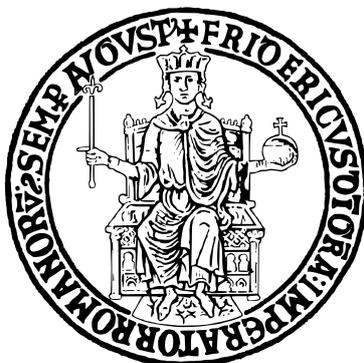


UNIVERSITÀ DEGLI STUDI DI NAPOLI
“FEDERICO II”



Scuola Politecnica e delle Scienze di Base

Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Magistrale in Fisica

Testing Non-local Gravity at galactic scales

Relatori:

Prof. Salvatore Capozziello

Dr. Filippo Bouchè

Candidato:

Ciro De Simone

Matr. N94000767

Anno Accademico 2022/2023

Contents

Introduction	1
1 Non-local Theories of Gravity	5
1.1 Locality and non-locality in Physics	5
1.1.1 Kinematical non-locality	7
1.1.2 Dynamical non-locality	9
1.2 Non-locality in Quantum Field Theory	12
1.2.1 Locality and causality	14
1.2.2 Locality and unitarity	15
1.3 Non-local Gravity Theories	16
1.3.1 Infinite Derivative Theories of Gravity	18
1.3.2 Integral Kernel Theories of Gravity	20
1.3.3 Non-local Teleparallel Gravity	22
2 The Deser-Woodard Model	26
2.1 The model and the localization of the action	26
2.1.1 Cosmological applications	28
2.2 The Noether Symmetry Approach	30
2.2.1 The distortion function	33
2.3 The weak field limit	35
3 Testing Non-local Gravity at galactic scales	40
3.1 Ultra-Diffuse Galaxies	40
3.1.1 Dark Matter density profile	42
3.1.2 Stellar density profile	45
3.2 Non-local corrections to the Newtonian potential	46
3.3 Kinematics of Ultra-Diffuse Galaxies	49
4 The statistical analysis: Methods and Results	52
4.1 Bayesian inference	52

4.2	Markov Chain Monte Carlo	55
4.3	Bayesian analysis of Non-local Gravity at galactic scales	59
4.4	Results	61
	Discussion and Conclusions	67
	Bibliography	70

Introduction

The Λ CDM model currently provides the most successful description of the gravitational and cosmological phenomena in the Universe. This paradigm is based on Einstein's theory of General Relativity (GR) and is characterized by two exotic fluids: the Dark Energy, which is associated to the expansion of the Universe and takes the form of a cosmological constant (Λ), and a non baryonic type of matter known as Cold Dark Matter (CDM). The Dark Energy should account for the $\sim 68\%$ of the energy density of the Universe, while the CDM and the baryonic matter should constitute the residual $\sim 26\%$ and $\sim 5\%$, respectively [1]. Other components, such as neutrinos and photons, only contribute subdominantly. A large variety of astrophysical and cosmological observations are in agreement with the Λ CDM model, ranging from Type Ia Supernovae to the Cosmic Microwave Background (CMB) radiation and the Large Scale Structure of the Universe.

However, it is worth pointing out that the Λ CDM model is affected by some shortcomings. The main examples are given by the huge discrepancy between the theoretical and observed value of the cosmological constant, as well as the fact that no particle candidate for Dark Matter has been observed at fundamental scales so far. These issues have motivated a growing interest toward extensions of the general relativistic description of the gravitational interaction, stemming from the possibility that GR is not valid at the scale and regimes where Dark Matter and Dark Energy need to be introduced, hence the theory itself has to be changed. It is clear that many ways to extend General Relativity are possible: some models involve the introduction of new scalar fields (scalar-tensor theories), or the replacement of the Ricci scalar R in the Einstein-Hilbert action with an higher-order function, like in $f(R)$ theories. Another interesting possibility is represented by Teleparallel Gravity, where the gravitational interaction can be described in terms of torsion rather than curvature, such that the metric tensor does not constitute the fundamental object of the theory.

Among these theories, non-local extensions of GR are drawing increasing attention. Non-locality has indeed played a significant role in the history of Physics: the first theories

to be developed were inherently non-local, and the success of Quantum Mechanics in the last century has given rise to a renewed interest toward non-local phenomena. On the one hand, Quantum Mechanics exhibits a so called kinematical non-locality, i.e. non-locality of the quantum mechanical states of the theory; on the other hand, Quantum Field Theory (QFT) exhibits a dynamical non-locality, i.e. non-locality of the action and equations of motion. Therefore, within the gravitational framework, the introduction of non-local terms in the Hilbert-Einstein Lagrangian could be an effective approach rather to bridge GR and QFT, or to fix the Λ CDM shortcomings. Indeed, a crucial feature of these Non-local Gravity theories is that non-local corrections can be significant both in the ultraviolet and infrared regime.

In this master thesis, a specific Non-local Gravity model, known as the Deser-Woodard model, will be investigated. It was first introduced almost two decades ago in order to explain the late-time acceleration of the Universe. Then, it has been tested in several bounded gravitational systems, ranging from the orbits of the star S2 around the Galactic centre [2] to the gravitational lensing effects in massive galaxy clusters [3]. Therefore, this work aims at performing a completely new test of the Deser-Woodard model by leveraging the data from Ultra-Diffuse galaxies (UDGs). These galaxies represent the ideal scenario for testing extended theories of gravity as well as the Dark Matter hypothesis, since the UDGs exhibit a wide range of behaviours with respect to their Dark Matter content. In fact, recent observations suggest that some of these galaxies are made up almost entirely by Dark Matter, while others seem to completely lack any Dark Matter contribution. The non-local Deser-Woodard model will be then tested in two different configuration: one case in which the non-local effects only play the role of Λ , and another case in which the non-local corrections to the Newtonian potential act as both the dark fluids.

In this thesis, the attention is focused on the Ultra-Diffuse galaxy DrangonFly44 (DF44), in the Coma cluster. Observations carried out in 2016 [4] indicated that this galaxy might be dominated by Dark Matter for 99% of its total mass. Therefore, DF44 emerges as the perfect system to test Non-local Gravity as a theory that might be able to explain the CDM along with the Dark Energy.

Throughout our analysis, the Dark Matter component has been modeled by a generalized Navarro-Frenk-White (gNFW) profile, while the stellar component by a Sérsic profile. The statistical analysis is carried out in the Bayesian framework, using Monte Carlo Markov Chain techniques, implemented through the software Wolfram Mathematica. This approach aims at estimating the best ranges for the parameters of the non-local model, and also at comparing the Deser-Woodard model with General Relativity.

The thesis is organized as follows:

- Chapter 1 is devoted to a description of the concept of locality and non-locality in Physics, drawing the distinction between kinematical and dynamical non-locality and analyzing carefully two examples from Quantum Mechanics: the phenomenon of entanglement as an evidence for kinematical non-locality and an analogue of the so called Aharonov-Bohm effect for the dynamical non-locality. Moreover, I will present two dynamical non-local modifications of GR: the Infinite Derivative theories of Gravity (IDG) and Integral Kernel theories of Gravity (IKG). It is also interesting to investigate how a non-local modification of GR is possible in the framework of the Teleparallel formulation of General Relativity. Finally, I will investigate the relation between non-locality and the notions of causality and unitarity.
- In Chapter 2 the Deser-Woodard model will be described in detail, together with the so called localization of the action. This procedure allows to recast this model in terms of a biscalar-tensor theory. In this chapter I will also outline the so called Noether Symmetry Approach, a theoretical selection criterion for gravity models, and some cosmological applications of the Deser-Woodard model related to the accelerated expansion of the Universe. In the last part of the chapter I will present the solution of the field equations of the Deser-Woodard model which shows how the non-local degrees of freedom are encoded in two characteristic lengths that specify the magnitude of the non-local corrections.
- In Chapter 3 I will outline the general features of Ultra-Diffuse Galaxies and the modelization of the matter profile of the galaxy in terms of a stellar and Dark Matter component. It is also important to notice that in non-local gravity Gauss' theorem does not hold, hence the gravitational potential derived for a point particle source has to be extended in order to take into account the mass distribution of systems like Ultra-Diffuse galaxies. Finally, I will derive the theoretical prediction for the dispersion velocity in the Deser-Woodard model, which constitute the physical observable that will be compared with the experimental data.
- In the first section of Chapter 4 I will delineate the statistical technique employed for testing the Deser-Woodard model. In particular, the attention is focused on the general features of Bayesian inference and Markov Chain Monte Carlo methods. I will also describe the Metropolis-Hastings algorithm and the effectiveness of this approach for the purposes of our analysis. Finally, the last section is devoted to a discussion of the results obtained in the statistical analysis.

The conclusions will be finally drawn, focusing on the main results and the future prospects in the framework of Non-local Gravity.

Chapter 1

Non-local Theories of Gravity

1.1 Locality and non-locality in Physics

The fundamental idea behind the principle of locality is that an object can be influenced only by its immediate neighborhood.

This principle is incorporated in all modern theories of Physics, from Quantum Field Theory (QFT) to General Relativity (GR). However, it is worth noticing that in the history of Physics, the first theories conceived were inherently non-local. This is the case for Newton's law of universal gravitation or Coulomb's law of electric force. Those theories were formulated in terms of a so called "action at a distance", the idea that an object can be affected instantaneously by another object spatially separated from it. Notice also that the notion of action at a distance was already difficult to accept for Newton, who famously said [5]:

"It is inconceivable that inanimate Matter should, without the Mediation of something else, which is not material, operate upon, and affect other matter without mutual Contact. . . That Gravity should be innate, inherent and essential to Matter, so that one body may act upon another at a distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking can ever fall into it. Gravity must be caused by an Agent acting constantly according to certain laws; but whether this Agent be material or immaterial, I have left to the Consideration of my readers."

Technically, Newton's law of gravitation as well as Coulomb's force are described by the Poisson's equation, a second order linear partial differential equation that encodes the notion of action at a distance.

This problem persisted for a long period of time, until Faraday and Maxwell introduced the notion of a field. The crucial idea for the gravitational field, for example, is that a particle of mass M produces a gravitational field at every point of space and time that can act on other massive particles, and vice versa.

The same results holds for electromagnetism since Maxwell's equations account for all electromagnetic phenomena in terms of electric and magnetic fields. Thus, the notion of field encapsulates the principle of locality and identifies a physical entity which carries energy and momentum.

Despite the fact that all modern theories incorporate the principle of locality, there are many interesting phenomena in the electrodynamics of media that are intrinsically non-local. For instance, in the treatment of electrodynamics of bulk matter it is sometimes necessary to add non-local constitutive equations in order to specify the properties of the materials under investigation.

More generally, the dynamics of the electric and magnetic fields can be non-local, i.e. expressed by integral or integro-differential equations which cannot be written in terms of differential equations of integer degree. Non-locality in time has to be considered when the fields change more rapidly than the characteristic period of oscillation of the medium. Similarly, spatial non-locality has to be considered if the fields vary rapidly in space.

The interest toward Non-local theories of Gravity is also motivated by modern research in the field of Quantum Gravity. In fact, there are many indications suggesting that one of the fundamental assumptions in modern physics such as locality, causality, unitarity, etc. has to be abandoned in order to find a consistent theory valid at Planck scale.

An example of how Quantum Gravity may affect locality is related to the localizability problem [6]. Let us consider the process of measurement of the position of a particle, where a probe (e.g. a photon) has to interact with the particle. As long as the probe is a classical object there are no limitations on how accurate the measurement can be made. Nonetheless, if we wish to measure the position of the particle with accuracy δx at quantum scale it is necessary to take into account Heisenberg's uncertainty principle, which implies that the energy necessary to perform this measurement has to be of order $(\delta x)^{-1}$. This thought experiment is known as the "Heisenberg microscope" and shows that any increase in the accuracy of the measurement of position requires a corresponding increase in the energy of the probe. If the energy of the probe is extremely high then it is impossible to neglect gravitational effects. For instance, it turns out that if the probe is a quantum particle of mass m then the uncertainty on the position measurement δx is proportional to $m^{-1/2}$. This implies that if we want to reduce the uncertainty on the position measurement we need to increase the value of the mass of the probe but we have

a fundamental limitation from GR. In fact, extremely massive objects may lead to the formation of a black hole hence there is an upper bound on the mass of the probe and consequently a lower bound on the measurement uncertainty δx , which is typically of the order of the Planck length $L_P = \sqrt{\hbar G_N/c^3} \approx 10^{-35}$ m.

For the purposes of this thesis, it is crucial to distinguish between two types of non-locality: kinematical non-locality and dynamical non-locality. The former refers to the states of a theory while the latter refers to the interactions, hence to the form of the action and equations of motion. Non-locality emerges naturally in Quantum Mechanics and can be considered one of the main issues in the attempt to merge the formalism of QFT and GR. In particular, all the fundamental interactions in QFT exhibit dynamical non-locality as soon as their one-loop effective actions are considered.

Quantum Mechanics exhibits both kinds of non-locality: kinematical non-locality is apparent from phenomena such as entanglement; while dynamical non-locality can be identified in the so called Aharonov-Bohm effect. It is also worth pointing out that kinematical non-locality does not imply dynamical non-locality and vice versa.

1.1.1 Kinematical non-locality

Entanglement is one of the most surprising consequences of Quantum Mechanics and is strictly connected to the notion of kinematic non-locality. In order to understand the phenomenon of entanglement it is crucial to introduce the notion of “Bell experiment” [7], which can be outlined as follows.

Let Alice and Bob be two observers separated by an arbitrary distance. Now suppose that two systems, which may have previously interacted (for example they were produced by the same source), are now spatially separated and each one of them can be measured by Alice or Bob.

The two observers can perform many different measurements on each system, e.g. a measurement of position or spin. An example of this type of system may be a pair of spin 1/2 particles produced by the same source, and whose spin in different directions is measured. Let x denote the measurement done by Alice and y the measurement done by Bob.

Once the measurements are performed on the two systems, they yield outcomes a and b respectively. Notice that the actual values assigned to the measurement choices x, y and outcomes a, b are purely conventional; they are only labels to distinguish the different possibilities. Furthermore, repeating the same procedure many times, even when the same choices of measurements x and y are made, results in outcomes a and b that can be significantly different. Thus, the possible outcomes of the experiments follow a probability distribution $p(ab|xy)$, which can depend on the particular experiment being performed.

A fair estimate of such probabilities can be obtained by repeating the experiment a sufficient number of times and collecting the observed data.

In general, it is found that

$$p(ab|xy) \neq p(a|x)p(b|y), \tag{1.1}$$

implying that the outcomes of the experiments are not statistically independent from each other.

The crucial observation in this scenario is that, in principle, the two systems may be separated by a large distance, they may even be space-like separated.

Notice that the results obtained so far are not in contradiction with the notion of locality, since there is still the possibility that this correlation is the result of some dependence relation between the two systems that was established when they previously interacted. Hence there is no evidence of a direct influence of one system on the other.

This observation is particularly relevant for the understanding of what is meant in general by locality in physics.

The assumption of locality implies that there exists a set of past factors, described by some variable λ , that have an influence on both outcomes, and which accounts for the observed dependence between a and b .

In this context it is clear that once the effect of those factors is taken into account, then the two outcomes should be completely decoupled, that is

$$p(ab|xy, \lambda) = p(a|x, \lambda) p(b|y, \lambda), \tag{1.2}$$

and the probabilities for a and b factorize.

This property implies that the probability for a (or b) only depends on the set of past variables λ and on the local measurement x (or y), but not on the distant measurement and outcome. In addition, notice that the set of past variables λ does not have to be constant for all runs of the experiment, because it may depend on physical quantities that are not fully under control. The variability of λ can be taken into account by introducing a probability distribution $q(\lambda)$, which results in the following expression for the probability of the outcome a and b :

$$p(ab|xy) = \int_{\Lambda} d\lambda q(\lambda) p(a|x, \lambda) p(b|y, \lambda) \tag{1.3}$$

assuming that Λ is the range of possible values of the set of variables λ and that the probability distribution $q(\lambda)$ does not depend on the measurements x and y .

The existence of a decomposition like (1.3) is usually associated to the so called local hidden-variable or local realistic models. It is also the notion of locality used in the context of Bell experiments, called local causality rather than locality by Bell.

Notice also that the only assumption in equation (1.3) is locality, there are no restrictions on the physical laws governing this causal relation. In relativistic terms, it is required that events in one region of spacetime should not influence events in space-like separated regions.

The results obtained so far have to be compared with the predictions of quantum mechanics. In a paper published in 1964 [8], J.Bell considers exactly the scenario outlined above: measurements performed independently on the two separated particles of an entangled pair. He proved that the assumption that the outcomes depend upon a set of variable λ implies a mathematical constraint (Bell inequalities) on how the outcomes of the two measurements are correlated.

Bell also showed that quantum mechanics predicts correlations that violate this inequality, proving the kinematical non-locality of quantum mechanics. The violation of Bell inequalities were first confirmed experimentally by Freedman and Clauser in 1972, then by Aspect in 1982 and in many other experiments. This result has deeply influenced our understanding of Physics, from the foundational perspective to quantum information science.

1.1.2 Dynamical non-locality

A good starting point to study dynamical non-locality is the so called Aharonov-Bohm (AB) effect, a feature of Quantum Mechanics that has no analog in classical physics.

This effect is characterized by a measurable phase shift in the wave function of a charged particle that rotates around a region inaccessible to the particle, where an electromagnetic field is located. It means that the motion of the particle happens in a region where the electric and magnetic fields are zero, but the scalar and vector potential are nonzero.

The non-locality discovered in the Aharonov-Bohm effect is completely different from the violation of the Bell inequalities [9]. In fact, the AB effect shows a dynamical type of non-locality which directly involves the equations of motion.

It is of paramount importance to understand the previous statement with regard to the Schrödinger equation, the basic equation of motion of Quantum Mechanics. This equation is completely local since it contains only quantities that are local in space and time such as the potential $V(r)$ and the wave function $\psi(r, t)$, and a finite number of its derivatives. The crucial point is that the dynamical locality of Quantum Mechanics cannot be inferred from the Schrödinger equation, since it refers to the wave function $\psi(x, t)$ which is neither directly observable nor gauge invariant.

In order to check for the validity of the locality principle, the evolution of observable

physical variables, such as position and momentum and their equations of motion, should instead be analyzed.

Let us consider the well known double-slit experiment. A source emits electrons toward a screen with two narrow slits, behind which there is a second screen covered with electron detectors. Electrons leave the source one by one and they end up at different points on the second screen. Three experimental scenarios are quite interesting in this context, in fact, we can consider the case where alternatively only one the two slits is open and the case where both slits are open. We are mainly interested in the number of electrons that reach a position x on the screen, denoted $N_1(x)$, $N_2(x)$ and $N_{12}(x)$ respectively, over a fixed amount of time.

The predictions of quantum mechanics strongly differ from their classical analogues, since $N_{12}(x) \neq N_1(x) + N_2(x)$. Moreover, the number of particles that reach the screen when both the slits are open, $N_{12}(x)$, exhibits an interference pattern.

In the framework of Quantum Mechanics this experiment is explained via a dual process. First of all, the electron is described by a wave function which propagates from the source to the screen. The wave function passes through the two slits and exhibits an interference pattern at the screen.

Notice that the propagation of the wave function from the source to the screen is viewed as an ordinary wave interference phenomenon, not very different from interference of sound waves for example. Here the AB effect plays a crucial role, since the phase that governs the interference is local classically, but non-local quantum mechanically. More precisely, the wave function $\psi(r)$ of the electron can be written as a superposition of two terms, each describing a wave propagating through one of the slits:

$$\psi(r) = \frac{1}{\sqrt{2}}(e^{i\alpha_1} \psi_1(r) + e^{i\alpha_2} \psi_2(r)). \quad (1.4)$$

The interference pattern that emerges at the screen depends on the relative phase given by $\alpha = \alpha_2 - \alpha_1$.

The difference with classical wave interference is that classically, it is possible to measure the relative phase as well as the individual phases of the waves (for instance in the case of water waves), while quantum mechanically α_1 and α_2 cannot be observed. Hence the individual phases of the two wavepackets cannot be measured locally, they simply have no observational meaning.

Let us consider for simplicity a one-dimensional situation: two non overlapping wavepackets $\psi_1(x)$ and $\psi_2(x)$. They are similar to the wavepackets describing the electron immediately after passing through the two slits in the double-slit experiment.

Let the wave function at $t = 0$ be

$$\psi(x) = \frac{1}{\sqrt{2}}(\psi_1(x) + e^{i\alpha} \psi_2(x)) \quad (1.5)$$

where α is the relative phase. It is clear that the interference pattern due to the overlap of the two wavepackets will depend on this relative phase.

Let us analyze which observables encode the information about this phase α at $t = 0$. The position operator x has an average given by:

$$\bar{x} = \int \psi^*(x) x \psi(x) dx, \quad (1.6)$$

which does not depend on the relative phase α since the wave functions are non-overlapping. Analogously any power of the operator x has average that does not depend on α .

In the case of the momentum operator p the average \bar{p} is given by:

$$\bar{p} = \int \psi^*(x) (-i\hbar) \frac{\partial}{\partial x} \psi(x) dx, \quad (1.7)$$

similarly to the position operator, only the cross terms depend on α , hence \bar{p} does not depend on α either. Furthermore, it can be shown that also $\overline{x^n p^k}$ do not depend on α , for any n and k .

It seems that there is a paradox since every observable related to the particle can be constructed from the operators x and p , but the moments of x and p do not depend on α , thus no observable can depend on α . The contradiction is that at a later time, when the two wavepackets overlap, the interference pattern does depend on the relative phase, so this information must be encoded somewhere also at $t = 0$.

It turns out that the information is contained in the “shift operator”, $e^{ipL/\hbar}$. This operator is responsible for the shift of the wave function of a particle by a finite distance L . Taking the two wavepackets to be identical up to a shift, $\psi_2(x) = \psi_1(x + L)$, the average of the shift operator is $e^{-i\alpha}/2$, that clearly depends on α .

We can conclude that the information about the interference is contained in the average of the shift operator, not in the averages of any power of position or momentum.

Notice that, technically, the shift operator is non-Hermitian and therefore is not an observable. The actual observables are its Hermitian components, $\sin(pL/\hbar)$ e $\cos(pL/\hbar)$. In what follows the shift operator will be considered for mathematical simplicity since the same results hold for its Hermitian component.

Let us investigate the properties of the shift operator. First, it is clearly non-local: it shifts the wave function of a particle by a finite distance. Second, it is a purely quantum operator: it contains \hbar in such a manner that it is not defined in the classical limit of $\hbar \rightarrow 0$.

Let us also investigate how the shift operator evolves over time. The position and momentum operators x and p evolve according to the Heisenberg equations of motion

$$\frac{dx}{dt} = \frac{p}{m} \quad , \quad \frac{dp}{dt} = -\frac{dV}{dx} \quad , \quad (1.8)$$

which is very similar to the classical evolution of position and momentum. However, in Quantum Mechanics x and p are not represented by real numbers, rather by self-adjoint operators.

The consequences of this observation become clear when looking at the evolution of the shift operator. Classically, the time evolution of $e^{ipL/\hbar}$ is

$$\frac{de^{ipL/\hbar}}{dt} = \frac{de^{ipL/\hbar}}{dp} \frac{dp}{dt} = -i\frac{L}{\hbar} e^{ipL/\hbar} \frac{dV}{dx}. \quad (1.9)$$

In Quantum Mechanics, however, this procedure cannot be applied since we are dealing with operators and not ordinary functions, hence the chain rule for differentiation cannot be used. Instead, we have to solve directly the corresponding Heisenberg equation

$$i\hbar \frac{de^{ipL/\hbar}}{dt} = [e^{ipL/\hbar}, H], \quad \text{where } H = \frac{p^2}{2m} - V(x), \quad (1.10)$$

which leads to

$$\frac{de^{ipL/\hbar}}{dt} = -i\frac{L}{\hbar} e^{ipL/\hbar} \frac{V(x+L) - V(x)}{L}. \quad (1.11)$$

The shift operator has a local equation of motion in classical mechanics, but the equation becomes non-local in quantum mechanics, because it depends on the potential difference between two remote places, $V(x+L)$ and $V(x)$.

This thought experiment has the same structure as the scalar AB effect, which predicts a phase shift for a wave function due to a scalar potential, but in the context of the double-slit experiment.

This argument gives an interesting perspective also on the relation between locality and causality that will be analyzed in more detail later. The idea is that non-local interactions do not violate necessarily causality because of the Heisenberg's uncertainty principle. Vice versa, we can think of the uncertainty relations as a mean for allowing non-local interactions in nature.

1.2 Non-locality in Quantum Field Theory

In QFT a theory is generally considered local if the Lagrangian density \mathcal{L} is a function of the fields and their derivatives up to finite order at a single point in spacetime, i.e.

$$\mathcal{L} = \mathcal{L}(\phi, \partial\phi, \dots, \partial^n\phi). \quad (1.12)$$

This notion is strongly related to causality and it is imposed in order to avoid any action at a distance [10].

The definition (1.12) shows that a straightforward generalization to non-local actions in QFT is possible simply by adding to \mathcal{L} non-local operators such as

$$\mathcal{L} = \mathcal{L}(\phi, \partial\phi, \dots, \partial^n\phi, \square^{-1}\phi, e^{\square/M_s^2}\phi, \dots), \quad (1.13)$$

where M_s is a characteristic non-local energy scale.

Non-locality in this case is due to integral operator such as \square^{-1} or non-polynomial operators such as e^{\square/M_s^2} that involve an infinite number of derivatives. Furthermore, these types of non-locality naturally emerge in QFT when one considers one-loop effective actions: functionals of the vacuum expectation value of the field, obtained by integrating out the quantum fluctuations.

An example of non-locality in the quantum effective action is given by quantum electrodynamics [11], where integrating out the quantum fluctuations associated to the electron and considering only the terms that involve the photon results in the following effective action

$$\Gamma_{QED}[A_\mu] = -\frac{1}{4} \int d^4x \left[F_{\mu\nu} \frac{1}{e^2(\square)} F^{\mu\nu} + \mathcal{O}(F^4) \right]. \quad (1.14)$$

In the limit $|\square/m_e^2| \gg 1$, the term $1/e^2(\square)$ can be written as

$$\frac{1}{e^2(\square)} \simeq \frac{1}{e^2(\mu)} - \frac{1}{12\pi^2} \log\left(-\frac{\square}{\mu^2}\right), \quad (1.15)$$

where μ is the renormalization scale and $e(\mu)$ is the renormalized electric charge at the scale μ . Notice that the logarithm of the d'Alembert operator is a non-local operator which can be represented by the integral

$$\log\left(-\frac{\square}{\mu^2}\right) \equiv \int_0^\infty dm^2 \left[\frac{1}{m^2 + \mu^2} - \frac{1}{m^2 - \square} \right], \quad (1.16)$$

in this case the non-locality of the effective action reflects the running of the coupling constant.

As far as gravity is concerned, there is a 2D gravity action that is interesting for the development of this thesis. The 2D action for the gravitational interaction, including a cosmological constant Λ , is given by

$$S = \frac{1}{2\kappa^2} \int d^2x \sqrt{-g} (2R - \Lambda) + S^{(m)}, \quad (1.17)$$

where $S^{(m)}$ is the action associated to matter and $\kappa^2 = 8\pi G_N/c^4$.

It is crucial to notice that in a two dimensional spacetime the Einstein-Hilbert action is

a topological invariant and gives no contribution to the dynamics. However, the gravitational dynamics is not trivial at quantum level and the quantum effective action associated to the quantum fluctuations of the matter fields is the so called Polyakov quantum effective action in two dimensions:

$$\Gamma = -\frac{N-25}{96\pi} \int d^2x \sqrt{-g} R \square^{-1} R - \Lambda \int d^2x \sqrt{-g} \quad (1.18)$$

where N is an integer related to the matter action $S^{(m)}$.

The effective action (1.18) is intriguing because the non-locality is due to the inverse of the \square operator, and a similar action will be introduced later in the context of the Deser-Woodard model. Finally, the emergence of non-locality in fundamental interaction is an additional motivation for the interest toward non-local descriptions of Gravity.

1.2.1 Locality and causality

A question that naturally arises once non-local operator are involved in a theory is whether or not the principle of causality is valid [12].

Causality is in fact a fundamental requirement in all modern theories of Physics and it is crucial to understand how this notion is affected by the introduction of integral operators such as \square^{-1} . For sake of simplicity, it is enough to focus on a non-local action for a scalar field $\phi(x)$ [11] such as

$$S = \int d^4x \phi(x) \square^{-1} \phi(x) \quad (1.19)$$

where the term $\square^{-1}\phi$ can be expressed in terms of the Green's function associated to \square^{-1} introduced before. Let us now consider the variation of the above action with respect to the scalar field ϕ in order to determine the field equations of the theory and make some important observations on the notion of causality:

$$\begin{aligned} \frac{\delta S}{\delta \phi(y)} &= \frac{\delta}{\delta \phi(y)} \int d^4x \phi(x) (\square^{-1} \phi)(x) = \frac{\delta}{\delta \phi(y)} \int d^4x d^4z \phi(x) G(x, z) \phi(z) = \\ &= \int d^4z G(y, z) \phi(z) + \int d^4x \phi(x) G(x, y) = \int d^4x [G(y, x) + G(x, y)] \phi(x). \end{aligned} \quad (1.20)$$

Therefore, the application of the variational principle results in equations of motion where the Green's function is symmetrized.

It appears that these equations violate the principle of causality, since both the retarded and advanced Green's functions contribute. It means that the behavior of the field at a point x depends on the value of field at points y that belong to the future light cone of x , $y^0 > x^0$.

One possible way of removing those ambiguities consists of assuming that the non-local action is not fundamental. In this approach the non-local action is taken as a quantum effective action whose functional argument is the mean quantum field and any advanced Green's function has to be replaced by hand with a retarded Green's function.

It is worth pointing out that this breakdown of causality does not immediately lead to inconsistency in the calculation of scattering amplitudes or $\langle in|out \rangle$ matrix elements, because of the Feynman diagrammatic techniques involved. Furthermore, the only physically observable quantities are bilinear combinations of the scattering amplitudes, and they can be always written as expectation values of quantum operators in the initial state $|in\rangle$.

1.2.2 Locality and unitarity

It is also interesting to investigate how locality is related to another crucial notion in QFT such as unitarity [10]. In fact, unitarity is associated to the conservation of probability and sets many relevant constraints. For instance, unitarity imposes that the states in the Hilbert space transform through unitary representations of the Poincaré group. It also affects the form of the interaction terms in the Lagrangian, since the S -matrix must be unitary, i.e. $S^\dagger S = S S^\dagger = \mathbb{1}$, where $\mathbb{1}$ is the identity operator.

Let us consider the following Lagrangian density for two scalar fields ϕ and π of mass M and m respectively

$$\mathcal{L} = -\frac{1}{2}\phi(\square + M^2)\phi - \frac{1}{2}\pi(\square + m^2)\pi + \frac{\lambda}{2}\phi\pi^2. \quad (1.21)$$

In order to study the relation between unitarity and non-locality, let us integrate out a field, that is, substitute one of the fields in (1.21) with its classical expectation value given by the solution of the equations of motion. A similar procedure is applied in the framework of the weak interaction where the 4-Fermi theory is recovered by integrating out the W and Z bosons.

For the field ϕ the equations of motion are given by

$$-(\square + M^2)\phi + \frac{\lambda}{2}\pi^2 = 0, \quad (1.22)$$

hence the action (1.21) has the following non-local form

$$\mathcal{L}_{non-local} = -\frac{1}{2}\pi(\square + m^2)\pi + \frac{\lambda^2}{8}\pi^2\frac{1}{\square + M^2}\pi^2, \quad (1.23)$$

where non-locality is due to the inverse of the d'Alembert operator \square .

It is also worthwhile to mention that it is possible to obtain a local expression of (1.23) in the limit $\square \ll M^2$ as

$$\mathcal{L}_{local} = -\frac{1}{2}\pi(\square + m^2)\pi + \frac{\lambda^2}{8}\left(\frac{\pi^4}{M^2} - \pi^2\frac{\square}{M^4}\pi^2 + \dots\right). \quad (1.24)$$

The implication of this result for unitarity is that at energies close to M^2 there is an apparent pole instead of the scalar field ϕ which has been integrated out. In particular, unitarity would be violated if the particle ϕ had actually been removed from the Hilbert space when we integrated it out, because it results in a violation of the so called optical theorem.

The upshot is that, in the regime of energies greater than M^2 , a different effective description is necessary, where the particle associated to the pole is given a corresponding field. Furthermore, this example shows the emergence of a relation between non-locality and the loss of unitarity.

1.3 Non-local Gravity Theories

In this section two models of Non-local Gravity will be analyzed in detail: the Infinite Derivative theories of Gravity (IDGs) and Integral Kernel theories of Gravity (IKGs) [13]. In addition, an approach to Non-local gravity based on the Teleparallel formulation of GR will be presented.

The IDGs involve analytic transcendental functions of the d'Alembert operator

$$\square = g_{\mu\nu} \nabla^\mu \nabla^\nu = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} g^{\mu\nu} \partial_\nu), \quad (1.25)$$

where ∇^μ is the covariant derivative and g is the determinant of the metric; while the IKGs are mainly constructed from the inverse d'Alembert operator \square^{-1} .

An interesting feature of the IDGs is that they provide renormalizable and unitary quantum gravity theories. They can also be used to address the black hole and Big Bang singularity problem in GR. On the other hand, IKGs deal with infrared (IR) quantum corrections coming from the formulation of QFT in curved spacetime. They were firstly considered in order to study the late-time cosmic expansion of the Universe without invoking Dark Energy contributions. Despite these features, there is no local or non-local theory capable of solving all the large-scale structure issues and fitting all the available observations obtained so far.

In order to understand the features of non-local theories of gravity, it is worthwhile to recall briefly the properties of General Relativity [14].

This theory offers a metric description of the gravitational interaction where spacetime is a dynamical object described by a pseudo-Riemannian manifold. The conceptual foundation of General Relativity is the Equivalence Principle. In its Weak formulation this principle states that the inertial mass of an object, which measures the resistance of the object to acceleration, coincides numerically with the gravitational mass, i.e. the mass

measured using the force and acceleration due only to gravity.

On the other hand, the Strong Equivalence Principle states that, at any spacetime point, it is possible to choose a local inertial reference frame such that, in a sufficiently small region surrounding that point, all physical laws take the same form they would take in absence of gravity, namely the form prescribed by Special Relativity. Another crucial notion in GR is given by the principle of general covariance, which affirms that the form of the physical laws is invariant under arbitrary differentiable coordinate transformations (diffeomorphisms).

The field equations of GR can be obtained from the Einstein-Hilbert action which is local:

$$\mathcal{S} = \frac{1}{16\pi G_N} \int d^4x \sqrt{-g} (R - 2\Lambda) + \mathcal{S}^{(M)} \quad (1.26)$$

where R is the Ricci scalar, Λ is the cosmological constant and $\mathcal{S}^{(M)}$ is the standard matter action. The variation of the action (1.26) with respect to the metric results in the field equations

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G_N}{c^4} T_{\mu\nu}, \quad (1.27)$$

where $R_{\mu\nu}$ is the Ricci tensor, $g_{\mu\nu}$ is the metric tensor and $T_{\mu\nu}$ is the stress-energy tensor. The field equations of GR constitute a system of ten independent second order partial differential equations, whose solutions are the components of the metric tensor $g_{\mu\nu}$.

It is crucial to notice that, despite the fact that the metric tensor has 10 independent components, the degrees of freedom of the gravitational field are two. This results from the gauge symmetry associated to the choice of a coordinate system and the Bianchi identities for the Riemann curvature tensor $R_{\alpha\beta\mu\nu}$:

$$R_{\alpha\beta\mu\nu} + R_{\alpha\mu\nu\beta} + R_{\alpha\nu\beta\mu} = 0, \quad (1.28)$$

$$\nabla_\gamma R_{\alpha\beta\mu\nu} + \nabla_\mu R_{\alpha\beta\nu\gamma} + \nabla_\nu R_{\alpha\beta\gamma\mu} = 0. \quad (1.29)$$

The contraction of the second Bianchi identity shows that the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - Rg_{\mu\nu}/2$ has zero divergence:

$$\nabla^\mu G_{\mu\nu} = \nabla^\mu \left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \right) = 0. \quad (1.30)$$

From the field equations (1.27), since the divergence of the metric tensor is equal to zero, it is clear that this condition implies

$$\nabla^\mu T_{\mu\nu} = 0, \quad (1.31)$$

which is a conservation law for the stress-energy tensor.

1.3.1 Infinite Derivative Theories of Gravity

In order to analyze the main features of IDGs, it is useful to start from the following infinite derivatives Lorentz-invariant action [15]:

$$\mathcal{S} = \frac{1}{2} \int d^4x d^4y \phi(x) \mathcal{K}(x-y) \phi(y) - \int d^4x V(\phi), \quad (1.32)$$

where ϕ is a scalar field, $\mathcal{K}(x-y)$ is an operator depending on the distance $(x-y)$ between two points through a generic function of the d'Alembert operator \square and $V(\phi(x))$ is a local potential term.

It is worth mentioning that the operator \square has the dimension of an inverse length squared, hence a dimensionless quantity can be obtained as \square/M_s^2 , where M_s is an appropriate mass scale. For sake of brevity, only \square will be used in the following instead of \square/M_s^2 .

Notice also that the operator $\mathcal{K}(x-y)$ signals the non-locality and the kinetic term can be written as follows

$$\begin{aligned} S_K &= \frac{1}{2} \int d^4x d^4y \phi(x) \mathcal{K}(x-y) \phi(y) = \\ &= \frac{1}{2} \int d^4x d^4y \phi(x) \int \frac{d^4k}{(2\pi)^4} F(-k^2) e^{ik(x-y)} \phi(y) = \\ &= \frac{1}{2} \int d^4x d^4y \phi(x) F(\square) \int \frac{d^4k}{(2\pi)^4} e^{ik(x-y)} \phi(y) = \\ &= \frac{1}{2} \int d^4x \phi(x) F(\square) \phi(x), \end{aligned} \quad (1.33)$$

using the Fourier transform $F(-k^2)$ of $\mathcal{K}(x-y)$ and the integral representation of the Dirac delta. The general form of the operator $\mathcal{K}(x-y)$ is

$$\mathcal{K}(x-y) = F(\square) \delta^{(4)}(x-y), \quad (1.34)$$

and a factorization of $F(\square)$ yields

$$F(\square) = e^{-\gamma(\square)} \prod_{i=1}^N (\square - m_i^2), \quad (1.35)$$

with $\gamma(\square)$ being an entire function, so it can be represented as a power series. By means of a Fourier transformation, it is possible to show that ghosts appear when $N > 1$. For this reason, hereafter we focus on the $N = 1$ choice, where unitarity is preserved.

It is also important to understand why infinite order derivative operators have a non-local nature. This property is apparent from the notion of discrete derivatives: in order to evaluate the discrete version of the first derivative of a function it is necessary to evaluate the function in two adjacent points on a lattice. Hence to calculate the second derivative

three points are necessary and so on. Consequently, as the order of derivation increases also the number of points where the function has to be evaluated increases. It is clear that, when the order of derivation is infinite, the number of points where the function has to be evaluated is also infinite and non-locality emerges.

The most general action made of functions of the d'Alembert operator, which is ghost-free and quadratic in the curvature [16][17], must contain infinite covariant derivatives and reads

$$\mathcal{S} = \frac{1}{2\kappa} \int d^4x \sqrt{-g} [R + \alpha (RF_1(\square)R + R_{\mu\nu}F_2(\square)R^{\mu\nu} + R_{\mu\nu\rho\sigma}F_3(\square)R^{\mu\nu\rho\sigma})], \quad (1.36)$$

where the $F_i(\square)$ are transcendental analytic functions of \square .

An interesting example of non-local action is a subcase of (1.36), which also provides GR at the zero order:

$$\mathcal{S} = \frac{1}{2\kappa} \int d^4x \sqrt{-g} \left(R - G_{\mu\nu} \frac{e^{H(-\square)} - 1}{\square} R^{\mu\nu} \right) + \mathcal{S}^{(m)}, \quad (1.37)$$

where $H(-\square)$ is an analytic function of \square and $\mathcal{S}^{(m)}$ is the usual matter action. The field equations, up to order $O(R^2)$ read

$$G_{\mu\nu} + O(R^2) = \kappa e^{-H(-\square)} T_{\mu\nu}^{(m)}, \quad (1.38)$$

and they are equivalent to GR at the lowest order of the Taylor expansion.

In spherical symmetry, equation (1.38) yields regular black holes without singularities, while in an homogeneous and isotropic universe, it admits bouncing cosmological solutions.

The mechanism aimed at avoiding classical singularities basically consists in a non-local smearing of the point-like source of the Schwarzschild background [18], which automatically implies that the metric can no longer be considered as a vacuum solution. In the non-local region, where $r < 2/M_s$, the effects of non-locality start being relevant, while as soon as the radius start to increase, the solution approaches the Schwarzschild one. As mentioned above, the theory does not contain singularities for any value of r , including $r = 0$.

Another interesting example is given by the following action [19]:

$$\mathcal{S} = \frac{1}{2\kappa} \int d^4x \sqrt{-g} \left(R - G_{\mu\nu} \frac{V_2^{-1} - 1}{\square} R^{\mu\nu} + \frac{R}{2} \frac{V_0^{-1} - V_2^{-1}}{\square} R \right), \quad (1.39)$$

where

$$V_2^{-1} \equiv e^{H_2(-\square)} p^{(n_2)}(-\square), \quad (1.40)$$

$$V_0^{-1} - V_2^{-1} \equiv \frac{1}{3} [e^{H_0(-\square)}(1 + \square) - e^{H_2(-\square)}], \quad (1.41)$$

with H_i and $p^{(n_2)}$ being analytic functions. It turns out that this model is unitary, renormalizable and represents the maximal UV-completion for Starobinsky Gravity.

It is also important to mention how the principle of causality is affected in the case of IDG theories. In [15] it is shown that non-local interactions imply a violation of causality confined in a region $l_s = 1/M_s$ in coordinate space given by the scale of non-locality. In terms of momentum k , causality is violated for $k^2 > M_s^2$.

1.3.2 Integral Kernel Theories of Gravity

The other family of non-local theories of gravity is represented by IKGs. In this case, the non-locality is due to transcendental functions of the Ricci scalar, such as $\square^{-1}R$, which can be expressed as an integral operator:

$$\square^{-1}R(x) = \int d^4x G(x, x')R(x'), \quad (1.42)$$

where $G(x, x')$ is the Green's function associated to \square^{-1} .

The integral nature of the \square^{-1} operator gives rise to long range non-localities, hence it can be used to address the IR shortcomings of GR. This type of operator appears at fundamental level when non perturbative methods are applied to QFT on curved spacetime. In addition, the action is generally covariant and the stress-energy tensor is conserved if the following boundary conditions are imposed [20]:

$$G(x, y) = 0 \Big|_{t_0} = 0, \quad \partial_0 G(x, y) \Big|_{t_0} = 0, \quad (1.43)$$

which imply that the non-local effects are significant from a specific time t_0 .

An example of IKG action is obtained in [21]:

$$S = - \int d^4x \sqrt{-g} [V(x) + V(x)(\square - V)^{-1}V(x)] + \frac{1}{6}\Sigma, \quad (1.44)$$

where $V(x)$ is a generic potential and Σ is a boundary term that contains the non-local operator \square^{-1} defined as follows

$$\begin{aligned} \Sigma = & \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} [R - R_{\mu\nu}\square^{-1}G^{\mu\nu} + \frac{1}{2}R(\square^{-1}R^{\mu\nu})(\square^{-1}R_{\mu\nu}) + \\ & - R^{\mu\nu}(\square^{-1}R_{\mu\nu})(\square^{-1}R) - 2(\square^{-1}R^{\mu\nu})(\nabla_\mu \square^{-1}R^{\alpha\beta})(\nabla_\nu \square^{-1}R_{\alpha\beta}) + \\ & - 2(\nabla^\mu \square^{-1}R^{\nu\alpha})(\nabla_\nu \square^{-1}R_{\mu\alpha})\square^{-1}R + \square^{-1}R^{\mu\nu}(\nabla_\mu \square^{-1}R)(\nabla_\nu \square^{-1}R) + \\ & + \mathcal{O}(R_{\mu\nu}^4)]. \end{aligned} \quad (1.45)$$

It turns out that this action is quite complicated to treat analytically: the associated equations of motion are highly nonlinear and of high order, hence it is intricate to find exact solutions. Nonetheless, similar actions have been analyzed in the literature such as the Deser-Woodard model, that will be used in this thesis, for which the action reads

$$\mathcal{S} = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} R [1 + f(\square^{-1}R)] + \mathcal{S}^{(m)}, \quad (1.46)$$

where $f(\square^{-1}R)$ is a generic function of $\square^{-1}R$.

Among the possible IKG models, a prominent role both at theoretical and cosmological level is played by the so called ‘‘RR-model’’ [22]:

$$S_{NL} = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} \left(R - \frac{m^2}{6} R \square^{-2} R \right), \quad (1.47)$$

where the non-local operator is \square^{-2} instead of \square^{-1} of the Deser-Woodard model.

One of the main features of this model is that the coefficient of the non-local term $\square^{-2}R$ depends on a mass parameter m . This model can be written in a local form through the so called localization procedure [23] by setting

$$U = -\square^{-1}R \quad \text{and} \quad S = -\square^{-1}U = \square^{-2}R, \quad (1.48)$$

here U and S are two auxiliary scalar fields.

This condition can be imposed in the action (1.47) by means of two Lagrange multipliers ξ_1 and ξ_2 , such that

$$\mathcal{S}_{NL} = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} \left[R \left(-\frac{m^2}{6} S \right) - \xi_1 (\square U + R) - \xi_2 (\square S + U) \right] + \mathcal{S}^{(m)}. \quad (1.49)$$

The variation with respect to the metric tensor yields the following equations of motions

$$G_{\mu\nu} = \frac{m^2}{6} K_{\mu\nu} + 8\pi G_N T_{\mu\nu}, \quad (1.50)$$

where the tensor $K_{\mu\nu}$ satisfies the condition $\nabla^\mu K_{\mu\nu} = 0$ and depends on the metric and the scalar fields U and S :

$$\begin{aligned} K_{\mu\nu} = & 2S G_{\mu\nu} - 2\nabla_\mu \partial_\nu S - 2U g_{\mu\nu} + g_{\mu\nu} \partial_\rho S \partial^\rho U \\ & - \frac{1}{2} g_{\mu\nu} U^2 - (\partial_\mu S \partial_\nu U + \partial_\nu S \partial_\mu U). \end{aligned} \quad (1.51)$$

This approach is intriguing because it shows that non-local operators allows to introduce a mass term without breaking the gauge invariance of the massless theory both in electromagnetism and linearized massive gravity. Furthermore, this model can be applied to Cosmology in the so called degravitation mechanism [24].

1.3.3 Non-local Teleparallel Gravity

It is worthwhile to mention an approach to Non-local theories of Gravity based on the so called Teleparallel formulation of General Relativity [25]. Similarly to GR, Teleparallel Gravity offers a geometric description of the gravitational interaction, where the gravitational field is described in terms of torsion rather than curvature.

The crucial idea is the introduction of a tetrad field [26]: a set of four vector fields e_a defined on the entire spacetime manifold M , such that for every point of the manifold the vectors of the tetrad field form a basis for the tangent space $T_p M$ at that point.

If the coordinate induced basis for the tangent space at a given point of the manifold is denoted as $\{\partial_\mu\}$, the tetrad vectors at the point p are given by

$$e_a = e_a^\mu \partial_\mu. \quad (1.52)$$

A tetrad determines also a co-tetrad at a given point p , a set of 1-forms related to the coordinate induced dual base $\{dx^\mu\}$ by

$$e^a = e^a_\mu dx^\mu, \quad (1.53)$$

such that

$$e^\mu_a e^\nu_b = \delta_{ab}, \quad (1.54)$$

where latin a greek indices are used to label the coordinates on the manifold and the tangent space respectively.

The tetrad field has to satisfy an orthonormality condition with respect to the metric tensor $g_{\mu\nu}$, namely

$$g_{\mu\nu} e^\mu_a e^\nu_b = \eta_{ab}, \quad (1.55)$$

where $\eta_{ab} = \text{diag}(+1, -1, -1, -1)$ is the Minkowski metric.

Notice that of the sixteen degrees of freedom of the orthonormal tetrad field, ten are fixed by the metric and the other six specify the tetrad field with respect to a fiducial orthonormal frame field. This is an essential property of this formulation and encodes the notion that at each event x , the different orthonormal tetrad fields are related to one another by an element of the local Lorentz group, which can be characterized by three boost speeds and three rotation angles.

In Teleparallel Gravity, Einstein's theory of General Relativity is recast as the gauge theory of the Abelian group of spacetime translations, indicated as $GR_{||}$. Hence it bears some resemblance to electrodynamics, which is the gauge theory of the Abelian $U(1)$ group. It is important to notice that the gauge approach gives rise to spacetime theories

with curvature and torsion.

There is a spectrum of such theories: on the one hand there are theories based on a pseudo-Riemannian manifold with curvature but no torsion such as General Relativity; on the other hand there are theories with torsion and no curvature.

In particular, Teleparallel Gravity is characterized by two connections: the Levi-Civita connection that emerges naturally in GR from the metric

$$\Gamma_{\alpha\beta}^{\mu} = \frac{1}{2}g^{\mu\eta}(\partial_{\alpha}g_{\beta\eta} + \partial_{\beta}g_{\eta\alpha} - \partial_{\eta}g_{\alpha\beta}), \quad (1.56)$$

which is torsion free and metric compatible; and an additional connection, the so called Weitzenböck connection:

$$\hat{\Gamma}_{\alpha\beta}^{\mu} = e_c^{\mu} \partial_{\alpha} e^c_{\beta}, \quad (1.57)$$

which is not symmetric, hence torsion is nonzero, but it is curvature free and metric compatible. The torsion tensor $C_{\alpha\beta}$ is then given by

$$C_{\alpha\beta}^{\mu} = \hat{\Gamma}_{\alpha\beta}^{\mu} - \hat{\Gamma}_{\beta\alpha}^{\mu} \quad (1.58)$$

and plays a crucial role in Teleparallel Gravity.

The upshot of these considerations is that at each event in spacetime, the gravitational field is characterized by both torsion and curvature. They are complementary representations of the gravitational field.

The tetrad formulation of GR dates back to Einstein's attempt to unify the gravitational and electromagnetic interactions. In fact, it turns out that the field equations of Teleparallel Gravity are very similar to Maxwell's equations of electromagnetism.

This resemblance is extremely interesting for a non-local extension of GR [27], since in the electrodynamics of non-local media Maxwell's equations are unchanged but a non-local constitutive relation has to be introduced.

This observation motivates the construction of a non-local theory of gravity in close analogy to the non-local electrodynamics of media, where the role of the torsion tensor is very similar to that of the electromagnetic tensor $F_{\mu\nu}$ and non-locality is added to $GR_{||}$ through a constitutive equation. This approach has also been used to address the Dark Matter problem in GR as can be seen in [28].

The field equations of GR without a cosmological constant can indeed be written in the framework of $GR_{||}$ as

$$\frac{\partial}{\partial x^{\nu}} \mathcal{H}^{\mu\nu}_c = \sqrt{-g} (T_c^{\mu} + E_c^{\mu}), \quad (1.59)$$

where $\mathcal{H}^{\mu\nu}_c$ is the so called auxiliary field strength, constructed from the torsion tensor $C_{\alpha\beta}^{\mu}$; while T_c^{μ} and E_c^{μ} are related to the energy-momentum tensor of matter and gravitational field respectively.

The equations (1.59) are very similar to the inhomogeneous Maxwell's equations

$$\partial_\nu(\sqrt{-g}H^{\mu\nu}) = \frac{4\pi}{c}\sqrt{-g}j^\mu, \quad (1.60)$$

if $H^{\mu\nu}$ is the magnetic field strength and j^μ is the electromagnetic four-current.

In the electrodynamics of non-local media it is necessary to add a non-local constitutive equation between $H^{\mu\nu}$ and $F^{\mu\nu}$ in order to solve the field equations. Analogously, in the framework of non-local $GR_{||}$ it is necessary to add a non-local constitutive equation between $\mathcal{H}^{\mu\nu}_c$ and the torsion tensor $C_{\alpha\beta}{}^\mu$.

In conclusion, it is interesting to notice that the torsion tensor satisfies an equation that is similar to the homogeneous Maxwell's equation

$$F_{[\mu\nu,\rho]} = 0, \quad (1.61)$$

based on its definition (1.58).

Chapter 2

The Deser-Woodard Model

This chapter is devoted to a description of the so called Deser-Woodard model [29]. This model can be considered as a particular case of the IKGs presented in the previous chapter, which is quite interesting from a cosmological standpoint. It will be shown that the presence of the operator \square^{-1} can be used to explain the late-time cosmic acceleration of the Universe. In Section (2.2) I will outline the so called Noether Symmetry Approach for the selection of theories of gravity. This method will be then applied to the Deser-Woodard model in order to understand how the term $\square^{-1}R$ modifies the Einstein-Hilbert action. In the last section the attention is focused on the Weak-Field limit of the Deser-Woodard model, which shows how the Newtonian limit for the gravitational potential is recovered at lowest order of the theory and how it is modified by non-locality.

2.1 The model and the localization of the action

The Deser-Woodard (DW) model is a non-local extension of GR where the non-locality is induced by the operator \square^{-1} acting on the Ricci scalar. The action of this model reads:

$$\mathcal{S} = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} R [1 + f(\square^{-1}R)] + \mathcal{S}^{(M)} \quad (2.1)$$

where R is the Ricci scalar, $\kappa^2 = 8\pi G_N$, $\mathcal{S}^{(M)}$ is the standard matter action and $f(\square^{-1}R)$ is an arbitrary function of the non-local term $\square^{-1}R$ called *distortion function*.

The variation of the action (2.1) with respect to the metric tensor yields:

$$G_{\mu\nu} + \Delta G_{\mu\nu} = \kappa^2 T_{\mu\nu}^{(M)}, \quad (2.2)$$

where the correction to the field equations of GR is given by

$$\begin{aligned} \Delta G_{\mu\nu} = & (G_{\mu\nu} + g_{\mu\nu}\square - \nabla_\mu \nabla_\nu)[f + \square^{-1}(Rf')] + \\ & + \left[\delta_\mu^{(\rho} \delta_\nu^{\sigma)} - \frac{1}{2}g_{\mu\nu}g^{\rho\sigma} \right] \partial_\rho(\square^{-1}R) \partial_\sigma[\square^{-1}(Rf')], \end{aligned} \quad (2.3)$$

$$f' \equiv \frac{\partial f}{\partial(\square^{-1}R)}. \quad (2.4)$$

It is possible to obtain a local representation of (2.1) through the so called localization procedure [23]. Two auxiliary scalar fields ϕ and ξ are introduced as follows:

$$\begin{aligned} \mathcal{S} &= \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} [R(1 + f(\phi)) + \xi(\square\phi - R)] + S^{(m)} = \\ &= \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} [R(1 + f(\phi) - \xi) - \nabla_\alpha \xi \nabla^\alpha \phi] + S^{(m)}, \end{aligned} \quad (2.5)$$

where in the last passage a total derivative has been integrated out.

Notice that the Deser-Woodard model, which is a metric non-local theory of gravity, has been recast as a biscalar-tensor theory.

Varying the action (2.5) with respect to the metric tensor yields the field equations for this localized version of the Non-local Gravity model:

$$(1 + f(\phi) - \xi) G_{\mu\nu} + \frac{1}{2} g_{\mu\nu} \nabla^\alpha \xi \nabla_\alpha \phi = \kappa^2 T_{\mu\nu}^M + \nabla_\mu \xi \nabla_\nu \phi + (\nabla_\mu \nabla_\nu - g_{\mu\nu} \square)(f(\phi) - \xi), \quad (2.6)$$

where $G_{\mu\nu}$ is the Einstein tensor while $T_{\mu\nu}^M$ is the stress energy tensor associated to matter. The variation of the action (2.5) with respect to the scalar fields ϕ and ξ shows that those fields satisfy the Klein-Gordon equations:

$$\square\phi = R \implies \phi = \square^{-1}R, \quad (2.7)$$

$$\square\xi = -R \frac{df}{d\phi}. \quad (2.8)$$

Those equations imply also that the scalar field ξ has the role of a Lagrange multiplier, fixing the constraint necessary to recover the action (2.1).

Another interesting equation is the trace of (2.6) which reads

$$[1 + f(\phi) - \xi - 3f'(\phi)] R = -\kappa^2 T^M + \nabla_\alpha \xi \nabla^\alpha \phi + 3f''(\phi) \nabla_\alpha \phi \nabla^\alpha \phi, \quad (2.9)$$

where T^M is the trace of the stress-energy tensor.

It is worthwhile to mention that the equation (2.9) can be used to express the Ricci scalar with respect to the scalar fields ϕ and ξ and their derivatives.

$$R = \frac{-\kappa^2 T^M + \nabla_\alpha \xi \nabla^\alpha \phi + 3f''(\phi) \nabla_\alpha \phi \nabla^\alpha \phi}{1 + f(\phi) - \xi - 3f'(\phi)}. \quad (2.10)$$

Furthermore, specific forms of the distortion function can be found through the so called *Noether Symmetry Approach*, requiring that the theory is invariant under point transformations.

2.1.1 Cosmological applications

The Deser-Woodard model is a phenomenological model that was first proposed in order to explain the current acceleration of the expansion of the Universe in terms of non-local gravitational effects [29]. Indeed, this model allows a delayed response to cosmic events, such as the transition from radiation to matter dominance. Through this delayed response, the Deser-Woodard model accounts for the late time accelerated expansion of the Universe without invoking any unknown form of Dark Energy. Therefore, the fine-tuning related to the value of Λ is avoided.

The best way to identify these properties is to consider a homogeneous, isotropic and spatially flat metric

$$ds^2 = dt^2 - a^2(t)[dx^2 + dy^2 + dz^2], \quad (2.11)$$

such that the deceleration parameter has the form

$$q(t) \equiv -\frac{a\ddot{a}}{\dot{a}^2} = -1 - \frac{\dot{H}}{H^2} \quad \text{where} \quad H(t) \equiv \frac{\dot{a}}{a}. \quad (2.12)$$

Assuming that the scale factor $a(t)$ can be expressed as $a(t) \approx t^s$, the Hubble and deceleration parameters become

$$q(t) = \frac{1-s}{s} \quad \text{and} \quad H(t) \equiv \frac{s}{t}, \quad (2.13)$$

where $s = 1/2$ corresponds to radiation dominance, while $s = 2/3$ to matter dominance. Furthermore, in this framework the action of the operator \square^{-1} on the Ricci scalar can be written as

$$\square^{-1}R(t) = -\int_0^t dt' \frac{1}{a^3(t')} \int_0^{t'} dt'' a^3(t'')R(t''). \quad (2.14)$$

Assuming that the transition from the radiation to matter dominance occurred at a particular time $t = t_{eq}$, which is justified by numerical results, the expression (2.14) becomes

$$\square^{-1}R(t) = -\frac{6s(2s-1)}{3s-1} \left\{ \ln\left(\frac{t}{t_{eq}}\right) - \frac{1}{3s-1} + \frac{1}{3s-1} \left(\frac{t_{eq}}{t}\right)^{3s-1} \right\}, \quad (2.15)$$

where s is given by the scale factor $a(t)$ in the matter dominated era.

In addition, if we choose the value of $s = 2/3$ and $t_0 \approx 10^{10}$ years for the present time, the above formula becomes

$$\square^{-1}R(t_0)|_{s=2/3} \simeq -14. \quad (2.16)$$

The upshot of this calculation is that non-local corrections can explain the observed acceleration of the expansion of the Universe without resorting to a cosmological constant, because the non-local correction is of the order of magnitude expected for the fine tuning of the cosmological parameters in the Λ CDM model.

In the evaluation of any extended theory of gravity such as $f(R)$ or $f(\square^{-1}R)$ it is of paramount importance the so called reconstruction problem [30]. In Cosmology it refers to the issue of characterizing the parameters of a given model such that the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric is supported. In the case of the DW model, we want to determine if there exists a distortion function which reproduces the Λ CDM expansion without a cosmological constant.

The field equations of the DW model in for the FLRW metric take the form

$$3H^2 + \Delta G_{00} = 8\pi G_N \rho, \quad (2.17)$$

$$-2\dot{H} - 3H^2 + \frac{1}{3a^2} \delta^{ij} \Delta G_{ij} = 8\pi G_N P, \quad (2.18)$$

where ρ is the matter energy density and P is the pressure. Those quantities are related by the energy-momentum conservation law

$$\dot{\rho} + 3H(P + \rho) = 0. \quad (2.19)$$

The explicit form of the non-local contributions ΔG_{00} and ΔG_{ij} depends on the Hubble parameter H and the non-local term $\square^{-1}R$:

$$\begin{aligned} \Delta G_{00} = & [3H^2 + 3H\partial_t] \{f(\square^{-1}R) + \square^{-1}[Rf'(\square^{-1}R)]\} + \\ & + \frac{1}{2} \partial_t(\square^{-1}R) \partial_t(\square^{-1}[Rf'(\square^{-1}R)]), \end{aligned} \quad (2.20)$$

$$\begin{aligned} \Delta G_{ij} = & - [2\dot{H} + 3H^2 + 2H\partial_t + \partial_t^2] \{f(\square^{-1}R) + \square^{-1}[Rf'(\square^{-1}R)]\} g_{ij} + \\ & + \frac{1}{2} \partial_t(\square^{-1}R) \partial_t(\square^{-1}[Rf'(\square^{-1}R)]) g_{ij}. \end{aligned} \quad (2.21)$$

The general reconstruction procedure starts by introducing the function

$$F(\square^{-1}R) = f(\square^{-1}R) + \square^{-1}(R f'(\square^{-1}R)), \quad (2.22)$$

and noticing that by taking the difference of the equations (2.20) and (2.21) the function F satisfies the following ordinary differential equation

$$\ddot{F} + 5H\dot{F} + (6H^2 + 2\dot{H})(F + 1) = 8\pi G_N(\rho - P). \quad (2.23)$$

Once the matter content of the Universe and the function $a(t)$ have been chosen, we have to solve the above equation for the function F and invert equation (2.22) to obtain an ordinary differential equation for the distortion function $f(\square^{-1}R)$.

By setting $X = \square^{-1}R$ and $Y = X + 16.5$, this approach leads to the following expression for the distortion function $f(X)$

$$f(X) = 0.245 [\tanh(0.35 Y + 0.032 Y^2 + 0.003 Y^3) - 1]. \quad (2.24)$$

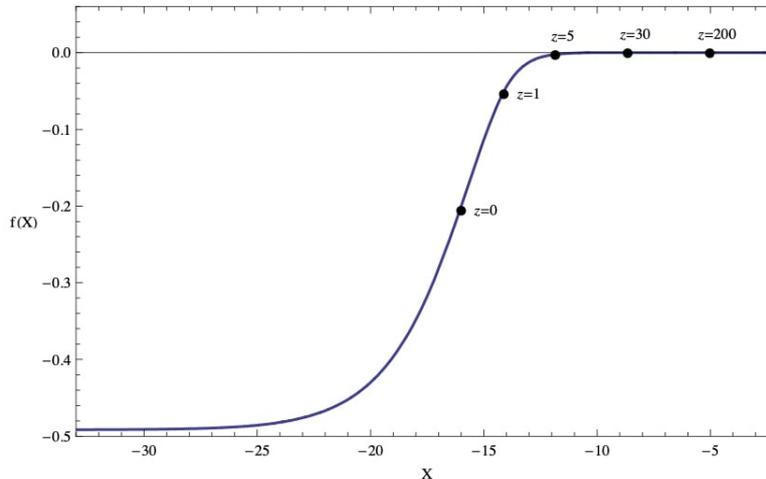


Figure 2.1: Reconstructed distortion function for the Λ CDM background evolution without cosmological constant [30].

2.2 The Noether Symmetry Approach

The Noether Symmetry Approach is a geometric criterion to select theories of gravity based on the spacetime symmetries [31]. This method can be employed to solve the field equations of different theories of gravity and constrain arbitrary functions in the action. In many attempts to extend General Relativity, the Ricci scalar R is replaced by some functions $f(R)$, $f(T)$, $f(\square^{-1}R)$, $f(\mathcal{G})$ etc, where T is the torsion invariant and \mathcal{G} is the Gauss-Bonnet scalar. Those theories provide physically reliable models only when confronted with data, in other words phenomenological constraints are required. As a matter of fact, it is also possible to find theoretical constraints searching for symmetries, such as in the Noether Symmetry Approach.

It is crucial to point out that this approach is applied to dynamical systems described by second order differential equations of the form

$$\ddot{x}^\mu + \Gamma_{\nu\eta}^\mu \dot{x}^\nu \dot{x}^\eta = F^\mu, \quad (2.25)$$

where $\Gamma_{\nu\eta}^\mu$ are the affine connection coefficients obtained from the metric tensor $g_{\mu\nu}$, while F^μ is a C^∞ vector field on spacetime which is associated to the force field of the dynamical system. In particular, the idea is to express the symmetries of the equation (2.25) in terms of the symmetries of the metric. This is extremely useful, since it is much easier to find the symmetries of the metric, the associated integral of motions and reduce the order of the field equations.

First of all, let us introduce the crucial notion of point transformation, namely the map-

ping of points (x, y) into points (\bar{x}, \bar{y}) where x is the independent variable and y is the dependent variable.

A class of one parameter point transformations can be written as

$$\bar{x} = \bar{x}(x, y, \epsilon) \quad \bar{y} = \bar{y}(x, y, \epsilon), \quad (2.26)$$

where ϵ is a parameter. The transformations have to satisfy the following properties:

- they are invertible,
- repeated application yields a transformation of the same family,
- the transformation corresponds to the identity if, for example, $\epsilon = 0$

$$\bar{x}(x, y, 0) = x, \quad \bar{y}(x, y, 0) = y. \quad (2.27)$$

Starting from the one parameter point transformations (2.26), \bar{x} and \bar{y} can be expanded around $\epsilon = 0$, which corresponds to the identity, resulting in

$$\bar{x}(x, y, \epsilon) = x + \epsilon \left. \frac{\partial \bar{x}}{\partial \epsilon} \right|_{\epsilon=0} + \dots = x + \epsilon \xi(x, y) + \dots \quad (2.28)$$

$$\bar{y}(x, y, \epsilon) = y + \epsilon \left. \frac{\partial \bar{y}}{\partial \epsilon} \right|_{\epsilon=0} + \dots = y + \epsilon \eta(x, y) + \dots \quad (2.29)$$

The vector

$$\mathbf{X} = \xi(x, y) \frac{\partial}{\partial x} + \eta(x, y) \frac{\partial}{\partial y} \quad (2.30)$$

is called the infinitesimal generator of the transformations and corresponds to the tangent vector to the group orbits at the point (x, y) .

In order to investigate how differential equations are affected by these transformations, it is necessary to prolong the transformations to the derivatives of the dependent variable y . In particular, the prolongation of an infinitesimal generator generalizes this object in order to account for the transformation of the independent variable, the dependent variable and its derivatives. Furthermore, the derivative of \bar{y} can be expressed as

$$\bar{y}'(x, y, y', \epsilon) = \frac{d\bar{y}(x, y, \epsilon)}{d\bar{x}(x, y, \epsilon)} = \frac{y'(\partial \bar{y} / \partial y) + \partial \bar{y} / \partial x}{y'(\partial \bar{x} / \partial y) + \partial \bar{x} / \partial x}, \quad (2.31)$$

$$\bar{y}''(x, y, y', y'', \epsilon) = \frac{d\bar{y}'}{d\bar{x}}, \quad (2.32)$$

clearly this method can be used to obtain derivatives of any order of \bar{y} .

Replacing the Taylor expansions (2.28) and (2.29) into the equations (2.31) and (2.32) results in

$$\bar{y}'' = y'' + \epsilon \left(\frac{d\eta}{dx} - y' \frac{d\xi}{dx} \right) + \dots = y'' + \epsilon \eta^{[1]} + \dots, \quad (2.33)$$

$$\bar{y}^{(n)} = y^{(n)} + \epsilon \left(\frac{d\eta^{(n-1)}}{dx} - y^{(n)} \frac{d\xi}{dx} \right) + \dots = y^{(n)} + \epsilon \eta^{[n]} + \dots, \quad (2.34)$$

where

$$\eta^{[n]} = \frac{d\eta^{(n-1)}}{dx} - y^{(n)} \frac{d\xi}{dx} \quad (2.35)$$

is the n^{th} prolongation function of η . Finally, the n^{th} prolongation of the generator \mathbf{X} is

$$\mathbf{X}^{[n]} = \mathbf{X} + \eta^{[1]} \frac{\partial}{\partial y'} + \dots + \eta^{[n]} \frac{\partial}{\partial y^{(n)}}. \quad (2.36)$$

As an example, let us start from the above expression for the n^{th} prolongation of the generator and consider a point-like Lagrangian density $\mathcal{L} = \mathcal{L}(t, q^i(t), \dot{q}^i(t))$, where q^i are the coordinates, t is the time and \dot{q}^i refers to the derivative of the coordinates q^i with respect to time.

The first prolongation of Noether's vector becomes

$$\mathbf{X}^{[1]} = \xi(t, q) \frac{\partial}{\partial t} + \eta^i(t, q^k) \frac{\partial}{\partial q^i} + (\dot{\eta}^i - \dot{\xi} \dot{q}^i) \frac{\partial}{\partial \dot{q}^i}. \quad (2.37)$$

The first Noether's theorem can be stated as follows: the one parameter group of point transformations generated by \mathbf{X} is a one parameter group of Noether point symmetries for the dynamical system described by \mathcal{L} if and only if there exists a function $g(t, q^k(t))$ such that

$$\mathbf{X}^{[1]} \mathcal{L} + \dot{\xi} \mathcal{L} = \dot{g}, \quad (2.38)$$

whose associated first integral of motion is

$$I(t, q^i, \dot{q}^i) = \xi \left(\dot{q}^i \frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \mathcal{L} \right) - \eta^i \frac{\partial \mathcal{L}}{\partial \dot{q}^i} + g. \quad (2.39)$$

In conclusion, let us outline a general procedure that allows to find the Noether symmetries, since it will be later employed to the specific case of the Deser-Woodard model.

It is clear from equation (2.30) that the Noether symmetries are determined once the coefficients of the generator \mathbf{X} are known, consistently with the equation (2.38).

If the Lagrangian density $\mathcal{L}(t, q^i, \dot{q}^i)$ describes a generic dynamical system, then

- write an ansatz for the spacetime metric, then the form of the generator \mathbf{X} defined on the configuration space is given,
- expand the equation (2.38) to obtain a polynomial depending on the coefficients of the generator, $\xi(t, q^i)$ and $\eta^i(t, q^k)$, and products of generalized velocities \dot{q}^k ,
- since the unknown coefficients ξ and η^i depend only on (t, q^k) , in order for the polynomial to vanish, the coefficients of the products of the generalized velocities have to vanish. Thus we end up with a set of partial differential equations for ξ and η , which can be solved in a straightforward way in many cases,

- finally, once the generator is known, it is possible to find the first integral of motion $I(t, q^i, \dot{q}^i)$ given in equation (2.39) and reduce the order of the Euler-Lagrange equations; depending on the number of symmetries, one can achieve the complete integrability of the dynamical system

As a final comment, it is interesting to notice that there are also other criteria to search for symmetries of differential equations. For instance, some methods involve the equations of motions, rather than the Lagrangian or the Hamiltonian of the system. Another type of symmetries is given by Cartan symmetries, which are point transformations with generators in the tangent bundle that leave the Cartan 1-form invariant. It is also worth mentioning that the Noether Symmetry Approach has interesting applications in Cosmology [32].

2.2.1 The distortion function

The Noether Symmetry Approach can be applied to the Deser-Woodard model (2.1) following the procedure outlined in the previous section [2].

Since the Ultra Diffuse Galaxies are almost spherically symmetric systems, the background spacetime is chosen as spherically symmetric, hence the line element is given by

$$ds^2 = e^{\nu(t,r)} dt^2 - e^{\lambda(t,r)} dr^2 - r^2 d\Omega^2, \quad (2.40)$$

where $\nu(t, r)$ and $\lambda(t, r)$ are two arbitrary functions.

Notice that it is not known whether the Birkhoff's theorem is valid or not in Non-local Gravity, consequently the functions λ and ν depend both on the radial coordinate r and time t .

The next step is to substitute the metric (2.40) into the action (2.1). Thus, the following point-like Lagrangian density is obtained:

$$\begin{aligned} \mathcal{L} = e^{-\frac{1}{2}(\lambda+\nu)} [& - e^{\nu} r^2 \nu_r \phi_r f'(\phi) + e^{\lambda} r^2 \lambda_t \phi_t f'(\phi) + \\ & - 2 e^{\nu} f(\phi) (e^{\lambda} + r \lambda_r - 1) - 2 e^{\lambda+\nu} + 2 e^{\nu} + e^{\nu} r^2 \xi_r \phi_r + \\ & + e^{\nu} r^2 \nu_r \xi_r - e^{\lambda} r^2 \xi_t \phi_t - e^{\lambda} r^2 \lambda_t \xi_t + \\ & + 2 e^{\nu} \xi (e^{\lambda} + r \lambda_r - 1) - 2 e^{\nu} r \lambda_r], \end{aligned} \quad (2.41)$$

where the derivative with respect to the variables r and t has been denoted by the subscript. Furthermore, the canonical generator of the point transformation (i.e. the Noether vector) takes the form

$$X = \xi^t(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial t} + \xi^r(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial r} + \quad (2.42)$$

$$\begin{aligned}
 & + \eta^\nu(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \nu} + \eta^\lambda(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \lambda} + \\
 & + \eta^\phi(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \phi} + \eta^\xi(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \xi},
 \end{aligned}$$

where the functions ξ^t , ξ^r , η^ν , η^λ , η^ϕ and η^ξ are the coefficients of the Noether vector X . Once the point-like Lagrangian density and the Noether vector are known, the dynamical system (2.41) satisfies the required symmetry if the following condition holds:

$$X^{[1]} \mathcal{L} + \mathcal{L} \left(\frac{d\xi^t}{dt} + \frac{d\xi^r}{dr} \right) = \frac{dh^t}{dt} + \frac{dh^r}{dr}, \quad (2.43)$$

where h^t and h^r have the role of the function g introduced in the previous section and depend on the variables $(t, r, \lambda, \phi, \xi)$.

Notice also that $X^{[1]}$ is the first prolongation of the Noether vector X , hence it can be expressed as

$$\begin{aligned}
 X^{[1]} = & X + \alpha^{[1]}(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial r_t} + \beta^{[1]}(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \nu_t} + \\
 & \gamma^{[1]}(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \lambda_t} + \delta^{[1]}(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \phi_t} + \\
 & \epsilon^{[1]}(t, r, \nu, \lambda, \phi, \xi) \frac{\partial}{\partial \xi_t},
 \end{aligned} \quad (2.44)$$

having introduced the coefficients $\alpha^{[1]}$, $\beta^{[1]}$, $\gamma^{[1]}$, $\delta^{[1]}$ and $\epsilon^{[1]}$.

The expansion of the condition (2.43) results in 75 equations with 9 unknown variables: the coefficients of the Noether vector $\{\xi^t, \xi^r, \eta^\nu, \eta^\lambda, \eta^\phi, \eta^\xi\}$, the unknown functions $\{h^t, h^r\}$ and the form of the distortion function $f(\phi)$.

The specific forms of the distortion function obtained by solving the above-mentioned equations are:

$$f(\phi) = c_4 + c_3 \phi \quad (2.45)$$

$$f(\phi) = c_4 + \frac{c_5}{c_1} e^{c_1 \phi}, \quad (2.46)$$

with the associated Noether vectors

$$X = (c_1 t + \xi^t(r)) \frac{\partial}{\partial t} - 2c_1 \frac{\partial}{\partial \nu} + (c_2 + 2c_1) \frac{\partial}{\partial \phi} + c_3 (c_2 + 2c_1) \frac{\partial}{\partial \xi}, \quad (2.47)$$

$$X = (c_2 t + \xi^t(r)) \frac{\partial}{\partial t} - \frac{c_3}{2} r \frac{\partial}{\partial r} + (2c_2 + 2c_3) \frac{\partial}{\partial \nu} + c_1 c_3 \frac{\partial}{\partial \phi} + c_3 (\xi - c_4 - 1) \frac{\partial}{\partial \xi}, \quad (2.48)$$

where c_n , n from 1 to 5, is an integration constant.

In conclusion, the Noether Symmetry Approach selects a linear and exponential form for the distortion function. Those forms are already known in the literature: the exponential distortion function has been used to describe inflation [23], whereas the linear distortion function has been applied to solve the problem of unboundedness of euclidean gravity [33].

2.3 The weak field limit

The objective of this section is to study the Weak-Field limit of the Deser-Woodard model [2]. In particular, once the distortion function has been chosen, it is possible to derive the Non-local gravity potential, as well as the expressions for the scalar fields ϕ and ξ , in the Weak Field limit.

From General Relativity it is known that the Newtonian potential for time-like particles can be obtained by expanding the g_{00} component of the metric to the second order.

Furthermore, the Post-Newtonian (PN) limit can be obtained by expanding the components of the metric as

$$g_{00} \sim \mathcal{O}(6), \quad g_{0i} \sim \mathcal{O}(5), \quad g_{ij} \sim \mathcal{O}(4). \quad (2.49)$$

Notice, however, that the leading order of the PN limit does not require an expansion to $\mathcal{O}(6)$. In order to determine the Weak-Field limit of the DW model, it is sufficient to consider the gravitational field produced by a point-like source, assuming also that the metric is static and spherically symmetric.

It is worth noticing that the Birkhoff's theorem is not expected to hold in Non-local Gravity, nonetheless it is reasonable to assume that as a first approximation in the weak field limit the metric is static.

In this framework, the metric is then given by

$$ds^2 = A(r) dt^2 - B(r) dr^2 - r^2 d\Omega^2. \quad (2.50)$$

Since the coefficients of the metric depend only on the radial coordinate, the scalar fields inherit the isometries of the metric: $\phi = \phi(r)$ and $\xi = \xi(r)$.

In order to solve the field equations, let us assume the following expansions for the metric coefficients and scalar fields

$$A(r) = 1 + \frac{1}{c^2} \Phi(r)^{(2)} + \frac{1}{c^4} \Phi(r)^{(4)} + \frac{1}{c^6} \Phi(r)^{(6)} + \mathcal{O}(8), \quad (2.51)$$

$$B(r) = 1 + \frac{1}{c^2} \Psi(r)^{(2)} + \frac{1}{c^4} \Psi(r)^{(4)} + \mathcal{O}(6), \quad (2.52)$$

$$\phi(r) = \phi_0 + \frac{1}{c^2} \phi(r)^{(2)} + \frac{1}{c^4} \phi(r)^{(4)} + \frac{1}{c^6} \phi(r)^{(6)} + \mathcal{O}(8), \quad (2.53)$$

$$\xi(r) = \xi_0 + \frac{1}{c^2} \xi(r)^{(2)} + \frac{1}{c^4} \xi(r)^{(4)} + \frac{1}{c^6} \xi(r)^{(6)} + \mathcal{O}(8), \quad (2.54)$$

where ϕ_0 and ξ_0 are the constant background values of each field. Those constants are chosen to be equal to $\xi_0 = e^{\phi_0}$ and $\phi_0 = 0$ to recover the usual coupling of the Newton's constant with the Ricci scalar. In addition, we choose the exponential form for the distortion function:

$$f(\phi) = 1 + e^\phi \quad (2.55)$$

where the integration constants in (2.46) have been set equal to one, $c_4 = c_5 = c_1 = 1$ for simplicity. As reported in [2], once the distortion function has been chosen, it is possible to derive the 00- and 11-components of the equation (2.6) and the two equations (2.7) and (2.8) for the scalar fields.

The solutions of those equations can be expressed as

$$A(r) = 1 - \frac{2G_N M \phi_c}{c^2 r} + \frac{G_N^2 M^2}{c^4 r^2} \left[\frac{14}{9} \phi_c^2 + \frac{18 r_\xi - 11 r_\phi}{6 r_\xi r_\phi} r \right] + \quad (2.56)$$

$$- \frac{G_N^3 M^3}{c^6 r^3} \left[\frac{50 r_\xi - 7 r_\phi}{12 r_\phi r_\xi} \phi_c r + \frac{16 \phi_c^3}{27} - \frac{r^2 (2 r_\xi^2 - r_\phi^2)}{r_\xi^2 r_\phi^2} \right],$$

$$B(r) = 1 + \frac{2G_N M \phi_c}{3 c^2 r} + \frac{G_N^2 M^2}{c^4 r^2} \left[\frac{2 \phi_c^2}{9} + \left(\frac{3}{2 r_\xi} - \frac{1}{r_\phi} \right) r \right], \quad (2.57)$$

$$\phi(r) = \frac{4G_N M \phi_c}{3 c^2 r} - \frac{G_N^2 M^2}{c^4 r^2} \left[\left(\frac{11}{6 r_\xi} - \frac{1}{r_\phi} \right) r - \frac{2 \phi_c^2}{9} \right] + \quad (2.58)$$

$$- \frac{G_N^3 M^3}{c^6 r^3} \left[\frac{r^2}{r_\phi^2} - \left(\frac{25}{12 r_\xi} - \frac{7}{6 r_\phi} \right) \phi_c r - \frac{4 \phi_c^3}{81} \right],$$

$$\xi(r) = 1 + \frac{G_N^2 M^2}{c^4 r^2} \left[\frac{2 \phi_c^2}{3} - \left(\frac{13}{6 r_\xi} - \frac{1}{r_\phi} \right) r \right] + \quad (2.59)$$

$$+ \frac{G_N^3 M^3}{c^6 r^3} \left[\frac{20 \phi_c^3}{27} - \left(\frac{1}{r_\xi^2} - \frac{1}{r_\phi^2} \right) r^2 - \left(\frac{131}{36 r_\xi} + \frac{1}{6 r_\phi} \right) \phi_c r \right].$$

It is crucial to notice that there are 3 parameters in these expressions: a dimensionless constant ϕ_c and two characteristic lengths r_ϕ and r_ξ . The constant ϕ_c modifies the gravitational coupling, i.e. $G_{eff} = G_N \phi_c$, and it will be set equal to 1 in order to recover the Newtonian limit at the lowest order. On the other hand, the two lengths scales r_ϕ and r_ξ emerge in the $\mathcal{O}(4)$ order and they are related to the two scalar fields.

In the Weak-Field limit the relation between the metric coefficient g_{00} and the gravitational potential $\Phi(r)$ is given by

$$g_{00} = 1 + \frac{2 \Phi(r)}{c^2}, \quad (2.60)$$

and comparing this result with equation (2.56), the gravitational potential in the Deser-Woodard model becomes

$$\Phi(r) = - \frac{G_N M \phi_c}{r} + \frac{G_N^2 M^2}{2 c^2 r^2} \left[\frac{14}{9} \phi_c^2 + \frac{18 r_\xi - 11 r_\phi}{6 r_\xi r_\phi} r \right] + \quad (2.61)$$

$$- \frac{G_N^3 M^3}{2 c^4 r^3} \left[\frac{50 r_\xi - 7 r_\phi}{12 r_\phi r_\xi} \phi_c r + \frac{16 \phi_c^3}{27} - \frac{r^2 (2 r_\xi^2 - r_\phi^2)}{r_\xi^2 r_\phi^2} \right],$$

which can be expressed as the sum of three contributions

$$\Phi(r) = \Phi_0(r) + \Phi_1(r) + \Phi_2(r) \quad (2.62)$$

where

$$\Phi_0(r) = -\frac{G_N M \phi_c}{r} \quad (2.63)$$

$$\Phi_1(r) = \frac{G_N^2 M^2}{2 c^2 r^2} \left[\frac{14}{9} \phi_c^2 + \frac{18 r_\xi - 11 r_\phi}{6 r_\xi r_\phi} r \right] \quad (2.64)$$

$$\Phi_2(r) = -\frac{G_N^3 M^3}{2 c^4 r^3} \left[\frac{50 r_\xi - 7 r_\phi}{12 r_\phi r_\xi} \phi_c r + \frac{16 \phi_c^3}{27} - \frac{r^2 (2 r_\xi^2 - r_\phi^2)}{r_\xi^2 r_\phi^2} \right] \quad (2.65)$$

The solutions of the non-local field equations (2.56) and (2.57) show a modification of the general relativistic predictions which depends on two characteristic lengths r_ϕ and r_ξ . Those length scales characterize the regime where the non-local effects become significant and one can wonder whether they are related to the mass of the gravitational systems under investigation. Furthermore, these lengths may play a crucial role in recovering the predictions of GR, since they fix the magnitude of the non-local corrections.

It is indeed expected that the effects of non-local gravity become unimportant at Solar System scale, and more generally at all scales where GR has been probed. Deser and Woodard proposed in [34] a so called ‘‘screening mechanism’’ based on the idea that the term $\square^{-1}R$ has different sign when evaluated in the cosmological framework or for bound systems such as the Solar System. They argued that the distortion function could be suitably set to zero at the scales of bound systems, while still being relevant cosmologically. A subsequent study reported in [35] showed a critical flaw in this argument, proving that there is no such sign change, hence the screening mechanism remains an open problem of the DW model.

Let us note that the problem of the screening mechanism also affects other theories such $f(R)$ theories, where a chameleon mechanism [36] is usually employed. The crucial feature of this approach is the introduction of a scalar field coupled to matter; hence this scenario is particularly interesting in Non-local Gravity where the localization procedure allows to encapsulate the non-local contributions in two scalar fields which may depend on the matter content of different astrophysical systems.

An example of such a screening mechanism is the so called Vainshtein mechanism [37]. This paradigm introduces a typical length scale (Vainshtein radius) such that the contributions of some of the degrees of freedom of the theory are suppressed for distances lower than this radius. In this way non-local contributions become unimportant in the case of the Solar System where GR has been extensively tested. It is important to notice that the

Vainshtein mechanism involves non linear effects, as well as the fact that the Vainshtein radius depends on the type of gravitational source and model.

The DW theory has been already tested at different astrophysical scales. For instance, in [38] this model has been tested in the case of elliptical galaxies without assuming the existence of Dark Matter. The authors were able to identify constraints on the parameters of the non-local theory and recover the so called fundamental plane of the elliptical galaxies, an empirical relation between the central velocity dispersion σ_0 , the effective radius R_e , and the mean effective surface brightness I_e . It was found that the best agreement between the theoretical predictions and the observed data was obtained for values of r_ϕ and r_ξ between 0.001 kpc and 0.01 kpc.

The Deser-Woodard theory has also been tested at galaxy cluster scale [3], by comparing the gravitational lensing observations from the CLASH program with the theoretical predictions of the DW model. In this case the clusters of galaxies contain Dark Matter and the statistical analysis aimed at constraining the parameters of the density profile of Dark Matter as well as the non-local length scales of the model. The authors were able to set lower bounds on the two non-local lengths $r_\phi > 10^{-5}$ kpc and $r_\xi > 10^{-5}$ kpc.

Finally, it is worth mentioning a test of DW model at the scale of the star S2 [2] that orbits the Galactic center. In particular, the orbit of S2 has been simulated for the DW gravitational potential and compared to the observations by New Technology Telescope and Very Large Telescope. It was found that for values of $r_\xi > 10^{-10}$ kpc and 10^{-9} kpc $< r_\phi < 10^{-6}$ kpc the Deser-Woodard theory can reproduce exactly the orbit of S2.

In all those cases, the authors tested the reliability of the Deser-Woodard theory at different astrophysical scales.

Chapter 3

Testing Non-local Gravity at galactic scales

In this chapter I will present the general features of the Ultra-Diffuse galaxies, analyzing in detail the modelling of the Dark Matter and stellar matter density profile that will be used for the statistical analysis. In section (3.2) I will generalize the Deser-Woodard potential (2.61) to the extended mass distribution of the Ultra-Diffuse galaxy, while in the last section I will derive the dispersion velocity σ which will act as the kinematical observable in the statistical analysis.

3.1 Ultra-Diffuse Galaxies

Ultra-Diffuse galaxies (UDG) can be considered as a special type of dwarf spheroidal and dwarf elliptical galaxies [39] characterized by an extremely low surface brightness $\mu_0 = 24-26$ mag/arcsec⁻² (in the g photometric band) and half-light radii $R_{eff} = 1.5-4.6$ kpc. In particular, their size is similar to spiral galaxies such as the Milky Way, while their luminosity is almost indistinguishable from elliptical galaxies. Another striking feature of some UDGs is the high number of globular clusters compared to their luminosity, usually from 5 to 7 times more than other galaxies with the same luminosity. Furthermore, UDGs have been observed in different types of environments, from galaxy groups to voids, and this poses some problems for their formation process. It is possible that UDGs formed through the process of tidal stripping [40] or from Dark Matter halos characterized by high angular momentum [41]. Another interesting formation scenario was proposed in [42] where UDGs would have formed about 12 billion years ago, and that a burst of star formation induced by ram pressure between 200 million and 1 billion years ago would lead to a halt in the process of star formation.

In recent years there is a growing interest toward UDGs because they exhibit a wide range of features in terms of their Dark Matter content. In fact, observations from the Dragonfly Array Telescope suggest that some of those galaxies are dominated almost entirely by Dark Matter (DragonFly44), while others seem to completely lack any Dark Matter contribution (NGC1052-DF2 and NGC1052-DF4). This feature raises a number of questions regarding the phenomenon of structure formation, i.e. the process by which small densities of Dark Matter overcome the cosmological expansion, and clump up to form Dark Matter halos which serve as sites of formation. Those properties make UDGs the ideal testbed for both the hypothesis of Dark Matter and extended theories of gravity.

Let us consider two specific examples of UDGs: DragonFly44 and NGC1052-DF2. Those galaxies have similar contents of baryonic matter but their Dark Matter content is completely different.

The galaxy DragonFly44 (DF44) belongs to the Coma cluster at a distance of ~ 100 Mpc and some observations in 2016 indicated the presence of a population of approximately 100 globular clusters (although observations in subsequent years suggest that the number may be overestimated [43]). This number of globular clusters points toward a Dark Matter halo of mass $\sim 10^{11}M_{\odot}$, corresponding to 99% of the total mass of the galaxy.

NGC1052-DF2, on the other hand, sits at the opposite end of the spectrum. Observed in 2018, this galaxy in the constellation Cetus seems to contain very low amount of Dark Matter. In fact, the dynamics of this galaxy can be successfully explained by taking into account only the baryonic matter.

Those two galaxies have also been analyzed in other gravitational scenarios, such as Modified Newtonian Dynamics (MOND) without assuming Dark Matter [44] or Degenerate Higher-Order Scalar Tensor (DHOST) theories [39].

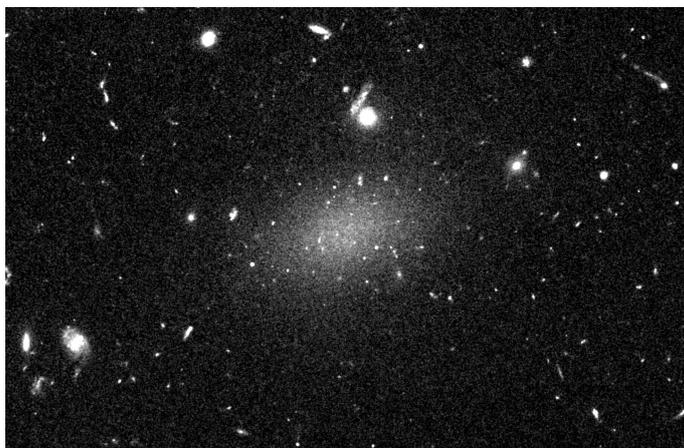


Figure 3.1: Ultra-Diffuse galaxy DF44, credits: Teymoor Saifollahi and NASA/HST (HST Proposal 14643, PI: van Dokkum)

3.1.1 Dark Matter density profile

There are many astrophysical and cosmological observations that support the existence of a new type of matter known as Dark Matter (DM). The best indication toward the existence of Dark Matter comes from the rotation curves of spiral galaxies [45]. In this case, it is easy to obtain a relation between the rotational velocity of objects that are gravitationally bound to the galaxy and their distance from the galactic centre.

It turns out that the best tracer for Dark Matter is indeed represented by stars, since the rate of collision between stars is extremely low, hence their motion is determined primarily by the gravitational interaction. Applying Gauss' Law to a galaxy results in the following expression for the velocity of the stars:

$$v_c(r) = \sqrt{\frac{G_N M(r)}{r}}, \quad (3.1)$$

where $M(r)$ is the mass contained within the radius r and G_N is Newton's gravitational constant. This result implies that in the outskirts of a galaxy the velocity is expected to decay as $v \propto r^{-1/2}$, since in this region the mass M does not increase significantly.

On the other hand, the observations carried out by Vera Rubin [46] on the rotation curves of spiral galaxies, showed that the velocity decreases very slowly with the radius (Figure 3.2), not as expected by (3.1).

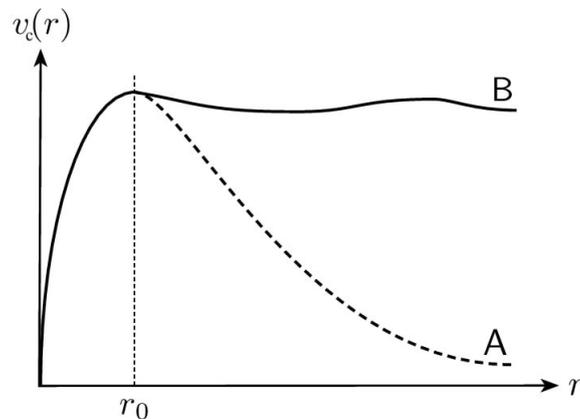


Figure 3.2: expected (A) and observed (B) profile of the rotation curve, r_0 is the radius from which the mass of the galaxy does not increase significantly [47].

From (Figure 3.2) it is clear that there is a radius r_0 from which the rotation curve becomes almost flat.

$$v_c(r) = v_c(r_0) = \text{const} \quad \text{for } r > r_0. \quad (3.2)$$

Additional studies and observations showed that the flatness in the rotation curve could not be accounted for by modifying the relative weight of the diverse galactic components,

rather it could be explained by adding a new matter component with a different spatial distribution.

Equation (3.1) evidently shows that the flatness of the rotation curve can be obtained if $M(r) \propto r$, which is exactly the mass distribution expected for a self-gravitational gas of non-interacting particles. Hence the rotation curves can be used to infer the Dark Matter density

$$\rho(r) \propto \frac{M(r)}{r^3} \sim \frac{1}{r^2}. \quad (3.3)$$

This distribution is ascribed to Dark Matter, more precisely to Dark Matter halos that surround the galaxy. The implicit assumption made here is that DM is distributed in a spherically symmetric halo about the center of the galaxy, in contrast to baryonic matter that is concentrated in the disk. This implies that DM is not dissipative, while baryonic matter dissipates energy through self interaction and thus collapse into a disk.

In general, it is expected that Dark Matter is made up of electrically neutral particles, otherwise they would scatter light and an electromagnetic effect would have been observed.

It is also crucial to mention the importance of numerical simulations in order to study the structure formation in the Early Universe, when Dark Matter is expected to play a significant role. In this context, it was found that Dark Matter has to be non-relativistic (cold) at the epoch of structure formation.

There is also evidence that Dark Matter is non baryonic from the studies of the Cosmic Microwave Background (CMB) in Cosmology [48]. In fact, the best piece of evidence for the existence of Dark Matter comes from the CMB, the residual background photons from the Last Scattering Surface, at a temperature of roughly $2.7K$. Around this mean temperature, there are very small fluctuations ($\Delta T/T \sim 10^{-4/5}$) that contain information about the amount of matter and energy in the Universe. They suggest that only 4 – 5% of the total energy budget of the Universe is made out of ordinary, baryonic matter. The abundance of Dark Matter is instead estimated around 27% of total energy budget, the remaining being associated to the Dark Energy.

Finally, Dark Matter has to be almost collisionless, since observations of cluster collisions allow to set an upper bound to the self-interaction of Dark Matter particles which is extremely low.

The Dark Matter halo is considered a key constituent of most galaxies. The halo is usually described in terms of the so called Navarro-Frenk-White profile [49]:

$$\rho_{gNFW}(r) = \frac{\rho_s}{\frac{r}{r_s} \left(1 + \frac{r}{r_s}\right)^2}, \quad (3.4)$$

which is characterized by a length scale r_s and a density scale ρ_s specific of the galaxy under investigation. This profile is obtained from numerical simulation of structure formation and can be applied to a wide range of galaxies, showing that the Dark Matter density scales as $1/r$ for $r < r_s$, and $1/r^3$ for $r > r_s$.

In the case of Ultra-Diffuse galaxies, it is interesting to consider a generalized Navarro-Frenk-White (gNFW) profile [39] for the Dark Matter density:

$$\rho_{gNFW}(r) = \rho_s \left(\frac{r}{r_s} \right)^{-\gamma} \left(1 + \frac{r}{r_s} \right)^{\gamma-3}, \quad (3.5)$$

where γ is an additional parameter related to the slope of the profile near the maximum. This model is a generalization of the NFW profile, which is recovered when $\gamma = 1$.

In this thesis we are going to consider a reparametrized version of (3.5), replacing the two scale parameters $\{\rho_s, r_s\}$ with the concentration parameter c_{200} and the virial mass M_{200}

$$c_{200} = \frac{r_{200}}{r_s} \quad \text{and} \quad M_{200} = \frac{4\pi}{3} 200 \rho_c r_{200}^3, \quad (3.6)$$

where ρ_c is the critical density of the Universe

$$\rho_c = \frac{3H^2(z)}{8\pi G_N}, \quad (3.7)$$

written in terms of the Hubble parameter $H(z)$ as a function of the redshift z which for DF44 corresponds to $z = 0.023156$ (the redshift of the Coma cluster). Notice also that M_{200} corresponds to the mass enclosed in a sphere of radius r_{200} with mean density equal to 200 times ρ_c . Finally, the density scale ρ_s depends on c_{200} as

$$\rho_s = \frac{200}{3} \rho_c(z) \frac{(3-\gamma)(c_{200})^\gamma}{{}_2F_1[3-\gamma, 3-\gamma, 4-\gamma, -c_{200}]} \quad (3.8)$$

where ${}_2F_1(a, b, c, z)$ is the hypergeometric function.

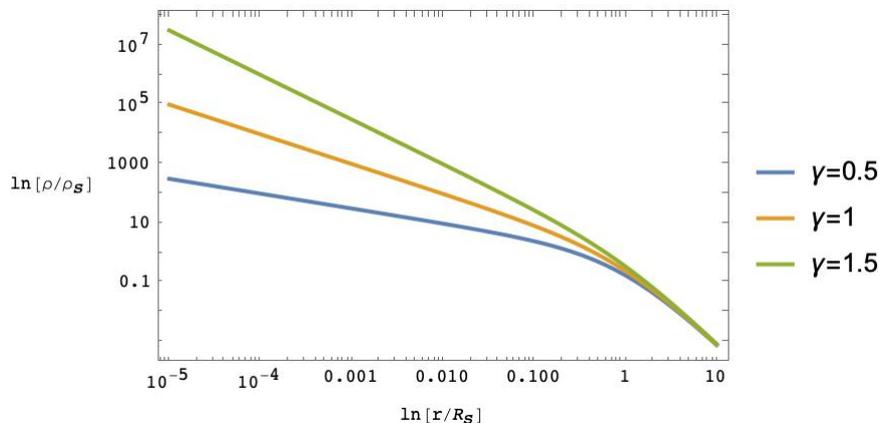


Figure 3.3: Generalized NFW profile for different values of γ in the range $r/R_s \in [10^{-5}, 10]$

3.1.2 Stellar density profile

The intensity of a UDG can be modeled using a single Sérsic profile [50], which relates the intensity of the galaxy to the distance from its center R

$$I(R) = I_0 \exp \left[- \left(\frac{R}{a_s} \right)^{1/n} \right], \quad (3.9)$$

where I_0 is the intensity at $R = 0$, a_s is a length scale known as the Sérsic parameter and n is the so called Sérsic index, which is related to the shape of the density profile.

The Sérsic parameter is usually written in terms of the half-to-light radius R_{eff} , i.e. the radius of the sphere that contains half of the luminous flux of the galaxy, as

$$a_s(D) = \frac{R_{eff}(D)}{(2n - 0.33)^n}. \quad (3.10)$$

In this thesis we are interested in the Ultra-Diffuse galaxy DF44 which is similar to an elliptical galaxy, hence the half-to-light-radius has to be modified in order to take into account this feature. Given the semi-minor axis b and semi-major axis a of the galaxy, the half-to-light radius becomes

$$R'_{eff} = R_{eff} \sqrt{\frac{b}{a}}. \quad (3.11)$$

Once the intensity $I(R)$ is known, the luminosity density profile $l(r)$ can be obtained in principle by deprojection. However, there is no exact expression of $l(r)$ for a generic Sérsic profile, rather the following analytical approximation holds

$$l(r, D) = l_1(D) \tilde{l} \left(\frac{r}{a_s(D)} \right), \quad (3.12)$$

where $\tilde{l}(r/a_s)$ is the dimensionless function

$$\tilde{l} \left(\frac{r}{a_s} \right) = \left(\frac{r}{a_s} \right)^{p_n} \exp \left[- \left(\frac{r}{a_s} \right)^{1/n} \right] \quad (3.13)$$

and $p_n \approx 1 - 0.6097/n + 0.05463/n^2$ is a function of the Sérsic index. In addition, l_1 also depends on n as

$$l_1(D) = \frac{L_{tot}(D)}{4\pi n \Gamma(3 - p_n) n a_s^3(D)} \quad (3.14)$$

where $\Gamma(3 - p_n)$ is the Gamma function and L_{tot} is the total luminosity of the galaxy.

It is worthwhile to mention that the total luminosity has to be considered as a function of the distance

$$L_{tot}(D) = 10^{-0.4(m_{V606} - \mu(D) - M_{\odot, V606})}, \quad (3.15)$$

if $m_{V_{606}}$ is the apparent magnitude of the galaxy, $M_{\odot, V_{606}}$ is the absolute magnitude of the Sun in the V_{606} photometric band and $\mu(D)$ is the distance modulus

$$\mu(D) = 5 \log_{10} D + 25, \quad (3.16)$$

where the distance D is measured in Mpc.

In conclusion, the profile density for the stellar component ρ_* can be obtained by multiplying the luminosity density profile by the mass-to-light ratio Υ_* of the galaxy

$$\rho_*(r, D) = \Upsilon_* l(r, D). \quad (3.17)$$

In the specific case of the galaxy DF44 the Sérsic index is $n = 0.94$, the axis ratio is $b/a = 0.68$ and the central surface brightness is $\mu_0 = 24.1 \text{ mag arcsec}^{-2}$.

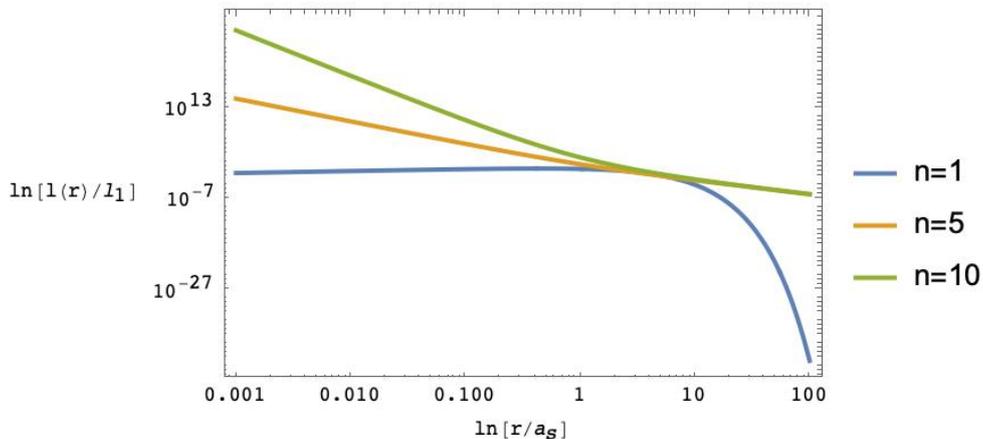


Figure 3.4: Sérsic profile for different values of the index n in the range $r/a_s \in [10^{-3}, 100]$

3.2 Non-local corrections to the Newtonian potential

The gravitational potential (2.61) is valid only for point-like mass distributions and cannot be employed to describe extended objects such as Ultra-Diffuse galaxies. Furthermore, Gauss' theorem does not hold in Non-local gravity, thus the expression (2.61) has to be modified in order to take into account the mass distribution of the galaxy. Let us distinguish in equation (2.61) three components in terms of their dependence on $G_N M$:

$$\Phi_0(r) = -\frac{G_N M}{r}, \quad (3.18)$$

$$\Phi_1(r) = \frac{G_N^2 M^2}{2 c^2 r^2} \left(\frac{14}{9} + \frac{18 r_\xi - 11 r_\phi}{6 r_\xi r_\phi} r \right), \quad (3.19)$$

$$\Phi_2(r) = -\frac{G_N^3 M^3}{2 c^4 r^3} \left[\frac{50 r_\xi - 7 r_\phi}{12 r_\phi r_\xi} r + \frac{16}{27} - \frac{r^2 (2 r_\xi^2 - r_\phi^2)}{r_\xi^2 r_\phi^2} \right]. \quad (3.20)$$

Considering the order of magnitude of the mass of the galaxy DF44 of $\sim 10^{11}M_{\odot}$ and the typical distances \sim kpc of the globular clusters with respect to the centre of the galaxy DF44, we can obtain the following estimates for the three contributions to the gravitational potential $\Phi(r)$.

$$\mathcal{O}\left(\frac{G_N M}{r}\right) \sim 10^{-28} \text{ kpc}^2/\text{s}^2, \quad (3.21)$$

$$\mathcal{O}\left(\frac{G_N^2 M^2}{2 c^2 r^2}\right) \sim 10^{-33} \text{ kpc}^2/\text{s}^2, \quad (3.22)$$

$$\mathcal{O}\left(\frac{G_N^3 M^3}{2 c^4 r^3}\right) \sim 10^{-39} \text{ kpc}^2/\text{s}^2. \quad (3.23)$$

For the purposes of this thesis, we are going to retain only the first two contributions and neglect the third, given the difference in order of magnitude.

Since Gauss' theorem is not valid, the gravitational potential at a given point r on a spherical surface is not just proportional to the mass enclosed by that surface. Rather, both the matter inside and outside the spherical surface contribute. This observation implies that the extended gravitational potential can be obtained by integrating over the entire mass distribution of the galaxy, hereafter indicated as $\rho(r')$, where the primed variables (r', θ', φ') identify the points where the mass distribution is located.

In this approach, an extended mass distribution is treated as an ensemble of infinitesimal mass element dM which contribute as point-like masses

$$dM = \rho(r') r'^2 dr' \sin \theta' d\theta' d\varphi'. \quad (3.24)$$

It is worthwhile to notice that the non-local contribution to the potential Φ_0 is quadratic in the mass M , hence the infinitesimal contribution is proportional to $2M(r')dM$, where the mass M at a point r' is

$$M(r') = 4\pi \int_0^{r'} dr'' r''^2 \rho(r''), \quad \text{for } r' < r \quad (3.25)$$

$$M(r') = 4\pi \int_{r'}^{\infty} dr'' r''^2 \rho(r'') \quad \text{for } r' > r \quad (3.26)$$

while in the second expression we are assuming that the mass distribution extends up to ∞ .

Finally, the radius r that appears in Φ_0 and Φ_1 has to be substituted by $|\mathbf{r} - \mathbf{r}'|$ (where \mathbf{r} specifies the point where we evaluate the gravitational potential and \mathbf{r}' specifies a point where the mass distribution is nonzero) in order to take into account every contribution $M(r')$. If we assume spherical symmetry, the term $|\mathbf{r} - \mathbf{r}'|$ can be expressed as

$$|\mathbf{r} - \mathbf{r}'| = (r^2 + r'^2 - 2rr' \cos \theta')^{1/2}, \quad (3.27)$$

since it is possible to choose the azimuthal angles φ and φ' to be equal.

The extended gravitational potential are given by

$$\begin{aligned} \Phi_0^{ext}(r) = & \int_0^r dr' \rho(r') r'^2 \int_0^\pi d\theta' \sin \theta' \int_0^{2\pi} d\varphi' \Phi_0(\mathbf{r} - \mathbf{r}') + \\ & + \int_r^\infty dr' \rho(r') r'^2 \int_0^\pi d\theta' \sin \theta' \int_0^{2\pi} d\varphi' \Phi_0(\mathbf{r} - \mathbf{r}'), \end{aligned} \quad (3.28)$$

$$\begin{aligned} \Phi_1^{ext}(r) = & 8\pi \left[\int_0^r dr' \rho(r') r'^2 \int_0^{r'} dr'' \rho(r'') r''^2 \int_0^\pi d\theta' \sin \theta' \int_0^{2\pi} d\varphi' \Phi_1(\mathbf{r} - \mathbf{r}') + \right. \\ & \left. + \int_r^\infty dr' \rho(r') r'^2 \int_{r'}^\infty dr'' \rho(r'') r''^2 \int_0^\pi d\theta' \sin \theta' \int_0^{2\pi} d\varphi' \Phi_1(\mathbf{r} - \mathbf{r}') \right]. \end{aligned} \quad (3.29)$$

Once the density profiles of the Dark Matter and stellar component are chosen, the extended integrals can be computed for all the cases of interest using Wolfram Mathematica. It should be noted that some of the extended potentials do not have an analytic form, thus they have to be solved through tabular methods in order to obtain differentiable expressions that can be used to determine the velocity dispersion in the last part of the thesis. The tabular method is based on the idea of decomposing the space of the parameters on which the integral depends in a very high number of points, and evaluating the integral numerically for each one of those points. Then, it is possible to use Wolfram Mathematica to obtain a smooth interpolation of those points.

One last remark is necessary about the integral over the external mass distribution from r (or r') to ∞ . It is reasonable to believe that the mass distribution extends up to a finite distance, rather than infinity, hence we have set a cutoff at 50 kpc based on the properties of the galaxy DF44. In fact, 50 kpc is almost half of the distance between DF44 and the centre of the Coma cluster to which it belongs and more than twice the estimated size of the Milky Way's Dark Matter halo.

The extension of the Newtonian potential Φ_0 has an analytic expression which for the generalized NFW profile reads:

$$\begin{aligned} \Phi_{0,DM}^{ext}(r, \rho_s, r_s, \gamma) = & 4\pi G_N \rho_s r_s^2 \left(\frac{1 - r^{2-\gamma} (r + r_s)^{-2+\gamma}}{\gamma - 2} + \right. \\ & \left. + \frac{(-1)^\gamma r_s \text{Beta}[-r/r_s, 3 - \gamma, \gamma - 2]}{r} \right), \end{aligned} \quad (3.30)$$

where Beta is the incomplete beta function, a generalization of Euler Beta function. In the case of the stellar matter distribution the extended potential becomes:

$$\begin{aligned} \Phi_{0,STAR}^{ext}(r, D) = & l_1(D)2\pi G_N \left(0.001 D^2 + 0.27 D^{0.4} r^{3/2} {}_1F_1(1.49, 2.49, -49.2 r/D) + \right. \\ & \left. + 0.17 D^{0.4} r^{3/2} {}_1F_1(2.43, 3.43, -49.2 r/D) \right) \end{aligned} \quad (3.31)$$

where ${}_1F_1$ is the confluent hypergeometric function.

3.3 Kinematics of Ultra-Diffuse Galaxies

In order to test the compatibility of the Deser-Woodard model with the observational data from DF44, we will study the so called radial velocity dispersion σ_r . This quantity characterizes the kinematic of the galaxy and provides information on the statistical dispersion of radial velocity of globular clusters in DF44 about the mean [39].

Assuming spherical symmetry and dynamical equilibrium, the radial dispersion velocity can be obtained from the Jeans equation

$$\frac{d(l(r) \sigma_r^2)}{dr} + \frac{\beta(r)}{r} l(r) \sigma_r^2 = l(r) \frac{d\Phi(r)}{dr}, \quad (3.32)$$

where the anisotropy parameter $\beta(r)$ measures the radial or tangential behavior of the globular clusters' orbits in DF44

$$\beta(r) = 1 - \frac{\sigma_t^2(r)}{\sigma_r^2(r)}, \quad (3.33)$$

such that the system is completely isotropic if $\beta(r) = 0$, completely tangential if $\beta(r) \rightarrow -\infty$ and purely radial if $\beta(r) = 1$. Notice also that the right-hand side of (3.32) contains the gravitational potential $\Phi(r)$ of the theory under investigation.

The general solution for the Jeans equation in the framework of GR is given by

$$\sigma_r^2(r) = \frac{1}{l(r)f(r)} \int_r^\infty ds f(s) l(s) \frac{M(s)}{s^2}, \quad (3.34)$$

where $f(r)$ is a function of the anisotropy parameter $\beta(r)$.

It is crucial to understand that the observed quantity is not σ_r^2 , rather its projection along the line-of-sight, i.e. the direction between the observer and the observed object, defined as follows:

$$\sigma_{los}^2(R) = \frac{2}{I(R)} \left[\int_R^\infty dr r \frac{l(r) \sigma_r^2}{\sqrt{r^2 - R^2}} - R^2 \int_R^\infty dr \beta(r) \frac{l(r) \sigma_r^2}{r \sqrt{r^2 - R^2}} \right], \quad (3.35)$$

which depends on the 2D projected radius R and the surface brightness of the galaxy $I(R)$. For the purposes of the thesis, it is helpful to use a more compact definition of σ_{los}^2

$$\sigma_{los}^2(R) = \frac{2G_N}{I(R)} \int_R^\infty dr K\left(\frac{r}{R}\right) l(r) \frac{M(r)}{r}, \quad (3.36)$$

where the Kernel function $K(r/R)$ depends on the parametrization of the anisotropy parameter $\beta(r)$ [50].

Once the general form of the radial dispersion velocity along the line-of-sight in GR is known, we can generalize this result to a different gravity theory such as the Deser-Woodard model. This can be achieved by recasting the equation for the derivative of the DW potential as

$$\frac{d\Phi}{dr} = \frac{G_N M_{eff}(r)}{r^2} \quad (3.37)$$

where M_{eff} is an effective mass that contains the non-local contributions, hence it is also a function of the two characteristic non-local lengths r_ϕ and r_ξ .

The final expression for $\sigma_{los}^2(R)$ becomes

$$\sigma_{los}^2(R) = \frac{2G_N}{I(R)} \int_R^\infty dr K\left(\frac{r}{R}\right) l(r) \frac{M_{eff}(r)}{r}. \quad (3.38)$$

Since it is impossible to effectively measure the anisotropy profile $\beta(r)$, in the statistical analysis that follows we will choose a particular radial anisotropy profile:

$$\beta(r) = \beta_0 + (\beta_\infty - \beta_0) \frac{r}{r + r_a}, \quad (3.39)$$

where β_0 and β_∞ characterize the profile in the limits for $r \rightarrow 0$ and $r \rightarrow \infty$, while r_a is a scale parameter.

Chapter 4

The statistical analysis: Methods and Results

In this chapter the attention is focused on testing the Deser-Woodard model in the case of the Ultra-Diffuse galaxy DF44.

First, I will outline the general features of Bayesian statistics as a fundamental tool for parameters estimation and model comparison. In addition, I will discuss the implementation of those methods through a particular Markov Chain Monte Carlo algorithm known as the Metropolis-Hastings algorithm and present the results of the statistical analysis.

4.1 Bayesian inference

The measurement process in Physics is strictly related to the idea of statistical inference. Once the data have been recorded, the problem is to determine the probabilities of each of the causes which may have produced the phenomenon under investigation.

In the framework of Bayesian statistics [51], where the concept of probability is associated to a measure of the degree of belief that an event will occur, the inference is based on Bayes' theorem.

Let us indicate the data as d , the collection of parameters of the model under study as $\theta = \{\theta_1, \dots, \theta_m\}$ and any other information on the model considered as I .

Bayes' theorem states that the conditional probability $p(\theta|d, I)$ of the parameters θ , given the data d and the model I is

$$p(\theta|d, I) = \frac{p(d|\theta, I) p(\theta|I)}{p(d|I)}, \quad (4.1)$$

where $p(\theta|d, I)$ is the posterior distribution (or posterior), $p(\theta|I)$ is the prior distribution (or prior), $p(d|\theta, I)$ is the likelihood and $p(d|I)$ is a normalization factor called evidence.

The two crucial elements for statistical inference are: the prior $p(\theta|I)$, which corresponds to the probability distribution of the parameters θ before the knowledge of the data d ; the posterior $p(\theta|d, I)$, which is the probability distribution of the parameters after gaining the knowledge of the data d . Thus, Bayes' theorem shows that the probability of a given hypothesis (in this case the parameters θ) has to be updated when new information is available. An essential feature of Bayes' theorem is that the posterior converges to a probability distribution which is independent of the prior as large data sets become available.

The initial choice of the prior is done by the experimentalist, not by the theory, and has to reflect the current state of knowledge on the parameters of the model. For instance, if some parameters have physical meaning that constraints their possible values (e.g. mass is positive quantity) it is possible to encode this information in the prior.

A standard choice for the prior is a uniform distribution in the parameters θ on their domain Ω_θ

$$p(\theta|d) = \begin{cases} \text{constant}, & \text{for } \theta \in \Omega_\theta \\ 0 & \text{otherwise.} \end{cases} \quad (4.2)$$

It is worthwhile to notice that under a change of variable $Y = Y(\theta)$ the prior becomes

$$p(Y|d) = p(\theta|d) \left| \det \left(\frac{\partial \theta}{\partial Y} \right) \right|, \quad (4.3)$$

which is not flat if the change of variable is non-linear, a feature that can be problematic for a high dimensional parameter space. This method was proposed by Laplace in 1812 and aims at giving equal probability to all possible alternatives (principle of indifference).

As far as the likelihood $p(d|\theta, I)$ is concerned, it is usually constructed based on how the data are collected. Furthermore, from equation (4.1) it is clear that if asymptotically the posterior probability is independent of the prior,

$$p(\theta|d, I) \propto p(d|\theta, I), \quad (4.4)$$

and the posterior coincides with the likelihood. Thus the maximum likelihood method is recovered in Bayesian statistics if the mean value of $p(\theta|d, I)$ is equal to its maximum. The form of the likelihood function also provides a criterion for selecting the prior, in particular it is possible to choose the prior such that the prior and the likelihood are conjugate, which means that the posterior and the prior belong to the same distributional family. For example the gaussian distribution has a conjugate which coincide with itself (selfconjugate), while the conjugate prior to the Poisson distribution is the Gamma distribution.

Once the posterior has been calculated and more information is available, the posterior becomes the prior and Bayes' theorem can be applied to determine the updated posterior distribution. The strength of the posterior, namely how much the probability distribution of the model parameters is determined by the data, can be checked by carrying out a sensitivity analysis, i.e. by changing the prior in a consistent way and showing that the posterior does not change.

The last ingredient of the Bayes' theorem is the Bayesian evidence, which constitutes an essential tool for comparing different models as well as evaluating the performance of a given model [52]. The above-mentioned evidence is the normalization factor of the posterior distribution, thus it can be obtained by integrating the posterior over the range of the parameters Ω_θ

$$p(d|I) = \int_{\Omega_\theta} p(d|\theta, I) p(\theta|I) d\theta. \quad (4.5)$$

Clearly the evidence only depends on the data d and the model I under investigation. Notice that the posterior distribution is usually not known in a closed form, hence the evidence must be computed using numerical methods based on the Markov Chain Monte Carlo approach such as the nested sampling, that allows to express the multi-dimensional integral (4.5) as a one dimensional integral.

The role of the evidence in model comparison can be immediately understood in light of Bayes' theorem, which allows to express the probability of the model I given the data d from the evidence $p(d|I)$ by inverting the order of conditioning:

$$p(I|d) \propto p(d|I) p(I), \quad (4.6)$$

where $p(I)$ is the probability of the model I . Given a set of data d , it is possible to compare two models I_i and I_j by taking the ratio of the probabilities of the two models obtained from equation (4.6):

$$\frac{p(I_i|d)}{p(I_j|d)} = \mathcal{B}_j^i \frac{p(I_i)}{p(I_j)}, \quad (4.7)$$

defining the Bayes factor \mathcal{B}_j^i as the ratio of the evidences of two models

$$\mathcal{B}_j^i = \frac{p(d|I_i)}{p(d|I_j)}. \quad (4.8)$$

If the Bayes factor is greater (lower) than 1 then the model I_i is favoured (disfavoured) over the model I_j by the data. In order to assess the strength of the evidence, the Bayes factor is usually compared to an empirical scale known as the Jeffreys' scale [53] (Table 4.1). The values of $|\ln \mathcal{B}_j^i| = 1.0, 2.5, 5.0$ correspond respectively to odds of $\sim 3 : 1, 12 : 1, 150 : 1$.

Table 4.1: Jeffreys' scale.

Strength of evidence	
$\ln \mathcal{B}_j^i < 0$	the reference model I_j is favoured
$0 < \ln \mathcal{B}_j^i < 1.0$	inconclusive evidence for model I_i or I_j
$1.0 < \ln \mathcal{B}_j^i < 2.5$	substantial evidence for model I_i
$2.50 < \ln \mathcal{B}_j^i < 5.0$	strong evidence for model I_i
$\ln \mathcal{B}_j^i > 5.0$	decisive evidence for model I_i

It is interesting to notice that Bayesian statistics naturally embodies an Occam's razor effect, the idea that if two hypotheses are equally successful in explaining the observed data then the simpler hypothesis should always be favoured.

In light of equation (4.8) it is possible to notice that the Bayes factor, which depends on the data, reflects this particular criterion. In fact, let us assume that the model I_j is more complex (meaning that it has more parameters than I_i or a wider parameter range) than the model I_i . It is clear that complex models are naturally more flexible, i.e. they can accommodate a larger range of possible data by adjusting or tuning the parameters, which implies that the probability profile $p(d|I_j)$ is much broader than the probability $p(d|I_i)$.

This argument shows that if the model I_i fits the data sufficiently well, the corresponding probability $p(I_i|d)$ will be greater than the probability $p(I_j|d)$ associated to the more complicated model. In conclusion, if we assume that the prior probabilities of the two models $p(I_i)$ and $p(I_j)$ are very similar, the simpler model I_i will always be favoured because $p(d|I_i) > p(d|I_j)$.

4.2 Markov Chain Monte Carlo

The name Markov Chain Monte Carlo (MCMC) refers to a class of algorithms used in statistics to sample from a target probability distribution [54]. Those techniques are extremely useful in Bayesian inference, where the objective is to obtain information on the unknown posterior distribution.

A Markov chain can be thought of as a collection of random variables $\{\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(n)}\}$ that satisfies the so called Markov property, that is, the probability of the $(n+1)$ -th element only depends on the value of the n -th element. Notice that each of the ele-

ments $\theta^{(n)}$ is a realization of a random variable θ which is usually multi-dimensional, i.e. $\theta = \{\theta_1, \dots, \theta_m\}$. Furthermore, the elements of the chain are generated in a probabilistic fashion, with a transition probability $T(\theta^{(n)}, \theta^{(n+1)})$ of moving from the element $\theta^{(n)}$ to $\theta^{(n+1)}$ of the chain.

For the purposes of this thesis, the random variable θ contains the parameters of the gravitational model under study. The probability distribution of those parameters can be effectively sampled by Markov chains because of the following condition: a Markov chain converges to a stationary state where the elements of the chain are random variables distributed according to the target distribution (in our case the posterior distribution $p(\theta|d)$), if the chain is irreducible and has no cycles. Irreducible means that there is a non zero probability of moving between any two elements in a finite number of steps, while a Markov chain has a cycle when it reaches a given element iteratively.

A Markov chain can be obtained by means of the following sufficient condition known as the detailed balance condition:

$$p(\theta^{(n)}|d) T(\theta^{(n)}, \theta^{(n+1)}) = p(\theta^{(n+1)}|d) T(\theta^{(n+1)}, \theta^{(n)}), \quad (4.9)$$

which can also be written as an equality between the ratio of the posterior and transition probabilities

$$\frac{T(\theta^{(n)}, \theta^{(n+1)})}{T(\theta^{(n+1)}, \theta^{(n)})} = \frac{p(\theta^{(n+1)}|d)}{p(\theta^{(n)}|d)}. \quad (4.10)$$

The samples from the posterior distribution can be used to estimate the expectation value of any function of the parameters through Monte Carlo techniques, e.g. the posterior mean can be expressed as

$$E[\theta] = \int p(\theta|d) \theta d\theta \approx \frac{1}{N} \sum_{n=0}^{N-1} \theta^{(n)}, \quad (4.11)$$

and the same approach can be applied to any function of the parameters θ .

One of the advantages of the MCMC method is the possibility to evaluate the marginal probability $p(\theta_j|d)$ with respect to a given parameter θ_j

$$p(\theta_j|d) = \int p(\theta|d) \prod_{i \neq j}^m d\theta_i, \quad (4.12)$$

without computing the integral, rather by dividing the range of θ_j in a series of bins and counting the number of samples in each bin. This result is interesting because marginalization is frequently used in Bayesian statistics to deal with nuisance parameters, i.e. parameters that can influence the data but are of no interest for the statistical analysis. In particular, one can infer the joint probability of the parameters of interest and the nuisance parameters, and then integrate over the latter. A feature worth mentioning is

that the computational time and the number of dimensions of the parameter space are related approximately linearly.

Another crucial notion of Bayesian statistics is the so called “credible interval”, which identifies a region that contains a percentage of the probable values. For instance, an $\alpha\%$ credible interval corresponds to the central part of the posterior distribution that contains $\alpha\%$ of the values. This concept is similar to the “confidence region” usually adopted in the frequentistic approach, with the crucial difference that in this case the parameters are fixed and a $\alpha\%$ confidence region means that, once repeated samples have been collected, there is a $\alpha\%$ probability that the confidence region calculated for each one of those samples contains the true value of the parameter.

The MCMC algorithm that will be employed in this thesis is the so called Metropolis-Hastings algorithm and consists of the following steps:

- select a random point $\theta^{(0)}$, associated to a posterior probability $p_0 = p(\theta^{(0)}|d)$, as the first element of the chain;
- consider a proposal distribution $q(\theta^{(i)}, \theta^{(j)})$, which might be a Gaussian of fixed width centered around $\theta^{(i)}$, and draw a candidate point of the chain $\theta^{(c)}$ from the distribution $q(\theta^{(0)}, \theta^{(c)})$;
- the point $\theta^{(c)}$ has an associated probability $p_c = p(\theta^{(c)}|d)$ and is accepted as a new element of the chain with probability

$$\alpha(\theta^{(c)}, \theta^{(0)}) = \min \left(\frac{p_c q(\theta^{(0)}, \theta^{(c)})}{p_0 q(\theta^{(c)}, \theta^{(0)})}, 1 \right), \quad (4.13)$$

which has a simpler form in the original Metropolis algorithm, where the proposal distribution q is symmetric

$$\alpha(\theta^{(c)}, \theta^{(0)}) = \min \left(\frac{p_c}{p_0}, 1 \right); \quad (4.14)$$

- $\theta^{(c)}$ can be accepted or rejected based on the following criterion: generate a random number u from the uniform distribution $[0, 1)$, and accept $\theta^{(c)}$ if $u < \alpha$;
- If the point $\theta^{(c)}$ is accepted, it is added to the chain and the procedure can be repeated substituting $\theta^{(0)}$ with $\theta^{(c)}$ in the first step. If the point $\theta^{(c)}$ is rejected, a copy of the element $\theta^{(0)}$ will be added to the chain and the algorithm can be applied again.

The crucial feature of the Metropolis-Hastings algorithm is that it is not necessary to know exactly the target distribution that we want to sample, rather it is enough to know

a function that is proportional to the target distribution. In our case, we only need the unnormalized posterior, i.e. the product of the likelihood and the prior.

Notice also from (4.14) that the candidate point can be accepted or not, despite $p_c > p_0$, which leads us to the topic of acceptance rate: the probability that a new element is accepted into the chain. It is possible to show that the optimal acceptance rate is approximately 25%. It is clear that the proposal distribution q strongly affects the construction of the Markov chain: if the scale of q is too small compared to the scale of the target distribution then the exploration of the parameter space will be restricted to a local region, while a large scale corresponds to a lower acceptance rate and a slow growth of the chain.

The Metropolis-Hastings algorithm is a general procedure for constructing a Markov chain that converges to the target distribution, but since the chain is started at a random point the initial elements of the chain are not samples of the target distribution and they have to be discarded. In fact, the posterior probability corresponding to the initial points will generally be small and progressively increase as the chain approaches the region where there is a better fit to the data. This initial part of the Markov chain is called “burn in” or “warm up”, and can be identified by studying the evolution of the chain in parameter space as a function of the number of steps. More specifically, the chain starts sampling from the posterior distribution when we observe a flattening of the natural logarithm of the posterior as a function of the number of steps (Figure 4.1).

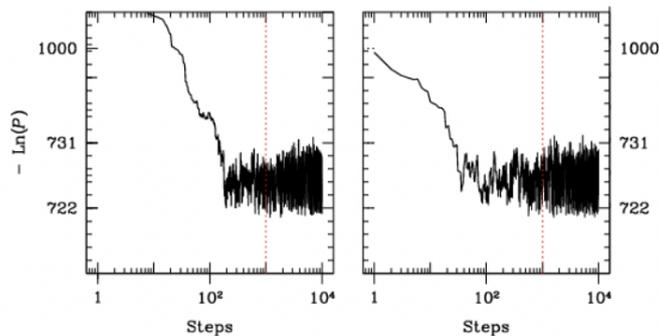


Figure 4.1: Burn in period [52].

Two features of the MCMC algorithms must be highlighted: first, the MCMC algorithm is local and it may remain trapped inside a local region of the target distribution, thus being unable to explore effectively the parameter space; second, the elements of the Markov chain are correlated to one another, which is not a problem in itself for the statistical inference of parameters but in some cases it is useful to have uncorrelated samples. This result can be achieved through the so called thinning technique, which consists of retaining only one element of the chain every $N \sim 10^3$.

4.3 Bayesian analysis of Non-local Gravity at galactic scales

The goal of our statistical analysis is to obtain a sample of the posterior probability distribution $p(\theta|d)$ of the parameters of General Relativity and the Deser-Woodard model in the various scenarios outlined before. This goal can be reached by minimizing a χ^2 employing the Metropolis-Hastings algorithm, implemented through the software Wolfram Mathematica.

The data for DF44 are reported in Table (4.2) [55], where the radial circular velocities of the globular clusters of the galaxy v_i , as well as the corresponding dispersion velocities σ_i are divided into nine bins with their corresponding uncertainties. From those data the χ^2 reads [39]

$$\chi^2(\theta) = \sum_{i=1}^9 \left(\frac{\sigma_{eff,i} - \sigma_{los,i}(\theta)}{\delta\sigma_{eff,i}} \right)^2 + \sum_{i=1}^9 \ln(2\pi\delta\sigma_{eff,i}^2), \quad (4.15)$$

where $\sigma_{los,i}(\theta)$ is the theoretical prediction for the dispersion velocity, $\sigma_{eff,i} = (\sigma_i^2 + v_i^2)^{1/2}$ is the effective velocity dispersion and $\delta\sigma_{eff,i}$ is the corresponding uncertainty. Notice that equation (4.15) differs from the usual definition of χ^2 which is employed to compare the theoretical predictions of a model with the actual observations. The reason is that the equation (4.15) is obtained from the logarithm of the likelihood function $p(d|\theta, I)$. This function is approximated as a gaussian distribution, hence the first term in our definition of χ^2 can be obtained from the exponent of the gaussian distribution while the second term is due to the normalization constant.

Table 4.2: Velocity dispersion profile.

R (kpc)	v (km/s)	σ (km/s)
0.23	$0.2^{+2.2}_{-2.6}$	$26.1^{+4.4}_{-3.5}$
0.49	$-3.4^{+2.6}_{-2.8}$	$26.7^{+4.1}_{-3.4}$
0.79	$-3.1^{+2.2}_{-2.5}$	$26.5^{+4.4}_{-2.9}$
1.13	$-0.7^{+2.0}_{-1.8}$	$31.8^{+3.3}_{-2.9}$
1.53	$0.5^{+2.0}_{-2.6}$	$29.1^{+3.4}_{-2.4}$
1.94	$0.7^{+2.0}_{-2.0}$	$29.5^{+3.0}_{-2.6}$
2.55	$-0.4^{+2.3}_{-2.0}$	$29.3^{+3.1}_{-2.5}$
3.62	$0.3^{+2.3}_{-2.5}$	$34.4^{+3.8}_{-3.6}$
5.13	$5.9^{+3.8}_{-4.1}$	$40.2^{+7.9}_{-7.6}$

The number of parameters contained in θ depends on the particular model under investigation. For example, when we study the kinematics of DF44 in the framework of GR without assuming any Dark Matter contribution, there are five parameters $\theta = \{D, \Upsilon_*, \beta_0, \beta_\infty, r_a\}$. On the other hand, when we address the same problem in the framework of the Deser-Woodard model assuming a nonzero Dark Matter contribution, there are ten parameters $\theta = \{D, \Upsilon_*, \beta_0, \beta_\infty, r_a, M_{200}, c_{200}, \gamma, r_\phi, r_\xi\}$.

It is important to notice that the χ^2 that was actually minimized using the Metropolis-Hastings algorithm includes also some priors on the parameters of the theory as reported in Table (4.3). Notice that we set flat priors on the logarithm of the non-local parameters of the Deser-Woodard model r_ϕ and r_ξ over a large range $(-10, 5)$. It is also worth mentioning that some of the parameters are associated to more than one prior; the reason is that, as previously mentioned, priors can also contain information on the sign of the parameters, for instance on the distance D there is one gaussian prior based on the observations and a uniform prior over the range $(0, \infty)$ because the distance is a positive quantity. The same argument holds for the other parameters, a remark can be made for the parameter γ that is associated to a uniform prior over the interval $(0, 3)$ where the upper limit is due to convergence of the extended potentials calculated in section 3.2.

Table 4.3: Prior distributions used throughout our statistical analysis. \mathcal{N} refers to normal distributions, $\log \mathcal{N}$ to log-normal distributions, and \mathcal{U} to flat priors. Moreover, f_c denotes the function $f_c(M_{200})$ from the $c - M$ relation of [56].

Parameters	Priors	References
D (Mpc)	$\mathcal{N}(102, 14)$	[57]
	$\mathcal{U}(0, \infty)$	–
Υ_*	$\log \mathcal{N}(1.5, 0.1)$	[58, 59]
	$\mathcal{U}(0, 20)$	–
$1 - \beta_i$	$\log \mathcal{N}(0, 0.5)$	–
	$\mathcal{U}(-10, 1)$	–
c_{200}	$\log \mathcal{N}(f_c, 0.16)$	[56]
	$\mathcal{U}(0, 20)$	–
$\log M_{200}$ (M_\odot)	$\mathcal{U}(0, 20)$	–
γ	$\mathcal{U}(0, 3)$	–
$\log r_\phi$ (kpc)	$\mathcal{U}(-10, 5)$	–
$\log r_\xi$ (kpc)	$\mathcal{U}(-10, 5)$	–

In order to remove the burn-in and improve the exploration of the parameter space, the following three-step procedure has been employed:

- first of all, run an exploratory Markov Chain of 10000 points using arbitrary values for the parameters $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ of a given model and choosing as proposal distribution a multivariate gaussian distribution with $n \times n$ covariance matrix C

$$C = \begin{pmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0 & 0.01 \end{pmatrix} \quad (4.16)$$

- run an additional Markov Chain of 10000 points using as matrix C the covariance matrix of the previous sample. The starting point of the Markov Chain corresponds to the parameters for which the first Markov Chain has obtained a minimum of the χ^2 ;
- run a final Markov Chain of 50000 points applying the same method of the second chain. The sample of the target distribution obtained from this chain will be used for the statistical analysis.

This process has been applied to the Deser-Woodard model and to General Relativity. Once the final Markov Chain has been obtained, it is possible to gather information on the parameters of the models such as their marginal posterior distribution.

4.4 Results

The results of the statistical analysis are reported in Table (4.4), which contains the best estimates or ranges of the parameters for the different models analyzed, as well as the Bayes factors. This coefficient has been calculated for the different gravitational scenarios by taking as a reference model I_j the case of General Relativity with Dark Matter, in fact the corresponding Bayes factor is equal to 1 and its natural logarithm is zero.

First of all, in the GR case it is clear that the baryonic component alone cannot explain the kinematics of DF44 since the Bayes factor strongly disfavours this scenario with respect to the case of GR with the additional Dark Matter component. This result is also supported by the dispersion velocity profiles shown in the left panels of Figure (4.3). The additional degrees of freedom associated to the DM indeed allow a significantly better fit, showing a very good agreement with the observational data points. Moreover, the results

of our analysis in the framework of GR are in perfect agreement with the previous results in literature, that stand for a dominant DM contribution to the kinematical features of DF44 [39].

In our best-fit scenario, based on GR plus a dominant Dark Matter component, it is clear that the favoured model for the velocity dispersion is the one dominated by the tangential component, since the parameters β_0 and β_∞ have negative values. Notice also that the parameter r_a remains unconstrained.

Furthermore, as far as the Dark Matter density profile is concerned, the value of γ is lower than one, exhibiting a departure from a standard Navarro-Frenk-White profile. Again, our results are completely compatible with previous analysis like the one carried out in [39].

Table 4.4: Results from the statistical analysis of DF44. Unconstrained parameters are labelled with U . The Bayes factor has been computed using the standard Λ CDM paradigm as reference model, where the role of Λ is played by the non-local gravitational corrections, while the CDM component is retained.

	Galaxy parameters		Anisotropy parameters			gNFW parameters			Non-local parameters		Bayes factor
	D	Υ_*	β_0	β_∞	r_a	c_{200}	$\log M_{200}$	γ	$\log r_\phi$	$\log r_\xi$	$\ln \mathcal{B}_j^i$
	(Mpc)				(kpc)		(M_\odot)		(kpc)	(kpc)	
GR: stars	$112.46^{+12.66}_{-12.23}$	$2.15^{+0.43}_{-0.33}$	< -7.73	$-0.28^{+0.69}_{-1.08}$	U	–	–	–	–	–	$-4.29^{+0.04}_{-0.04}$
GR: stars + DM	$99.92^{+14.32}_{-14.62}$	$1.61^{+0.43}_{-0.33}$	> -4.31	> -1.95	U	$10.28^{+4.62}_{-3.75}$	$10.31^{+1.05}_{-0.62}$	< 0.62	–	–	0
NL: stars	$111.38^{+12.78}_{-12.30}$	$2.14^{+0.40}_{-0.31}$	< -7.78	$-0.24^{+0.65}_{-1.10}$	U	–	–	–	$-4.08^{+4.00}_{-3.63}$	U	$-5.30^{+0.06}_{-0.04}$
NL: stars + DM	$100.33^{+14.10}_{-13.92}$	$1.65^{+0.46}_{-0.38}$	$-2.70^{+2.12}_{-3.98}$	$-0.78^{+1.00}_{-2.40}$	U	> 4.80	> 10.25	< 0.54	$-0.51^{+2.79}_{-3.34}$	> -3.97	$-1.41^{+0.05}_{-0.04}$

Let us now turn to the Non-local Gravity scenario. It is worthwhile to compare both the non-local cases with the corresponding analyses in GR. It can be immediately seen that the results are very similar and that non-local corrections do not provide significant changes to the estimates obtained in the framework of GR. However, notice from the Bayes factors that the non-local gravity model is slightly disfavoured with respect to GR. This result can be clearly understood in light of the previous discussion regarding the Occam's razor effect in Bayesian statistics. Even though the fits in the non-local framework are slightly better than those in the GR case, as shown in the right panels of Figure (4.3), this improvement is not sufficient to justify the additional complexity of the Deser-Woodard

model with respect to GR.

The almost equal constraints on the free parameters of Table (4.4) are due to the subdominance of the non-local corrections to the Newtonian potential. The contribution of $\Phi_1(r)$ can be evaluated from the expression (2.64), once the potential has been extended to the case of Dark Matter and stellar matter distributions. Let us first consider the case of non-local gravity with only the stellar contribution: substituting the median value of the parameters reported in Table (4.4) and a value of r of the order $\sim \text{kpc}$ clearly shows that the non-local correction is subdominant, in fact

$$\Phi_1 \sim 10^{-38} \text{ kpc}^2/\text{s}^2, \quad (4.17)$$

while the standard GR contribution is

$$\Phi_0 \sim 10^{-31} \text{ kpc}^2/\text{s}^2. \quad (4.18)$$

Similarly, in the case of non-local gravity with the contributions of both stellar matter and Dark Matter, the non-local correction is of the order

$$\Phi_1 \sim 10^{-35} \text{ kpc}^2/\text{s}^2, \quad (4.19)$$

compared to the standard Newtonian term

$$\Phi_0 \sim 10^{-30} \text{ kpc}^2/\text{s}^2. \quad (4.20)$$

Table (4.4) also shows that in the non-local case with only the stellar component the best estimates for the non-local length r_ϕ is significantly lower than the one obtained when there is also a Dark Matter contribution. This is due to the fact that in the first case the non-local correction, which is inversely proportional to the non-local radius, has to compensate the contribution of Dark Matter. However, as previously shown, the non-local correction is still unable to substitute the Dark Matter contribution, thus resulting in a poor fit of the observational data. As in the GR framework, the stellar scenario is therefore disfavoured with respect to the case in which the DM is considered. The Bayes factor indeed implies a strong evidence for the presence of Dark Matter, both in the GR and the Non-local case.

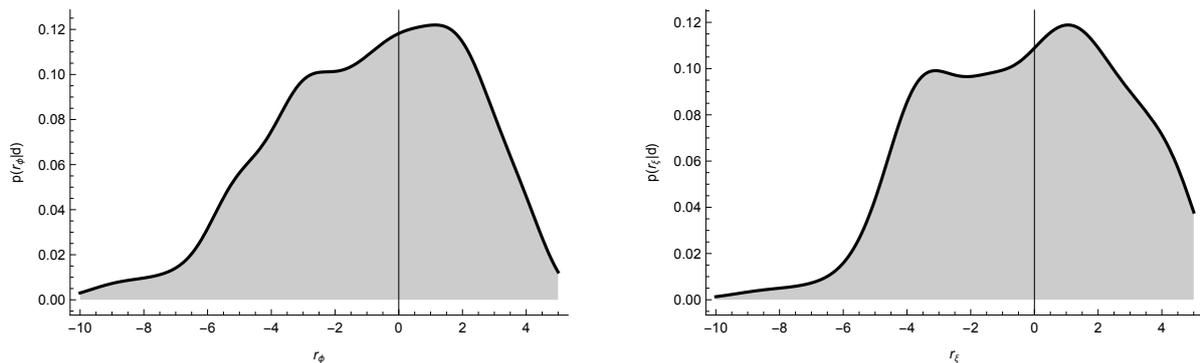
Our statistical analysis in the non-local gravity framework also allows to set constraints on the two non-local length scales r_ϕ and r_ξ that characterize the Deser-Woodard model. The non-local radius r_ϕ , associated to the dynamical scalar field ϕ , is better constrained than the length scale r_ξ , which is associated to the Lagrange multiplier. The latter remains unconstrained in the case of the stellar fit, while a lower bound can be set when the DM contribution is taken into account. In the best-fit non-local scenario, when DM is

added to the modelling of DF44, the constraints on the non-local length scales therefore are

$$-3.85 < \log r_\phi < 2.30 \quad \text{and} \quad \log r_\xi > -3.97, \quad (4.21)$$

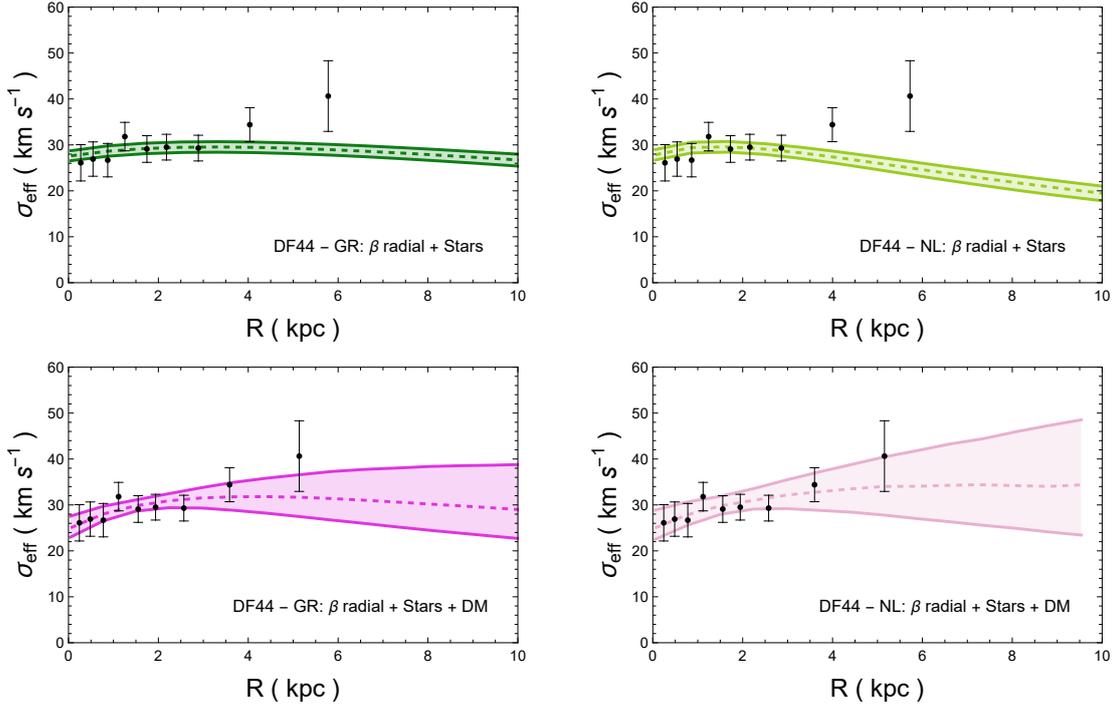
In this case (NL:Stars+DM) it is also interesting to look at the marginal posterior distributions associated to the parameters r_ϕ and r_ξ that can be obtained using the approach presented in section (4.2). In particular, figure (4.2) shows a smooth kernel distribution (i.e. a nonparametric representation of a probability density function characterized by a smoothing kernel and a bandwidth) for the marginal posteriors of r_ϕ and r_ξ based on the results attained from the Metropolis-Hastings algorithm. While r_ϕ shows a peaked posterior, thus allowing to set proper but loose constraints, the length scale r_ξ results in an highly irregular marginalized posterior, and only a lower bound can be therefore set.

Figure 4.2: Smooth kernel distributions of $p(r_\phi|d)$ and $p(r_\xi|d)$ with a gaussian smoothing kernel and using the standard deviation as bandwidth



Overall, the statistical analysis shows that the Deser-Woodard model offers a good explanation for the accelerated expansion of the Universe but the corrections to the Newtonian potential at the galactic scale of DF44 are not sufficient to replace the effect of Dark Matter.

Figure 4.3: Velocity dispersion profiles of DF44 in the framework of General Relativity (left panels) and Non-local gravity (right panels). Black dots and bars are the observational data σ_{eff} from [55], with uncertainties. Colored dashed lines and shaded regions respectively are the median and the 1σ credible region of the σ_{los} profile, derived from equation. (3.38).



Discussion and Conclusions

In this work we have performed a new test of the Deser-Woodard model in the case of the Ultra-Diffuse galaxy DF44. We have compared this model with GR and tested two gravitational scenarios: one where the Deser-Woodard model mimics the effects of Dark Energy at cosmological scale and the other where this model also replaces the Dark Matter. The comparison between the observational data and the theoretical predictions has been realized using the velocity dispersion along the line-of-sight as our reference physical observable. For what concern the theoretical predictions, it has been necessary to extend the non-local potential to the extended matter density profiles that models the Ultra-Diffuse Galaxies. Therefore, to the best of our knowledge, the extension of the non-local potential has been calculated for the first time both for the Sèrsic profile, modelling the stellar component, and the generalized Navarro-Frenk-White profile, modelling the Dark Matter distribution. This extension has been performed in the Weak-Field limit, through tabular methods. The theoretical prediction for the velocity dispersion was then obtained using the Jeans equation once we had specified the anisotropy profile $\beta(r)$.

The statistical analysis has shown that the Deser-Woodard model offers a good fit to the observational data when it acts as a Dark Energy model, namely when it is used to account for the accelerated cosmic expansion, while retaining the Dark Matter component to account for the clustering properties of the Universe. Indeed, the corrections to the Newtonian potential induced by the non-local model cannot explain the kinematics of DF44 only in terms of the stellar component, rather a significant amount of Dark Matter is required. More specifically, when a Dark Matter component is added to the model it is possible to fit the observational data, even though the Bayesian parameters slightly favour the GR case. A full Occam's razor effect is here at play: the non-local corrections, despite being subdominant compared to the GR contribution, allow a non-negligible improvement of the fit. However, this gain is too slight to support the additional complexity of the non-local gravity model. Nonetheless, in this particular case we were able to set a constraint on the non-local length scale r_ϕ related to the dynamical scalar field ϕ and a lower bound on the non-local length r_ξ associated to the Lagrange multiplier ξ .

As the number of model parameters is big with respect to the observational data points, the analysis carried out in this thesis can be improved by using an additional prior to break the degeneracy between some of the parameters. The stellar-to-halo mass relation (SHMR) [60], which account for the co-evolution of the stellar content of a galaxy and its Dark Matter halo, might be a robust information to add *a priori* to our analysis setup. In fact, it is expected both from theoretical and numerical consideration that the stellar mass of a galaxy is strongly related to the halo mass as a result of the galaxy formation process. Furthermore, it would be interesting to consider a different anisotropy profile for the velocity dispersion, with $\beta(r)$ constant. Indeed, this profile cannot be directly deduced from observations as the instrumental resolution is not high enough to discern the variation of the stellar kinematics over the galaxy radius.

It is also worthwhile to remember that the extended non-local potentials have been obtained by interpolation once the parameter space has been suitably tabulated, hence the results obtained so far can actually be sharpened by a denser sampling of the parameter space. This approach may result in better constraints on the non-local parameters.

The natural continuation of this work is to test the Deser-Woodard model at other astrophysical scales in order to study if there is a functional relation between the non-local radii and the masses of different astrophysical systems. In fact, some tests of this model have already been performed at the scale of the S2 star[2], which orbits the Galactic Centre, and at galaxy cluster scale [3]. The work presented in this thesis can be considered as a test of the Deser-Woodard model at an intermediate astrophysical scale, aimed at finding additional constrains on the non-local lengths and studying the mass distribution of the Ultra-Diffuse galaxy DF44. More specifically, the information gathered in all those tests, and additional ones that will be realized in the upcoming future, can be employed to investigate a possible functional relation between the non-local radii and the gaps observed among the typical masses and scales of galaxies, clusters of galaxies and superclusters. This line of research may then allow to address the problem of hierarchical structure formation theoretically in the framework of Non-local theories of Gravity.

Another interesting development of this work is to consider different types of Non-local theories of Gravity, replacing the operator \square^{-1} with other non-local operators that may emerge from QFT or Teleparallel Gravity. To this extent the Dark Matter hypothesis, as well as possible alternatives like Non-local theories of Gravity, can be effectively tested at those astrophysical scales using MCMC methods and computer simulations.

Bibliography

- [1] N. Aghanim et al. “Planck 2018 results. VI. Cosmological parameters”. In: *Astron. Astrophys.* 641 (2020). [Erratum: *Astron. Astrophys.* 652, C4 (2021)], A6. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910). arXiv: [1807.06209](https://arxiv.org/abs/1807.06209) [[astro-ph.CO](#)].
- [2] K. F. Dialektopoulos, D. Borika, S. Capozziello, V. Borika Jovanović, and P. Jovanović. “Constraining nonlocal gravity by S2 star orbits”. In: *Phys. Rev. D* 99.4 (2019), p. 044053. DOI: [10.1103/PhysRevD.99.044053](https://doi.org/10.1103/PhysRevD.99.044053). arXiv: [1812.09289](https://arxiv.org/abs/1812.09289) [[astro-ph.GA](#)].
- [3] Filippo Bouchè, Salvatore Capozziello, Vincenzo Salzano, and Keiichi Umetsu. “Testing non-local gravity by clusters of galaxies”. In: *Eur. Phys. J. C* 82.7 (2022), p. 652. DOI: [10.1140/epjc/s10052-022-10586-5](https://doi.org/10.1140/epjc/s10052-022-10586-5). arXiv: [2205.03216](https://arxiv.org/abs/2205.03216) [[gr-qc](#)].
- [4] Pieter van Dokkum, Roberto Abraham, Jean Brodie, Charlie Conroy, Shany Danieli, Allison Merritt, Lamiya Mowla, Aaron Romanowsky, and Jielai Zhang. “A High Stellar Velocity Dispersion and ~ 100 Globular Clusters for the Ultra-diffuse Galaxy Dragonfly 44”. In: *Astrophys. J. Lett.* 828.1 (2016), p. L6. DOI: [10.3847/2041-8205/828/1/L6](https://doi.org/10.3847/2041-8205/828/1/L6). arXiv: [1606.06291](https://arxiv.org/abs/1606.06291) [[astro-ph.GA](#)].
- [5] John Henry. “Newton and Action at a Distance”. In: *The Oxford Handbook of Newton* (2019), 1–34. DOI: [10.1093/oxfordhb/9780199930418.013.17](https://doi.org/10.1093/oxfordhb/9780199930418.013.17).
- [6] Sergio Doplicher, Klaus Fredenhagen, and John E. Roberts. “The Quantum structure of space-time at the Planck scale and quantum fields”. In: *Commun. Math. Phys.* 172 (1995), pp. 187–220. DOI: [10.1007/BF02104515](https://doi.org/10.1007/BF02104515). arXiv: [hep-th/0303037](https://arxiv.org/abs/hep-th/0303037).
- [7] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. “Bell nonlocality”. In: *Reviews of modern physics* 86.2 (2014), p. 419. DOI: <https://doi.org/10.1103/RevModPhys.86.419>.
- [8] J. S. Bell. “On the Einstein Podolsky Rosen paradox”. In: *Physics Physique Fizika* 1 (3 1964), pp. 195–200. DOI: [10.1103/PhysicsPhysiqueFizika.1.195](https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195). URL: <https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195>.
- [9] Sandu Popescu. “Dynamical quantum non-locality”. In: *Nature Physics* 6.3 (2010), pp. 151–153. DOI: [10.1038/nphys1619](https://doi.org/10.1038/nphys1619).

-
- [10] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014.
- [11] Enis Belgacem, Yves Dirian, Stefano Foffa, and Michele Maggiore. “Nonlocal gravity. Conceptual aspects and cosmological predictions”. In: *JCAP* 03 (2018), p. 002. DOI: [10.1088/1475-7516/2018/03/002](https://doi.org/10.1088/1475-7516/2018/03/002). arXiv: [1712.07066](https://arxiv.org/abs/1712.07066) [hep-th].
- [12] Andrei O. Barvinsky. “Serendipitous discoveries in nonlocal gravity theory”. In: *Phys. Rev. D* 85 (2012), p. 104018. DOI: [10.1103/PhysRevD.85.104018](https://doi.org/10.1103/PhysRevD.85.104018). arXiv: [1112.4340](https://arxiv.org/abs/1112.4340) [hep-th].
- [13] Salvatore Capozziello and Francesco Bajardi. “Nonlocal gravity cosmology: An overview”. In: *Int. J. Mod. Phys. D* 31.06 (2022). DOI: [10.1142/S0218271822300099](https://doi.org/10.1142/S0218271822300099). arXiv: [2201.04512](https://arxiv.org/abs/2201.04512) [gr-qc].
- [14] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973.
- [15] Luca Buoninfante, Gaetano Lambiase, and Anupam Mazumdar. “Ghost-free infinite derivative quantum field theory”. In: *Nucl. Phys. B* 944 (2019), p. 114646. DOI: [10.1016/j.nuclphysb.2019.114646](https://doi.org/10.1016/j.nuclphysb.2019.114646). arXiv: [1805.03559](https://arxiv.org/abs/1805.03559) [hep-th].
- [16] Tirthabir Biswas, Alexey S. Koshelev, and Anupam Mazumdar. “Gravitational theories with stable (anti-)de Sitter backgrounds”. In: *Fundam. Theor. Phys.* 183 (2016), pp. 97–114. DOI: [10.1007/978-3-319-31299-6_5](https://doi.org/10.1007/978-3-319-31299-6_5). arXiv: [1602.08475](https://arxiv.org/abs/1602.08475) [hep-th].
- [17] Tirthabir Biswas, Anupam Mazumdar, and Warren Siegel. “Bouncing universes in string-inspired gravity”. In: *JCAP* 03 (2006), p. 009. DOI: [10.1088/1475-7516/2006/03/009](https://doi.org/10.1088/1475-7516/2006/03/009). arXiv: [hep-th/0508194](https://arxiv.org/abs/hep-th/0508194).
- [18] Luca Buoninfante, Alexey S. Koshelev, Gaetano Lambiase, and Anupam Mazumdar. “Classical properties of non-local, ghost- and singularity-free gravity”. In: *JCAP* 09 (2018), p. 034. DOI: [10.1088/1475-7516/2018/09/034](https://doi.org/10.1088/1475-7516/2018/09/034). arXiv: [1802.00399](https://arxiv.org/abs/1802.00399) [gr-qc].
- [19] Fabio Briscese, Leonardo Modesto, and Shinji Tsujikawa. “Super-renormalizable or finite completion of the Starobinsky theory”. In: *Phys. Rev. D* 89.2 (2014), p. 024029. DOI: [10.1103/PhysRevD.89.024029](https://doi.org/10.1103/PhysRevD.89.024029). arXiv: [1308.1413](https://arxiv.org/abs/1308.1413) [hep-th].
- [20] Aindriu Conroy, Tomi Koivisto, Anupam Mazumdar, and Ali Teimouri. “Generalized quadratic curvature, non-local infrared modifications of gravity and Newtonian potentials”. In: *Class. Quant. Grav.* 32.1 (2015), p. 015024. DOI: [10.1088/0264-9381/32/1/015024](https://doi.org/10.1088/0264-9381/32/1/015024). arXiv: [1406.4998](https://arxiv.org/abs/1406.4998) [hep-th].

-
- [21] A. O. Barvinsky. “Aspects of Nonlocality in Quantum Field Theory, Quantum Gravity and Cosmology”. In: *Mod. Phys. Lett. A* 30.03n04 (2015), p. 1540003. DOI: [10.1142/S0217732315400039](https://doi.org/10.1142/S0217732315400039). arXiv: [1408.6112](https://arxiv.org/abs/1408.6112) [hep-th].
- [22] Michele Maggiore and Michele Mancarella. “Nonlocal gravity and dark energy”. In: *Phys. Rev. D* 90.2 (2014), p. 023005. DOI: [10.1103/PhysRevD.90.023005](https://doi.org/10.1103/PhysRevD.90.023005). arXiv: [1402.0448](https://arxiv.org/abs/1402.0448) [hep-th].
- [23] Shin’ichi Nojiri and Sergei D. Odintsov. “Modified non-local-F(R) gravity as the key for the inflation and dark energy”. In: *Phys. Lett. B* 659 (2008), pp. 821–826. DOI: [10.1016/j.physletb.2007.12.001](https://doi.org/10.1016/j.physletb.2007.12.001). arXiv: [0708.0924](https://arxiv.org/abs/0708.0924) [hep-th].
- [24] Nima Arkani-Hamed, Savas Dimopoulos, Gia Dvali, and Gregory Gabadadze. “Non-local modification of gravity and the cosmological constant problem”. In: (Sept. 2002). arXiv: [hep-th/0209227](https://arxiv.org/abs/hep-th/0209227).
- [25] Yi-Fu Cai, Salvatore Capozziello, Mariafelicia De Laurentis, and Emmanuel N. Saridakis. “f(T) teleparallel gravity and cosmology”. In: *Rept. Prog. Phys.* 79.10 (2016), p. 106901. DOI: [10.1088/0034-4885/79/10/106901](https://doi.org/10.1088/0034-4885/79/10/106901). arXiv: [1511.07586](https://arxiv.org/abs/1511.07586) [gr-qc].
- [26] M. Nakahara. *Geometry, topology and physics*. 2003.
- [27] Bahram Mashhoon. *Nonlocal Gravity*. Vol. 167. Oxford University Press, 2017.
- [28] Friedrich W. Hehl and Bahram Mashhoon. “Nonlocal Gravity Simulates Dark Matter”. In: *Phys. Lett. B* 673 (2009), pp. 279–282. DOI: [10.1016/j.physletb.2009.02.033](https://doi.org/10.1016/j.physletb.2009.02.033). arXiv: [0812.1059](https://arxiv.org/abs/0812.1059) [gr-qc].
- [29] Stanley Deser and Richard P Woodard. “Nonlocal cosmology”. In: *Physical Review Letters* 99.11 (2007). DOI: <https://doi.org/10.1103/PhysRevLett.99.111301>.
- [30] C. Deffayet and R. P. Woodard. “Reconstructing the Distortion Function for Non-local Cosmology”. In: *JCAP* 08 (2009), p. 023. DOI: [10.1088/1475-7516/2009/08/023](https://doi.org/10.1088/1475-7516/2009/08/023). arXiv: [0904.0961](https://arxiv.org/abs/0904.0961) [gr-qc].
- [31] Konstantinos F. Dialektopoulos and Salvatore Capozziello. “Noether Symmetries as a geometric criterion to select theories of gravity”. In: *Int. J. Geom. Meth. Mod. Phys.* 15.suppl01 (2018), p. 1840007. DOI: [10.1142/S0219887818400078](https://doi.org/10.1142/S0219887818400078). arXiv: [1808.03484](https://arxiv.org/abs/1808.03484) [gr-qc].
- [32] S. Capozziello, R. De Ritis, C. Rubano, and P. Scudellaro. “Noether symmetries in cosmology”. In: *Riv. Nuovo Cim.* 19N4 (1996), pp. 1–114. DOI: [10.1007/BF02742992](https://doi.org/10.1007/BF02742992).
- [33] Christof Wetterich. “Effective nonlocal Euclidean gravity”. In: *General Relativity and Gravitation* 30.1 (1998), pp. 159–172. DOI: [10.1023/A:1018837319976](https://doi.org/10.1023/A:1018837319976).

-
- [34] S. Deser and R. P. Woodard. “Observational Viability and Stability of Nonlocal Cosmology”. In: *JCAP* 11 (2013), p. 036. DOI: [10.1088/1475-7516/2013/11/036](https://doi.org/10.1088/1475-7516/2013/11/036). arXiv: [1307.6639](https://arxiv.org/abs/1307.6639) [[astro-ph.CO](#)].
- [35] Enis Belgacem, Andreas Finke, Antonia Frassino, and Michele Maggiore. “Testing nonlocal gravity with Lunar Laser Ranging”. In: *JCAP* 02 (2019), p. 035. DOI: [10.1088/1475-7516/2019/02/035](https://doi.org/10.1088/1475-7516/2019/02/035). arXiv: [1812.11181](https://arxiv.org/abs/1812.11181) [[gr-qc](#)].
- [36] Salvatore Capozziello and Shinji Tsujikawa. “Solar system and equivalence principle constraints on $f(R)$ gravity by chameleon approach”. In: *Phys. Rev. D* 77 (2008), p. 107501. DOI: [10.1103/PhysRevD.77.107501](https://doi.org/10.1103/PhysRevD.77.107501). arXiv: [0712.2268](https://arxiv.org/abs/0712.2268) [[gr-qc](#)].
- [37] A. I. Vainshtein. “To the problem of nonvanishing gravitation mass”. In: *Phys. Lett. B* 39 (1972), pp. 393–394. DOI: [10.1016/0370-2693\(72\)90147-5](https://doi.org/10.1016/0370-2693(72)90147-5).
- [38] Duško Borka, Vesna Borka Jovanović, Salvatore Capozziello, and Predrag Jovanović. “Velocity distribution of elliptical galaxies in the framework of Non-local Gravity model”. In: *Adv. Space Res.* 71 (2023), pp. 1235–1244. DOI: [10.1016/j.asr.2022.08.060](https://doi.org/10.1016/j.asr.2022.08.060). arXiv: [2209.01696](https://arxiv.org/abs/2209.01696) [[gr-qc](#)].
- [39] Enrico Laudato and Vincenzo Salzano. “DHOST gravity in ultra-diffuse galaxies—Part II: NGC 1052-DF4 and Dragonfly 44”. In: *Eur. Phys. J. C* 83.5 (2023), p. 402. DOI: [10.1140/epjc/s10052-023-11564-1](https://doi.org/10.1140/epjc/s10052-023-11564-1). arXiv: [2211.08839](https://arxiv.org/abs/2211.08839) [[gr-qc](#)].
- [40] Timothy Carleton, Raphaël Errani, Michael Cooper, Manoj Kaplinghat, Jorge Peñarrubia, and Yicheng Guo. “The formation of ultra-diffuse galaxies in cored dark matter haloes through tidal stripping and heating”. In: *Monthly Notices of the Royal Astronomical Society* 485.1 (2019), pp. 382–395. DOI: [10.1093/mnras/stz383](https://doi.org/10.1093/mnras/stz383). URL: <https://doi.org/10.1093/mnras/stz383>.
- [41] N. C. Amorisco and A. Loeb. “Ultradiffuse galaxies: the high-spin tail of the abundant dwarf galaxy population”. In: *Monthly Notices of the Royal Astronomical Society: Letters* 459.1 (2016), pp. L51–L55. DOI: [10.1093/mnrasl/slw055](https://doi.org/10.1093/mnrasl/slw055). URL: <https://doi.org/10.1093/mnrasl/slw055>.
- [42] Kirill A. Grishin, Igor V. Chilingarian, Anton V. Afanasiev, Daniel Fabricant, Ivan Yu. Katkov, Sean Moran, and Masafumi Yagi. “Transforming gas-rich low-mass disk galaxies into ultra-diffuse galaxies by ram pressure”. In: *Nature Astronomy* 5.12 (2021), pp. 1308–1318. DOI: [10.1038/s41550-021-01470-5](https://doi.org/10.1038/s41550-021-01470-5). URL: <https://doi.org/10.1038/s41550-021-01470-5>.

- [43] Teymoor Saifollahi, Ignacio Trujillo, Michael A Beasley, Reynier F Peletier, and Johan H Knapen. “The number of globular clusters around the iconic UDG DF44 is as expected for dwarf galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 502.4 (2020), pp. 5921–5934. DOI: [10.1093/mnras/staa3016](https://doi.org/10.1093/mnras/staa3016). URL: <https://doi.org/10.1093/mnras/staa3016>.
- [44] Hosein Hagi, Vahid Amiri, Akram Hasani Zonoozi, Indranil Banik, Pavel Kroupa, and Moritz Haslbauer. “The Star Formation History and Dynamics of the Ultra-diffuse Galaxy Dragonfly 44 in MOND and MOG”. In: *Astrophys. J. Lett.* 884.1 (2019), p. L25. DOI: [10.3847/2041-8213/ab4517](https://doi.org/10.3847/2041-8213/ab4517). arXiv: [1909.07978](https://arxiv.org/abs/1909.07978) [astro-ph.GA].
- [45] Mariangela Lisanti. “Lectures on Dark Matter Physics”. In: *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*. 2017, pp. 399–446. DOI: [10.1142/9789813149441_0007](https://doi.org/10.1142/9789813149441_0007). arXiv: [1603.03797](https://arxiv.org/abs/1603.03797) [hep-ph].
- [46] Vera C. Rubin and W. Kent Ford Jr. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: *Astrophys. J.* 159 (1970), pp. 379–403. DOI: [10.1086/150317](https://doi.org/10.1086/150317).
- [47] A Araujo, DF López, and JG Pereira. “De Sitter-Invariant Special Relativity and Galaxy Rotation Curves”. In: *Gravitation and Cosmology* 25 (2019), pp. 157–163. DOI: <https://doi.org/10.1134/S0202289319020026>.
- [48] Anne M. Green. “Dark matter in astrophysics/cosmology”. In: *SciPost Phys. Lect. Notes* 37 (2022), p. 1. DOI: [10.21468/SciPostPhysLectNotes.37](https://doi.org/10.21468/SciPostPhysLectNotes.37). arXiv: [2109.05854](https://arxiv.org/abs/2109.05854) [hep-ph].
- [49] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. “The Structure of cold dark matter halos”. In: *Astrophys. J.* 462 (1996), pp. 563–575. DOI: [10.1086/177173](https://doi.org/10.1086/177173). arXiv: [astro-ph/9508025](https://arxiv.org/abs/astro-ph/9508025).
- [50] Gary A. Mamon and Ewa L. Lokas. “Confronting lambda-CDM with the optical observations of elliptical galaxies. 2. Weighing the dark matter component”. In: *Mon. Not. Roy. Astron. Soc.* 363 (2005). [Addendum: *Mon. Not. Roy. Astron. Soc.* 370, 1582 (2006)], pp. 705–722. DOI: [10.1111/j.1365-2966.2005.09400.x](https://doi.org/10.1111/j.1365-2966.2005.09400.x). arXiv: [astro-ph/0405491](https://arxiv.org/abs/astro-ph/0405491).
- [51] Nelson Christensen and Renate Meyer. “Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements”. In: *Classical and Quantum Gravity* (June 2000). arXiv: [astro-ph/0006401](https://arxiv.org/abs/astro-ph/0006401).
- [52] Roberto Trotta. “Bayesian Methods in Cosmology”. In: Jan. 2017. arXiv: [1701.01467](https://arxiv.org/abs/1701.01467).

- [53] Harold Jeffreys. *The Theory of Probability*. Oxford Classic Texts in the Physical Sciences. 1939. ISBN: 978-0-19-850368-2, 978-0-19-853193-7.
- [54] Joanna Dunkley, Martin Bucher, Pedro G. Ferreira, Kavilan Moodley, and Constantinos Skordis. “Fast and reliable mcmc for cosmological parameter estimation”. In: *Mon. Not. Roy. Astron. Soc.* 356 (2005), pp. 925–936. DOI: [10.1111/j.1365-2966.2004.08464.x](https://doi.org/10.1111/j.1365-2966.2004.08464.x). arXiv: [astro-ph/0405462](https://arxiv.org/abs/astro-ph/0405462).
- [55] Pieter van Dokkum et al. “Spatially-resolved stellar kinematics of the ultra diffuse galaxy Dragonfly 44. I. Observations, kinematics, and cold dark matter halo fits”. In: *Astrophys. J.* 880 (2019), p. 91. DOI: [10.3847/1538-4357/ab2914](https://doi.org/10.3847/1538-4357/ab2914). arXiv: [1904.04838](https://arxiv.org/abs/1904.04838) [[astro-ph.GA](https://arxiv.org/abs/astro-ph.GA)].
- [56] Camila A. Correa, J. Stuart B. Wyithe, Joop Schaye, and Alan R. Duffy. “The accretion history of dark matter haloes – III. A physical model for the concentration–mass relation”. In: *Mon. Not. Roy. Astron. Soc.* 452.2 (2015), pp. 1217–1232. DOI: [10.1093/mnras/stv1363](https://doi.org/10.1093/mnras/stv1363). arXiv: [1502.00391](https://arxiv.org/abs/1502.00391) [[astro-ph.CO](https://arxiv.org/abs/astro-ph.CO)].
- [57] Bjarne Thomsen, William A. Baum, Mark Hammergren, and Guy Worthey. “The distance to the coma cluster from surface brightness fluctuations”. In: *Astrophys. J. Lett.* 483 (1997), pp. L37–L40. DOI: [10.1086/310735](https://doi.org/10.1086/310735). arXiv: [astro-ph/9704121](https://arxiv.org/abs/astro-ph/9704121).
- [58] Asher Wasserman et al. “Spatially Resolved Stellar Kinematics of the Ultra-diffuse Galaxy Dragonfly 44. II. Constraints on Fuzzy Dark Matter”. In: *The Astrophysical Journal* 885.2 (2019), p. 155. DOI: [10.3847/1538-4357/ab3eb9](https://doi.org/10.3847/1538-4357/ab3eb9). arXiv: [1905.10373](https://arxiv.org/abs/1905.10373) [[astro-ph.GA](https://arxiv.org/abs/astro-ph.GA)].
- [59] Edward N Taylor et al. “Galaxy And Mass Assembly: Stellar Mass Estimates”. In: *Mon. Not. Roy. Astron. Soc.* 418 (2011), p. 1587. DOI: [10.1111/j.1365-2966.2011.19536.x](https://doi.org/10.1111/j.1365-2966.2011.19536.x). arXiv: [1108.0635](https://arxiv.org/abs/1108.0635) [[astro-ph.CO](https://arxiv.org/abs/astro-ph.CO)].
- [60] Aldo Rodríguez-Puebla, Joel R. Primack, Vladimir Avila-Reese, and S. M. Faber. “Constraining the galaxy–halo connection over the last 13.3 Gyr: star formation histories, galaxy mergers and structural properties”. In: *Monthly Notices of the Royal Astronomical Society* 470.1 (2017), pp. 651–687. DOI: [10.1093/mnras/stx1172](https://doi.org/10.1093/mnras/stx1172). URL: <https://doi.org/10.1093/mnras/stx1172>.