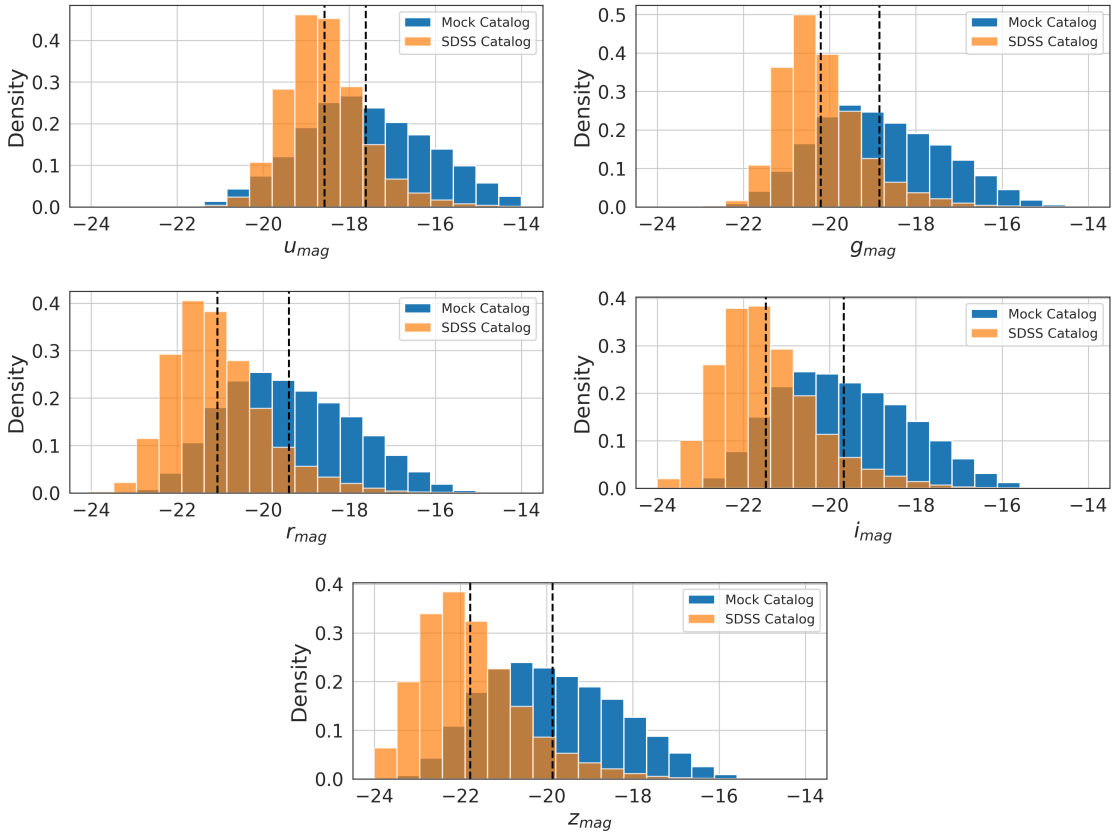

Figure 3.15: Histograms for the ugriz-photometryc system magnitudes for both, simulated and observational data. Vertical dashed lines show the mean value of each distribution. These histograms clearly show that within the mock sample, the distribution of magnitudes for galaxies significantly differs from that for the SDSS sample. This suggests that the algorithms shall be able to obtain generalized predictions given the extended range of features of the SDSS sample and the combinations between them.

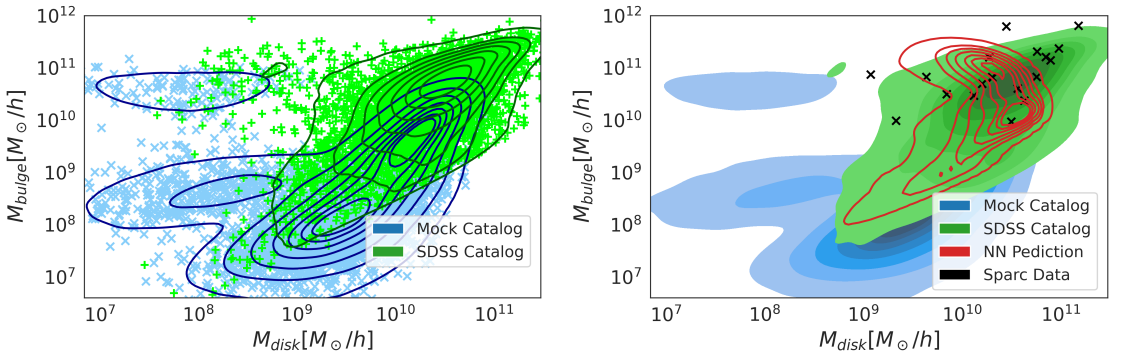

Figure 3.16: KDE plots of the Bulge-Disk decomposition. (Left) The distribution for both simulated (blue) and observed (green) data are shown as well as the solid contour levels. We observe a trimodal distribution for the mock catalog, whereas the observations show a bimodal distribution. (Right) We show the contour levels from the previous figure and in addition, we report the prediction for the NN for observational data (red) and the galaxies selected from the SPARC database (black crosses).

Let us stress that these scaling relations are essential for building mock catalogs and therefore machine learning algorithms might be a powerful tool for identifying such patterns and correlations. For deriving the results shown in this section, we only use the NN algorithm, since it was the only method capable of generalizing the data properly. This behavior is expected given that the architecture of other algorithms is designed to work properly only within the training data ranges. In other words, in contrast to other algorithms, NN is able to extrapolate and not only interpolate.

Mass-magnitude relation

In fig. 3.14 we show the distributions projected in $M_\star$- r-magnitude, $M_{\rm disk}$- r-magnitude and $M_{\rm bulge}$- r-magnitude planes. In all cases distributions for three datasets are plotted: firstly, that from the original Mock catalog in color blue, secondly that for the SDSS from the original catalog in green and the third one corresponds to predictions for the SDSS galaxies in color red.

In fig. 3.14 panel (a), the scaling relation between $M_{\rm disk} - r_{\rm mag}$ is shown. It can be noticed that a large fraction of the SDSS sample lies inside the region corresponding to the Mock catalog distribution, except galaxies with masses above $10^{11} M_\odot/h$. NN predictions for SDSS galaxies agree with observational data from the original catalog. Dashed lines represent the best fit for the three sets of data shown in the figure.

For data derived from the mock the best linear fit is

$$M_\star = -0.427 r_{\rm mag} + 1.370, \tag{3.11}$$

whilst for the SDSS catalog Mendel et al. (2014) the best fit corresponds to the following relation:

$$M_\star = -0.461 r_{\rm mag} + 0.916, \tag{3.12}$$

Finally, for the NN prediction, we have:

$$M_\star = -0.457 r_{\rm mag} + 0.954. \tag{3.13}$$

As can be noticed, the best-fit slopes for each dataset only show slight differences.

For $M_{\rm disk}$ in fig. 3.14 panel (b), the SDSS galaxies reach higher values than the mock catalog, so it clearly confirms that the neural network is capable to generalize values beyond the training data. We also observe a small blob separated from the largest one, since the values for the magnitude in that region correspond to large $M_\star$ values above $10^{10} M_\odot/h$ and by following eq. (3.8), we can

conclude that these galaxies have large values for the bulge in order to compensate for the stellar mass. This sort of galaxies are usually dubbed as cD−like galaxies (central dominant) Guo et al. (2011); Oemler Jr (1976). Interestingly the KNN algorithm correctly avoids predicting values in this region since cD−like galaxies are absent within the SDSS catalog. For $M_{\rm bulge}$ in 3.14 panel (c) the generalization is actually more evident. In this case, the relationship is not as simple as a linear regression as in the $M_\star - r_{\rm mag}$, but it shows the capability of NN to predict more complex relations.

As we can see, the regions embedding the distribution for the virial mass (and consequently the baryonic mass) corresponding to the mock catalog sample, are delimited by the numerical resolution of the Millenium simulation. However, the observations can go beyond those limits, showing higher masses and different magnitudes outside those regions.

Histograms describing the distribution of the absolute magnitudes in different bands in the ugriz-photometric system, for both SDSS and the mock catalog, are shown in Fig. 3.15. Distributions of magnitudes u, g, r, i, and z used for the training are reported, in Figs. 3.15. The dashed lines correspond to the mean values for each catalog. We can observe that in all cases the mean value for the observed galaxies is shifted to the left for high luminosities. The combination of these values in each band gives rise to values for baryonic masses higher than the mock catalogs as shown in fig. 3.14.

Bulge-disk relation

The bulge-disk decomposition is delimited by the eq. (3.8). At the same time, $M_\star$ can be directly determined given the magnitude, by using a scaling relation as that shown in Fig. 3.14. Thus, for a specific value for $M_{\rm bulge}$, its counterpart, $M_{\rm disk}$, will only take values within certain intervals, and vice versa.

Figure 3.16 (left) shows masses of the bulge and the disk of galaxies from the mock and SDSS catalogs whereas all information is contained inside the contour levels. As in the previous sections, the observations lie inside a region beyond the mock catalog data.

Nevertheless, we observe an interesting behavior for both datasets: The mock catalog shows a trimodal distribution while the SDSS data has a bimodal one. The most prominent region, for $M_{\rm disk} > 10^9 M_{\odot/h}$, corresponds to low values for $M_{\rm bulge}$, and it is associated to disk-dominated galaxies. The second region, for $M_{\rm bulge} > 10^{10} M_{\odot/h}$ is the bulge-dominated region Conselice (2006). This sort of morphology arises in both observed and simulated galaxies. Although disk-dominant galaxies are more abundant in both cases, within the SDSS data, the number of bulge-dominant galaxies is too small to be considered in the contour levels.

The third region in the $M_{\rm bulge} - M_{\rm disk}$ plane reveals an interesting fact and it is only visible for the synthetic data and corresponds to galaxies with low disk and bulge masses. Nevertheless, these galaxies are not reported in the SDSS catalog. This discrepancy suggests that there may be an observational bias. This difference may be due to the fact that current telescopes might not be able to detect the low-luminosity galaxies that actually appear in the numerical simulations.

Fig. 3.16 (right) shows data from both catalogs together with the NN prediction. Additionally, crosses in that panel correspond to data for galaxies from SPARC. For this last sample, the relationship between disk and bulge is nonlinear and not readily fitted with an analytic function as it happens with scaling relations derived in 3.3.3, nevertheless, the machine learning algorithm is able to make good predictions. Furthermore, it is interesting to note that the NN algorithm gives rise to mass predictions consistent with the SDSS distribution and does not predict bulge-dominant galaxies as expected.

Halo abundance matching

As mentioned in sec. 3.3.2, there exists a correlation between the total stellar mass and the total mass of the galaxy, which includes not only the baryonic contribution but also that of dark matter.

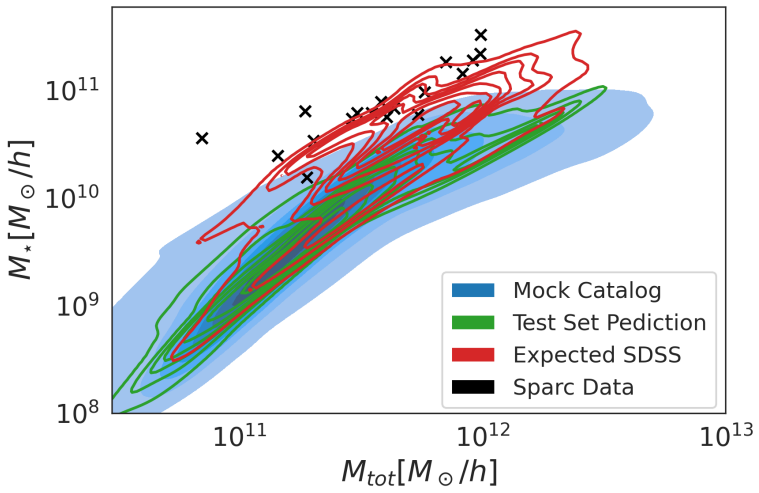

Figure 3.17: Predicted halo abundance matching relationship between M_{tot} and $M_\star$ for both the mock catalog test set and the SDSS Mendel catalog. We observe the prediction for the SDSS galaxies in red curves, which follows the same behavior as the synthetic catalog. Moreover, the crosses represent the galaxies in SPARC, which lie within the regions delimited by the predictions.

It is important to stress that this scale relation represents one of the most important keys when synthetic catalogs are constructed and it has not been widely explored from observations. It is well known that determining the total mass of a galactic system is a difficult task and a variety of assumptions are usually taken into account in order to circumvent our lack of knowledge about dark matter and how it is related to visible mass. In this sense, machine learning algorithms can be useful to reconstruct the total mass of galaxies in a model-independent way and, moreover, they may provide a tool to draw the total mass distribution at large scales within mock catalogs without imposing unnecessary assumptions.

In this work, we present a prediction for halo abundance matching considering galaxies within the SDSS dataset. In this case, we do not have information about the total mass of the galaxies in this dataset. In contrast, within the SPARC information, we can derive $M_{\rm tot}$ from the total rotation curve. We can observe this is not a linear relation, instead, the stellar mass increases faster than the total mass, but once $M_\star > 10^{11} M_\odot/h$, becomes approximately constant. This region in the parameter space corresponds to galaxies with a large portion of dark matter.

Some of the galaxies within SPARC are also contained in the SDSS sample. However, the way in which the stellar mass was determined in both cases is different. Interestingly, the SPARC data fall well inside the range of predicted values in the $M_\star - M_{\rm tot}$ plane for SDSS, red curves.

And both observational datasets show higher values for $M_\star$ than those in the mock catalog, as in the previous relations. This corresponds to another observational bias, due to the actual telescopes being capable of observing more luminous objects than those reported in the synthetic data, see Fig. 3.15. However, the $M_{\rm tot}$ values predicted for SDSS are within the same regions delimited by the mock catalog. This result is interesting since this is the only mass that presents this behavior.

Chapter 4

Small-Scale Simulations for Ultra-Light Scalar Dark Matter

This chapter focuses on small-scale simulations dedicated to exploring the properties and behavior of ultra-light scalar dark matter. Ultra-light scalar dark matter refers to a hypothetical type of dark matter particle characterized by its extremely low mass and its association with scalar fields. These small-scale simulations play a crucial role in understanding the dynamics and distribution of this form of dark matter. We present the results obtained in this project from different approaches, ranging from convergence tests to realistic simulations. To carry out this study, we rely on the scientific article titled "Solving the Schrödinger Poisson System using the coordinate Adaptive Moving Mesh method" Munive-Villa et al. (2022), which is presented as one of the results obtained within the framework of this research project. The article contains information about methods for solving systems of partial differential equations and results on the use of a coordinate transformation using the AMM method.

4.1 Ultra-Light Scalar Field Dark Matter

Ultra-light scalar dark matter is an intriguing proposal that challenges the paradigm of cold dark matter. In this section, we will explore the SFDM (Scalar Field Dark Matter) model, which postulates an ultra-light scalar field as the main component of dark matter. We will investigate the relativistic Lagrangian of the scalar field, its description in the non-relativistic and weak field limit, which corresponds to the Schrödinger-Poisson system, and in the case of self-interaction, the Gross-Pitaevskii-Poisson system is obtained[1]. Finally, we will review the effect of considering more than one scalar field in this model.

4.1.1 SFDM Model

In the context of string theory, it has been discovered that the incomplete decay of the inflaton field provides the necessary conditions to unify the description of dark matter, dark energy, and inflation Liddle and Ureña López (2006). As a result, a candidate for dark matter emerges with a lower mass limit of $m > 10^{-23}$ eV. Unlike the axions proposed by Peccei-Quinn Peccei and Quinn (1977), which are reviewed in Section 1.5.3, this dark matter candidate is not intended to solve the strong CP problem Callan et al. (1976). This is a subtle difference; however, Peccei-Quinn axions are ruled out for mass ranges $m < 10^{-5}$, as can be seen in Figure 1.12. Nevertheless, the analogy with axions is the reason why this candidate can also be found in the literature as an axion-like

[1]The Gross-Pitaevskii equation is also known as the nonlinear Schrödinger equation

particle. Let us then consider a scalar field ϕ as the candidate for dark matter. The action that describes the SFDM model can be expressed as Hui et al. (2017):

$$S = \int \frac{d^4x}{\hbar c^2}\sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\frac{m^2c^2}{\hbar^2}\phi^2\right],$$

the equation of motion for this action is given by the Klein-Gordon equation[2] Furthermore, if we consider a perturbation in the metric of the form

$$ds^2 = \left(1+\frac{2\Phi}{c^2}\right)c^2dt - a^2\left(1-\frac{2\Phi}{c^2}\right)d\boldsymbol{r}^2, \tag{4.1}$$

assume that $\ddot{\psi} \ll mc^2\left|\dot{\psi}\right|/\hbar$, and describe the field ϕ in terms of a complex field ψ as:

$$\phi = \sqrt{\frac{\hbar^3 c}{2m}}(\psi \mathrm{e}^{-imc^2t/\hbar}) + \text{c.c},$$

we obtain the following equation

$$i\hbar\left(\dot{\psi}+\frac{3}{2}H\psi\right) = \left(-\frac{\hbar^2}{2ma^2}\nabla^2 + m\Phi\right)\psi, \tag{4.2}$$

if we consider $a = 1$ and $\dot{a} = 0$, we recover the Schrödinger equation. Now, the analogy with cold dark matter arises when considering the Madelung transformation Madelung (1927)[3]

$$\psi = \sqrt{\frac{\rho}{m}}\mathrm{e}^{i\theta},$$
$$\boldsymbol{u} = \frac{\hbar}{am}\nabla\theta,$$

where the fluid equations for the SFDM model are given by

$$\begin{aligned}&\dot{\rho} + 3H\rho + \frac{1}{a}\nabla\cdot(\rho\boldsymbol{u}) = 0,\\ &\dot{\boldsymbol{u}} + H\boldsymbol{u} + \frac{1}{a}(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\frac{1}{a}\nabla\Phi + \frac{\hbar^2}{2a^3m^2}\nabla\left(\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}}\right).\end{aligned} \tag{4.3}$$

Notice that in comparison with equation (2.5), this system introduces a pressure term

$$P_Q = -\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}},$$

which is known as the *quantum pressure* and plays a crucial role in cosmological evolution. For instance, in Kopp et al. (2017), it is shown that ΛCDM is recovered when this term is approximately zero (dust-like matter). As mentioned throughout this section, the ΛCDM model is a strong candidate for explaining the large-scale background cosmology, and the SFDM model, with its quantum origin and analogy to ΛCDM, becomes an interesting alternative. Moreover, in the early epochs when perturbations were small, the SFDM model can provide a suitable description. In Ureña-López (2019), an extensive review of the model is presented, and in Figure 4.1, it can be seen how the SFDM model is compatible with the standard ΛCDM model, indicating its ability to fit the large-scale physics of the universe.

[2]For a detailed study, one can refer to, for example, Section 3.5 of Cottingham and Greenwood (2007).

[3]For an English version, please refer to `https://neo-classical-physics.info/uploads/3/4/3/6/34363841/madelung_-_hydrodynamical_interp..pdf`

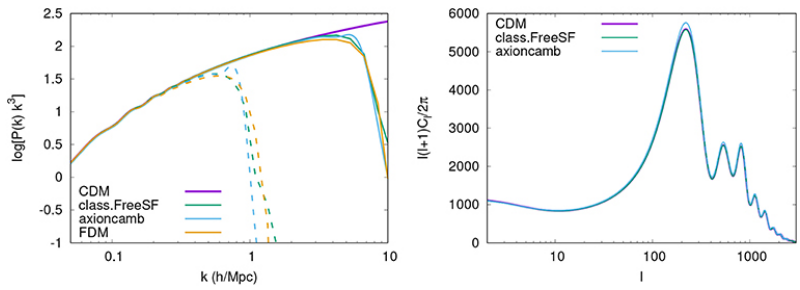

Figure 4.1: Mass Power Spectrum (MPS) (LEFT) and CMB Anisotropies (RIGHT) obtained for the SFDM model using the modified Boltzmann codes CLASS.FreeSF Ureña-López and Gonzalez-Morales (2016) and axioncamb Hlozek et al. (2015). The CLASS.FreeSF code solves the polar form of the perturbed Klein-Gordon equation, while axioncamb solves its equivalent fluid form. In general, SFDM appears indistinguishable from CDM, although there are some small discrepancies in the results generated by the Boltzmann codes. Image taken from Ureña-López (2019).

Self-Interacting Scalar Dark Matter

For a non-massive scalar field, the action is given by the following equation:

$$S = \int \frac{d^4x}{\hbar c^2}\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi.$$

In this case, a $\phi \to \phi + C$ symmetry is present, where C is a constant. However, in particle physics[4] and in the study of quantum gravity Hui et al. (2017), history leads us to suspect that these symmetries are not exact but are broken above a certain energy threshold. If we consider a field A with a period of 2π, the symmetry $A \to A + 2\pi$ exists, and the action can be expressed as:

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}F^2 g^{\mu\nu}\partial_\mu A\partial_\nu A - \mu^4(1-\cos A)\right].$$

By considering the first terms of the Taylor series for the equation of motion of this action, and assuming that $\phi = FA$ can be expressed in terms of the complex field ψ as before, we obtain the following equation:

$$i\hbar\left(\dot{\psi} + \frac{3}{2}H\psi\right) = \left(-\frac{\hbar^2}{2ma^2}\nabla^2 + m\Phi - \frac{4\pi\hbar^2\left|\lambda_s\right|}{m^2}\left|\psi\right|^2\right)\psi + \mathcal{O}(\left|\psi\right|^4\psi). \tag{4.4}$$

The cubic term $\left|\psi\right|^2\psi$ is considered the self-interaction term, and for the perturbation in the form (4.1), we obtain the Gross-Pitaevskii equation. The mass for this scalar is $m = \mu^2/F$. Tests of this model have been conducted in Mocz et al. (2023).

Multi-Field Scalar Dark Matter

In addition to the previous models, we will explore the idea of having multiple scalar fields as components of dark matter. We will investigate the implications of these multi-field models on the formation of structures at large scales and discuss their relevance in the study of dark matter.

4.2 Numerical Solutions of the Schrödinger-Poisson System

In this section, we define the Schrödinger-Poisson system using equations (4.4) and (2.3). We consider the parameters $a = 1$, $H = 0$, and $\delta\rho_m = \rho_m - \langle\rho_m\rangle$. This approximation is valid for

[4]The unification of the electroweak force is a clear example of this phenomenon.

small-scale simulations, where the effects of cosmic expansion are negligible compared to the size of the simulated system.

$$\begin{aligned} i\hbar\dot{\psi} &= \left(-\frac{\hbar^2}{2m}\nabla^2 + m\Phi - \frac{4\pi\hbar^2\,|\lambda_s|}{m^2}\,|\psi|^2\right)\psi, \\ \nabla^2\phi &= 4\pi G\left(\rho_m - \langle\rho_m\rangle\right). \end{aligned} \tag{4.5}$$

4.2.1 Schrödinger-Poisson System in Code Units

In order to carry out accurate numerical simulations of the Schrödinger-Poisson system, it is crucial to consider a system in *code units*. These normalized units allow computational calculations to be performed without taking physical units into account, facilitating the implementation and understanding of the algorithms used.

We begin by analyzing the Poisson equation (4.5), which describes the relationship between the gravitational potential and the mass density. To achieve normalization, we consider the definition of the De Broglie wavelength $\lambda_B = \hbar/mv$, where $\hbar$ is the reduced Planck constant and m is the mass of the particles involved. Additionally, we introduce a representative density $\rho_s = v^2/G\lambda_B^2$, where v is the characteristic velocity of the system and G is the gravitational constant. We rewrite the Poisson equation using these normalized quantities:

$$\nabla^2\frac{\phi}{v^2} = \frac{4\pi}{\lambda_B^2}\frac{|\psi|^2}{\rho_s}.$$

We observe that the potential ϕ has units of energy density $[\phi] = \mathrm{km}^2/\mathrm{s}^2$, which leads us to define the dimensionless variable $\tilde{\phi} = \phi/v^2$. Furthermore, we perform the change of coordinates $\tilde{x} \to x/\lambda_B$ and absorb the factor $\sqrt{\rho_s}$ into the wave function, obtaining $\tilde{\psi} = \psi/\sqrt{\rho_s}$. With these transformations, the dimensionless Poisson equation takes the form:

$$\nabla_{\tilde{x}}^2\tilde{\phi} = 4\pi|\tilde{\psi}|^2. \tag{4.6}$$

Now we focus on the Schrödinger equation. In addition to considering the change of variables in this equation, we consider the dimensionless time coordinate $\hat{t} \to vt/\lambda_B$. Applying this transformation, we obtain the following dimensionless form of the Schrödinger equation:

$$i\psi_{\hat{t}} = \left(-\frac{1}{2}\nabla_{\tilde{x}}^2 + \tilde{\Phi} - 4\pi\tilde{\lambda}_s|\tilde{\psi}|^2\right)\psi, \tag{4.7}$$

where we have introduced the dimensionless parameter

$$\tilde{\lambda}_s = \frac{\lambda_B^2\,|\lambda_s|}{m\rho_s},$$

which represents the nonlinear interaction among the particles in the system. Finally, equations (4.6) and (4.7) form a coupled set of dimensionless equations, allowing us to carry out efficient numerical simulations without taking physical units into account. These code units provide a solid basis for algorithm implementation and result analysis in the context of the Schrödinger-Poisson system. From now on, unless necessary, the tilde and hat notation will be omitted in the dimensionless variables. It is essential to mention that one of the great advantages of working with a dimensionless system is that the results are valid for different models. In the context of SFM, a model is defined by the value of the scalar mass m. It is important to highlight that this change of units in the variables ($\boldsymbol{x}$, t, ψ, ϕ) is equivalent to considering $m = 1$, $\hbar = 1$, and $G = 1$. However, knowing how this change is made is important because the transformation will allow us to return to the physical units, in Table 4.1, some values are shown for $v = 10$ km/s. The table illustrates that a value of $m_{\mathrm{sfdm}} \sim 10^{-22}$ eV is appropriate for simulating galaxies. However, it is common to utilize a combination of the parameters m_{sfdm} and v to identify specific structures with desired masses for simulating dark matter haloes.

m_{sfdm} [eV]	λ_B [kpc]	ρ_s [$M_\odot$/kpc^3]	τ [Gyr]
10^{-21}	0.19	6.33×10^8	0.02
10^{-22}	1.92	6.33×10^6	0.19
10^{-23}	19.17	6.33×10^4	1.87
10^{-24}	191.72	6.33×10^2	18.75

Table 4.1: Values of the variables as a function of the scalar mass m_{sfdm} for $v = 10$ km/s. As mentioned in the text, these transformations allow us to go from code units to physical units. Furthermore, τ is useful for defining the simulation time.

4.2.2 Discretization of the Schrödinger-Poisson System

To numerically solve the Schrödinger-Poisson system, we first need to consider a discrete spatial domain. The domain of a simulation in d dimensions is defined as $\Omega \subset fancyR^d$, and its boundary is defined as $\partial\Omega$. If we define the simulation space as

$$\vec{x} \in \Omega = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d] \subset \mathbb{R}^d,$$

the discrete domain is defined as

$$\vec{x}_h \in \Omega_h = [a_1, a_1 + h_1, \ldots, b_1] \times [a_2, a_2 + h_2, \ldots, b_2] \times \ldots \times [a_d, a_d + h_d, \ldots, b_d],$$

where h_i is the spatial step in the i-th axis. For this domain, the boundaries are defined as

$$\partial\Omega_h = \{\vec{x}_h \in \Omega_h \mid (x_1 = a_1 \vee x_1 = b_1) \vee (x_2 = a_2 \vee x_2 = b_2) \vee \ldots \vee (x_d = a_d \vee x_d = b_d)\}.$$

Figure 4.2 shows an example of discretizing the domain $\vec{x} \in \Omega = [-1, 1] \times [0, 1] \subset \mathbb{R}^2$, where the boundaries are shown in black and correspond to $\partial\Omega_h = \{\vec{x}_h \in \Omega_h \mid (x = -1 \vee x = 1) \vee (y = 0 \vee y = 1)\}$.

The Schrödinger-Poisson system is a coupled system of differential equations for ϕ and ψ. The system is highly nonlinear as it involves terms like $\phi\psi$ and $|\psi|^2 \psi$, which cannot be discretized in the context of finite differences. To address this problem, we assume that the terms accompanying ψ do not correspond to the current time step $n+1$, but rather to the previous time step n, effectively linearizing the system.

The spatial discretization is done as in the Poisson equation in Section A.3.1, where the Laplacian operator is discretized. As we can see in equations (4.7) and (A.19), the only relevant spatial difference operator in the system is the Laplacian. The potential term and the nonlinear term need to be rewritten as

$$\phi + 4\pi |\lambda_s| |\psi|^2 \rightarrow \boldsymbol{\phi} + 4\pi |\lambda_s| \|\boldsymbol{\psi}\|^{\mathbb{2}},$$

where $\boldsymbol{\phi}$ and $\|\boldsymbol{\psi}\|^{\mathbb{2}}$ represent diagonal matrices with entries corresponding to the vectors ϕ and $|\psi|^2$, respectively. In this way, the matrix M required to solve the time-dependent equation (A.21) is given by

$$M = \frac{i}{2}(M_{xx} + M_{yy} + M_{zz}) - i\left(\boldsymbol{\phi} + 4\pi |\lambda_s| \|\boldsymbol{\psi}\|^{\mathbb{2}}\right). \tag{4.8}$$

In this equation, the challenge of nonlinearity in the Schrödinger-Poisson system becomes apparent. To perform the matrix factorization as described in Section A.3, it is necessary to assume that ϕ and $|\psi|^2$ are known, since, as discussed in Section A.5, the temporal evolution requires knowledge of the right-hand side of the Schrödinger equation (4.6), which depends on ψ^n. The same applies to the matrix (4.8); we only know the coefficients at the n-th time step. This mismatch can be considered as an approximation to the right-hand side of the Schrödinger equation, and this approximation improves as $\Delta t \rightarrow 0$. This is the discretization approach used for the Schrödinger equation only. For the Poisson equation, we follow the same procedure as in Section A.3.1.

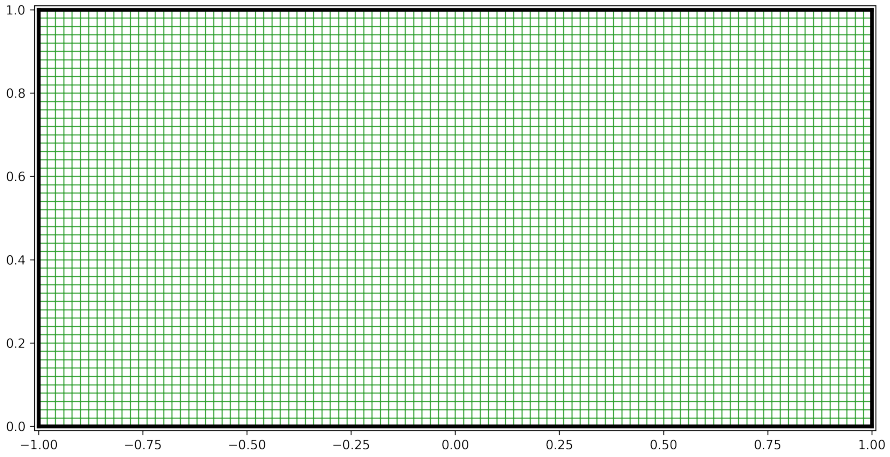

Figure 4.2: Mesh discretizing the domain $\Omega = [-1, 1] \times [0, 1]$ *with* 51×101 *points. The green region represents the domain* Ω_h*, and the black lines represent the boundaries* $\partial\Omega_h$*.*

The temporal evolution is performed on the Schrödinger equation as shown in Section A.5. In this work, $\theta = 1/2$ has been used as it proves to be the most stable method. On the other hand, the Poisson equation needs to be solved at each time step using the solution of the Schrödinger equation ψ^n from the current step. The solution of the Poisson equation becomes ϕ^n, and with this pair of solutions, we compute the matrix M associated with the Schrödinger equation to find ψ^{n+1}.

4.2.3 Initial Conditions

Stable Solutions of the SP system:

To obtain a stable solution, we consider spherical symmetry and solve the time-independent Schrödinger equation. First, we apply the Madelung transformation with the Ansatz $\psi = \varphi(r)\mathrm{e}^{i\gamma mt/\hbar}$. Utilizing the defined transformation and adimensional units, we have $\hat{\psi} = \hat{\varphi}\mathrm{e}^{-i\gamma t}$. Thus, the time-independent Schrödinger-Poisson system is given by

$$\begin{aligned} \nabla^2\varphi &= 2(\phi - \gamma)\varphi, \\ \nabla^2\phi &= 4\pi\varphi^2. \end{aligned} \tag{4.9}$$

Using spherical symmetry, we have obtained the ground and first states depicted in Fig. 4.4. The boundary conditions imposed are as follows:

$$\begin{aligned} \varphi(0) = 1 \quad & \varphi_r|_{r\to\infty} = 0, \\ \phi(r\to\infty) = -\frac{M}{r} \quad & \phi_r|_{r\to 0} = 0. \end{aligned}$$

This clarifies the specific boundary conditions applied in the system. From Fig. 4.4, it can be observed that the minimum required volume for an isolated configuration in the ground states is 24^3 units of volume in code units. This requirement enables us to conduct simulations with a resolution of $\Delta x < 0.1$. However, for realistic simulations, the box size should be larger, which would result in a decrease in resolution. As we can see in the equation (4.9), the solutions can be found if we have the eigenvalue $\phi - \gamma$. Nevertheless, this value is not scalar in the numerical context since we define ϕ for every mesh point. Due to this fact, we have to find a eigenvector instead. The methodology employed for this equation involves fixing ϕ_0 and randomly varying the γ parameter until the solution converges, indicated by $\phi_h < \infty\ \forall\phi_i \in \phi_h$.

This empirical approach relies on trial and error to determine the appropriate parameter values that lead to convergence. Once we have achieved convergence we have to find an appropriate value of ϕ, since $\lambda = \phi - \gamma$ must be constant for every ϕ_i, we can fix the value of λ and then find several configurations (ϕ_0, γ) that give rise to solutions of the equation (4.9). Figure 4.3 displays the outcome of this parameter variation for the first excited mode as depicted in Figure 4.4. This computational approach is commonly referred to as the shooting method, wherein we assign a boundary condition $\phi(r = 0)$ and subsequently strive to obtain a suitable fit for the other boundary condition, $\phi(r \to \infty)$.

These solutions denoted as ψ_i and ϕ_i (where i corresponds to the excited level of the system), can be utilized to establish initial conditions for our simulations. Let us designate ψ_0 and ϕ_0 as the core solutions of the system described by equations (4.9). Since these solutions exhibit stability in isolation, it is crucial to maintain a certain distance between the cores. For instance, in Figure 4.4, we observe that a distance of $r = 8$ is sufficient to consider it as the asymptotic limit of $r \to \infty$. Beyond this limit, we consider the cores to be effectively isolated. Furthermore, it is possible to give an initial momentum $\boldsymbol{p} = (p_x, p_y, p_z)$ to the core by means of the momentum operator that is defined by the product $\psi \times \mathrm{e}^{i\boldsymbol{p}\cdot\boldsymbol{x}}$.

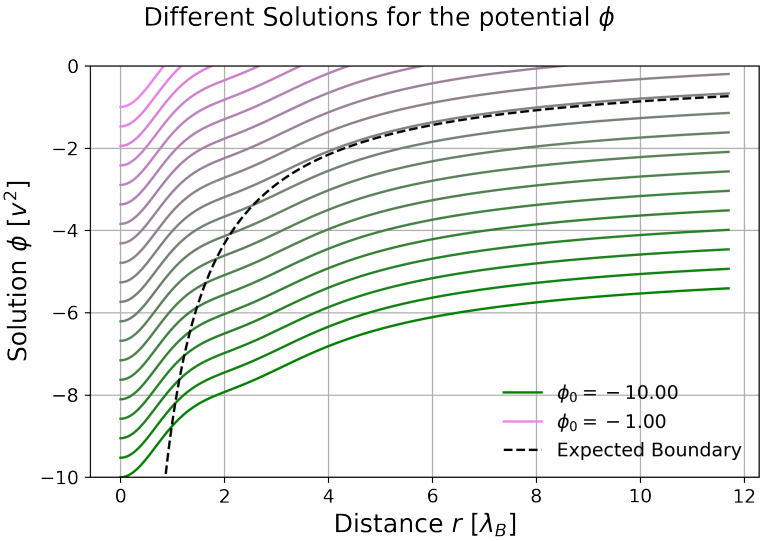

Figure 4.3: Convergence analysis using the shooting method for the potential for the first excited mode. The image illustrates the process of varying the value of $\phi(r=0)$ to achieve the condition $\phi(r\to\infty) = -M/r$ while keeping fixed γ in (4.9). Different solutions for the potential ϕ are plotted with varying initial values $\phi(r=0)$, indicated by different colors. The expected boundary, represented by a dashed black line, is shown for reference. The total mass and the corresponding potential energy are calculated for each solution. The convergence of the shooting method is analyzed by comparing the obtained solutions with the expected boundary.

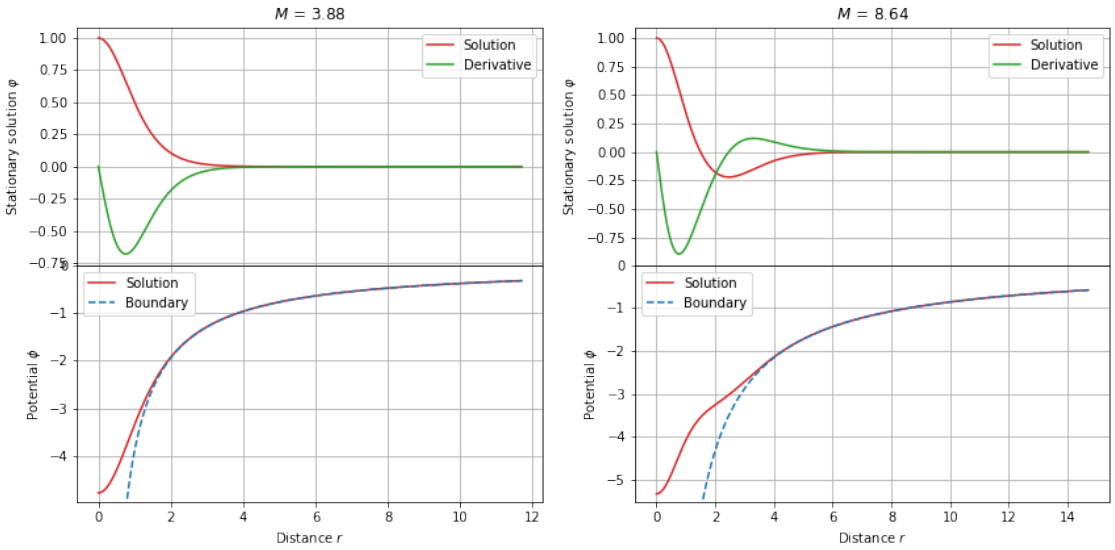

*Figure 4.4: (**Left**) Ground state of the soliton. (**Right**) The first excited state of the soliton. **Above** the amplitude φ and its derivative. **Below** the potential density and the imposed boundary in the infinity. Both solutions were obtained assuming spherical symmetry in the time-independent Scrödinger-Poisson system.*

4.3 Stability Tests of the Solutions

To verify the approximations considered in this section, we will examine specific cases such as the harmonic oscillator, for which we have the analytical temporal evolution. Additionally, we will solve the time-independent Schrödinger-Poisson system to find the eigenvalues of the self-gravitating system. Although we do not have an analytical form for the temporal evolution of the system, we know that the solution is stationary, specifically $\langle|\psi(\boldsymbol{x},t)|^2\rangle = \langle|\psi(\boldsymbol{x},0)|^2\rangle$. We will use this solution and evaluate the capability of this algorithm to preserve this condition. This solution will be related to dark matter halos. Finally, we will perform convergence tests to demonstrate that the algorithm converges to expected solutions as Δx, Δy, and $\Delta z \to 0$.

Harmonic Oscillator

As an illustrative example, Let us consider the harmonic oscillator. The initial condition for this system is given by

$$\Psi_{n_x,n_y,n_z}(\mathbf{r},0) = \frac{\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{3}{4}} e^{-\frac{m\omega}{2\hbar}\boldsymbol{r}^2}}{\sqrt{2^{n_x+n_y+n_z}n_x!n_y!n_z!}} H_{n_x}\left(\sqrt{\frac{m\omega}{\hbar}}\boldsymbol{r}\right) H_{n_y}\left(\sqrt{\frac{m\omega}{\hbar}}\boldsymbol{r}\right) H_{n_z}\left(\sqrt{\frac{m\omega}{\hbar}}\boldsymbol{r}\right), \quad (4.10)$$

where ω represents the angular frequency of the harmonic oscillator, $\mathbf{r}$ is the position vector in the three-dimensional space, and n_x, n_y, n_z are the quantum numbers associated with the excitation of the system. The term H_n denotes the Hermite polynomial of order n. This initial condition describes a stationary solution of the Schrödinger equation, implying that the probability density $|\Psi(\mathbf{r},t)|^2$ remains constant over time. This initial condition is a stationary solution of the Schrödinger equation ($|\Psi(\boldsymbol{r},\boldsymbol{t})| = |\Psi(\boldsymbol{r},\boldsymbol{t}=\mathbf{0})|$) when

$$V = \frac{1}{2}m\omega\boldsymbol{r}^2, \qquad \lambda = 0.$$

For this system, the energy is calculated using the formula

$$E_n = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega,$$

and consequently, the temporal evolution of equation (4.10) can be calculated using the formula

$$\Psi_{n_x,n_y,n_z}(\mathbf{r},t) = \Psi_{n_x,n_y,n_z}(\mathbf{r},0)\mathrm{e}^{-iE_nt/\hbar}.$$

With this formula, we can compare the numerical solution using the formula

$$\mathrm{Err} = \left|\Psi_{\mathrm{Ana}} - \Psi_{\mathrm{Num}}\right|. \quad (4.11)$$

This exercise is performed in $d = 2$ dimensions. Figure 4.5 shows the evolution of the Schrödinger equation using the Crank-Nicolson method (A.21) with $\theta = 1/2$. The left panels display the real part of the numerical solution, while the right panels show the error calculated using formula (4.11) for different time steps. The figure demonstrates that the error is on the order of $\Delta x\Delta y \sim 10^{-4}$, as expected. Another important point to note is that the fixed boundary $\Psi = 0 \quad \vec{\boldsymbol{x}} \in \partial\Omega$ does not affect the solution, as the error in that region is less than 10^{-6}, which was truncated for the sake of the graph.

4.3.1 Ground state test

A simple test that can be performed on the Schrödinger-Poisson system is to consider stationary oscillation modes such as spherical harmonics or spherically symmetric solutions. The standard

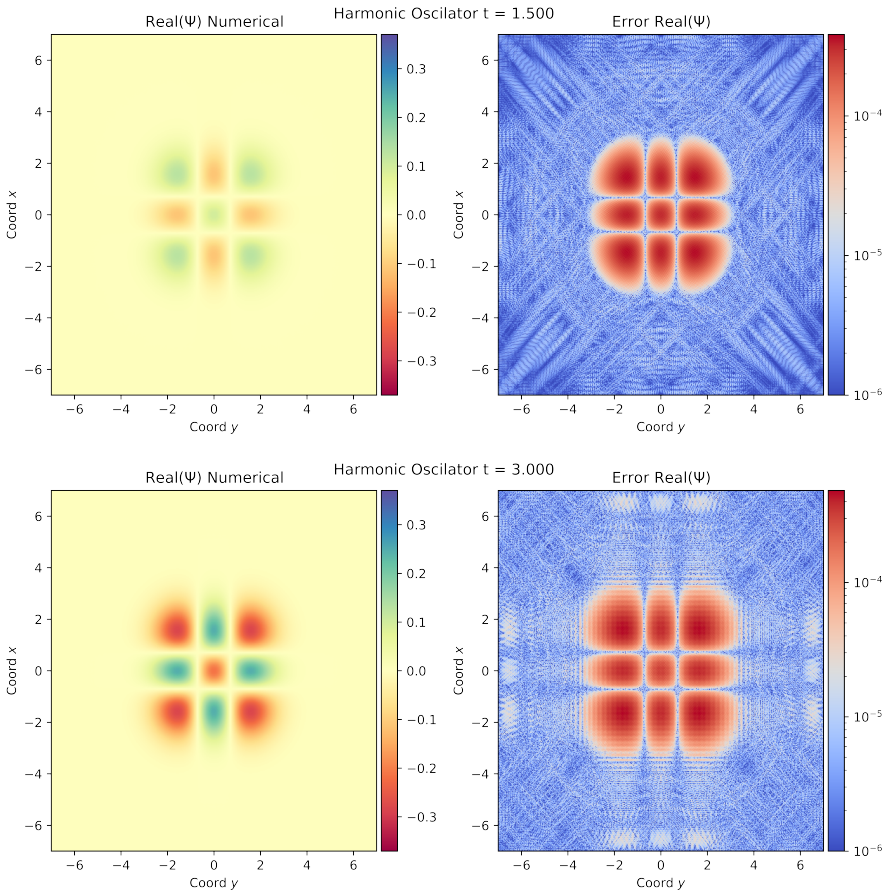

*Figure 4.5: Real part of the numerical solution of the harmonic oscillator (**Left**) in two dimensions with $n_x = n_y = 2$ for a domain with 512^3 points and a time step of 10^{-3}. The simulation considered $\omega = \hbar = m = 1$. The Schrödinger equation (4.5) is solved using equation (A.21) with $\theta = 1/2$. The solution is shown at different times. (**Right**) The error, obtained by comparing the absolute difference between the analytical and numerical solutions, is displayed. The maximum error is approximately 10^{-4}.*

test is conducted on the spherically symmetric ground state. The idea is that since the system is stationary, the initial density should remain invariant.

However, due to numerical errors, the density may vary. As mentioned before, we expect that as $\Delta x \to 0$, Err $\to 0$. In other words, the numerical error should disappear as the precision increases. We performed tests on the ground state (see Figure 4.4), using a domain $[-20, 20]^3 \subset \mathbf{R}^3$ with $N = 100^3$, 200^3, and 300^3 points in the grid. Figure 4.6 shows the results considering two different types of boundary conditions: Dirichlet and periodic.

For periodic boundary conditions, the modified Poisson equation (4.5) needs to be solved, while for Dirichlet boundary conditions, the regular Poisson equation is used. Regarding the boundary conditions for ψ, we consider $\psi(\boldsymbol{r}) = 0$ on the boundary $\boldsymbol{r} \in \partial\Omega$, and for the potential ϕ, we consider $\phi = -M/r_c$, where $M = \int \psi\psi^* dxdydz$ and r_c is the distance from the maximum point of the soliton denoted by (x_c, y_c, z_c) to the boundary $\boldsymbol{r}_c = \boldsymbol{r} - (x_c, y_c, z_c)$ for $\boldsymbol{r} \in \partial\Omega$. This example serves to demonstrate the convergence of the algorithm described above.

In Figure 4.6, the convergence with respect to resolution on the density can be observed, along with the stability of the mass. The latter is important because, in a closed system, this parameter should remain invariant. However, in numerical simulations, this can change due to numerical errors.

4.3.2 Boosted configuration

In this exercise, we take one core initially at $r = (20, 0, 0)\lambda_B$ and apply an initial momentum of $p = \pm(p_x, 0, 0)$ for $p_x = -1$ to make it move. The resulting snapshots are shown in Figure 4.7. With this experiment, we aim to demonstrate that both systems can be equalized in terms of units and that the dynamics are equivalent in the simplest case as well as the steady-state case.

4.4 Some Test Simulations

4.4.1 Head-on Collisions

To simulate head-on collisions, we will utilize the core defined as the base state in Figure 4.4. In this experiment, we consider two cores initially separated by a distance of $r = 12\lambda_B$ and positioned at $r = \mp(6, 0, 0)$. This initial configuration ensures that the cores are initially isolated from each other. Furthermore, we impart an initial momentum of $p = \pm(p_x, 0, 0)$, where p_x takes the values 1, 2, and 3, to make the cores approach and collide. The results of these simulations are shown in Figures 4.8, 4.9, and 4.10. In 4.9, and 4.10 we omitted the initial condition since is the same in all three cases.

In Merger Case 1 (Figure 4.8), characterized by a slow merger, the energy of the system is negative. This negative energy indicates that the cores are gravitationally bound, and most of the matter remains confined within the potential well, as expected.

In Merger Case 2 (Figure 4.9), the energy of the system is positive, indicating that the cores are not gravitationally bound. The kinetic energy of the merging objects surpasses their gravitational attraction, allowing the cores to overcome the gravitational force and potentially move away from each other's influence. This scenario suggests the possibility of solitonic behavior. However, despite the positive energy, the overall energy levels of the system remain low, resulting in significant perturbations of the cores.

In Merger Case 3 (Figure 4.10), characterized by a greater positive energy, solitonic behavior is observed. However, for the SFDM (Scalar Field Dark Matter) case, substantial perturbations are noticeable. These perturbations arise due to the presence of a quantum pressure term intrinsic to SFDM, which restricts the occurrence of a "Dry merger" case characterized by minimal perturbations.

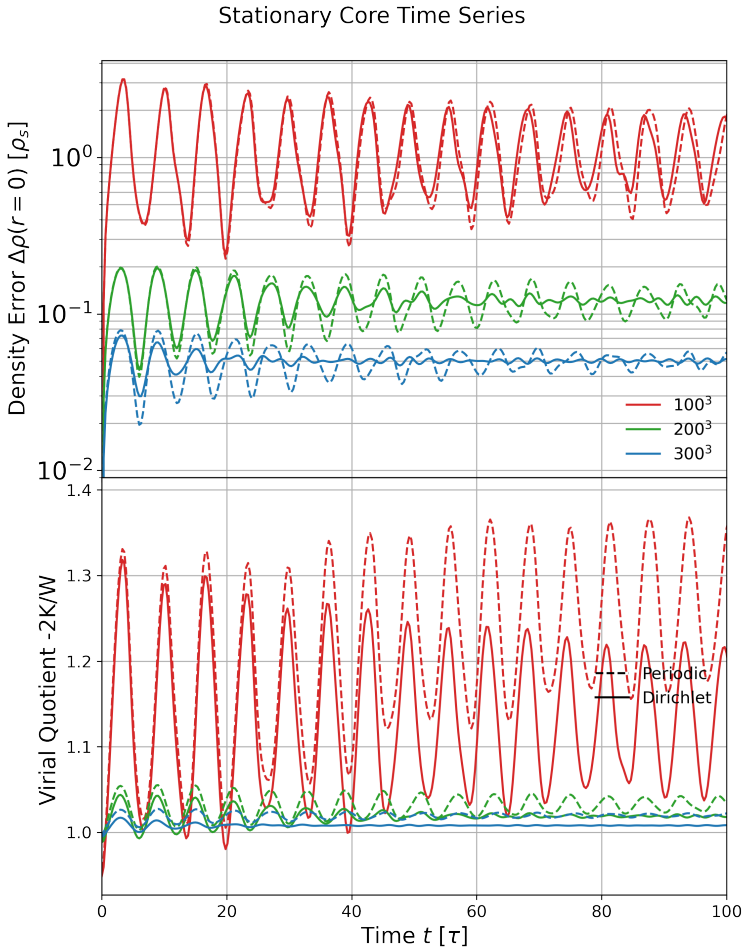

Figure 4.6: Temporal evolution of the error value in density at $\boldsymbol{r} = 0$, mass M, and virial energy $2K+W$ is depicted in this image. It demonstrates that as we increase the number of points in the grid from $N = 100^3$ to 200^3 to 300^3, the solution approaches the desired value of $\Delta\rho(\boldsymbol{r} = 0) = 0$. The mass remains constant regardless of resolution, and the Virial Energy tends towards one, as expected for a stationary system. The dashed lines correspond to the periodic boundary conditions, whilst the solid lines represent the periodic Dirichlet conditions.

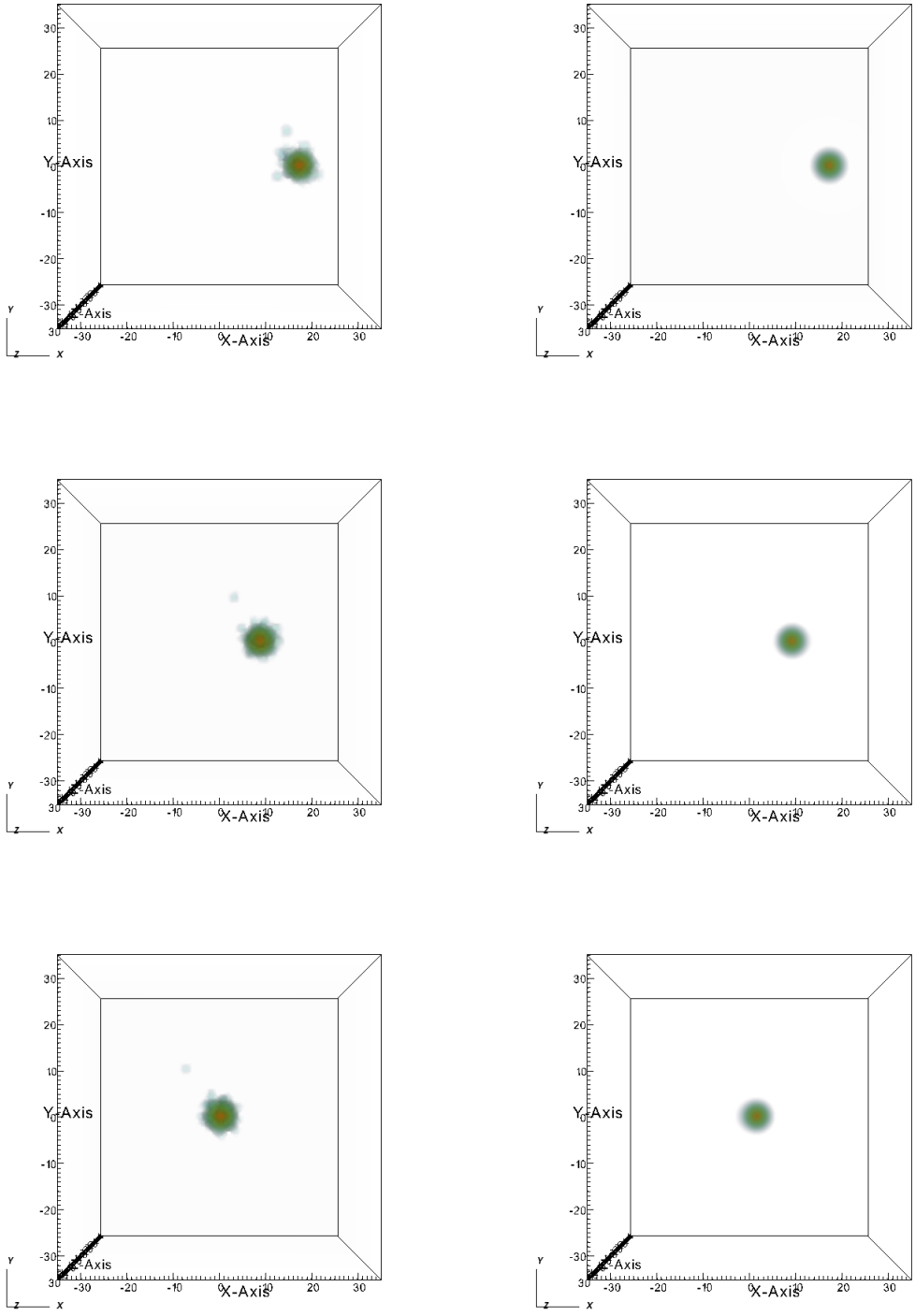

*Figure 4.7: The simplest case apart from the steady one. Here both models are equivalent, this test is made to verify that units are being managed properly. (**Left**) Gadget simulations, (**Right**) SFD simulations. We show the initial, central, and last snapshots of the simulated halos. We computed the density of the Halos using the Coarse-Grained B.*

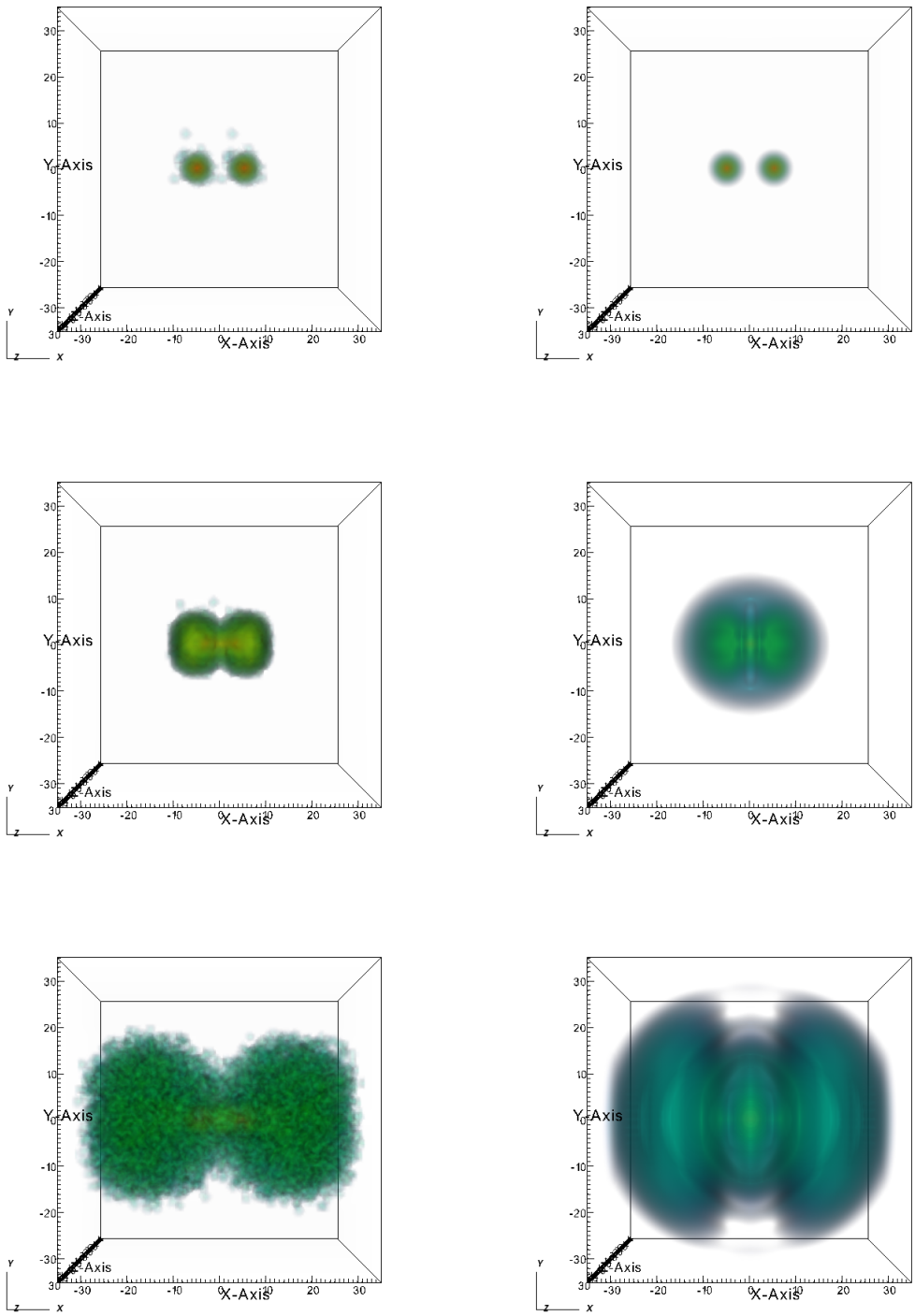

Figure 4.8: Merger Case 1: Test of a slow merger. In this case, the energy is negative, indicating that the cores are gravitationally bound. Most of the matter remains contained in the potential well, as expected.

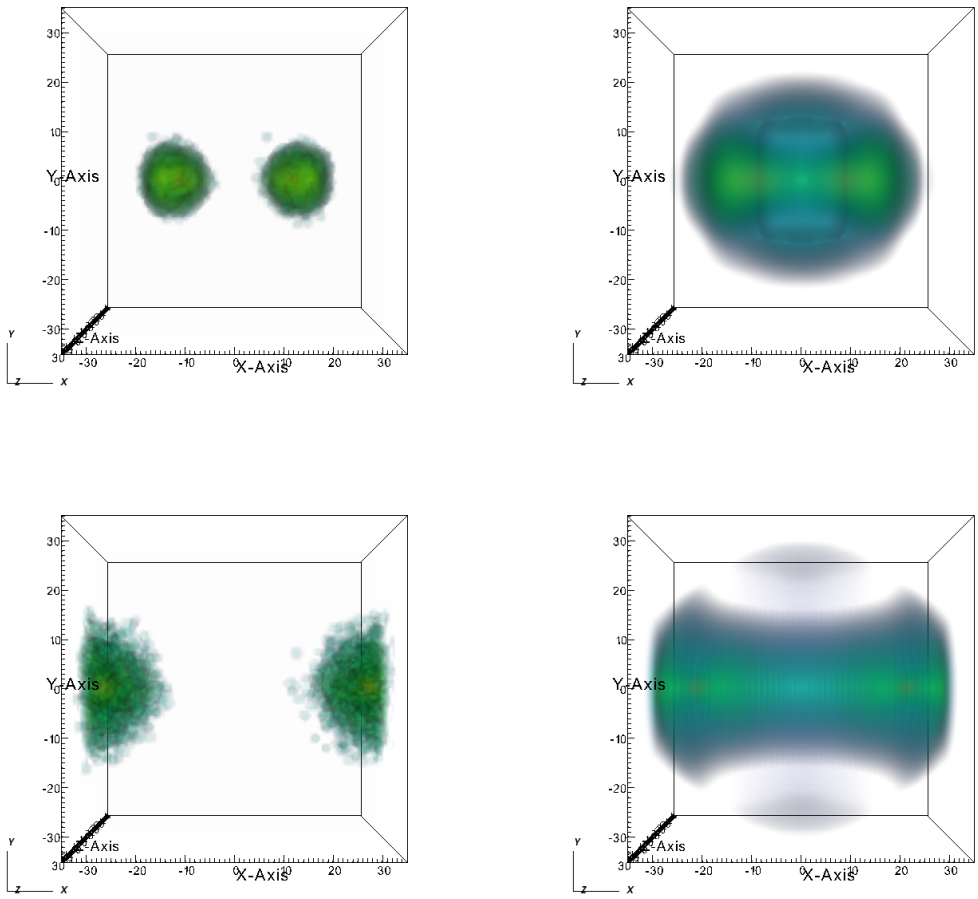

Figure 4.9: Merger Case 2: In this case, the energy is positive, indicating that the cores are not gravitationally bound and the kinetic energy of the merging objects exceeds their gravitational attraction. As a result, the cores have enough kinetic energy to overcome the gravitational force and potentially escape from each other's influence. This scenario suggests the possibility of solitonic behavior. However, it is important to note that despite the positive energy, the overall energy levels of the system remain low, and as a consequence, the cores become perturbed.

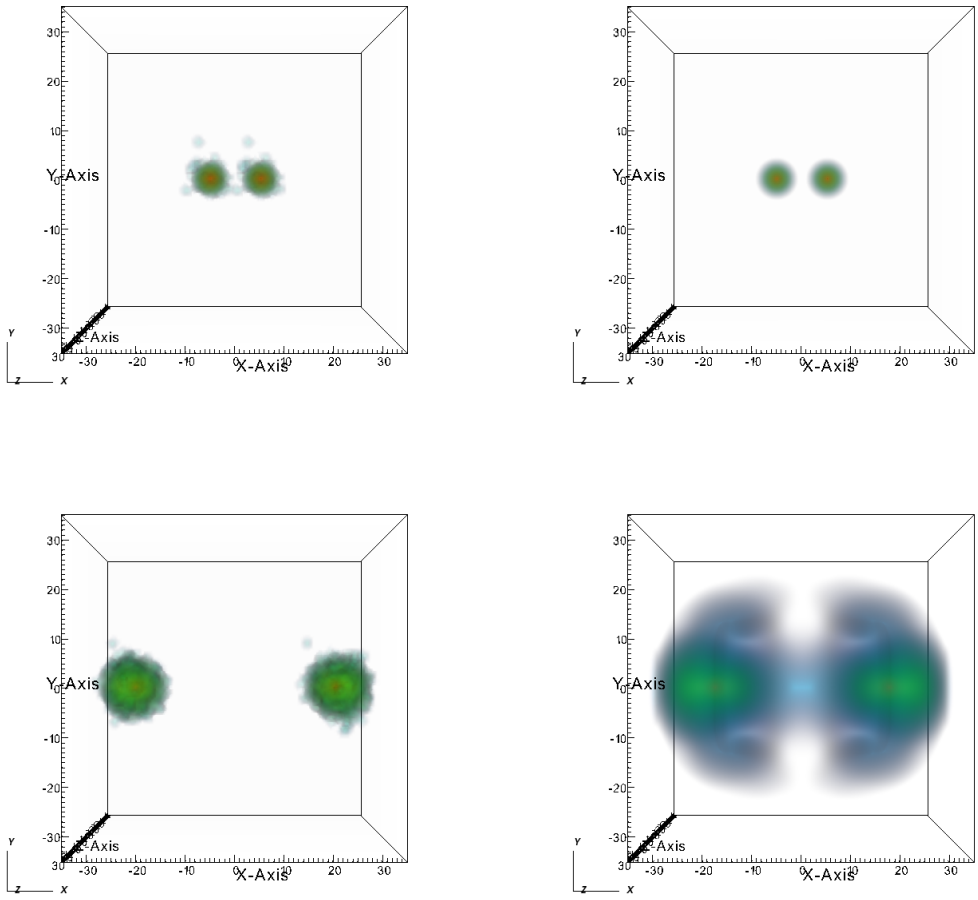

Figure 4.10: Merger Case 3: In this case, the energy is positive and greater, indicating solitonic behavior. However, for the SFDM (Scalar Field Dark Matter) case, we observe significant perturbations. These perturbations arise due to the presence of a quantum pressure term inherent to SFDM, which restricts the occurrence of a "Dry merger" case.

4.4.2 Off-center Collisions

Similar to the previous case, we examine off-center collisions between two cores. In this exercise, the cores are initially positioned at $r = \mp(6, 1, 0)$, slightly off-center. Additionally, an initial momentum of $p = \pm(1, 0, 0)$ is applied to bring the cores closer and induce a collision. The results of these simulations are presented in the following graphs.

Studying off-center mergers and analyzing the impact of quantum pressure on the dynamics of dark matter models offers valuable insights into the formation and evolution of structures in the universe. Further analysis and simulations are necessary to explore the intricate interplay between gravitational forces, quantum pressure, and their effects on the merger process. For instance, it is essential to investigate mergers between realistic density profiles and study the resulting rotational curves.

The figures in Figure 4.11 depict the outcomes of the off-center collisions. This merger scenario is characterized by negative energy, suggesting that the potential well retains most of the mass. As expected, the presence of tidal forces is more pronounced in the SFDM case due to the influence of quantum pressure.

Due to the symmetry along the z-axis, we expect the merger to occur at $r = 0$. At $r = 0$, the density exhibits core-like behavior in the center due to the quantum pressure effect. Figure 4.12 illustrates this phenomenon in the case of SFDM.

4.4.3 Binary Configuration

In this section, we consider a binary system consisting of two cores initially positioned at $r = \mp(6, 0, 0)$. To simulate circular motion in a binary configuration, we apply an initial momentum of $p = \pm(0, v_c, 0)$, where v_c is calculated using the formula:

$$v_c = \sqrt{\frac{M}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}},$$

Here, M is calculated as $4\pi \int \rho r^2 dr$ for a spherically symmetrical soliton.

The results of the simulation are presented in Figure 4.13. This binary system exhibits negative energy, indicating that the system is gravitationally bound. However, the dynamics, in this case, are somewhat different compared to the previous scenarios, as the cores initially move perpendicular to each other. We expect the system to undergo oscillations before eventually collapsing.

Figure 4.13 illustrates the evolution of the binary system. The left column shows the results from the Gadget code simulation, while the right column displays the corresponding results from our SFDM code. Despite the initial perpendicular motion, the cores exhibit oscillatory behavior, indicating the influence of gravitational attraction. Over time, the cores are expected to approach each other and experience a collapse due to their mutual gravitational interaction.

Additionally, Figure 4.14 provides further insights into the system. The top panel shows the mass of the system as a function of time. Due to the presence of an absorbing potential near the boundaries, the mass gradually disappears from the system. As the impact of this effect becomes more significant, the accuracy of the solutions decreases.

The bottom panel of Figure 4.14 presents the virial quotient, which serves as a measure of the system's stability. A value of 1 indicates complete steadiness, while $2K/W > 1$ indicates increasing perturbations. By monitoring the virial quotient, we can assess the level of perturbations and fluctuations in the system.

4.4.4 Halo in an External Gravitational Field

The case of an external potential Φ_{External} halo is treated in the Schrödinger equation (4.12) as an additional term, alongside the self-gravity term Φ.

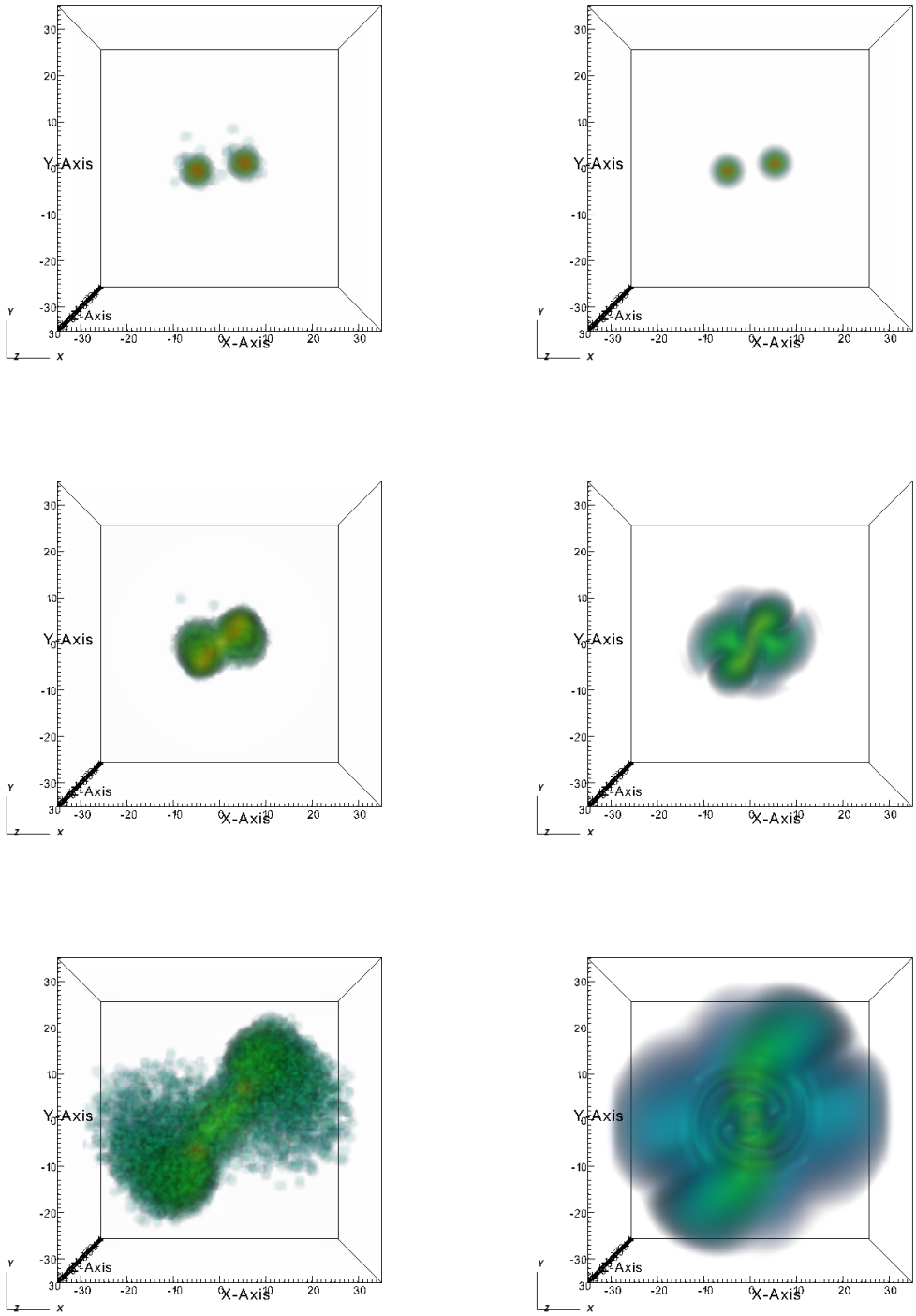

Figure 4.11: Merger Case Off-center. This merger has negative energy, we are expecting that the potential well retains most of the mass. In the figures, we can see how tidal forces are stronger for the SFDM case due to the quantum pressure.

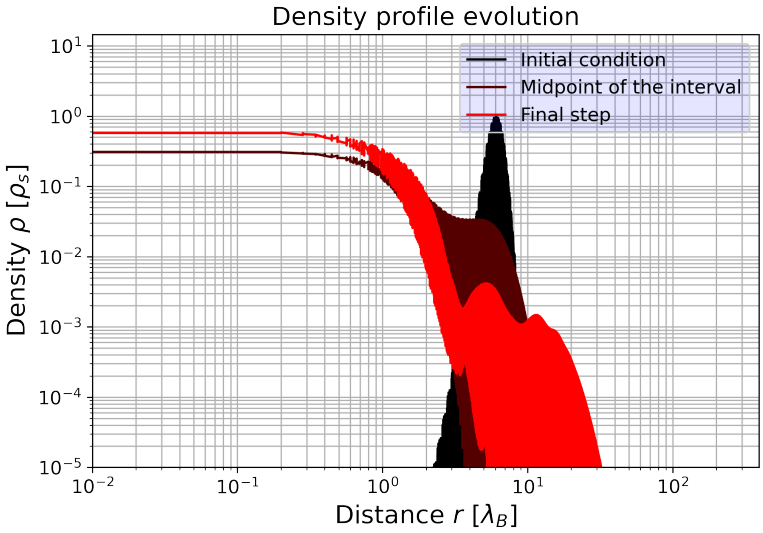

Figure 4.12: Density evolution during the merger. The density rapidly starts to form a core in the center, while all the matter outside follows an exponentially decreasing NFW-like density profile.

$$i\dot{\psi} = \left(-\frac{1}{2}\nabla^2 + (\Phi + \Phi_{\text{External}}) - 4\pi\tilde{\lambda}_s\,|\psi|^2\right)\psi. \tag{4.12}$$

For this problem, we consider a core as a test particle at an initial position of $r = (-7, 0, 0)$ and give it an initial momentum of $p = (0, 2, 0)$. This core is embedded in a Host halo with a Zhao profile defined by the parameters

$$\begin{aligned} r_s &= 5, \\ \log 10\rho_s &= 7, \\ \alpha, \beta, \gamma &= 2, 10, 0. \end{aligned}$$

In the figure 4.15 we show the solution of the Poisson equation given the density defined above. Note we have considered the limit of infinity to be satisfied in order to have confidence in the solution near the center. On the other hand, we have considered the same boundary conditions as for the ground state case 4.4.

Additionally, in Figure 4.16, we illustrate the displacement of the core around the external potential V_e, which is determined based on the maximum value of the density. However, it is important to note that this value may become inaccurate when the soliton falls into the center of the potential, as it can be heavily influenced by tidal forces and experience complete disruption of the core.

Due to considering a *sponge potential*, the mass will gradually disappear as we approach the boundaries, as shown in Figure 4.17, illustrating this edge effect.

The evolution of the halo is described in Figure 4.18, which shows three main steps of the simulation: the initial, intermediate, and final steps. As the configuration falls into potential, the profile becomes more extended. Nevertheless, the model predicts a core shape in the center in contrast to CDM models.

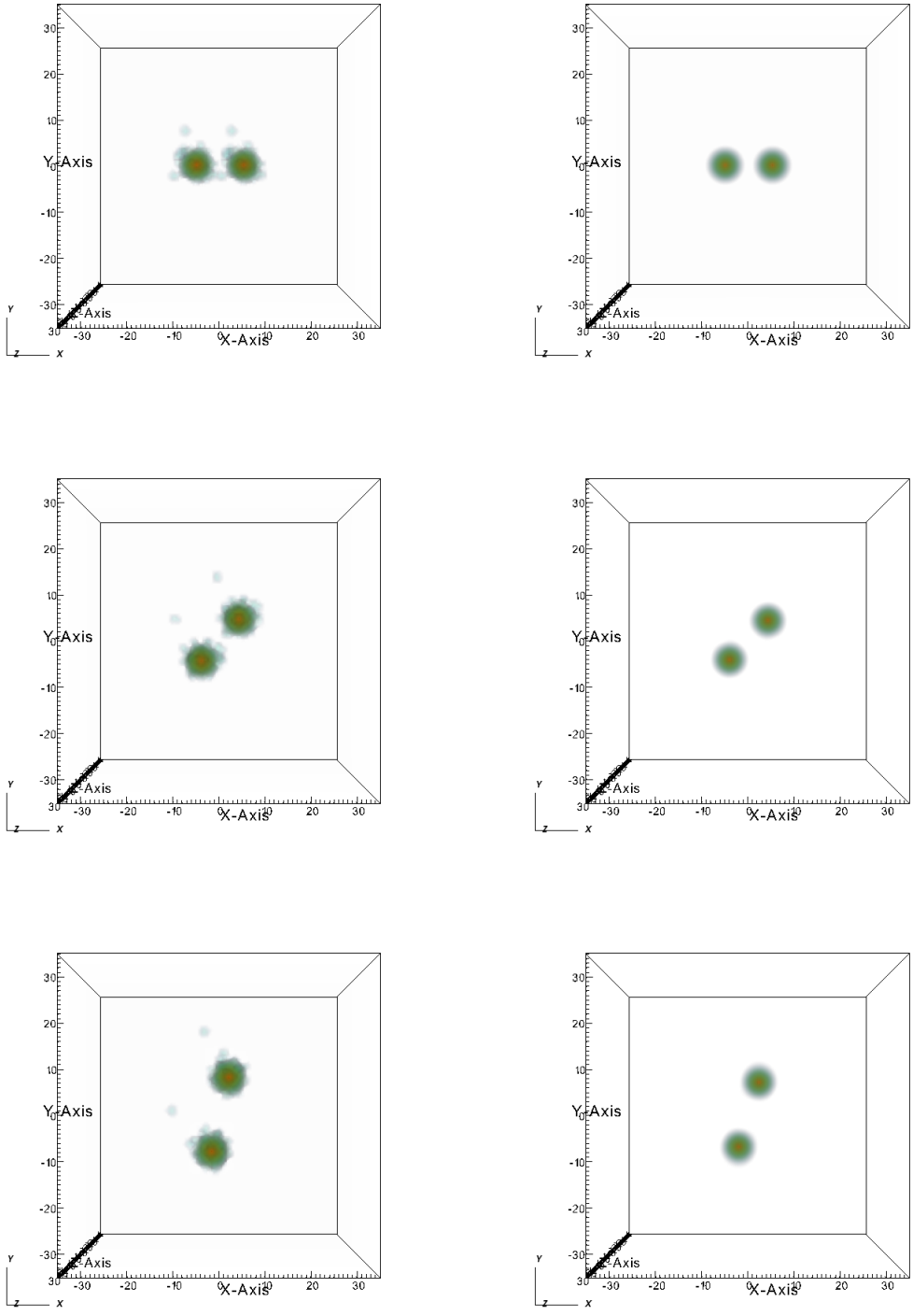

Figure 4.13: Binary system. This case has negative energy which means that the system is gravitationally bounded. However, the dynamics here are somewhat different than in previous cases since the cores are moving perpendicularly at the beginning. We are expecting to stay oscillating for some time before collapsing.

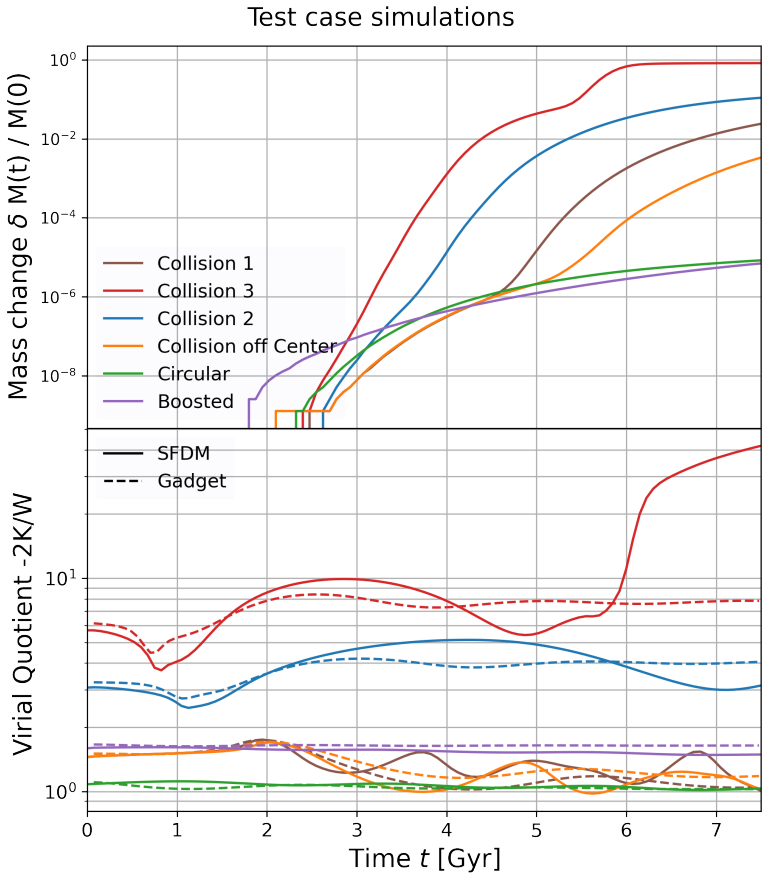

Figure 4.14: ***Above:*** *Evolution of the system's mass over time. The mass of the system gradually diminishes due to the presence of an absorbing potential near the boundaries. As this effect becomes more significant, the accuracy of the solutions decreases.* ***Below:*** *Virial quotient as a measure of the system's stability. A value of 1 indicates complete steadiness, while $2K/W > 1$ signifies that the system is perturbed.*

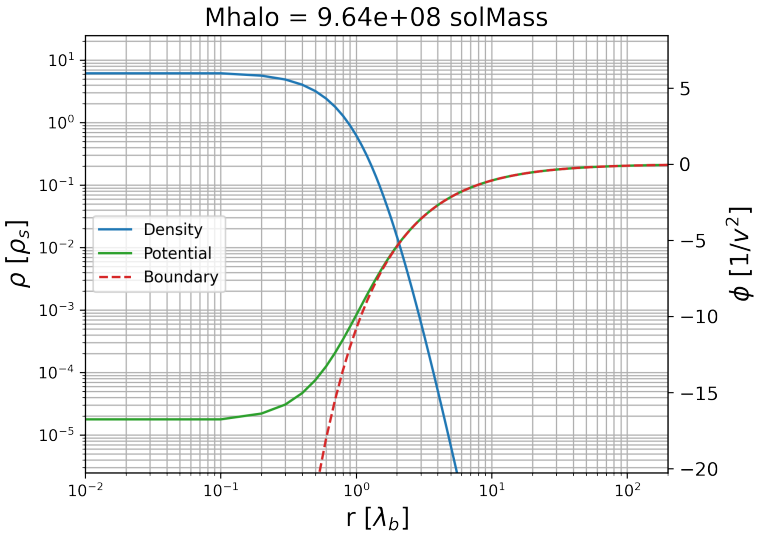

Figure 4.15: Host halo created solving the Poisson equation (4.6) *for a density given by the Zhao prescription* (3.1)*. The plot shows the limit to the infinity in the potential ϕ.*

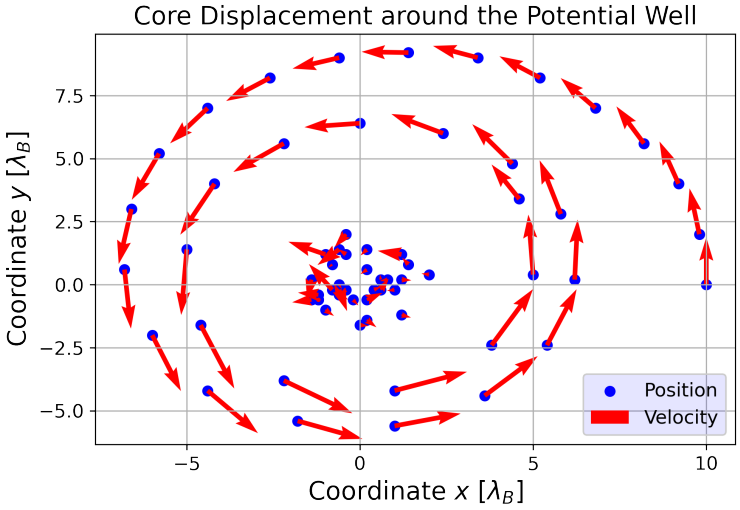

Figure 4.16: Plane $z = 0$ of the simulation. Positions of the density maximum and the expectation values of the momentum $\vec{p} = (\langle p_x \rangle, \langle p_y \rangle)$ represented by points and arrows, respectively. The magnitude of the momentum is represented by the size of the arrows.

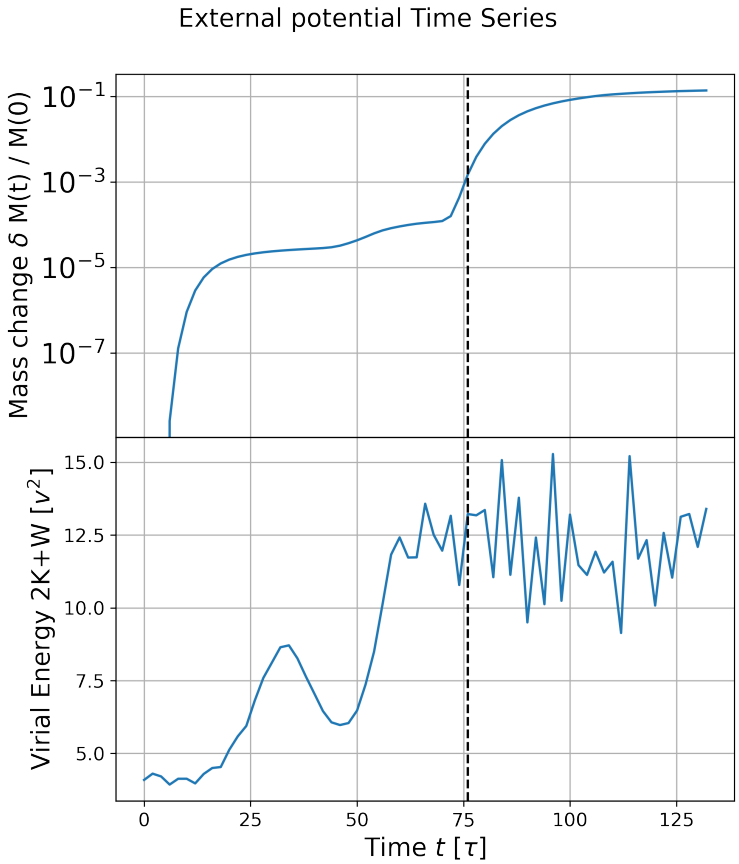

Figure 4.17: ***Top*** *Mass loss due to edge effects. When the soliton falls into the potential well, tidal forces eject material out of the system, reaching the boundaries and causing mass loss. When the system has lost too much matter, the results will be considered inaccurate. Bottom The virial coefficient, which represents a measure of the system's stability. When it tends to one, we say that the system is relaxing. In this example, despite the mass loss, the system is dynamically active. The black dashed line represents the limit where the relative percentage change is 1%.*

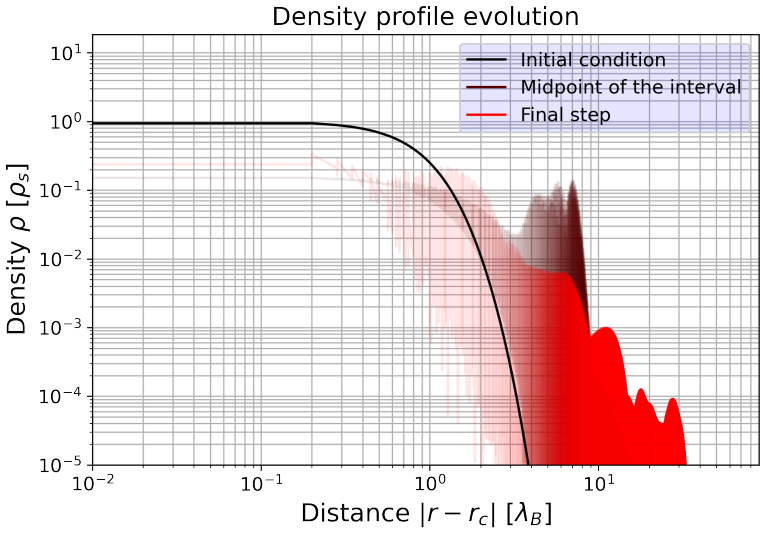

Figure 4.18: Evolution of the density measured from the central point of the maximum density. This test demonstrates the presence of the Core-NFW Schive et al. (2014a) profile behavior of the evolution.

4.5 Resimulations

Resimulations play a crucial role in ΛCDM cosmology as they provide a valuable tool for studying the formation and evolution of large-scale structures in the universe. These resimulations involve the recreation of the processes involved in the formation of galaxies and cosmic structures using advanced computational techniques. They are based on the ΛCDM model, which explains the accelerated expansion of the universe through the existence of dark energy (Λ) and cold dark matter (CDM).

In a ΛCDM resimulation, a sufficiently large volume is simulated to capture the relevant scales of structure formation. Initial conditions are generated to reflect the observed density fluctuations in the early universe. By combining particle simulations with hydrodynamics techniques, the evolution of both dark matter and baryonic matter can be simulated over time. These resimulations enable the study of cluster formation, cosmic filaments, and other large-scale structures, as well as the exploration of the interactions and distribution of dark matter and baryonic matter throughout the universe.

On the other hand, resimulations using Initial Conditions (ICS) derived from ΛCDM at specific redshifts are employed to study the evolution of Scalar Field Dark Matter (SFDM) within the framework of the ΛCDM model. SFDM proposes that dark matter is composed of an ultra-light scalar field, offering an alternative to the concept of cold dark matter. By utilizing initial conditions derived from ΛCDM simulations at a specific redshift, these resimulations allow for the study of the evolution of scalar dark matter through advanced numerical methods.

To create the initial conditions, we used a cosmological ΛCDM simulation and identified halos using the Rockstar code Behroozi et al. (2012a). We traced back the merger history for the selected halo using the ConsitetnTrees code Behroozi et al. (2012b). Specifically, we focused on identifying massive halos at $z = 0$ with masses $m \geq 1 \times 10^{11}\ M_{\odot}$. Additionally, we ensured that all halos in the merger history were traced until $z = 1$, which is the redshift at which the initial conditions for the SFDM model were created. These resimulations span from $t \sim 7.817$ Gyr/h to the present day.

In order to create appropriate initial conditions, it was necessary to match the mass of the particles with the mass of the core. This was achieved by setting $m_{\rm sfm} = 4.78 \times 10^{-23}$ eV and

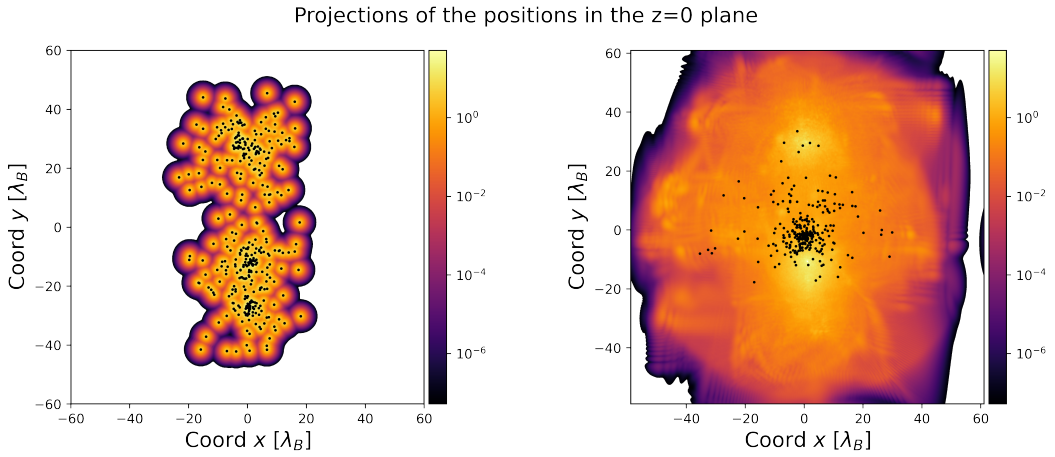

*Figure 4.19: This figure presents an example of a resimulation, where the red dots represent the N-body particles from the cosmological simulation at $z = 1$ (**Left**) and $z = 0$ (**Right**). The colormap illustrates density created by replacing the particles with solitons in the SFDM model. We compared the evolution in both the ΛCDM and SFDM cases.*

$v_{\rm sfdm} = 10$ km/s. Figure 4.19 (left) illustrates the process of creating initial conditions for the SFDM model, where the particles are replaced by solitons. Additionally, the velocity of each particle was set using the momentum operator

$$\psi_i \to \mathrm{e}^{i\boldsymbol{p}_i \cdot \boldsymbol{x}},$$

for each i-th particle, taking into account that in the adimensional system, this operator becomes $\boldsymbol{p}_i \sim \boldsymbol{v}/v_{\rm sfdm}$. Once the system evolves over time, the results of the SFDM and ΛCDM cases need to be compared. Figure 4.19 shows a comparison between these simulations, highlighting the differences that may arise. It should be noted that the differences depicted in the figure are primarily due to the resolution of the SFDM case, which utilizes only 350^3 mesh points. A higher-resolution version is currently being run to obtain more accurate results.

Chapter 5

Conclusions and Perspectives

We have presented the results of an algorithm that solves the 3D Schrödinger-Poisson system, which constitutes a powerful tool for the study of complex systems, such as compact groups of galaxies. Furthermore, we have showcased the results obtained by applying a variation to the ΛCDM model, considering dark matter as an imperfect fluid. This consideration has a direct impact on primordial perturbations and is reflected in the population of observed halos and galaxies in the present universe.

Additionally, we have explored two approaches to working with observational data. On one hand, we have focused on extracting the dark matter halo from observational data of the velocity map of a spiral galaxy. On the other hand, we have analyzed the limitations of a deep neural network when exposed to observed and estimated scaling relations using fictitious mock catalogs.

These results are of great significance as they contribute to the advancement of knowledge in the fields of cosmology and particle physics. The study of complex systems and the consideration of dark matter as an imperfect fluid allow for a deeper understanding of the structure and evolution of the universe. Furthermore, the approaches employed for working with observational data offer new perspectives and challenges in the analysis of large datasets.

Looking ahead, there are several avenues for future research. One direction is to further refine the algorithm for solving the Schrödinger-Poisson system, exploring more efficient numerical methods and considering additional physical effects. Another direction is to expand the analysis of observational data, incorporating more diverse datasets and developing advanced statistical techniques to extract meaningful information. Furthermore, the investigation of the impact of imperfect dark matter fluid models on larger cosmological scales and their implications for structure formation and galaxy evolution remains an exciting area of exploration.

Overall, the findings presented in this work open up new opportunities for investigating the nature of dark matter, understanding the complex dynamics of galaxies and galaxy clusters, and advancing our knowledge of the fundamental principles underlying the universe.

During this research, sufficient tools and databases have been discovered to study the dynamics of dark matter directly from nearby galaxies, as exemplified by the MaNGA catalog. Understanding the properties of dark matter halos in nearby galaxies is important because it allows for comparisons with cosmological simulations and enables a detailed study of the evolution of dark matter halos at small scales using the SFDM model.

Furthermore, the moving mesh method has various applications, with the simplest being analogous to mesh refinement at different levels. This is a particular case of the AMM method that can be used in cosmological simulations, akin to the well-known "zoom simulations." Moreover, the AMM method can refine not only based on regular patterns but also according to different functions, providing a plethora of possibilities.

The approach taken in this research is fruitful for different areas of physics investigation, as the phenomenon of dark matter has been highly attractive in recent years. Additionally, technology

has evolved at such a rapid pace that it is highly efficient to use simple algorithms such as finite differences and obtain precise and rapid results with the use of graphics processing units (GPUs).

In general, the code developed for this project fulfills two important conditions. Firstly, it is capable of reproducing analytical solutions within a certain order of error. Secondly, it demonstrates convergence as we decrease $\Delta x \to 0$.

Furthermore, the code is compatible with simple and isolated cases in conjunction with the N-body Gadget code, such as the case of the initial momentum configuration. An important aspect to highlight is that the code can currently be executed on a GPU or in parallel for a single node. The execution speed on a GPU is up to nine times faster for the same conditions, but there is a limitation in VRAM memory. The simulations we report were performed on a GPU with 350 grid points per axis, resulting in a requirement of approximately 75-80 GB of memory. However, we have successfully conducted simulations with up to 512 grid points per axis efficiently on a CPU.

This code demonstrates the capability of performing simulations for a wide range of initial conditions, as long as the boundaries are sufficiently far away, as in that case, the simulations for ΛCDM and SFD models are compatible. In this chapter, we studied the limitations of the code in simulating the dynamics of dark matter halos. However, for galactic-scale simulations, we expect mergers to occur with negative energies for two main reasons. Firstly, mergers from cosmological resolutions ultimately form halos, where gravitational potential energy dominates. Secondly, the mergers considered in observations are believed to be stable systems, such as compact groups Hickson (1997). Moreover, codes that search for halos assume that the identified halos are stationary at all times. These conclusions are necessary to propose a methodology for applying the resimulation technique in the SFDM model for comparison with ΛCDM.

On the other hand, the adaptive mesh refinement method provides the possibility of solving the problem with extended boundaries while refining the central region, as seen in Section A.6. In this case, solutions of sufficient quality are obtained in the region of interest, while outside this refined domain, it can be considered as an extended boundary.

One of the main limitations of the code that solves the SP system is that it is not fully parallelized and heavily relies on the availability of hardware. However, it executes in parallel when solving the linear system of equations associated with the partial differential equation. In the future, the technique of GPU memory caching will be employed to utilize resources more efficiently and achieve high-resolution cosmological simulations. Additionally, the AMM method may exhibit degradation in the solutions of the Poisson equation when the created mesh is non-uniform or irregular. This would entail working with two different meshes for each equation of the SP system.

The next steps involve creating cosmological initial conditions to evolve the current code and implementing an algorithm for parallel execution across multiple GPUs to leverage available resources. Furthermore, significant progress has been made in creating initial conditions at small scales based on observations, paving the way for simulations based on real data. Additionally, the implementation of the Schwarzschild method to stabilize dark matter halos with Zhao-like target profiles is planned.

Appendix A

The Finite Difference Method

A.1 Infinitesimal Approximations

In the context of differential equations, derivatives of different orders arise, and they can be defined as a limit. For example, the derivative of a scalar function can be approximated from the left or from the right as follows:

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x) - f(x-\Delta x)}{\Delta x}.$$

We define the difference operators as in Karniadakis and Kirby II (2003) in the following way:

$$\begin{aligned}
\text{Forward Shift:} \quad & E^n f(x) \equiv f(x + n\Delta x). \\
\text{Forward Difference:} \quad & \delta^+ f(x) \equiv f(x+\Delta x) - f(x) = (E-1)f(x). \\
\text{Backward Difference:} \quad & \delta^- f(x) \equiv f(x) - f(x-\Delta x) = (1-E^{-1})f(x). \\
\text{Central Half Difference:} \quad & \delta f(x) \equiv f(x+\Delta x/2) - f(x-\Delta x/2) = (E^{1/2} - E^{1/2})f(x). \\
\text{Central Difference:} \quad & \bar{\delta} f(x) \equiv \frac{f(x+\Delta x) - f(x-\Delta x)}{2} = \frac{E - E^{-1}}{2} f(x). \\
\text{Average:} \quad & \mu f(x) \equiv \frac{f(x+\Delta x/2) + f(x-\Delta x/2)}{2} = \frac{E^{1/2} + E^{1/2}}{2} f(x). \\
\text{Derivative:} \quad & Df(x) \equiv \frac{\partial}{\partial x} f(x).
\end{aligned} \tag{A.1}$$

With these definitions, we can rewrite the derivative as

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{(E-1)f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(1-E^{-1})f(x)}{\Delta x}.$$

Then, by using a Taylor series expansion, we can approximate the shifts as

$$\begin{aligned}
E^{\pm 1} f(x) &= f(x \pm \Delta x) \\
&= f(x) \pm \Delta x f_x(x) + \frac{(\pm \Delta x)^2}{2} f_{xx}(x) + \cdots \\
&= \left[f(x) \pm \Delta x D + \frac{(\pm \Delta x D)^2}{2} + \cdots \right] f(x) \\
&= \mathrm{e}^{\pm \Delta x D} f(x).
\end{aligned} \tag{A.2}$$

Therefore, we can obtain the derivative operator D as

$$\Delta x D = \ln(E). \tag{A.3}$$

Recall that the natural logarithm can also be expressed as a Taylor series

$$\ln(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \cdots,$$

With this, we can write the forward, backward, and central derivatives using (A.1) as

$$\begin{aligned}
\text{Forward Difference:} \quad & \Delta x D = \ln(E) = \ln(1+\delta^+) = \delta^+ - \frac{\delta^{+2}}{2} + \frac{\delta^{+3}}{3} - \cdots. \\
\text{Backward Difference:} \quad & \Delta x D = \ln(E) = -\ln(1-\delta^-) = \delta^- + \frac{\delta^{-2}}{2} + \frac{\delta^{-3}}{3} + \cdots. \\
\text{Central Half Difference:} \quad & \Delta x D = 2\text{sinh}^{-1}\delta/2 = \delta - \frac{\delta^3}{24} + \frac{3\delta^5}{640} - \cdots.
\end{aligned} \tag{A.4}$$

For the central derivative, we need to solve for $\Delta x D$ by considering the same procedure as in (A.2)

$$\delta = E^{1/2} - E^{-1/2} = \text{e}^{\Delta x D/2} - \text{e}^{-\Delta x D/2} = 2\text{sinh}\frac{\Delta x D}{2}.$$

Finally, higher-order derivatives can be calculated as before

$$\left(\frac{\partial f(x)}{\partial x^n}\right)^n = D^n f(x),$$

the only difference compared to the previous procedure is that the resulting Taylor series of the first-order derivative needs to be multiplied N times. For example, the second derivative can be expressed as

$$\begin{aligned}
\text{Forward Difference:} \quad & \Delta x^2 D^2 = \delta^{+2} - \delta^{+3} + 11\delta^{+4}/12 - 5\delta^{+5}/6 + \cdots. \\
\text{Backward Difference:} \quad & \Delta x^2 D^2 = \delta^{-2} + \delta^{-3} + 11\delta^{-4}/12 + 5\delta^{-5}/6 + \cdots. \\
\text{Central Half Difference:} \quad & \Delta x^2 D^2 = \delta^2 - \delta^4/12 + \delta^6/90 - \delta^8/560 + \cdots.
\end{aligned} \tag{A.5}$$

A.2 Finite Differences

A finite difference is a numerical approximation of a derivative obtained by taking the difference between the values of a function at two nearby points. Instead of computing the exact derivative, this technique is used to obtain a numerical approximation of the derivative at a specific point. Let's consider a continuous function expressed as a series of points in a discrete domain, as shown in Figure A.1. These points are uniformly spaced along the domain. The approximation of derivatives is done using the difference formulas defined in A.1 truncated to the first term of the Taylor series. Let's define $f(x_i) = f_i$, $f(x_i \pm \Delta x) = f_{i\pm1}$, similarly, we can consider the N-dimensional case $f(x_i, y_j, z_k, \ldots) = f_{i,j,k,\ldots}$. Then, the approximations of derivatives using central differences are expressed as follows:

$$\begin{aligned}
\partial_x f_i &= \frac{f_{i+1} - f_{i-1}}{2\Delta x} + \mathcal{O}(\Delta x^2), \\
\partial_{xx} f_i &= \frac{f_{i+1} - 2f + f_{i-1}}{\Delta x^2} - \mathcal{O}(\Delta x^4), \\
\partial_{xy} f_i &= \frac{f_{i+1,j+1} - f_{i-1,j+1} - f_{i+1,j-1} + f_{i-1,j-1}}{4\Delta x \Delta y} + \mathcal{O}(\Delta x^2 \Delta y^2).
\end{aligned} \tag{A.6}$$

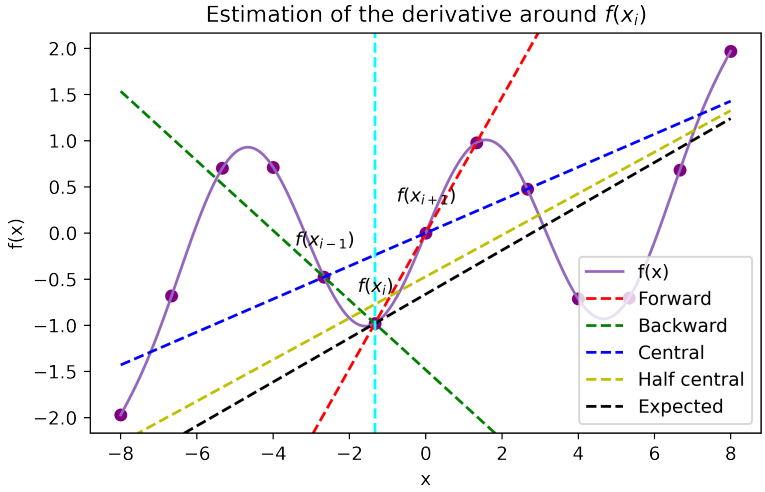

Figure A.1: Different estimations of the tangent line at the point $x_i, f(x_i)$ using a discrete domain and second-order approximation.

A.3 Matrix Factorization

Matrix factorization is a tool used to propose a system of linear equations whose solution approximates the solution of a differential equation. In this section, we will show how the process of discretizing a differential equation is carried out with different examples.

A.3.1 Poisson's Equation

Poisson's equation relates the spatial variation of a scalar field ϕ to the charge density or source that generates it. In other words, it describes how the field ϕ is distributed in a given domain when the acting sources are known. To solve Poisson's equation, numerical methods are commonly used since finding exact analytical solutions is difficult in many cases. Some common methods include finite difference methods, finite element methods, and spectral methods. These methods discretize the domain and approximate the differential equation with algebraic equations, which can be solved using numerical techniques. It is important to note that Poisson's equation is a linear equation, which means that if ϕ_1 and ϕ_2 are solutions of the equation with sources ρ_1 and ρ_2, respectively, then any linear combination of them is also a solution. This has important implications in terms of superposition and linearity of solutions. Poisson's equation is defined as follows:

$$\nabla^2 \phi = \rho. \tag{A.7}$$

The Laplacian operator ∇^2 can be written as

$$\nabla^2 = \partial_{xx} + \partial_{yy} + \partial_{zz}.$$

As we have seen before, these derivative operators can be discretized as shown in Equation A.6.

1D Case

Let's first consider the one-dimensional case, with a domain $x \in \Omega$ and Dirichlet boundary conditions given by

$$\phi = 0 \quad \text{for } x \in \partial\Omega,$$

where the discretization consists of $n-2$ points $x_i \in \Omega_h$. In this discrete domain, the function ϕ can be represented as the vector

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{n-1} \end{bmatrix}, \tag{A.8}$$

and similarly for ρ. With this construction, the ∂_{xx} operator becomes a tridiagonal matrix of size $(n-2) \times (n-2)$ given by

$$D_{xx} = \begin{bmatrix} -2 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -2 \end{bmatrix}, \tag{A.9}$$

thus, Poisson's equation (A.7) can be rewritten as the system of linear equations:

$$\frac{1}{\Delta x^2} D_{xx} \phi = \rho,$$

The solution approximation ϕ can be found using the inverse matrix as follows:

$$\phi = \Delta x^2 D_{xx}^{-1} \rho.$$

However, this is only valid when the value of ϕ at the boundary $x \in \partial\Omega_h$ is zero, i.e.:

$$\phi_0 = 0,$$
$$\phi_n = 0.$$

Additionally, computing the inverse matrix is computationally expensive. Later, we will explore different methods to approximate the solution ϕ.

2D Case

Once again, let's consider Dirichlet boundary conditions where

$$\phi = 0 \quad \text{for } x, y \in \partial\Omega.$$

The discretization of the system consists of $(n-2, m-2)$ points $x_i, y_j \in \Omega_h$. In this discrete domain, the function ϕ can be represented as the vector:

$$\phi = \begin{bmatrix} \phi_{1,1} \\ \phi_{1,2} \\ \vdots \\ \phi_{1,m-1} \\ \phi_{2,1} \\ \vdots \\ \phi_{n-1,m-1} \end{bmatrix},$$

We can simplify this vector as:

$$\phi = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{n-1} \end{bmatrix}, \tag{A.10}$$

where each block vector is given by:

$$\Phi_i = \begin{bmatrix} \phi_{i,1} \\ \phi_{i,2} \\ \vdots \\ \phi_{i,m-1} \end{bmatrix}. \tag{A.11}$$

With this construction, the $\partial_{xx} + \partial_{yy}$ operator can be represented as the sum of two matrices of size $(n-2)(m-2) \times (n-2)(m-2)$. Fortunately, using the simplification of equations (A.10), (A.11), and (A.6), we can easily describe the structure of these matrices. First, for Φ in equation (A.10), the second derivative in x can be written as:

$$M_{xx} = \frac{1}{\Delta x^2} \begin{bmatrix} -2 & \mathbb{1} & \mathbb{0} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \mathbb{1} & -2 & \mathbb{1} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} & -2 & \mathbb{1} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbb{0} & \mathbb{0} & \mathbb{0} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{1} & -2 \end{bmatrix},$$

where each of the matrices $\mathbb{k} = k\mathbb{1}$. Now, for each vector Φ_i like (A.10), the second derivative must be computed in y like in Equation (A.9), this can be represented as:

$$M_{yy} = \frac{1}{\Delta y^2} \begin{bmatrix} D_{yy} & \mathbb{0} & \mathbb{0} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & D_{yy} & \mathbb{0} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & \mathbb{0} & D_{yy} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbb{0} & \mathbb{0} & \mathbb{0} & \mathbb{0} & \dots & \mathbb{0} & \mathbb{0} & D_{yy} \end{bmatrix}.$$

We can propose the discretized form of the 2D Poisson equation. To simplify notation, we define the Kronecker product between two matrices A and B, denoted as $A \otimes B$. This product is defined as a block matrix in the following way:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}, \tag{A.12}$$

where A is an $m \times n$ matrix and B is a $p \times q$ matrix. The individual blocks of the resulting matrix are obtained by multiplying each element of matrix A by matrix B. Using this notation, the derivative matrices can be defined as:

$$M_{xx} = \frac{1}{\Delta x^2} D_{xx} \otimes \mathbb{1}_{m-2},$$
$$M_{yy} = \frac{1}{\Delta y^2} \mathbb{1}_{n-2} \otimes D_{yy}.$$

The difference between D_{xx} and D_{yy} is the size of each matrix, which is $(n-2) \times (n-2)$ and $(m-2) \times (m-2)$, respectively. Thus, the Poisson equation (A.7) can be rewritten as the system of linear equations:

$$(M_{xx} + M_{yy})\phi = \rho.$$

3D Case

The study in the 2D case allows us to easily extend the system to 3D. We begin by considering the same Dirichlet conditions:

$$\phi = 0 \quad \text{for } x, y, z \in \partial\Omega.$$

The discretization of the system consists of $(n-2) \times (m-2) \times (l-2)$ points $x_i, y_j, z_k \in \Omega_h$. In this discrete domain, the function ϕ is written as a vector:

$$\phi = \begin{bmatrix} \phi_{1,1,1} \\ \phi_{1,1,2} \\ \vdots \\ \phi_{1,1,l-1} \\ \phi_{1,2,1} \\ \vdots \\ \phi_{1,m-1,l-1} \\ \phi_{2,1,1} \\ \vdots \\ \phi_{n-1,m-1,l-1} \end{bmatrix}.$$

We can simplify this vector as:

$$\phi = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{l-1} \end{bmatrix}, \quad \Phi_i = \begin{bmatrix} \Phi_{i,1} \\ \Phi_{i,2} \\ \vdots \\ \Phi_{i,m-1} \end{bmatrix}, \quad \Phi_{i,j} = \begin{bmatrix} \phi_{i,j,1} \\ \phi_{i,j,2} \\ \vdots \\ \phi_{i,j,l-1} \end{bmatrix}.$$

The matrices are written as:

$$\begin{aligned} M_{xx} &= \frac{1}{\Delta x^2} D_{xx} \otimes \mathbb{1}_{m-2} \otimes \mathbb{1}_{l-2}, \\ M_{yy} &= \frac{1}{\Delta y^2} \mathbb{1}_{n-2} \otimes D_{yy} \otimes \mathbb{1}_{l-2}, \\ M_{zz} &= \frac{1}{\Delta z^2} \mathbb{1}_{n-2} \otimes \mathbb{1}_{m-2} \otimes D_{zz}. \end{aligned}$$

Thus, the Poisson equation (A.7) can be rewritten as the system of linear equations:

$$(M_{xx} + M_{yy} + M_{zz})\phi = \rho. \tag{A.13}$$

Example in 2D

Let's consider the analytical case with a given solution and source:

$$\begin{aligned} \phi &= \sin(2\pi x)\sin(2\pi y)\mathrm{e}^{-\left((x-\frac{1}{2})^2+(y-\frac{1}{2})^2\right)}, \\ \rho &= 4\mathrm{e}^{-\left((x-\frac{1}{2})^2+(y-\frac{1}{2})^2\right)} \left\{ \sin(2\pi x)\sin(2\pi y) \left[\left(x-\frac{1}{2}\right)^2 + \left(y-\frac{1}{2}\right)^2 - 2\pi^2 \right] \right. \\ &\quad \left. + \cos(2\pi x)\sin(2\pi y)\,\pi(1-2x) + \sin(2\pi x)\cos(2\pi y)\,\pi(1-2y) \right\}. \end{aligned} \tag{A.14}$$

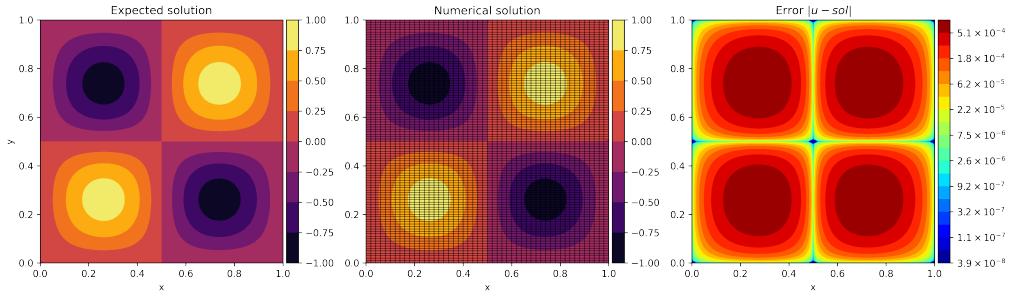

Figure A.2: Solution to the 2D Poisson equation with Dirichlet boundary conditions set to zero. The left panel shows the expected solution, the middle panel shows the solution obtained using the method described in section A.3.1 along with the discretization mesh of the domain, and the last panel shows the absolute difference between the analytical and numerical solutions.

In this example, we consider $\Omega = [0,1] \times [0,1]$ and a discretization with $n = 51$ and $m = 101$. This demonstrates that $n \neq m$ does not affect the solution in a square domain. In Figure A.2, the solution to the problem is shown by discretizing the source over the domain Ω_h and constructing the matrix as described in section A.3.1. This solution demonstrates that the error is of the order $\Delta x \Delta y \sim 2 \times 10^{-4}$. This shows that the approximation of the operator using the first term in the Taylor series is appropriate in this case. We can conclude that the error corresponds to the truncation error in the Taylor series, and if necessary, it can be reduced by making Δx and Δy smaller.

A.4 Boundary Conditions

Let D be the operator representing a linear partial differential equation of order n:

$$Du = f,$$

where the solution is found by solving the equation in the domain of dimension d, $\vec{x} \in \Omega \subset \mathbb{R}^d$, subject to certain boundary conditions. There are two main classes of boundary conditions: Dirichlet and Neumann. From these two, three subclasses emerge, namely Robin, Mixed, and Cauchy conditions, which we will briefly explain. First, we study Dirichlet boundary conditions, which are conditions imposed on the solution at the boundary $\partial\Omega$ as done in A.3.1, and are defined as:

$$u(\vec{x}) = g(\vec{x}) \quad \vec{x} \in \partial\Omega.$$

On the other hand, Neumann conditions are those imposed on the derivative of the solution and are defined as:

$$\frac{\partial u}{\partial \boldsymbol{n}} = f(\vec{x}) \quad \vec{x} \in \partial\Omega,$$

where $\boldsymbol{n}$ is the normal vector to the boundary. Mixed boundary conditions, also known as Robin conditions, are a linear combination of Dirichlet and Neumann conditions in different regions of the boundary $\Gamma \subset \partial\Omega$ and $\partial\Omega/\Gamma$:

$$\begin{aligned} u(\vec{x}) &= g(\vec{x}) \quad \vec{x} \in \Gamma, \\ \frac{\partial u}{\partial \boldsymbol{n}} &= f(\vec{x}) \quad \vec{x} \in \partial\Omega/\Gamma. \end{aligned} \tag{A.15}$$

Robin conditions are a linear combination of Dirichlet and Neumann conditions:

$$\frac{\partial u}{\partial \boldsymbol{n}} + \alpha u = \beta(\vec{x}) \quad \vec{x} \in \partial\Omega,$$

where α is the weighting factor on the Dirichlet conditions and $\beta(\vec{x})$ is the value of the boundary condition. Finally, Cauchy conditions are Dirichlet and Neumann conditions simultaneously on the boundary. This type of condition is useful in second-order equations.

A.4.1 Load Vector

When non-zero Dirichlet boundary conditions are imposed, the load vector arises as a consequence of the discretization process. We will briefly describe this behavior using the Poisson equation (A.7) in one dimension. We will use the discretization process of the Laplacian as described in equation (A.9). In contrast to the previous description of the vector ϕ in equation (A.8), we will include the boundaries:

$$\tilde{\phi} = \begin{bmatrix} \tilde{\phi}_0 \\ \tilde{\phi}_1 \\ \vdots \\ \tilde{\phi}_n \end{bmatrix},$$

where we denote the description with boundaries using a tilde. In this case, the matrix associated with the Laplacian needs to be modified in the first and last row, where we place a one on the diagonal and zeros elsewhere to consider the fixed boundary:

$$\tilde{D}_{xx} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}_{n\times n}.$$

However, the system of equations

$$\frac{1}{\Delta x^2}\tilde{D}_{xx}\tilde{\phi} = \tilde{\rho},$$

will be modified in two ways. First, we need to consider the boundary condition at the source, as the modification in the matrix $\tilde{D}_{xx}$ implies an identity relationship on both sides of the equation:

$$\tilde{\rho} = \begin{bmatrix} \tilde{\phi}_0 \\ \rho_1 \\ \vdots \\ \rho_{n-1} \\ \tilde{\phi}_n \end{bmatrix}.$$

Second, the terms in the matrix corresponding to the boundary (first and last rows) must be removed from the system of equations since both rows represent an identity:

$$\tilde{\phi}_i = \tilde{\phi}_i \quad \text{for } \boldsymbol{x}_i \in \partial\Omega_h.$$

To eliminate this pair of equations, we must perform the products involving the boundaries between the matrix $\tilde{D}_{xx}$ and the solution $\tilde{\phi}$. In the case of the Laplacian, these are the entries $(2,1)$ and $(n-1,n)$ of the matrix $\tilde{D}_{xx}$. This leads to a vector that we will call the load vector:

$$\tilde{\text{Load}} = \frac{1}{\Delta x^2} \begin{bmatrix} 0 \\ \phi_0 \\ \vdots \\ \phi_n \\ 0 \end{bmatrix}.$$

In this way, we will eliminate the equations corresponding to the boundary and recover the matrix (A.9) and the vector (A.8). Likewise, we will remove the tilde notation. Thus, the non-zero boundaries will be reflected on the right-hand side of the equation, resulting in

$$\frac{1}{\Delta x^2} D_{xx} = \rho - \text{Load}. \tag{A.16}$$

In summary, the procedure arithmetically coincides with moving the contribution from the boundaries to the right-hand side of the system of equations. This analysis must be done whenever nonzero Dirichlet boundary conditions are considered. The analysis can be easily simplified. First, let's define the rectangular and square matrices as follows:

$$D_{xx\ \text{Rec}} = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix}_{n-2,n},$$

$$D_{xx\ \text{Sq}} = \begin{bmatrix} -2 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -2 \end{bmatrix}_{n-2,n-2}.$$

From here, we can show that

$$\Delta x^2 \text{Load} = D_{xx\ \text{Rec}} \tilde{\phi} - D_{xx\ \text{Sq}} \phi.$$

The analysis can be complicated for higher dimensions. Fortunately, using the Kronecker product defined in (A.12), the analysis can be greatly simplified. Thus, in the 3D case, we can consider the system (A.13) to calculate the load vector easily:

$$\text{Load} = (M_{xx} + M_{yy} + M_{zz})_{\text{Rec}} \tilde{\phi} - (M_{xx} + M_{yy} + M_{zz})_{\text{Sq}} \phi.$$

This is computationally efficient since the matrix multiplication is usually more economical than traversing each element of the vector $\tilde{\phi}$ since the matrices are defined as sparse matrices, meaning that the zero entries of the matrix do not participate in the mathematical operations[1].

A.4.2 Periodic Boundary Conditions

Periodic boundary conditions are similar to Dirichlet boundary conditions, but the main difference is that the condition depends on the solution itself, so they cannot be treated in the same way. Periodic conditions, in the general 3D case, are defined as:

$$\begin{aligned} \phi_{0,j,k} &= \phi_{n,j,k}, \\ \phi_{i,0,k} &= \phi_{i,m,k}, \\ \phi_{i,j,0} &= \phi_{i,j,l}. \end{aligned} \tag{A.17}$$

[1]Note we can go further and consider a vector full of zeros for $\boldsymbol{x} \in \Omega$ and $g(x)$ for $\boldsymbol{x} \in \partial\Omega$. Then only the product with the rectangular matrix is needed

Here, it is explicitly assumed that the boundaries are located at $i, j, k = 0$ and $i, j, k = n, m, l$, respectively. It is possible to handle these conditions by modifying the matrix associated with the differential operator corresponding to the PDE in question. It is important to consider the fact that the boundaries (A.17) remove one variable at each point $\boldsymbol{x}_i \in \partial\Omega$. We will illustrate this analysis using the 1D Poisson equation (A.7).

$$D_{xx\ \text{Periodic}} = \begin{bmatrix} -2 & 1 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \dots & 0 & 1 & -2 \end{bmatrix}_{n-1,n-1} . \tag{A.18}$$

Notice that equation (A.18) considers a vector ϕ of the form:

$$\phi = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_{n-1} \end{bmatrix} .$$

In this way, by considering periodic boundaries where $\phi_0 = \phi_n$, it is directly reflected in the Laplacian operator. One consequence of considering a periodic domain is that the boundaries are somehow free, which leads to the existence of an equivalent family of solutions ϕ_{sol}, which generally differ by a constant value. To solve this issue, the mean value of the right-hand side Edwards et al. (2018) of the Poisson equation is subtracted:

$$\nabla^2\phi = \rho - < \rho > . \tag{A.19}$$

In other words, we are forcing the average of the boundaries to be zero, thereby eliminating the constant displacement. This approach to the problem is based on Gauss's theorem Dabo et al. (2008) and the fact that the surface integral of the gradient of the field around the perimeter of the simulation grid is identically zero when periodic boundary conditions are imposed. Equation (A.19), mentioned earlier, is a direct consequence of these fundamental principles. The analysis of periodic boundary conditions in 1D can be generalized to higher-dimensional problems using the Kronecker product (A.12).

A.5 Temporal Evolution

Partial differential equations (PDEs) are fundamental mathematical tools for describing a wide range of phenomena in science and engineering. In particular, time-dependent PDEs are used to model dynamic systems in which temporal evolution is of interest. Time-dependent PDEs are used to model heat diffusion, wave propagation, fluid flow, and many other dynamic phenomena. There are different types of time-dependent PDEs, such as the diffusion (or heat) equation, the wave equation, the transport equation, the Schrödinger equation, and the Navier-Stokes equation, to name a few. Each type of equation describes a specific phenomenon and has its own characteristics and solution methods. In the numerical solution of time-dependent partial differential equations, commonly used methods include explicit, implicit, and semi-implicit methods. Let's review these methods by considering a time-dependent partial differential equation of the form

$$\frac{\partial u}{\partial t} = F\left(u, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}, ...\right) .$$

Explicit Method

The explicit method is a widely used numerical technique for solving partial differential equations. This method is based on discretizing the temporal and spatial domains of the equation and using information from the previous time step to calculate the value at the current time step. To apply the explicit method, we discretize the temporal and spatial domains as done previously using finite difference methods:

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = F\left(u_i^n, \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}, \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}, ...\right).$$

Here, u_i^n represents the approximate value of u at time t_n and position x_i. We can solve this equation to obtain u_i^{n+1} in terms of the known values at time t_n. The explicit method is easy to implement and computationally efficient, but it can be unstable for certain types of equations, as its stability is limited by a stability condition known as the Courant–Friedrichs–Lewy (CFL) criterion. Assuming that the function F can be seen as a differential operator and can be expressed as a matrix M, the discretization of this equation can be expressed as a system of linear equations:

$$\mathbb{1}u^{n+1} = u^n + \Delta t M u^n.$$

Implicit Method

Unlike the explicit method, the implicit method uses information from the future time step to calculate the value at the current time step. In other words, the unknowns at the current time step are related to the unknowns at the future time step, leading to a system of nonlinear equations that must be solved. Using the same notation as before, we can write the discretized partial differential equation as:

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = F\left(u_i^{n+1}, \frac{u_{i+1}^{n+1} - u_{i-1}^{n+1}}{2\Delta x}, \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{\Delta x^2}, ...\right).$$

In this case, u_i^{n+1} is an unknown that we need to determine. The implicit method is more stable than the explicit method and can handle a wider range of problems. However, it is computationally more expensive due to the iterative resolution of the system of equations. Similarly, if we consider F as a matrix M, the discretization of this equation can be expressed as a system of linear equations:

$$\left(\mathbb{1} - \Delta t M\right) u^{n+1} = u^n.$$

Semi-Implicit Method

The semi-implicit scheme, also known as the Crank-Nicolson scheme or midpoint scheme, can be considered as a compromise between explicit and implicit schemes. In this scheme, the discrete term on the right-hand side of the PDE is the average of the evaluation at the previous time level, as in the explicit scheme, and the evaluation at the current time level, as in the implicit scheme.

$$\begin{aligned}\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{1}{2}\Bigg[&F\left(u_i^{n+1}, \frac{u_{i+1}^{n+1} - u_{i-1}^{n+1}}{2\Delta x}, \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{\Delta x^2}, ...\right) \\ &+ F\left(u_i^n, \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}, \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2}, ...\right)\Bigg].\end{aligned} \tag{A.20}$$

The semi-implicit method is particularly useful when the nonlinear part of the equation is dominant and can lead to numerical instabilities if treated explicitly. Again, if we consider F as a matrix M, the discretization of this equation can be expressed as a system of linear equations:

$$\left(\mathbb{1} - \frac{\Delta t}{2} M\right) u^{n+1} = \left(\mathbb{1} + \frac{\Delta t}{2} M\right) u^n.$$

In general, we can write the system of equations as:

$$(\mathbb{1} - \theta \Delta t M)\, u^{n+1} = (\mathbb{1} + (1-\theta)\Delta t M)\, u^n, \tag{A.21}$$

where θ is a parameter that controls the relative weighting between the explicit and implicit terms. For $\theta = 0$, the explicit method is recovered, for $\theta = 1$, the implicit method is obtained, and for $\theta = \frac{1}{2}$, the semi-implicit method is found. By adjusting the value of θ in the range of 0 to 1, we can smoothly transition from the explicit method to the implicit method. Therefore, Equation (A.21) provides us with the flexibility to choose a balance between explicitand implicit terms according to the needs of the particular problem. An additional advantage of this formulation is that it allows us to easily program all three methods using a single set of equations. This makes the implementation and comparison of the different methods simpler and more efficient.

A.6 Adaptive Moving Mesh Method

The adaptive moving mesh method (AMM) is a numerical technique used to solve partial differential equations (PDEs) in domains with complex geometries. This method involves the construction of a moving mesh using a transformation $\Omega_l \to \Omega_f$, where ξ, η, and $\kappa \in \Omega_l$ are the logical variables in a uniform domain, and x, y, $z \in \Omega_f$. Using this coordinate pair, any function $f : \Omega_f \to \mathbb{R}^n$ must be transformed to $g : \Omega_l \to \Omega_f \to \mathbb{R}^n$:

$$u = u(x, y, z) = u(x(\xi, \eta, \kappa), y(\xi, \eta, \kappa), z(\xi, \eta, \kappa)).$$

The Jacobian matrix

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_2}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_2} & \frac{\partial x_2}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_2} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix},$$

relates the partial derivatives of the physical coordinates (x, y, z) with respect to the logical coordinates (ξ, η, κ). Its inverse $(\boldsymbol{J}^{-1})$ relates the partial derivatives of the logical coordinates with respect to the physical coordinates. The determinant of the Jacobian is calculated as $J = \det(\boldsymbol{J}) = a_1 \cdot (a_2 \times a_3)$, where a_1, a_2, and a_3 are the column vectors of the matrix $\boldsymbol{J}$.

The gradient and divergence are expressed in terms of the logical and physical coordinates as follows:

$$\begin{aligned} \nabla u &= \sum_i a^i \frac{\partial u}{\partial \xi_i}, \\ \nabla \cdot \mathbf{v} &= \sum_i a^i \cdot \frac{\partial \mathbf{v}}{\partial \xi_i}, \end{aligned} \tag{A.22}$$

where they are applied to the function u and the vector field $\mathbf{v}$, respectively. These quantities are calculated as the sum of the products of the vectors a^i and the partial derivatives of u and $\mathbf{v}$ with respect to the logical coordinates ξ_i. The Laplacian is given by:

$$\nabla \cdot \nabla u = \sum_{i,j} a^i \cdot a^j \frac{\partial^2 u}{\partial \xi_i \partial \xi_j} + \sum_j \left(\sum_i a^i \cdot \frac{\partial a^j}{\partial \xi_i} \right) \frac{\partial u}{\partial \xi_j},$$

which is calculated as the sum of several terms involving second-order partial derivatives of u and the vectors a^i. In particular, the relation $a^i = \frac{1}{j} a_j \times a_k$ is used. For the equation u_t, a convention is introduced where u_t denotes the partial derivative of u with respect to time in the physical

domain, while $\dot{u}$ denotes the partial derivative of u with respect to time in the logical domain. The transformed equation for u_t is obtained using the expression:

$$u_t = \dot{u} + \sum_i \frac{\partial u}{\partial \xi_i} \frac{\partial \xi_i}{\partial t}.$$

The temporal derivatives of the coordinates in Ω_l and Ω_f are related through the following equations:

$$\dot{\boldsymbol{x}} = -\boldsymbol{J}\boldsymbol{\xi}_t,$$

$$\boldsymbol{\xi}_t = -\boldsymbol{J}^{-1}\dot{\boldsymbol{x}}.$$

In terms of these relationships, the equation for u_t transforms to:

$$u_t = \dot{u} - \sum_i \frac{\partial u}{\partial \xi_i} \sum_j \boldsymbol{J}^{-1}_{i,j} \dot{x}_j = \dot{u} - \nabla u \cdot \dot{\boldsymbol{x}}.$$

Finally, the equation for $\dot{\boldsymbol{x}}$ is presented, which depends on the components A_{ij} and B_i. This equation is solved together with the equation for $\dot{u}$, and the complete system of equations is:

$$\dot{u} = Mu + \nabla u \cdot \dot{\boldsymbol{x}} \tag{A.23}$$

$$\dot{\boldsymbol{x}} = \frac{1}{\tau\rho} \left[\sum_{i,j} A_{ij} \frac{\partial^2}{\partial \xi_i \partial \xi_j} + B_i \frac{\partial}{\partial \xi_i} \right] \boldsymbol{x}. \tag{A.24}$$

Appendix B

Coarse-graining method

This chapter is heavily based on the research conducted in Kopp et al. (2017), as it serves as the primary source for the comments and information presented here.

Cold dark matter (CDM) is an essential component of the widely accepted ΛCDM model, which provides an explanation for the structure and evolution of the universe. CDM plays a significant role in the formation of large-scale structures and the development of galaxies within halos. Scientists study CDM by modeling the dynamics of self-gravitating collisionless matter, which is believed to consist of undetected particle species. In the context of large-scale structure formation, CDM is referred to as "cold," indicating that its initial distribution in phase space can be described by smooth density and velocity fields. By gaining a better understanding of the behavior and properties of CDM, researchers can shed light on the nature of dark matter and its influence on the cosmic structure.

When studying CDM dynamics, the Poisson equation is commonly used to approximate the Einstein equations, justifying the utilization of Newtonian gravity and non-relativistic equations. Gravitational two-body collisions are suppressed due to the large number of particles, allowing the individual particle motion in phase space to be treated as collisionless. The time evolution of the phase space density is governed by the Vlasov equation, also known as the collisionless Boltzmann equation defined by the equation

$$\begin{aligned} \frac{\partial f}{\partial t} &= -\frac{\mathbf{u}}{a^2} \cdot \boldsymbol{\nabla}_x f + \boldsymbol{\nabla}_x \Phi \cdot \boldsymbol{\nabla}_u f \\ \Delta\Phi &= \frac{4\pi G \rho_0}{a} \left(\int d^3u f - 1 \right). \end{aligned} \tag{B.1}$$

By considering cold initial conditions and neglecting significant two-body collisions, the simplified framework of Newtonian gravity and the Vlasov equation (B.1) enables a manageable analysis of CDM dynamics, contributing to a deeper understanding of the formation and evolution of large-scale structures in the universe.

Solving the Vlasov (B.1) equation involves finding a solution to the coupled infinite hierarchy of equations for the cumulants of the phase space density. In the linear regime of large-scale structure formation, it is possible to truncate the Boltzmann hierarchy, leading to a simplified description known as the dust model. However, this approximation becomes inadequate as the density contrast enters the non-linear regime, resulting in the formation of caustics and the emergence of multiple streams and vorticity. At this stage, solving the complete Vlasov equation (B.1) becomes necessary.

N-body simulations are a popular method for approximating the solution to the Vlasov equation (B.1) and investigating the non-linear evolution of CDM. These simulations focus on the lowest cumulants of interest, such as density, velocity, and velocity dispersion, without requiring a detailed analysis of the fine-grained distribution of particles or the complete distribution function. However,

if the hierarchy cannot be truncated, calculating the cumulants through the high-dimensional phase space becomes necessary.

A general description of the method is given by defining the coarse-grained phase space density

$$\begin{aligned}\bar{f}(t,\boldsymbol{x},\boldsymbol{u}) &= \int \frac{d^3x'd^3u'}{(2\pi\sigma_x\sigma_u)^3}\exp\left\{-\frac{(\boldsymbol{x}-\boldsymbol{x}')^2}{2\sigma_x^2}-\frac{(\boldsymbol{u}-\boldsymbol{u}')^2}{2\sigma_u^2}\right\}f(t,\boldsymbol{x}',\boldsymbol{u}'),\\ &= \exp\left\{\frac{\sigma_x^2}{2}\Delta_x+\frac{\sigma_u^2}{2}\Delta_u\right\}\{f\},\end{aligned} \tag{B.2}$$

and it is obtained from f through a convolution with a Gaussian with widths σ_x and σ_u. The moments $\overline{M}^{(n)}$ of $\bar{f}$ are expressed in terms of the moments of f, that is, in terms of $M^{(n)}$

$$\begin{aligned}\bar{M}^{(0)} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta_x\right\}\{M^{(0)}\},\\ \bar{M}_i^{(1)} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta_x\right\}\{M_i^{(1)}\},\\ \bar{M}_{ij}^{(2)} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta_x\right\}\{M_{ij}^{(2)}\}+\sigma_u^2\bar{M}^{(0)}\delta_{ij},\end{aligned} \tag{B.3}$$

here, the moments are given in terms of the distribution function as $M_{i_1,\ldots,i_2}(\boldsymbol{x}) \equiv \int d^3u u_i\ldots u_{i_n}f(\boldsymbol{x},\boldsymbol{u})$, and can be derived in terms of the generating function

$$M_{i_1\ldots i_n}^{(n)}(\boldsymbol{x}) \equiv \int d^3u u_{i_1}\ldots u_{i_n}f(\boldsymbol{x},\boldsymbol{u}).$$

The first three moments represent the $n(t,\boldsymbol{x})$ is the density, $u_i(t,\boldsymbol{x})$ is the velocity of the fluid and Σ_{ij} is the mass-weighted velocity dispersion and are given by

$$\begin{aligned}M^{(0)} &= n,\\ M_i^{(1)} &= nu_i,\\ M_{ij}^{(2)} &= nu_iu_j+n\Sigma_{ij}.\end{aligned} \tag{B.4}$$

B.1 N-body case

The gravitational N-body problem is defined by the Hamiltonian system

$$\begin{aligned}E_N &= \sum_{i=1}^{N}\left(\frac{\boldsymbol{u}}{2a^2}-\frac{mG}{2a}\sum_{j=1,j\neq i}^{N}\frac{1}{|x_i-x_j|}\right),\\ \dot{\boldsymbol{x}} &= \frac{\partial E_N}{\partial \boldsymbol{u}_i},\quad \dot{\boldsymbol{u}}_i = -\frac{\partial E_N}{\partial \boldsymbol{x}_i},\end{aligned} \tag{B.5}$$

where a is the scale factor and E_N is the energy per particle, mass m, $\boldsymbol{x}_i$ is the comovil coordinate of the i-th particle asocieated with the congugated momentum $\boldsymbol{p}_i = m\boldsymbol{u}_i^2$. Defining the exact microscopic, or Klimontovich, phase space density

$$f_K(t,\boldsymbol{x},\boldsymbol{u}) = \frac{m}{\rho_0}\sum_{i=1}^{N}\delta_D[\boldsymbol{x}-\boldsymbol{x}_i(t)]\delta_D[\boldsymbol{x}-\boldsymbol{x}_i], \tag{B.6}$$

using the three-dimensional Dirac delta function δ_D , the N Hamiltonian equations (B.5) that determine the phase space trajectories $\{\boldsymbol{x}_i,\boldsymbol{u}_i\}$ can be neatly expressed in form of the Klimontovich

equation. The CDM phase space density f_c is the continuum limit of the Klimontovich phase space density that can be expressed by replacing (B.6) into (B.2) as

$$f_c(t,\boldsymbol{x},\boldsymbol{u}) = \int d^3q \delta_D[\boldsymbol{x}-\boldsymbol{X}_q]\delta_D[\boldsymbol{u}-\boldsymbol{U}_q]. \tag{B.7}$$

and the moments are expressed in term of the lagrangian coordinates $\boldsymbol{x}$, therefore $X(t,\boldsymbol{q})$ and $U(t,\boldsymbol{q})$ are the positions and velocities of the particles

$$\begin{aligned}
\bar{M}^{c(0)} &= \frac{1}{(2\pi)^3\sigma_x^3\sigma_u^3}\int d^3q \exp\left\{-\frac{|\boldsymbol{x}-\boldsymbol{X}(t,\boldsymbol{q})|}{2\sigma_x^2}\right\},\\
\bar{M}_i^{c(0)} &= \frac{1}{(2\pi)^3\sigma_x^3\sigma_u^3}\int d^3q \exp\left\{-\frac{|\boldsymbol{x}-\boldsymbol{X}(t,\boldsymbol{q})|}{2\sigma_x^2}\right\}U_i(t,\boldsymbol{q}),\\
\bar{M}_{ij}^{c(0)} &= \frac{1}{(2\pi)^3\sigma_x^3\sigma_u^3}\int d^3q \exp\left\{-\frac{|\boldsymbol{x}-\boldsymbol{X}(t,\boldsymbol{q})|}{2\sigma_x^2}\right\}U_i(t,\boldsymbol{q})U_j(t,\boldsymbol{q}) + \sigma_u^2\overline{M}^{c(0)}\delta_{ij},
\end{aligned} \tag{B.8}$$

particularly, for the case of N-body, the first moment of $\bar{f}_c$ can be expressed in a discrete way with a summation by means of the transformation $\int d^3q = \dfrac{m}{\rho_0}\sum_I$ as

$$\bar{n}(\boldsymbol{x}) = \frac{m}{\rho_0}\frac{1}{(2\pi)^{3/2}\sigma_x^3}\sum_I \exp\left\{-\frac{|\boldsymbol{x}-\boldsymbol{X}(\boldsymbol{q})|^2}{2\sigma_x^2}\right\}. \tag{B.9}$$

B.2 Schrödinger-method

It is convenient to define $\tilde{\hbar} \equiv \dfrac{\hbar}{m}$, then the non-relativistic Schrödinger-Poisson system in a ΛCDM universe with scale factor a reads

$$\begin{aligned}
i\tilde{\hbar}\frac{\partial\psi}{\partial t} &= -\frac{\tilde{\hbar}^2}{2a^2}\boldsymbol{\nabla}\psi + \Phi\psi\\
\boldsymbol{\nabla}^2\Phi &= \frac{4\pi G\rho_0}{a}(|\psi|^2-1),
\end{aligned} \tag{B.10}$$

It is convenient to obtain the Husimi representation ψ_H of the wave function ψ, which depends explicitly on the positions $\boldsymbol{x}$ and velocities $\boldsymbol{v}$ as follows Kopp et al. (2017)

$$\psi_H(t,\boldsymbol{x},\boldsymbol{u}) \equiv \int d^dx' K_H(\boldsymbol{x},\boldsymbol{x'},\boldsymbol{u})\psi(t,\boldsymbol{u}), \tag{B.11}$$

where

$$K_H(\boldsymbol{x},\boldsymbol{u}) \equiv \frac{\exp\left[-\dfrac{(\boldsymbol{x}-\boldsymbol{x}')^2}{4\sigma} - m\dfrac{i}{\tilde{\hbar}}\boldsymbol{u}\cdot\boldsymbol{x}'\right]}{(2\pi\tilde{\hbar})^{d/2}(2\pi\sigma_x^2)^{d/4}}. \tag{B.12}$$

analogously to the previous process, the coarse grained moments can be be computed using the transformation eq. (B.2), that is in terms of the moments for ψ, that is $M^{W(0)}$. The moments can expressed as follows

$$\begin{aligned}
M^{H(0)} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta\right\}\{M^{W(0)}\} \equiv n^H,\\
M_i^{H(1)} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta\right\}\{M_i^{W(1)}\} \equiv n^H u_i^H,\\
M_{ij}^{H} &= \exp\left\{\frac{\sigma_x^2}{2}\Delta\right\}\{M_{ij}^{W(2)}\} + \sigma_u^2 M^{H(0)}\delta_{ij} \equiv n^H u_i^H u_j^H + n^H\Sigma_{ij}^H.
\end{aligned} \tag{B.13}$$

Bibliography

Abdurro'uf, Accetta, K., and Aerts, e. a. (2022). The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data. *The Astrophysical Journals*, 259(2):35.

Adams, A. and Woolley, A. (1994). Hubble classification of galaxies using neural networks. *Vistas in Astronomy*, 38:273–280.

Andernach, H. and Zwicky, F. (2017). English and spanish translation of zwicky's (1933) the redshift of extragalactic nebulae.

Andredakis, Y. C., Peletier, R. F., and Balcells, M. (1995). The shape of the luminosity profiles of bulges of spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 275(3):874–888.

Angulo, R., Springel, V., White, S. D. M., Jenkins, A., Baugh, C. M., and Frenk, C. S. (2012). Scaling relations for galaxy clusters in the Millennium-XXL simulation. *Monthly Notices of the RAS*, 426(3):2046–2062.

Babcock, H. W. (1939). The rotation of the Andromeda Nebula. *Lick Observatory Bulletin*, 498:41–51.

BALL, N. M. and BRUNNER, R. J. (2010). DATA MINING AND MACHINE LEARNING IN ASTRONOMY. *International Journal of Modern Physics D*, 19(07):1049–1106.

Banerji, M., Lahav, O., Lintott, C. J., Abdalla, F. B., Schawinski, K., Bamford, S. P., Andreescu, D., Murray, P., Raddick, M. J., Slosar, A., Szalay, A., Thomas, D., and Vandenberg, J. (2010). Galaxy Zoo: reproducing galaxy morphologies via machine learning*. *Monthly Notices of the Royal Astronomical Society*, 406(1):342–353.

Baron, D. (2019). Machine learning in astronomy: a practical overview.

Barsanti, S., Owers, M. S., McDermid, R. M., Bekki, K., Bryant, J. J., Croom, S. M., Oh, S., Robotham, A. S. G., Scott, N., and van de Sande, J. (2021). The colors of bulges and disks in the core and outskirts of galaxy clusters. *The Astrophysical Journal*, 911(1):21.

Barton, E., Geller, M., Ramella, M., Marzke, R. O., and da Costa, L. N. (1996). Compact Group selection From Redshift Surveys. *The Astrophysical Journal*, 112:871.

Behroozi, P., Wechsler, R. H., Hearin, A. P., and Conroy, C. (2019). UniverseMachine: The correlation between galaxy growth and dark matter halo assembly from $z = 0 - 10$. *Monthly Notices of the Royal Astronomical Society*, 488(3):3143–3194.

Behroozi, P. S., Wechsler, R. H., and Wu, H.-Y. (2012a). THE ROCKSTAR PHASE-SPACE TEMPORAL HALO FINDER AND THE VELOCITY OFFSETS OF CLUSTER CORES. *The Astrophysical Journal*, 762(2):109.

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., Busha, M. T., Klypin, A. A., and Primack, J. R. (2012b). Gravitationally consistent halo catalogs and merger trees for precision cosmology. *The Astrophysical Journal*, 763(1):18.

Bennett, C. L., Larson, D., Weiland, J. L., Jarosik, N., Hinshaw, G., Odegard, N., Smith, K. M., Hill, R. S., Gold, B., Halpern, M., Komatsu, E., Nolta, M. R., Page, L., Spergel, D. N., Wollack, E., Dunkley, J., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., and Wright, E. L. (2013). Nine-year wilkinson microwave anisotropy probe (wmap) observations: Final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20.

Bertone, G., editor (2010). *Particle Dark Matter*. Cambridge University Press.

Binney, J. and Tremaine, S. (2008). *Galactic Dynamics: Second Edition*. Princeton University Press, rev - revised, 2 edition.

Blas, D., Lesgourgues, J., and Tram, T. (2011). The cosmic linear anisotropy solving system (CLASS). part II: Approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011(07):034–034.

Bose, S., Hellwing, W. A., Frenk, C. S., Jenkins, A., Lovell, M. R., Helly, J. C., and Li, B. (2015). The Copernicus Complexio: statistical properties of warm dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 455(1):318–333.

Bottema, R. and Pestaña, J. L. G. (2015). The distribution of dark and luminous matter inferred from extended rotation curves. *Monthly Notices of the Royal Astronomical Society*, 448(3):2566–2593.

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., and Lacey, C. G. (2006). Breaking the hierarchy of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 370(2):645–655.

Boylan-Kolchin, M., Bullock, J. S., and Kaplinghat, M. (2011). Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Monthly Notices of the Royal Astronomical Society: Letters*, 415(1):L40–L44.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

Bundy, K., Bershady, M. A., Law, D. R., Yan, R., Drory, N., MacDonald, N., Wake, D. A., Cherinka, B., Sánchez-Gallego, J. R., Weijmans, A.-M., Thomas, D., Tremonti, C., Masters, K., Coccato, L., Diamond-Stanic, A. M., Aragón-Salamanca, A., Avila-Reese, V., Badenes, C., Falcón-Barroso, J., Belfiore, F., Bizyaev, D., Blanc, G. A., Bland-Hawthorn, J., Blanton, M. R., Brownstein, J. R., Byler, N., Cappellari, M., Conroy, C., Dutton, A. A., Emsellem, E., Etherington, J., Frinchaboy, P. M., Fu, H., Gunn, J. E., Harding, P., Johnston, E. J., Kauffmann, G., Kinemuchi, K., Klaene, M. A., Knapen, J. H., Leauthaud, A., Li, C., Lin, L., Maiolino, R., Malanushenko, V., Malanushenko, E., Mao, S., Maraston, C., McDermid, R. M., Merrifield, M. R., Nichol, R. C., Oravetz, D., Pan, K., Parejko, J. K., Sanchez, S. F., Schlegel, D., Simmons, A., Steele, O., Steinmetz, M., Thanjavur, K., Thompson, B. A., Tinker, J. L., van den Bosch, R. C. E., Westfall, K. B., Wilkinson, D., Wright, S., Xiao, T., and Zhang, K. (2014). Overview of the sdss-iv manga survey: Mapping nearby galaxies at apache point observatory. *The Astrophysical Journal*, 798(1):7.

Callan, C., Dashen, R., and Gross, D. (1976). The structure of the gauge theory vacuum. *Physics Letters B*, 63(3):334–340.

Cautun, M., Benítez-Llambay, A., Deason, A. J., Frenk, C. S., Fattahi, A., Gómez, F. A., Grand, R. J. J., Oman, K. A., Navarro, J. F., and Simpson, C. M. (2020). The milky way total mass profile as inferred from Gaia DR2. *Monthly Notices of the Royal Astronomical Society*, 494(3):4291–4313.

Chollet, F. et al. (2015). Keras. `https://keras.io`.

Conselice, C. J. (2006). The fundamental properties of galaxies and a new galaxy classification system. *Monthly Notices of the Royal Astronomical Society*, 373(4):1389–1408.

Cottingham, W. N. and Greenwood, D. A. (2007). *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press.

Croton, D. J., Springel, V., White, S. D. M., De Lucia, G., Frenk, C. S., Gao, L., Jenkins, A., Kauffmann, G., Navarro, J. F., and Yoshida, N. (2006). Erratum: The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *Monthly Notices of the Royal Astronomical Society*, 367(2):864–864.

Côté, B., Martel, H., and Drissen, L. (2015). Cosmological simulations of the intergalactic medium evolution. ii. galaxy model and feedback. *The Astrophysical Journal*, 802:123.

Dabo, I., Kozinsky, B., Singh-Miller, N. E., and Marzari, N. (2008). Electrostatics in periodic boundary conditions and real-space corrections. *Phys. Rev. B*, 77:115139.

de Dios Rojas Olvera, J., Gómez-Vargas, I., and Vázquez, J. A. (2022). Observational cosmology with artificial neural networks. *Universe*, 8(2).

De Lucia, G. (2019). Lighting up dark matter haloes. *Galaxies*, 7(2).

De Lucia, G. and Blaizot, J. (2007). The hierarchical formation of the brightest cluster galaxies. *Monthly Notices of the Royal Astronomical Society (MNRAS)*, 375(1):2–14.

De Lucia, G., Kauffmann, G., Springel, V., White, S. D. M., Lanzoni, B., Stoehr, F., Tormen, G., and Yoshida, N. (2004). Substructures in cold dark matter haloes. *Monthly Notices of the Royal Astronomical Society (MNRAS)*, 348(1):333–344.

de Vaucouleurs, G. (1948). Recherches sur les Nebuleuses Extragalactiques. *Annales d'Astrophysique*, 11:247.

Di Matteo, T., Springel, V., and Hernquist, L. (2005). Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433(7026):604–607.

Díaz-Giménez, Eugenia and Zandivarez, Ariel (2015). Where are compact groups in the local universe? *A&A*, 578:A61.

Dimauro, P., Huertas-Company, M., Daddi, E., Pérez-González, P. G., Bernardi, M., Barro, G., Buitrago, F., Caro, F., Cattaneo, A., Dominguez-Sánchez, H., Faber, S. M., Häußler, B., Kocevski, D. D., Koekemoer, A. M., Koo, D. C., Lee, C. T., Mei, S., Margalef-Bentabol, B., Primack, J., Rodriguez-Puebla, A., Salvato, M., Shankar, F., and Tuccillo, D. (2018). A catalog of polychromatic bulge-disc decompositions of ∼17.600 galaxies in CANDELS. *Monthly Notices of the Royal Astronomical Society*, 478(4):5410–5426.

Díaz-Giménez, E. and Mamon, G. A. (2010). Compact groups from the Millennium Simulations – I. Their nature and the completeness of the Hickson sample. *Monthly Notices of the Royal Astronomical Society*, 409(3):1227–1243.

Edwards, F., Kendall, E., Hotchkiss, S., and Easther, R. (2018). PyUltraLight: a pseudo-spectral solver for ultralight dark matter dynamics. *Journal of Cosmology and Astroparticle Physics*, 2018(10):027–027.

Farhang, A., Khosroshahi, H. G., Mamon, G. A., Dariush, A. A., and Raouf, M. (2017). Evolution of compact and fossil groups of galaxies from semi-analytical models of galaxy formation. *The Astrophysical Journal*, 840(1):58.

Fields, B. D., Olive, K. A., Yeh, T.-H., and Young, C. (2020). Big-bang nucleosynthesis after planck. *Journal of Cosmology and Astroparticle Physics*, 2020(03):010–010.

Freeman, K. C. (1970). On the disks of spiral and s0 galaxies. *Astrophysical Journal, vol. 160, p. 811*, 160:811.

Geller, M. J. and Huchra, J. P. (1989). Mapping the universe. *Science*, 246(4932):897–903.

Genzel, R., Schreiber, N. M. F., Übler, H., Lang, P., Naab, T., Bender, R., Tacconi, L. J., Wisnioski, E., Wuyts, S., Alexander, T., Beifiori, A., Belli, S., Brammer, G., Burkert, A., Carollo, C. M., Chan, J., Davies, R., Fossati, M., Galametz, A., Genel, S., Gerhard, O., Lutz, D., Mendel, J. T., Momcheva, I., Nelson, E. J., Renzini, A., Saglia, R., Sternberg, A., Tacchella, S., Tadaki, K., and Wilman, D. (2017). Strongly baryon-dominated disk galaxies at the peak of galaxy formation ten billion years ago. *Nature*, 543(7645):397–401.

Ghahramany, N., Gharaati, S., and Ghanaatian, M. (2012). New approach to nuclear binding energy in integrated nuclear model. *Journal of Theoretical and Applied Physics*, 6(1):3.

Gingold, R. A. and Monaghan, J. J. (1977). Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, 181(3):375–389.

Guo, Q., White, S., Boylan-Kolchin, M., De Lucia, G., Kauffmann, G., Lemson, G., Li, C., Springel, V., and Weinmann, S. (2011). From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. *Monthly Notices of the Royal Astronomical Society (MNRAS)*, 413(1):101–131.

Guo, Q., White, S., Li, C., and Boylan-Kolchin, M. (2010). How do galaxies populate dark matter haloes? *Monthly Notices of the Royal Astronomical Society (MNRAS)*, 404(3):1111–1120.

Heisler, J., Tremaine, S., and Bahcall, J. N. (1985). Estimating the masses of galaxy groups: alternatives to the virial theorem. *The Astrophysical Journal*, 298:8–17.

Hickson, P. (1982). Systematic properties of compact groups of galaxies. *The Astrophysical Journal*, 255:382–391.

Hickson, P. (1997). Compact groups of galaxies. *Annual Review of Astronomy and Astrophysics*, 35(1):357–388.

Hickson, P., Mendes de Oliveira, C., Huchra, J. P., and Palumbo, G. G. (1992). Dynamical Properties of Compact Groups of Galaxies. *The Astrophysical Journal*, 399:353.

Hlozek, R., Grin, D., Marsh, D. J. E., and Ferreira, P. G. (2015). A search for ultralight axions using precision cosmological data. *Phys. Rev. D*, 91:103512.

Hogg, D. W., Blanton, M. R., Brinchmann, J., Eisenstein, D. J., Schlegel, D. J., Gunn, J. E., McKay, T. A., Rix, H.-W., Bahcall, N. A., Brinkmann, J., and Meiksin, A. (2004). The dependence on environment of the color-magnitude relation of galaxies. *The Astrophysical Journal*, 601(1):L29–L32.

Holmberg, E. (1958). A photographic photometry of extragalactic nebulae. *Meddelanden fran Lunds Astronomiska Observatorium Serie II*, 136:1.

Hopkins, P. F., Bundy, K., Croton, D., Hernquist, L., Keres, D., Khochfar, S., Stewart, K., Wetzel, A., and Younger, J. D. (2010). Mergers and Bulge Formation in ΛCDM: Which Mergers Matter? *The Astrophysical Journal*, 715(1):202–229.

Hu, W. and Dodelson, S. (2002). Cosmic microwave background anisotropies. *Annual Review of Astronomy and Astrophysics*, 40(1):171–216.

Hubble, E. (1929). A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the national academy of sciences*, 15(3):168–173.

Hui, L., Ostriker, J. P., Tremaine, S., and Witten, E. (2017). Ultralight scalars as cosmological dark matter. *Phys. Rev. D*, 95:043541.

III, J. R. G., Jurić, M., Schlegel, D., Hoyle, F., Vogeley, M., Tegmark, M., Bahcall, N., and Brinkmann, J. (2005). A map of the universe. *The Astrophysical Journal*, 624(2):463.

Ilić, S., Kopp, M., Skordis, C., and Thomas, D. B. (2021). Dark matter properties through cosmic history. *Phys. Rev. D*, 104:043520.

Johnston, E. J., Aragón-Salamanca, A., Merrifield, M. R., and Bedregal, A. G. (2012). Spectroscopic bulge–disc decomposition: a new method to study the evolution of lenticular galaxies. *Monthly Notices of the Royal Astronomical Society*, 422(3):2590–2599.

Kamdar, H. M., Turk, M. J., and Brunner, R. J. (2016). Machine learning and cosmological simulations – II. Hydrodynamical simulations. *Monthly Notices of the Royal Astronomical Society*, 457(2):1162–1179.

Karniadakis, G. E. and Kirby II, R. M. (2003). *Parallel Scientific Computing in C and MPI: A Seamless Approach to Parallel Algorithms and their Implementation*. Cambridge University Press.

Kent, S. M. (1985). CCD surface photometry of field galaxies. II. Bulge/disk decompositions. *The Astrophysical Journals*, 59:115–159.

King, I. R. (1966). The structure of star clusters. III. Some simple dynamical models. *Astronomical Journal*, 71:64.

Klypin, A. A., Trujillo-Gomez, S., and Primack, J. (2011). Dark matter halos in the standard cosmological model: Results from the bolshoi simulation. *The Astrophysical Journal*, 740(2):102.

Kopp, M., Skordis, C., and Thomas, D. B. (2016). Extensive investigation of the generalized dark matter model. *Phys. Rev. D*, 94:043512.

Kopp, M., Vattis, K., and Skordis, C. (2017). Solving the vlasov equation in two spatial dimensions with the schrödinger method. *Phys. Rev. D*, 96:123532.

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., Klypin, A. A., Gottlöber, S., Allgood, B., and Primack, J. R. (2004). The Dark Side of the Halo Occupation Distribution. *The Astrophysical Journal*, 609(1):35–49.

Kuiper, G. P. (1938). The empirical mass-luminosity relation. *The Astrophysical Journal*, 88:472.

Lelli, F., McGaugh, S. S., and Schombert, J. M. (2016). Sparc: Mass models for 175 disk galaxies with spitzer photometry and accurate rotation curves. *The Astronomical Journal*, 152(6):157.

Liddle, A. R., Pahud, C., and Ureña López, L. A. (2008). Triple unification of inflation, dark matter, and dark energy using a single field. *Phys. Rev. D*, 77:121301.

Liddle, A. R. and Ureña López, L. A. (2006). Inflation, dark matter, and dark energy in the string landscape. *Phys. Rev. Lett.*, 97:161301.

Liebert, J. and Probst, R. G. (1987). Very low mass stars. *Annual review of astronomy and astrophysics*, 25(1):473–519.

Lovell, M. R., Gonzalez-Perez, V., Bose, S., Boyarsky, A., Cole, S., Frenk, C. S., and Ruchayskiy, O. (2017). Addressing the too big to fail problem with baryon physics and sterile neutrino dark matter. *Monthly Notices of the Royal Astronomical Society*, 468(3):2836–2849.

Lucie-Smith, L., Peiris, H. V., Pontzen, A., and Lochner, M. (2018). Machine learning cosmological structure formation. *Monthly Notices of the Royal Astronomical Society*, 479(3):3405–3414.

López-Sánchez, J. N., Munive-Villa, E., Avilez-López, A., and Martínez-Bravo, O. M. (2022). Compact groups in GDM cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 515(3):3199–3211.

Madelung, E. (1927). Quantentheorie in hydrodynamischer form. *Zeitschrift for Physik*, 40(3-4):322–326.

Mahajan, S., Drinkwater, M. J., Driver, S., Hopkins, A. M., Graham, A. W., Brough, S., Brown, M. J. I., Holwerda, B. W., Owers, M. S., and Pimbblet, K. A. (2017). Galaxy and mass assembly (GAMA): blue spheroids within 87 mpc. *Monthly Notices of the Royal Astronomical Society*, 475(1):788–799.

Majumdar, D. (2014). *Dark Matter*. CRC Press.

McConnachie, A. W., Ellison, S. L., and Patton, D. R. (2008). Compact groups in theory and practice – I. The spatial properties of compact groups. *Monthly Notices of the Royal Astronomical Society*, 387(3):1281–1290.

Mendel, J. T., Simard, L., Palmer, M., Ellison, S. L., and Patton, D. R. (2014). VizieR Online Data Catalog: SDSS bulge, disk and total stellar mass estimates (Mendel+, 2014). *VizieR Online Data Catalog*, page J/ApJS/210/3.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Mocz, P., Fialkov, A., Vogelsberger, M., Boylan-Kolchin, M., Chavanis, P.-H., Amin, M. A., Bose, S., Dome, T., Hernquist, L., Lancaster, L., Notis, M., Painter, C., Robles, V. H., and Zavala, J. (2023). Cosmological structure formation and soliton phase transition in fuzzy dark matter with axion self-interactions. *Monthly Notices of the Royal Astronomical Society*, 521(2):2608–2615.

Munive-Villa, E., López-Sánchez, J. N., Avilez-López, A. A., and Guzmán, F. S. (2022). Solving the schrödinger-poisson system using the coordinate adaptive moving mesh method. *Phys. Rev. D*, 105:083521.

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. *The Astrophysical Journal*, 462:563.

Niemeyer, J. C. (2020). Small-scale structure of fuzzy and axion-like dark matter. *Progress in Particle and Nuclear Physics*, 113:103787.

Oemler Jr, A. (1976). The structure of elliptical and cd galaxies. *Astrophysical Journal, Vol. 209, p. 693-709*, 209:693–709.

Oh, S.-H., de Blok, W. J. G., Brinks, E., Walter, F., and Kennicutt, R. C. (2011). Dark and luminous matter in things dwarf galaxies. *The Astronomical Journal*, 141(6):193.

Peccei, R. D. and Quinn, H. R. (1977). CP conservation in the presence of pseudoparticles. *Phys. Rev. Lett.*, 38:1440–1443.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pérez-González, P. G., Rieke, G. H., Villar, V., Barro, G., Blaylock, M., Egami, E., Gallego, J., de Paz, A. G., Pascual, S., Zamorano, J., and Donley, J. L. (2008). The stellar mass assembly of galaxies fromz= 0 toz= 4: Analysis of a sample selected in the rest-frame near-infrared withSpitzer. *The Astrophysical Journal*, 675(1):234–261.

Pérez-Lorenzana, A., Montesinos, M., and Matos, T. (2008). Unification of cosmological scalar fields. *Phys. Rev. D*, 77:063507.

Planck Collaboration (2020). Planck 2018 results - vi. cosmological parameters. *A&A*, 641:A6.

Plummer, H. C. (1911). On the Problem of Distribution in Globular Star Clusters: (Plate 8.). *Monthly Notices of the Royal Astronomical Society*, 71(5):460–470.

Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E., and Primack, J. (2012). Halo concentrations in the standard Λ cold dark matter cosmology. *Monthly Notices of the Royal Astronomical Society*, 423(4):3018–3030.

Rafatullah, D. M., editor (2020). *Theory and Applications of Physical Science Vol. 3*. Book Publisher International (a part of SCIENCEDOMAIN International).

Reiprich, T. H. and Boehringer, H. (2002). The mass function of an x-ray flux-limited sample of galaxy clusters. *The Astrophysical Journal*, 567(2):716.

Rodríguez, J.-V., Rodríguez-Rodríguez, I., and Woo, W. L. (2022). On the application of machine learning in astronomy and astrophysics: A text-mining-based scientometric analysis. *WIREs Data Mining and Knowledge Discovery*, 12(5):e1476.

Rubin, V. C., Ford, W. K., J., and Thonnard, N. (1980). Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *The Astrophysical Journal*, 238:471–487.

Ruszkowski, M. and Springel, V. (2009). The role of dry mergers for the formation and evolution of brightest cluster galaxies. *The Astrophysical Journal*, 696(2):1094.

Sandage, A. (1961). *The Hubble Atlas of Galaxies*. Encyclopedia Britannica.

Sarkar, S. (1996). Big bang nucleosynthesis and physics beyond the standard model. *Reports on Progress in Physics*, 59(12):1493.

Schive, H.-Y. and Chiueh, T. (2018). Halo abundance and assembly history with extreme-axion wave dark matter at $z \geq 4$. *Mon. Not. Roy. Astron. Soc.*, 473(1):L36–L40.

Schive, H.-Y., Chiueh, T., and Broadhurst, T. (2014a). Cosmic Structure as the Quantum Interference of a Coherent Dark Wave. *Nature Phys.*, 10:496–499.

Schive, H.-Y., Liao, M.-H., Woo, T.-P., Wong, S.-K., Chiueh, T., Broadhurst, T., and Hwang, W. Y. P. (2014b). Understanding the Core-Halo Relation of Quantum Wave Dark Matter from 3D Simulations. *Phys. Rev. Lett.*, 113(26):261302.

Schneider, P. (2015). *Extragalactic Astronomy and Cosmology*. Springer Berlin Heidelberg.

Schwarzschild, M. (1979). A numerical model for a triaxial stellar system in dynamical equilibrium. *The Astrophysical Journal*, 232:236–247.

Smoot, G. F. (1999). COBE observations and results. *AIP Conference Proceedings*, 476(1):1–10.

Springel, V. (2005). The cosmological simulation code gadget-2. *Monthly Notices of the Royal Astronomical Society*, 364(4):1105–1134.

Springel, V., Frenk, C. S., and White, S. D. M. (2006). The large-scale structure of the universe. *Nature*, 440(7088):1137–1144.

Springel, V., Pakmor, R., Zier, O., and Reinecke, M. (2021). Simulating cosmic structure formation with the gadget-4 code. *Monthly Notices of the Royal Astronomical Society*, 506(2):2871–2949.

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435(7042):629–636.

Swinbank, M. (2017). Distant galaxies lack dark matter. *Nature*, 543(7645):318–319.

Tabor, M., Merrifield, M., Aragón-Salamanca, A., Fraser-McKelvie, A., Peterken, T., Smethurst, R., Drory, N., and Lane, R. R. (2019). SDSS-IV MaNGA: full spectroscopic bulge-disc decomposition of MaNGA early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 485(2):1546–1558.

Tegmark, M., Strauss, M. A., Blanton, M. R., Abazajian, K., Dodelson, S., Sandvik, H., Wang, X., Weinberg, D. H., Zehavi, I., Bahcall, N. A., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M. S., Berlind, A., Budavari, T., Connolly, A., Eisenstein, D. J., Finkbeiner, D., Frieman, J. A., Gunn, J. E., Hui, L., Jain, B., Johnston, D., Kent, S., Lin, H., Nakajima, R., Nichol, R. C., Ostriker, J. P., Pope, A., Scranton, R., Seljak, U. c. v., Sheth, R. K., Stebbins, A., Szalay, A. S., Szapudi, I., Xu, Y., Annis, J., Brinkmann, J., Burles, S., Castander, F. J., Csabai, I., Loveday, J., Doi, M., Fukugita, M., Gillespie, B., Hennessy, G., Hogg, D. W., Ivezić, i. c. v., Knapp, G. R., Lamb, D. Q., Lee, B. C., Lupton, R. H., McKay, T. A., Kunszt, P., Munn, J. A., O'Connell, L., Peoples, J., Pier, J. R., Richmond, M., Rockosi, C., Schneider, D. P., Stoughton, C., Tucker, D. L., Vanden Berk, D. E., Yanny, B., and York, D. G. (2004). Cosmological parameters from sdss and wmap. *Phys. Rev. D*, 69:103501.

Teyssier, R., Pires, S., Prunet, S., Aubert, D., Pichon, C., Amara, A., Benabed, K., Colombi, S., Refregier, A., and Starck, J.-L. (2009). Full-sky weak-lensing simulation with 70 billion particles. *A&A*, 497(2):335–341.

The Astropy Collaboration, Robitaille, Thomas P., Tollerud, Erik J., Greenfield, Perry, Droettboom, Michael, Bray, Erik, Aldcroft, Tom, Davis, Matt, Ginsburg, Adam, Price-Whelan, Adrian M., Kerzendorf, Wolfgang E., Conley, Alexander, Crighton, Neil, Barbary, Kyle, Muna, Demitri, Ferguson, Henry, Grollier, Frédéric, Parikh, Madhura M., Nair, Prasanth H., Günther, Hans M., Deil, Christoph, Woillez, Julien, Conseil, Simon, Kramer, Roban, Turner, James E. H., Singer, Leo, Fox, Ryan, Weaver, Benjamin A., Zabalza, Victor, Edwards, Zachary I., Azalee Bostroem, K., Burke, D. J., Casey, Andrew R., Crawford, Steven M., Dencheva, Nadia, Ely, Justin, Jenness, Tim, Labrie, Kathleen, Lim, Pey Lian, Pierfederici, Francesco, Pontzen, Andrew, Ptak, Andy,

Refsdal, Brian, Servillat, Mathieu, and Streicher, Ole (2013). Astropy: A community python package for astronomy. *A&A*, 558:A33.

Thomas, D. B., Kopp, M. I., and Markovivc, K. (2019). Using large scale structure data and a halo model to constrain Generalised Dark Matter. *Monthly Notices of the Royal Astronomical Society*, 830(2):155.

Tomozeiu, M., Mayer, L., and Quinn, T. (2016). Tidal stirring of satellites with shallow density profiles prevents them from being too big to fail. *The Astrophysical Journal Letters*, 827(1):L15.

Tully, R. B., Pierce, M. J., Huang, J.-S., Saunders, W., Verheijen, M. A. W., and Witchalls, P. L. (1998). Global extinction in spiral galaxies. *The Astronomical Journal*, 115(6):2264–2272.

Tzanavaris, P., Hornschemeier, A. E., Gallagher, S. C., Johnson, K. E., Gronwall, C., Immler, S., Reines, A. E., Hoversten, E., and Charlton, J. C. (2010). Ultraviolet+Infrared Star Formation Rates: Hickson Compact Groups with Swift and Spitzer. *The Astrophysical Journal*, 716(1):556–573.

Ureña-López, L. A. (2019). Brief review on scalar field dark matter models. *Frontiers in Astronomy and Space Sciences*, 6.

Ureña-López, L. A. and Gonzalez-Morales, A. X. (2016). Towards accurate cosmological predictions for rapidly oscillating scalar fields as dark matter. *Journal of Cosmology and Astroparticle Physics*, 2016(07):048.

Vale, A. and Ostriker, J. P. (2004). Linking halo mass to galaxy luminosity. *Monthly Notices of the Royal Astronomical Society (MNRAS)*, 353(1):189–200.

VanderPlas, J., Connolly, A. J., Ivezic, Z., and Gray, A. (2012). Introduction to astroml: Machine learning for astrophysics. In *2012 Conference on Intelligent Data Understanding*, pages 47–54.

Vavilova, I. B., Dobrycheva, D. V., Vasylenko, M. Yu., Elyiv, A. A., Melnyk, O. V., and Khramtsov, V. (2021). Machine learning technique for morphological classification of galaxies from the sdss - i. photometry-based approach. *A&A*, 648:A122.

Veltmaat, J., Schwabe, B., and Niemeyer, J. C. (2020). Baryon-driven growth of solitonic cores in fuzzy dark matter halos. *Phys. Rev. D*, 101:083518.

Venhola, Aku, Peletier, Reynier, Laurikainen, Eija, Salo, Heikki, Iodice, Enrichetta, Mieske, Steffen, Hilker, Michael, Wittmann, Carolin, Paolillo, Maurizio, Cantiello, Michele, Janz, Joachim, Spavone, Marilena, D´Abrusco, Raffaele, van de Ven, Glenn, Napolitano, Nicola, Verdoes Kleijn, Gijs, Capaccioli, Massimo, Grado, Aniello, Valentijn, Edwin, Falcón-Barroso, Jesús, and Limatola, Luca (2019). The fornax deep survey (fds) with vst - vi. optical properties of the dwarf galaxies in the fornax cluster. *A&A*, 625:A143.

Virbhadra, K. S. and Ellis, G. F. R. (2000). Schwarzschild black hole lensing. *Phys. Rev. D*, 62:084003.

Weijmans, A.-M. (2015). Manga: Mapping nearby galaxies at apache point observatory.

Wiens, C. D., Wenger, T. V., Tzanavaris, P., Johnson, K. E., Gallagher, S. C., and Xiao, L. (2019). The occurrence of compact groups of galaxies through cosmic time. *The Astrophysical Journal*, 873(2):124.

Workman, R. L. et al. (2022). Review of Particle Physics. *PTEP*, 2022:083C01.

Yoon, Y., Park, C., Chung, H., and Zhang, K. (2021). Rotation curves of galaxies and their dependence on morphology and stellar mass. *The Astrophysical Journal*, 922(2):249.

Yurin, D. and Springel, V. (2014). An iterative method for the construction of N-body galaxy models in collisionless equilibrium. *Monthly Notices of the Royal Astronomical Society*, 444(1):62–79.

Zhang, J., Liu, H., and Chu, M.-C. (2019). Cosmological simulation for fuzzy dark matter model. *Frontiers in Astronomy and Space Sciences*, 5.

Zhang, J., Tsai, Y.-L. S., Kuo, J.-L., Cheung, K., and Chu, M.-C. (2018). Ultralight axion dark matter and its impact on dark halo structure in n-body simulations. *The Astrophysical Journal*, 853(1):51.

Zhao, H. (1996). Analytical models for galactic nuclei. *Monthly Notices of the RAS*, 278(2):488–496.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. *The Astrophysical Journal*, 86:217.