Reconstructing Horndeski theories from cosmological observables

Joe Kennedy

Doctor of Philosophy
The University of Edinburgh
April 2019

Abstract

The nature of the accelerated expansion of the Universe remains one of the greatest challenges in modern physics. The simplest explanation is that the acceleration is driven by a cosmological constant. Large quantum corrections from the various matter fields in the Universe will contribute to the value of this constant. Unfortunately, these quantum effects lead to a discrepancy between the theoretical prediction of the rate of expansion and the observed rate by many orders of magnitude. Problems such as this have lead theorists to develop alternative models which can account for the accelerated expansion without a cosmological constant. These include the addition of an exotic matter species or even a modification to General Relativity itself. Many such theories introduce a *scalar field*, a concept which appears frequently in particle physics. For example, the Higgs particle is an excitation of a scalar field called the Higgs field which is a crucial component in the Standard Model of particle physics. Invoking a scalar field in cosmology adds an extra dynamical degree of freedom that can drive the accelerated expansion of the Universe, as well as introduce novel physical effects such as enhancing the clustering of matter. It is not a trivial task to include a scalar field into General Relativity as it can often lead to theoretical instabilities. There has recently been substantial interest in Horndeski theory, which is a general theory which couples the scalar field to gravity while avoiding theoretical issues. Subsets of Horndeski theory include a large range of common scalar field models such as quintessence. In order to study how the cosmological phenomenology of Horndeski theory differs from standard cosmology it is useful to have a generalised approach which enables the connection of theoretical predictions with observational data, without restricting to specific subclasses of models. The effective field theory of dark energy provides such a framework. However, the effective field theory of dark energy is purely phenomenological. In order to put constraints on Horndeski theory itself it is necessary to connect the constraints placed on the parameters in effective field theory with Horndeski theory. The aim of this thesis is to provide a method to connect constraints on cosmological parameters, soon to be measured to an unprecedented precision with the next generation of surveys, with Horndeski theory.

This thesis begins with an introduction to General Relativity and cosmology before discussing models which go beyond standard cosmology. A reconstruction which maps from the effective field theory of dark energy back to the space of covariant theories is then presented. This provides a method to connect constraints on phenomenological effective field theory parameters to covariant theories. We present many applications of this reconstruction. For example, we discuss how to map from frequently utilised observational parameters to an underlying Horndeski theory. This allows one to reconstruct, for example, a Horndeski theory which exhibits a weakening of the growth of structure relative to standard General Relativity. Extending these results into the nonlinear regime is then discussed. In principle this provides the necessary tools to systematically apply stringent tests to Horndeski theory with the next generation of cosmological surveys across a broad range of length scales.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in Refs. [1, 2] .

Parts of this work have been published in Ref. [3] (Submitted for publication at time of writing) .

(*Joe Kennedy, April 2019*)

Acknowledgements

A long project such as this could not have been completed without the support of a great number of people. It is too long a list to write down in full, but here I mention a few in particular.

First and foremost I would like to thank my family, in particular my parents, for providing me with the necessary creative and intellectual freedom to pursue the avenues that drew my interest from a young age. Their constant support and encouragement throughout my life has been invaluable.

The assistance of my supervisors Prof. Andy Taylor and Prof. Lucas Lombriser has been immeasurable throughout my PhD. I have thoroughly enjoyed the many hours that we have spent together cracking our heads against a wall to solve various problems that the research threw our way. Their patience and support were both crucial in developing my confidence as a researcher. I have learned so much from our collaboration and cannot thank you both enough.

I would also like to thank John Cain. As my secondary school physics teacher he did a fantastic job in providing the kindling which fuelled my interest in understanding the world around me, while also making sure I was focused enough to succeed academically. He is one of the greatest teachers I have ever known and this thesis would not have been completed without the inspiration he provided.

The community at the Edinburgh Royal Observatory is a special one. Every day that I have spent here has been a privilege, opening opportunities to discover so many aspects of the Universe in which we live from planet formation to cosmological physics. I wish to thank everyone who I have met through coffee breaks, collaborations or five-a-side football for providing an excellent environment in which to pursue this research. I'm quite sure it is the only academic community in the Galaxy that would allow The Darjeeling Feeling to perform an acoustic set during a Friday morning coffee break.

I would also like to thank Anthony Wood and Jana Pásztorová for being legendary flatmates throughout the completion of this PhD thesis. You both helped provide the healthy amount of insanity that is needed to maintain stamina throughout a PhD, and for that I am truly grateful.

Many thanks also to Ryan McManus, Adam Carnall, Dan Eastwood, Tom Kemp, Owen Turner, Fran Lane, Ben Morton, Alistair Bruce, Joey Faulkner and Alex Lewis for being excellent office mates throughout my time at the Institute for Astronomy. You all helped to provide welcome distractions when they were sorely needed and my work would not nearly have been so enjoyable without you.

Lay Summary

Humanity has always possessed a desire to understand the Universe in which we live. Over the past hundred years our comprehension of our place in the Universe has shifted drastically from the Ptolemaic model of a fixed Earth at the centre of everything. Progress has been made at an astonishing rate. Many of the questions that troubled us have been answered. But science never ends, and these solved problems have given rise to many more questions. Slowly but surely we are determining the right questions to ask but there is still a long way to go. We know that the Universe had a beginning, called the Big Bang, but we do not yet know whether it will have an end. We know that black holes exist, yet we do not know how to reconcile their existence with other fundamental physical laws. Contradictions in science can often generate significant breakthroughs. This thesis is a study of a profound contradiction which has persisted for at least 20 years.

The Universe is expanding. This has been known since the famous observations made by Edwin Hubble in 1929. However, the expansion is accelerating. As yet no-one understands why this is this case. The name given to our ignorance is dark energy. Why is the Universe accelerating? Dark energy. What is dark energy? It is causing the Universe to accelerate. The nature of dark energy, thought to make up around seventy percent of the total "stuff" in the Universe, remains a profound mystery and is a central topic of this thesis. Einstein famously invoked a concept called the cosmological constant in his theory of General Relativity before Hubble's discovery of the expanding Universe in order to achieve a static Universe. He later called this idea the greatest blunder of his life. It might be seen as a testament to Einstein's genius that even his mistakes seem to have importance. The cosmological constant is now back with a vengeance. When included in the Einstein equations it can also act to accelerate the universe, thus potentially explaining the underlying nature of dark energy.

This would surely have been the end of the story if it wasn't for quantum mechanics. Quantum mechanics describes the world at the smallest scales, predicting how electrons behave in atoms and the nature of particle interactions in accelerators such as the Large Hadron Collider. It also gives a prediction for the value of the cosmological constant. Unfortunately, this prediction does not match the observed rate of accelerated expansion by many orders of magnitude.

This contradiction between two of the most successful physical theories ever conceived, General Relativity and quantum mechanics, has lead to a great deal of new ideas which aim to explain the accelerated expansion of the Universe without necessarily invoking a cosmological constant. We will soon see an influx of new observations from cosmological surveys such as Euclid, which will measure the properties of our Universe to an unprecedented precision. It is hoped that these measurements can shed some light on the nature of the mechanism driving the accelerated expansion. Making measurements to an ever greater precision will not mean anything at all, unless there is a link between these measurements and the underlying theories. This link is what this thesis explores. It is no understatement to say that we could be on the cusp of a revolution in our understanding of the Universe. It is an open question whether the cosmological constant will do a better job at describing the observations than other theories. Or is there a new theory waiting in the wings poised to be revealed by the data? If it turns out to be the latter scenario, then this thesis provides a method to connect the observations with this new theory. Perhaps some day soon our understanding of not only cosmology, but of fundamental physics, will be transformed.

Contents

List of Figures

List of Tables

When you have eliminated the impossible whatever remains, however improbable, must be the truth.

Sherlock Holmes

Chapter 1

General Relativity and Cosmology

People assume that time is a strict progression of cause to effect, but actually from a nonlinear, non-subjective viewpoint, it is more like a big ball of wibbly-wobbly, timey-wimey stuff.

Doctor Who

1.1 Overview and motivation

What is the Universe made of? It is such a seemingly fundamental question that it is quite the embarrassment that, for the most part, nobody really knows. During the past fifty years there has been a major revolution in our understanding of what constitutes the majority of the "stuff" in the Universe. There was a time when humans thought that the Earth was situated at the centre of everything. This notion came to an abrupt end with Copernicus placing the Earth away from the centre of the Solar System. He was then later followed by Einstein, who argued that the very idea of *centre* was meaningless. Things were becoming complicated.

Atoms were then discovered and the baffling questions about nature's basic building blocks were thought to be overcome. Everything was made of indivisible entities, just as Leucippus and Democritus had suggested over two thousand years ago [10–12]. New discoveries in particle physics during the twentieth century radically altered this simplistic notion. By dividing atoms into electrons and nuclei, and subsequently nuclei into quarks and gluons, it became increasingly clear that there was far more depth and variety in the fundamental structures of the Universe. An extended family of composite and elementary particles revealed nature to be far more complicated than anyone imagined.[1]

Nature not behaving entirely as expected turned out to be the theme of twentieth century physics. It perhaps should have come as no surprise when it was then discovered that the contribution of every person, plant, planet and star both known and unknown, make up barely five per cent of the total "stuff" in the Universe.

For a start, much of the matter in the Universe does not interact with light. It is called dark matter and we infer that it must be there through the gravitational influence that it has on the rotation curves of galaxies [13], the dynamics of clusters of galaxies [14], how light is lensed as it traverses the Universe [15] and through the temperature patterns in the relic radiation left over from the Big Bang [4]. The nature of dark matter remains a pressing issue for modern astronomy and theoretical physics. It is often regarded to be a new massive elementary particle that is yet to be directly detected [16–18]. Alternatively, there has been revived interest in studying whether dark matter could be composed of low mass primordial black holes [19]. In either case, despite its fundamental nature remaining a mystery it can generally be considered as some species of ordinary matter which gravitates in accordance with the standard laws of gravity, namely the theory of General Relativity (GR) [20–22]. It simply doesn't interact with electromagnetic radiation and therefore we cannot see it directly. However, the majority of the Universe is made up of something which is so bizarre that it may result in a re-formulation of General Relativity itself.

Throw a tennis ball into the air and it continues to move upwards. For the briefest moment it then comes to a halt, before reversing its motion to travel back the way it came towards the ground. The name given to this everyday phenomenon is gravity. Simply put, any object with mass such as me, you, the tennis ball and the Earth pulls on every other object with mass. Despite being the oldest known force, gravity remains the least understood. By all means, our understanding of gravity

[1] An up to date list of particle discoveries can be obtained via the Particle Data Group at *pdg.lbl.gov/2018/listings/contents_listings.html*

has indeed considerably developed over time. An inconsistency between special relativity and Maxwell's theory of electromagnetism lead Einstein to develop GR. There has to date been no inconsistent measurement of GR with observations many of which have been either laboratory or Solar System tests [23]. From the perihelion of Mercury to the gravitational deflection of light it has proved to be remarkably successful and perhaps stands as the pinnacle of theoretical insight and predictive power. Gravitational waves stood as the last major prediction of the theory that remained to be directly detected until 2015 [24], 100 years after Einstein first formulated the theory. It has even found a commercial application in the everyday use of GPS navigation. So why is it also the least understood of all the forces in nature?

First and foremost GR cannot be the fundamental theory of gravity. The existence of singularities both at the Big Bang and in black holes indicate that it must break down when applied at high energies. There are entirely different physical laws that apply on very small scales which have a fundamentally different character to GR. These laws describe all known elementary particles in terms of quantum fields. They are discrete where GR is smooth and random where GR is deterministic. Reconciling the two theories into a quantum theory of gravity has proved profoundly challenging but progress is being made [25, 26].

Perhaps the most compelling reason hinting that GR needs developing comes from the fascinating observation that the Universe is accelerating in its expansion. Standard General Relativity says that the Universe should in a sense behave like the tennis ball thrown upwards on Earth. It will at first expand before the gravitational pull of all the matter in the Universe eventually leads to a brief halt and ends in a big crunch as the Universe collapses in on itself. This is not at all what is actually happening. The Universe is expanding at an accelerated rate. It is like throwing the tennis ball into the air and watching it keep on racing ever faster towards the sky. This is not how things are supposed to work. Understanding this strange aspect of nature remains one of the greatest challenges in physics. The driver of this acceleration was dubbed *dark energy* and it is a central topic of this thesis.

General Relativity has yet to be rigorously tested on cosmological scales. The next generation of cosmological surveys such as Euclid [27, 28] and LSST [29] aim to put constraints on GR at the largest scales to an unprecedented precision. Such surveys may provide the necessary signpost pointing towards a consistent theory beyond GR. This thesis explores how to connect the information that these surveys will provide with underlying physical theories. The eventual outcome will be to obtain a deeper understanding of fundamental physics from cosmological probes.

The structure of this thesis is as follows. Chapter 1 begins with an overview of General Relativity and important aspects of cosmology in order to set the scene for how the current consensus on our understanding of the Universe developed. Chapter 2 then provides some of the more technical background necessary to understand the ideas that go beyond General Relativity. We shall discuss a number of such models along with two important theorems that have motivated various research directions in achieving cosmic acceleration without a cosmological constant, namely the Weinberg no-go theorem and the Ostrogradsky theorem. Effective field theory is then introduced and discussed in some detail as it provides much of the theoretical backbone on which the rest of the thesis is based. Chapter 3 introduces the method of reconstructing physical theories from the effective field theory of dark energy. Chapter 4 discusses some applications of this reconstruction and shows how it can used to obtain theories which exhibit the desired cosmological phenomenology. Chapter 5 investigates the nonlinear freedom that is included in the reconstruction. In particular it will demonstrate how screening mechanisms can be incorporated into a reconstructed action, introduce a class of models which exhibit kinetic self-acceleration as well as extend the reconstruction to nonlinear scales. In principle this will allow one to reconstruct models beyond ΛCDM via constraints from a broad range of length scales.

1.2 General Relativity

1.2.1 What is gravity?

Consider again the tennis ball being tossed into the air. Throughout its motion it experiences a constant acceleration towards the ground. Isaac Newton's famous force law $F = m_I a$ links this acceleration with a force, a force which we call gravity. This force law was quantified by Newton by expressing the gravitational force between two masses m_1 and m_2 as being proportional their product and inversely proportional to the square of distance r between them with Newton's constant G being the proportionality constant

$$F = \frac{G m_1 m_2}{r^2} \,. \tag{1.1}$$

This beautifully compact equation is capable of describing a vast range of gravitational phenomena from the motion of planets, asteroids and comets around the Sun, the moon around the Earth and a tennis ball thrown into the air. Even with the tremendous success that Newton's law of gravitation acquired it was not complete. Eq. (1.1) is purely descriptive. It can do a fantastic job in predicting gravitational phenomena but it was not known why nature chose such a force law in the first place. In other words, there was no *mechanism*.

The crucial insight was provided by Albert Einstein three-hundred years later with his theory of General Relativity. This introductory section will give an overview of some of the key ideas and machinery that are needed to understand this new revolutionary view of the Universe. There are of course many excellent introductory reviews and books on the subject which the unfamiliar reader is encouraged to read [20–22, 30].

We shall begin with an observation that a gravitational force can seemingly be removed by a transformation to an accelerated frame of reference. In other words, in a sufficiently small region, there is no experiment that can determine whether an observer is in a non-inertial frame of reference or a gravitational field. Consider an object inertial mass m_I and gravitational mass m_g placed in an elevator which is itself situated in a gravitational field. The equation of motion is given by

$$m_I \frac{d^2 x}{dt^2} = \tilde{F} - m_g g \,. \tag{1.2}$$

For simplicity we have assumed that the mass can only move in the direction x aligned with the axis of the gravitational acceleration g. $\tilde{F}$ represents any other forces which may be acting on the mass, such as electromagnetic forces if it happens to be charged. The gravitational mass m_g can be thought of as the gravitational "charge" of the particle. By performing a coordinate transformation along the direction of motion in terms of the new coordinate x'

$$x' = x + \frac{1}{2} g t^2 \,, \tag{1.3}$$

the equation of motion becomes

$$m_I \frac{d^2x'}{dt^2} = \tilde{F} + (m_I - m_g)g\,. \tag{1.4}$$

Equation (1.4) shows that if the inertial mass m_I is equal to the gravitational mass m_g then the particle's equation of motion does not include any gravitational forces. Applying this argument in reverse, an inertial frame of reference described by a nonlinear coordinate system may appear to have gravitational forces even in the absence of mass. As an example, a particle which moves along a straight line in Cartesian coordinates appears to experience a repulsive force away from the origin when the same motion is expressed in polar coordinates. This repulsive force, known as centrifugal force, is simply an artefact of using a curved coordinates to describe linear motion. This notion of choosing "bad" coordinate systems to describe theories occurs time and time again in theoretical physics and it is important to understand whether a particular physical quantity really is physical or whether it can be removed by the freedom that exists in the theory. The equivalence between inertial mass and gravitational mass was well known before Einstein from Galileo's experiments rolling balls down inclined planes. The difference was that Einstein was the first person to take it seriously, elevating this equivalence into a principle called the *equivalence principle*. We shall discuss distinctions between different versions of this principle in Sec. 2.2.4.

A further motivation for Einstein was an apparent contradiction between Newtonian gravity and Maxwell's theory of electromagnetism. Maxwell's theory predicts that the electric force F_{E} experienced between two charges q_1 and q_2 placed at a distance r is given by Coulomb's law

$$F_{\mathrm{E}} = \frac{kq_1q_2}{r^2}\,, \tag{1.5}$$

where k is a constant of proportionality. The similarities between Eqs. (1.5) and (1.1) seem highly suggestive. Eq. (1.5) can be derived from Maxwell's electromagnetic field equations which describe all electromagnetic phenomena in terms of a vector field $A_\mu(x,t)$. The static limit of these field equations yields Eq. (1.5). If Coulomb's law can be derived from a field theory, then perhaps it is possible to derive the law of Newtonian gravity in Eq. (1.1) from a field theory as well. The importance of this fact was recognised by Newton, albeit without the mathematical machinery of a field to quantify his intuition. Newton's theory of gravity relied on "action at a distance" whereby the gravitational force experienced by a mass m_1 is exerted seemingly instantaneously when another mass m_2 is placed in its vicinity. This instantaneousness troubled Newton, noting that [31]

> *That Gravity should be innate ... 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.*

Fortunately, his readers included James Clerk Maxwell and Michael Faraday. Action at a distance is avoided in Maxwell's theory as all electromagnetic interactions are propagated via the electromagnetic field, or light. Einstein's special theory of relativity forbade any signals which travel faster than the speed of light. It was therefore of the utmost importance that a field theory for gravitation was developed which did not include action at a distance, and in the static limit reduced to Eq. (1.1). This field theory was provided by GR.

Before commencing our discussion of GR we stress an important point. It is always possible to perform a nonlinear coordinate transformation to remove the gravitational field *in a sufficiently small region.* For an extended mass m_1 which is near another mass m_2 the gravitational force experienced at one point of m_1 will be different to another position. A real gravitational force should induce tidal effects. There is no coordinate transformation which can remove tidal forces. Let us assume that there are two particles freely falling in a gravitational potential $\Phi(x_i)$ separated by a distance Δx^i. The relative acceleration experienced between the two objects is given by

$$\begin{aligned} \Delta\ddot{x}_i &= -\nabla_i\Phi(x_i+\Delta x_i)+\nabla_i\Phi(x_i)\,, \\ &= -\Delta x^j\nabla_i\nabla_j\Phi\,. \end{aligned} \tag{1.6}$$

Eq.(1.6) demonstrates that the relative acceleration experienced by two bodies is given by the second derivative of the gravitational potential. The fact that it is the second derivative was a key insight into connection between gravity and geometry.

1.2.2 Gravity and geometry

The observation that any local non-inertial frame of reference can be transformed to an inertial frame with a nonlinear coordinate transformation lends itself to a neat geometrical interpretation. It makes use of the framework of Riemannian geometry which provided Einstein with the mathematical machinery to construct GR. In this section we shall briefly review some of the key elements of Riemannian geometry and GR while referring the reader to Refs. [30, 32, 33] for a more complete discussion.

Riemannian geometry is a generalisation of Euclidean geometry to curved spaces. The main geometrical object in Riemannian geometry is a *manifold*. Formally, an n-dimensional manifold M is a collection of open sets U_α with a mapping ψ from each open set to n-dimensional Euclidean space $\mathbb{R}^n$ such that $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ which satisfies three properties:

- There is an open set U_α which encompasses every element of $p \in M$.
- Every open subset $U_\alpha \in M$ can be associated with a bijective map onto an open subset of $\mathbb{R}^n$ such that $\psi_\alpha : U_\alpha \to V_\alpha$ where $V_\alpha \in \mathbb{R}^n$.
- For any points in M which belong to the intersection of two open subsets of M there is a mapping from $\mathbb{R}^n \to \mathbb{R}^n$ which maps from $\psi_\alpha \left[U_\alpha \cap U_\beta\right] \to \psi_\beta \left[U_\alpha \cap U_\beta\right]$.

Each map ψ_α is called a chart or a coordinate system. The existence of a map from each point in a manifold to $\mathbb{R}^n$ expresses mathematically the notion that locally a manifold resembles flat space, despite potentially possessing non-zero curvature globally. It is this mathematical structure which enabled Einstein to formulate his physical intuition of the equivalence principle into a quantitative theory.

Let us review what is meant by flat space. In special relativity, the infinitesimal distance between two points in spacetime is given by the Minkowski line element

$$\begin{aligned} ds^2 &= \eta_{\mu\nu} dx^\mu dx^\nu \,, \\ &= -dt^2 + dx^2 + dy^2 + dz^2 \,, \end{aligned} \tag{1.7}$$

where $\eta_{\mu\nu} = (-1, 1, 1, 1)$. Two parallel straight lines in this space will never intersect, a statement which may not hold for a general surface. For example, the infinitesimal distance between any two points on a sphere of radius r is given by

$$ds^2 = r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \,, \tag{1.8}$$

where θ and ϕ are the polar and azimuthal angles respectively. Two parallel lines drawn from the equator to a pole will inevitably intersect. A sphere is of course a curved surface but it is important to properly define this notion. After performing a nonlinear coordinate transformation in Eq. (1.7) it may not be obvious at all that the space is flat. How is it possible to distinguish between a flat space written in a nonlinear coordinate system and a genuinely curved space like a sphere? This distinction is completely parallel to the question of how one can distinguish between non-inertial forces and gravitational forces as discussed in the previous section.

The line element on a general space can be written as

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \,. \tag{1.9}$$

where $g_{\mu\nu}(x)$ is called the metric tensor. It is a function of the coordinates and contains all the necessary information to describe the geometrical properties of the manifold. In particular, we shall describe how the specification of a metric tensor determines whether a space is flat or curved. Under a coordinate transformation $x^\mu \to \tilde{x}^\mu$ the components of the metric transform as

$$\tilde{g}_{\mu\nu} = \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma} \,. \tag{1.10}$$

Note that a tensor in itself is a geometrical object that is defined independently of the coordinate system used to describe it. A vector is a simple example of a tensor with one index. Under a coordinate transformation the components of the vector and the basis vectors both transform in such a way as to leave the vector itself unchanged. We shall briefly give a specific example by transforming flat three dimensional Euclidean space into spherical polar coordinates (r, θ, ϕ) related via

$$x = r \sin\theta \cos\phi \;, \;\; y = r \sin\theta \sin\phi \;, \;\; z = r \cos\theta \,. \tag{1.11}$$

Each component of the metric in the spherical polar coordinate system must then be computed using equation (1.10). For example, the g_{rr} component becomes

$$g_{rr} = \left(\frac{\partial x}{\partial r}\right)^2 g_{xx} + \left(\frac{\partial y}{\partial r}\right)^2 g_{yy} + \left(\frac{\partial z}{\partial r}\right)^2 g_{zz}\,,$$
$$= 1\,. \tag{1.12}$$

By computing the other components in the same manner the metric of flat space written in spherical polars is given by

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2\,. \tag{1.13}$$

Note the similarity of this equation with that in Eq. (1.8). Of course these two metrics describe spaces of different dimension but it demonstrates the care that must be taken when distinguishing between the metric describing a genuinely curved space, such as Eq. (1.8), and the metric of a flat space written in a nonlinear coordinate system, such as Eq. (1.13).

We shall now demonstrate that in a sufficiently small region the metric is equivalent to that of a flat space. This geometrical fact is the crucial ingredient in the formulation of the equivalence principle. For example, fixing the polar angle on the sphere to be $\theta = \pi/2$ the line element (1.8) becomes $r^2\left(d\theta^2 + d\phi^2\right)$. After re-scaling the coordinates by r so that $\tilde{\theta} = r\theta$ and $\tilde{\phi} = r\phi$, Eq. (1.8) becomes $d\tilde{\theta}^2 + d\tilde{\phi}^2$. So by considering the form of the metric around the local region $\theta = \pi/2$ one finds that the sphere is flat. For a general four-dimensional spacetime metric $g_{\mu\nu}$, locally around any point $x^\mu = 0$ we must have that $g_{\mu\nu} = \eta_{\mu\nu}$. The condition for this to hold is

$$\partial_\sigma g_{\mu\nu} = 0\,. \tag{1.14}$$

Eq. (1.14) is a compact way of writing forty equations, with four derivatives of ten independent components of the metric. At quadratic order, a general coordinate transformation from $x^\mu \to y^\mu$ can be expressed as

$$x^\mu = y^\mu + C^\mu_{\rho\sigma} y^\rho y^\sigma\,, \tag{1.15}$$

where we have assumed that locally the y coordinate system approximately equals the x coordinate system. The terms $C^\mu_{\rho\sigma}$ do not necessarily transform as tensors, but are simply numbers encoding all possible quadratic expansions between the coordinate systems. It is symmetric in ρ and σ and therefore has forty components in four dimensions. The condition for the metric to be locally flat then amounts to solving forty equations in Eq. (1.14) for forty unknowns in Eq. (1.15). It is not possible in general to extend this calculation to higher order by requiring higher derivatives of the metric to vanish, a fact which has important physical consequences.

A key concept in understanding the properties of curved spaces is the covariant derivative. This notion generalises the idea of a derivative to curved spaces and nonlinear coordinate systems. To highlight the necessity of generalising the notion of a derivative let us take an ordinary partial derivative of the components of a vector in a coordinate system $\tilde{x}^\mu$ and see how it transforms when written in terms of the coordinate system x^μ

$$\frac{\partial \tilde{V}^\mu}{\partial \tilde{x}^\nu} = \frac{\partial}{\partial \tilde{x}^\nu}\left(\frac{\partial \tilde{x}^\mu}{\partial x^\sigma} V^\sigma\right) . \tag{1.16}$$

Because the $\partial/\partial\tilde{x}^\nu$ operator acts on both terms inside the parentheses, there is an extra term which means it does not transform as a tensor. The partial derivative of a vector is therefore not an invariant notion and different observers using different coordinate systems will not agree on its value. The extra term that appears in Eq.(1.16) can be accounted for with the addition of another quantity in the definition of the derivative. The *covariant derivative* is defined as

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\sigma} V^\sigma . \tag{1.17}$$

where $\Gamma^\mu_{\mu\sigma}$ are the Christoffel connections. If equation (1.14) can hold in an arbitrary coordinate system as we have shown, then it must be the case that

$$\nabla_\mu g_{\rho\sigma} = 0 . \tag{1.18}$$

This condition can be used to derive the form of the Christoffel connections which are given by

$$\Gamma^\nu_{\mu\sigma} = \frac{1}{2} g^{\lambda\nu}\left(\partial_\mu g_{\lambda\sigma} + \partial_\sigma g_{\mu\lambda} - \partial_\lambda g_{\mu\sigma}\right) . \tag{1.19}$$

With this definition it is then a lengthy but straightforward exercise to check that under a coordinate transformation $\nabla_\mu V^\nu$ does indeed transform as a tensor. Note that the Christoffel connections $\Gamma^\nu_{\mu\sigma}$ can be defined independently of the metric. It is only after imposing the condition (1.18) that they relate directly to the metric.

Derivatives measure the rate at which a quantity changes. In order for this to make sense it is necessary to define what a quantity is changing with respect to. For the case of the covariant derivative in equation (1.17) it the rate of change of a vector along a curve on the manifold relative to whether it had been parallel transported along the curve. Consider a curve on a manifold parameterised by λ given by $x^\mu(\lambda)$. A vector V^ν is parallel transported along the curve if its direction does not change as it moves, or more quantitatively

$$\frac{dx^\mu}{d\lambda}\nabla_\mu V^\nu = 0\,, \tag{1.20}$$

which can be thought of as projecting the covariant derivative $\nabla_\mu V^\nu$ along the direction of the tangent vector to the curve $\frac{dx^\mu}{d\lambda}$. If the vector is itself the tangent vector such that $V^\nu = dx^\nu/d\lambda$ then the condition for parallel transport becomes

$$\frac{d^2x^\nu}{d\lambda^2} + \Gamma^\nu_{\mu\sigma}\frac{dx^\mu}{d\lambda}\frac{dx^\sigma}{d\lambda} = 0\,. \tag{1.21}$$

This is the *geodesic equation* which determines the motion of objects in a curved space in the absence of external forces. It can also be derived by minimising the action of a free particle in a curved space, given by the integral of the proper time along the path of the particle $S = \int d\tau$ (see for example Ref. [22]). Armed with the notion of a derivative on a curved manifold, we possess the machinery to determine whether a manifold is truly curved, or whether it is flat with a nonlinear coordinate system. The curvature of a space is quantified through the Riemann tensor which can be derived as follows. Consider the parallel transport of a vector around a closed loop on a curved space, as in Fig. 1.1. On a flat space the vector should not change direction when it returns to its starting position. On a curved space this does not happen. The degree to which the vector changes is quantified by the Riemann curvature tensor.[2] The commutator of two derivative operators acting on a vector quantifies how much a vector changes relative to whether it was parallel transported along x^μ followed by x^ν compared to x^ν followed by x^μ. This change is determined by the Riemann curvature tensor $R^\rho{}_{\sigma\mu\nu}$ defined by

$$[\nabla_\mu, \nabla_\nu]\, V^\rho = R^\rho{}_{\sigma\mu\nu} V^\sigma\,. \tag{1.22}$$

If this commutator is non-zero then the space is not flat. Explicitly $R^\rho{}_{\sigma\mu\nu}$ is given by

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}\,. \tag{1.23}$$

[2] Note that we neglect torsion.

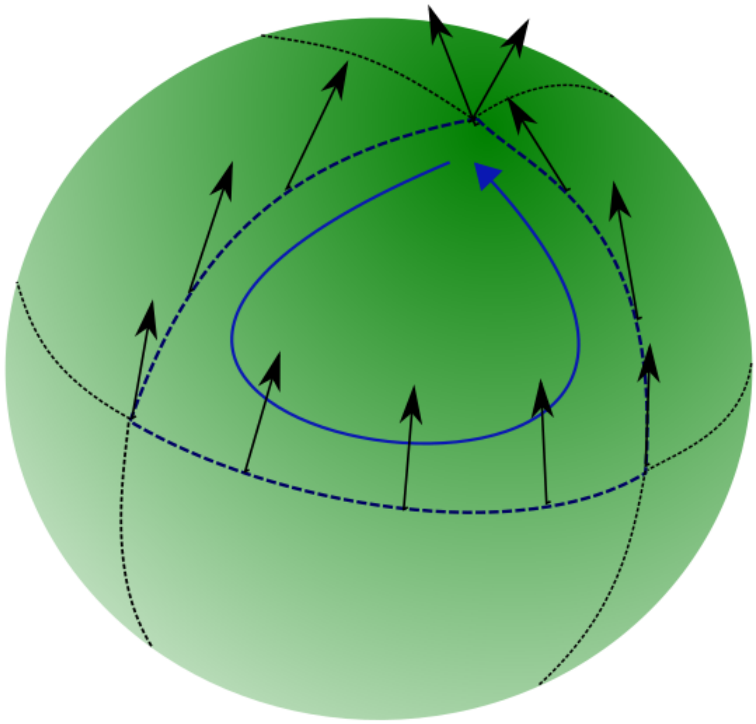

Figure 1.1 *Parallel transport of a vector along a closed curve on a sphere. The direction of the vector has changed when it returns to its starting point, a situation which would not occur on a flat space.*

Observe that the Riemann curvature tensor depends on the second derivatives of the metric. Now recall that the presence of a gravitational force depended on the second derivative of the gravitational potential, i.e. the presence of tidal forces. This was a crucial insight that enabled Einstein to make the connection between gravity and geometry. Gravity is not a force per-se, it is simply an artefact of the fact that spacetime can bend. A gravitational field does not exist on spacetime, as Maxwell's electromagnetic field does, the gravitational field is *itself spacetime*. They are one and the same. Not only does GR provide a mechanism for Newton's gravitational force, but it also gives the deeper insight that spacetime is also a field, just like the electromagnetic field. For this reason, GR is often considered among the most beautiful of physical theories.

There is one key element missing. Spacetime does not bend of its own accord. A source of stress-energy is required, such as a planet, star, galaxy or a student writing their PhD thesis. The stress energy tensor of a system of particles is a tensor with two spacetime indices $T_{\mu\nu}$. This quantity acts as the source for spacetime curvature. Taking the trace of $R^{\rho}{}_{\sigma\mu\nu}$ gives the Ricci tensor $R_{\mu\nu}$ which provides a natural ansatz for the relation between spacetime curvature and the stress-energy tensor given by $R_{\mu\nu} \propto T_{\mu\nu}$. Conservation of the stress-energy tensor requires that that $\nabla_\mu T^{\mu\nu} = 0$ and unfortunately $\nabla_\mu R^{\mu\nu} \neq 0$ so this ansatz cannot hold. There is however a specific combination of geometrical terms whose covariant derivative does vanish which can be determined from the Bianchi identity $\nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\right) = 0$. This can therefore allow us to write out a consistent relation between spacetime curvature and stress-energy as

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}\,, \tag{1.24}$$

where the constants on the right hand side are determined from the requirement that the theory possesses a valid Newtonian limit [21, 22]. The quantity $R = R^\mu{}_\mu$ is called the Ricci scalar. Eq. (1.24) is the final form of the Einstein equation relating the curvature of spacetime on the left hand side to the presence of a source of stress-energy on the right hand side.

It is also possible to derive Eq. (1.24) from an action principle. By varying the metric $g_\mu \to g_{\mu\nu} + \delta g_{\mu\nu}$ in the so-called Einstein-Hilbert action

$$S_{EH} = \frac{M_*^2}{2} \int d^4x \sqrt{-g}\, R\,, \tag{1.25}$$

and setting $\delta S_{EH} = 0$ the equation of motion (1.24) is obtained with $T_{\mu\nu} = 0$. Note we also write the gravitational constant G in terms of the Planck mass M_*, related via $M_*^2 = 1/8\pi G$ and we work in units where $c = \hbar = 1$ throughout this thesis. With the addition of a term describing the matter sector S_m in Eq. (1.25) the full Einstein equation (1.24) can be obtained with $T_{\mu\nu}$ determined from

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_m}{\delta g^{\mu\nu}}\,. \tag{1.26}$$

The Einstein equations predicts the existence of a wealth of new exotic phenomena previously unimaginable to 19th century physicists such as gravitational waves and the expansion of the Universe.[3] As this thesis is primarily concerned with the accelerated expansion of the Universe, the next section will apply GR to the Universe as whole.

[3]It is interesting to note that the first mention of an astronomical object resembling a black hole came in 1784 by John Michell, 131 years before General Relativity. He considered the situation of a massive body with a gravitational pull so great that not even light could not escape it [34]. Many thanks to Ed Copeland for pointing this out.

1.3 Cosmology

1.3.1 The background Universe

In this section we shall discuss the application of GR to cosmology. The first step is to apply the cosmological principle, which states that on large scales the Universe is homogeneous and isotropic. They are distinct concepts. For example, all points on the surface of a cylinder are equivalent, but there are two distinct directions, one leading to the end of the cylinder and the other back to the original position. It is therefore homogeneous but not isotropic. The size of the observable Universe is estimated to be around 3000Mpc and observations suggest that the cosmological principle holds above 100Mpc [35]. In fact, the isotropy of the Universe has recently been tested using the CMB to roughly one part in 10^5 [36]. For the purposes of the above analysis we shall assume that the cosmological principle does indeed hold on the largest scales.

In order to apply GR to the Universe it is necessary to find a solution to the Einstein equations which is both homogeneous and isotropic. The condition that the spatial metric be homogeneous and isotropic greatly reduces the number of admissible geometries to being either flat, a sphere with positive spatial curvature or a hyperboloid with constant negative spatial curvature. These spaces are maximally symmetric, in that they possess the symmetry group of every translation and rotation in the space. This discussion closely follows examples given in Ref. [37].

We shall now derive the form of the metric in a maximally symmetric space by embedding a three dimensional sphere of radius a in four dimensional Euclidean space. The defining equation of this surface is given by

$$x^2 + y^2 + z^2 + w^2 = a^2 \,. \tag{1.27}$$

By taking the differential of Eq. (1.27)

$$xdx + ydy + zdz + wdw = 0 \,, \tag{1.28}$$

and using Eq. (1.28) to eliminate dw we obtain the induced metric of the sphere in four dimensional Euclidean space

$$ds^2 = dx^2 + dy^2 + dz^2 + \frac{(xdx + ydy + zdz)^2}{a^2 - x^2 - y^2 - z^2}\,. \tag{1.29}$$

This expression can be greatly simplified by transforming to spherical polar coordinates $x = \tilde{r}\sin\theta\cos\phi$, $y = \tilde{r}\sin\theta\sin\phi$ and $z = \tilde{r}\cos\theta$ where it becomes

$$ds^2 = \left(1 - \frac{\tilde{r}^2}{a^2}\right)^{-1} d\tilde{r}^2 + \tilde{r}^2 d\theta^2 + \tilde{r}^2\sin^2\theta d\phi^2. \tag{1.30}$$

After re-defining the radial coordinate $\tilde{r}$ such that $r = \tilde{r}/a$ one obtains

$$ds^2 = a^2\left[\frac{dr^2}{1 - r^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \tag{1.31}$$

which is the metric of a space of constant positive curvature. By reversing the sign of a^2, or indeed by setting a^2 to zero in Eq. (1.27) one can repeat this calculation to obtain the metric of a three dimensional space of constant negative curvature or zero curvature. The general metric is given by

$$ds^2 = a^2\left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \tag{1.32}$$

where k is a constant that takes the value $+1$ in a positively curved space, -1 in a negatively curved space and is 0 in a flat space. Furthermore, it is useful to define a new coordinate χ such that

$$d\chi^2 = \frac{dr^2}{1 - kr^2}\,. \tag{1.33}$$

The relationship between r and χ then takes the form

$$r = S_k(\chi) = \begin{cases} \sinh\chi, & k = -1 \\ \chi, & k = 0 \\ \sin\chi, & k = +1 \end{cases}\,, \tag{1.34}$$

and Eq. (1.32) becomes

$$ds^2 = a^2\left(d\chi^2 + S_k^2(\chi)d\Omega^2\right), \tag{1.35}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the metric of a two dimensional sphere. Even though Eq. (1.35) was derived from embedding a constant curvature surface in four dimensional Euclidean space, it does not rely on this embedding for its definition. It is then straightforward to write down the metric of four dimensional Minkowski space with constant spatial curvature in the following way as

$$ds^2 = -dt^2 + a^2(t)\left(d\chi^2 + S_k^2(\chi)d\Omega^2\right) , \tag{1.36}$$

where the radius of curvature a has been promoted to a function of time $a(t)$. This is the *Friedman-Lemaître-Robertson-Walker* (FLRW) metric. It is the unique metric of a homogeneous and isotropic space with a time coordinate t, and is therefore a useful ansatz for the metric describing the Universe on the largest scales. One can also define the conformal time τ with $d\tau = dt/a(t)$ so placing the temporal and spatial coordinates on an equal footing. The metric in Eq. (1.36) then becomes

$$ds^2 = a^2(\tau)\left[-d\tau^2 + d\mathbf{x}^2\right] , \tag{1.37}$$

where we have written the spatial components of the metric in a general coordinate system $\mathbf{x}$.

Dynamics of the homogeneous Universe

Now that we have constructed the metric of a homogeneous and isotropic spacetime in Eq. (1.36) it is important to check that it is a solution to the Einstein equations. The only dynamical quantity that appears in the metric is the scale factor $a(t)$ which is related to the time dependent stress-energy content of the Universe through two equations called the Friedmann equations. See Ref. [38] for a more detailed derivation.

Computing the Christoffel symbols for the metric (1.36) one can obtain the corresponding Riemann curvature tensor from Eq. (1.23). For example, one can show that

$$\Gamma^0_{ij} = a^2 H h_{ij} \ , \ \Gamma^i_{0j} = H\delta^i_j \, , \tag{1.38}$$

where h_{ij} is the spatial metric and $H(t) \equiv \dot{a}/a$ is the Hubble parameter. Once every Christoffel connection has been computed the Ricci tensor and scalar can be determined. The components of the Ricci tensor are given by

$$R_{00} = -3\left(\dot{H} + H^2\right) , \tag{1.39}$$

$$R_{0i} = R_{i0} = 0 \, , \tag{1.40}$$

$$R_{ij} = a^2\left(3H^2 + \dot{H} + \tfrac{2k}{a^2}\right) h_{ij} \, , \tag{1.41}$$

and the Ricci scalar is found to be

$$R = 6\left(2H^2 + \dot{H} + \frac{k}{a^2}\right) . \tag{1.42}$$

The metric is of course just one ingredient in the Einstein equation. It is also necessary to include the contribution of the stress-energy tensor $T_{\mu\nu}$. On cosmological scales it is a good approximation to treat matter as a perfect fluid with an energy-momentum tensor of the form

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - p g_{\mu\nu} , \tag{1.43}$$

where u^μ is the four-velocity of a fluid element which is given by $u^\mu = (-1, 0, 0, 0)$ in the rest frame. The energy density ρ is related to the pressure p through the equation of state parameter

$$w = \frac{p}{\rho} . \tag{1.44}$$

Dark matter and baryons have an equation of state $w = 0$ and radiation has $w = 1/3$. We can now determine the background metric evolution when the Universe is dominated by a matter species with a general equation of state w. Plugging in the stress-energy tensor (1.43) into the right hand side of the Einstein equation and using the components the Ricci tensor for an FRW metric in equation (1.36) we find

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} , \tag{1.45}$$

$$3H^2 + 2\dot{H} = -8\pi G p - \frac{k}{a^2} . \tag{1.46}$$

Eliminating H from these equations it is possible to show that

$$\dot{\rho} + 3H(\rho + p) = 0 . \tag{1.47}$$

This is the continuity equation which can also be derived as the time component of stress-energy conservation $\nabla_\mu T^{\mu\nu} = 0$. Equations (1.45) and (1.46) are the Friedmann equations which can be solved to obtain $H(t)$ for a given $\rho(t)$. The continuity equation (1.47) can be solved for a general matter species with an equation of state w to give

$$\rho \propto a^{-3(1+w)} , \tag{1.48}$$

which shows that the energy density of matter scales as $\rho_m \propto a^{-3}$ and radiation as $\rho_r \propto a^{-4}$. It also shows that if there is a matter contribution which has a constant energy density $\dot{\rho} = 0$ this implies $w = -1$. As ρ must be positive this implies that a constant energy density is associated with negative pressure. The cosmological constant is by definition a constant energy density, and from Eq. (1.46) one can see that its associated negative pressure implies $\ddot{a} > 0$, i.e. accelerated expansion. Using equation (1.45) one can determine the energy density required to have a flat Universe with $k = 0$. This is called the *critical density* and is given by

$$\rho_{crit0} = \frac{3H_0^2}{8\pi G}\,, \tag{1.49}$$

where H_0 is the present day value of the Hubble parameter. Using this critical value we can define present day density parameters

$$\Omega_{i0} \equiv \frac{\rho_{i0}}{\rho_{crit0}}\,. \tag{1.50}$$

Now let us assume that the Universe is made up of matter, radiation, a cosmological constant $w = -1$ and spatial curvature. Equation (1.45) can then be expressed in terms of each Ω_{i0} in the following way

$$H^2/H_0^2 = \Omega_{m0}a^{-3} + \Omega_{r0}a^{-4} + \Omega_{\Lambda 0} + \Omega_{k0}a^{-2}\,, \tag{1.51}$$

where $\Omega_k = -k/H_0^2$ is the curvature density parameter. A useful way of interpreting equation (1.45) can be seen by re-writing equation (1.45) as

$$\dot{a}^2 \propto \Omega_{m0}a^{-1} + \Omega_{r0}a^{-2} + \Omega_{\Lambda 0}a^2 + \kappa\,, \tag{1.52}$$

where κ is a constant. Written in this way the Friedmann equation looks like a Hamiltonian for the scale factor with the first three terms on the right hand side interpreted as an effective potential $-\Phi_{eff}(a)$. If Ω_Λ is the dominant contribution then it reduces to $\dot{a}^2 \propto a^2$ which has accelerated expansion of $a(t)$ as a solution. In this way the dynamics of the background expansion of the Universe can be considered as being equivalent to a particle moving in a potential.

As the contribution from radiation is negligible at late times, determining the cosmology of the Universe reduces to determining where the Universe we live in happens to sit in the parameter space of Ω_{m0} and $\Omega_{\Lambda 0}$. The alternate possibilities are illustrated in Fig. 1.2. For example, the absence of spatial curvature implies that $\Omega_m + \Omega_\Lambda = 1$. Anywhere above (below) this line corresponds to an open (closed) Universe. As an open, closed and flat Universe all evolve differently, by determining the shape of the potential in equation (1.52) each point on this plane

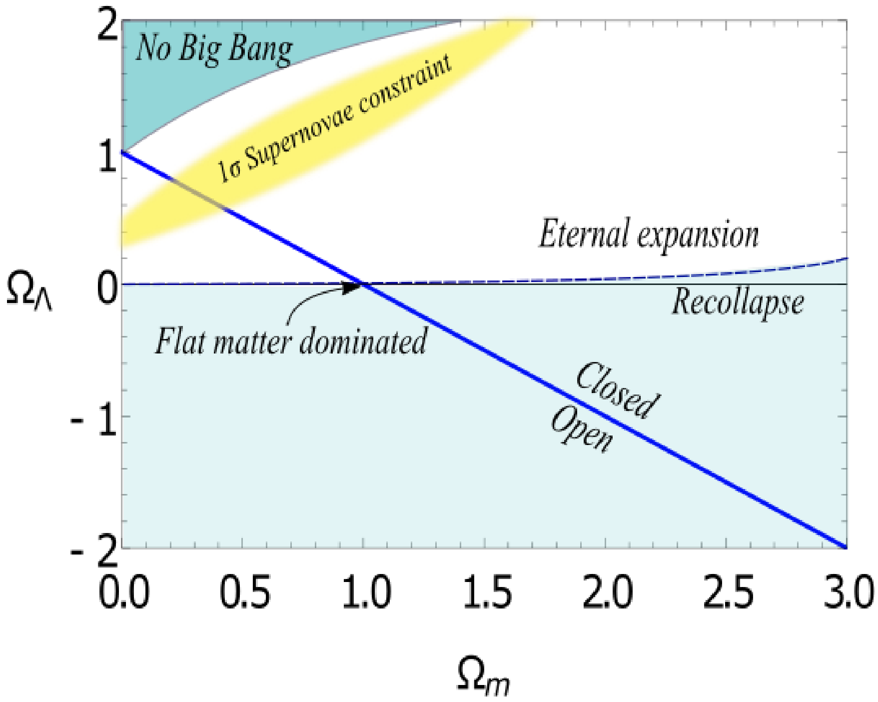

Figure 1.2 *The expansion history of the Universe is determined from points in the $(\Omega_\Lambda, \Omega_m)$ plane. The solid blue line $\Omega_\Lambda + \Omega_m = 1$ corresponds to a flat Universe, with all points above belonging to a closed Universe $k > 0$ and those below to an open Universe $k < 0$. All points above the dashed blue line correspond to a Universe which will never cease expanding whereas points below correspond to Universes which will reach a peak radius and then re-collapse. Combing the supernovae constraint with the CMB results of $k \approx 0$ results in a Universe with $\Omega_\Lambda \approx 0.7$ and $\Omega_m \approx 0.3$.*

geometry and the future evolution of the Universe. Also shown in figure 1.2 are the constraints arising from the type Ia Supernovae (see Sec.1.3.3). It is clear that the Supernovae result alone do not determine that the Universe is dominated by Ω_Λ. It could either be an open matter dominated Universe or a closed Universe with significant contributions from matter and a cosmological constant. Only when the spatial curvature of the Universe was determined to be flat from the cosmic microwave background radiation (see Sec. 1.3.3) was it clear that Universe we live in lay close to $\Omega_\Lambda \approx 0.7$ and $\Omega_m \approx 0.3$. As the density of baryons was determined from Big Bang Nucleosynthesis to be $\Omega_b \approx 0.04$ this also provided strong evidence that the majority of the matter was in the form of dark matter which does not interact with electromagnetic radiation.

The Friedmann equation (1.45) and (1.46) determine the background expansion history of the Universe. Departures from homogeneity and isotropy arising from the presence of structures in the Universe are discussed in the next section. We then conclude this introductory chapter with a discussion of the observational basis for the present consensus on the stress-energy content of the Universe.

1.3.2 The perturbed Universe

SVT decomposition

In the previous section we examined the dynamics of the Universe when it was treated as entirely homogeneous and isotropic. This assumption must break down at some length scale. The Universe has structure in the form of dark matter halos, filaments, galaxy clusters and super-clusters. In order to describe a Universe with this structure it is necessary to go beyond the assumptions of homogeneity and isotropy. This section discusses the important machinery that is needed to account for the presence of structure on top of a homogeneous and isotropic Universe. We do not discuss the highly nonlinear regime where models of spherical collapse are necessary (see, for example, Ref. [39] Sec.7.5) but instead restrict to a regime where structure can be treated as small perturbations on top of a background FLRW Universe.

The nonlinear nature of the Einstein equations suggests that this approach could be challenging. Fortunately, the story greatly simplifies at the level of first-order perturbations through a mathematical trick called a scalar-vector-tensor (SVT) decomposition. In this approach the various degrees of freedom separate into components which all evolve independently. Due to the importance of being able to perform such a decomposition for the formulation of linear perturbation theory, we shall now justify it in some detail. The proceeding discussion closely follows that presented in the appendix of Ref. [40].

Consider a general perturbation δQ written in Fourier space

$$\delta Q(t,\mathbf{k}) = \int d^3\mathbf{x}\ \delta Q(t,\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}\,. \tag{1.53}$$

We shall show that each Fourier mode of the perturbation evolves independently because of translational invariance. Let us assume that the evolution of the perturbation is determined by an equation, in this case the Einstein equation, that determines possible couplings between different modes as well as determining the time dependence

$$\delta Q_i(t_2,\mathbf{k}) = \sum_{j=1}^{N} \int d^3\bar{\mathbf{k}}\, T_{ij}(t_2,t_1,\mathbf{k},\bar{\mathbf{k}})\, \delta Q_j(t_1,\bar{\mathbf{k}})\,. \tag{1.54}$$

Here, T_{ij} is the matrix which involves the relation between the modes $\mathbf{k}$ and $\bar{\mathbf{k}}$ from time t_1 to t_2 which is determined from the evolution equation. Shifting the coordinates to a new frame $x^{i'} = x^i + \Delta x^i$ results in an extra phase factor, as can be seen from

$$\delta Q'(t,\mathbf{k}) = \int d^3\mathbf{x}' \delta Q(t,\mathbf{x}') e^{-i\mathbf{k}\cdot\mathbf{x}'}\,, \tag{1.55}$$

$$= e^{-i\mathbf{k}\cdot\Delta\mathbf{x}} \int d^3\mathbf{x}\ \delta Q(t,\mathbf{x}') e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \tag{1.56}$$

$$= e^{-i\mathbf{k}\cdot\Delta\mathbf{x}} \int d^3\mathbf{x}\ \delta Q(t,\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \tag{1.57}$$

where the final line was obtained through the invariance of the integral after the shift in coordinates. Therefore

$$\delta Q'(t,\mathbf{k}) = e^{-ik_j\Delta x^j}\,\delta Q(t,\mathbf{k})\,, \tag{1.58}$$

where we have written $\mathbf{k}$ and Δx in terms of their components k^i and Δx^i. Plugging this result in Eq. (1.54) the equation of motion transforms as

$$\delta Q_i'(t_2,\mathbf{k}) = \sum_{j=1}^{N} \int d^3\bar{\mathbf{k}}\, T_{ij}(t_2,t_1,\mathbf{k},\bar{\mathbf{k}})\, e^{-i(k_i-\bar{k}_i)\Delta x^i} \delta Q_j'(t_1,\bar{\mathbf{k}})\,. \tag{1.59}$$

This expression must be equal to the original expression with the transfer matrix also in the primed frame

$$\delta Q'_i(t_2, \mathbf{k}) = \sum_{j=1}^{N} \int d^3\bar{\mathbf{k}}\, T'_{ij}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}})\, \delta Q'_j(t_1, \bar{\mathbf{k}})\,. \tag{1.60}$$

As the equations of motion must be the same before and after the translation from the symmetry of the underlying action from which they were derived, such as the Einstein-Hilbert action in Eq. (1.25), it must be the case that

$$T'_{ij}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}) = T_{ij}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}})\, e^{-i(k_i - \bar{k}_i)\Delta x^i}\,. \tag{1.61}$$

This can only be true if the transfer matrix is zero, or if $k_i = \bar{k}_i$. In other words, the transfer matrix is diagonal and so on linear scales each Fourier mode evolves independently. Going beyond linear scales results in more complicated relations between Fourier modes where Eq. (1.61) does not hold. The modes become *coupled* on nonlinear scales.

Now that we have established that each mode of a linear perturbation evolves independently through translational invariance, let us explore a similar argument how this argument extends with another symmetry in the action, namely rotational invariance. If following a rotation by an angle ψ a perturbation changes by $e^{im\psi}$ then it is a perturbation of *helicity* m. Rotational invariance allows us to set the wavevector of the perturbation to be $\mathbf{k} = (0, 0, k)$ such that the only spatial dependence of the perturbation is along the x^3 axis with the factor e^{ikx^3}. Rotations around $\mathbf{k}$ are simpler in an alternate basis for the unit vectors in the two orthogonal directions in Fourier space $\mathbf{e}_1$ and $\mathbf{e}_2$ defined by

$$\mathbf{e}_\pm \equiv \frac{\mathbf{e}_1 \pm i\mathbf{e}_2}{\sqrt{2}}\,, \tag{1.62}$$

where the basis vector along the k_3 direction $\mathbf{e}_3$ is kept fixed. Under a standard rotation, for example with $x'_1 = x_1 \cos\psi + x_2 \sin\psi$ the new basis vectors transform as

$$\mathbf{e}'_\pm = e^{\pm i\psi}\mathbf{e}_\pm\,. \tag{1.63}$$

This gives the coordinate transformations

$$\frac{\partial x^{+\prime}}{\partial x^+} = e^{i\psi}\,,\ \frac{\partial x^{+\prime}}{\partial x^-} = 0\,,\ \frac{\partial x^{-\prime}}{\partial x^+} = 0\,,\ \frac{\partial x^{-\prime}}{\partial x^-} = e^{-i\psi}\,. \tag{1.64}$$

A tensor with a number of indices will then transform under a rotation as

$$T'_{i_1 i_2 \ldots i_n} = e^{i(n_+ - n_-)\psi} T_{i_1 i_2 \ldots i_n} \,. \tag{1.65}$$

For example we would have that

$$T'_{++} = e^{2i\psi} T_{++} \,, \tag{1.66}$$

showing that T_{++} is a helicity 2 object. Similarly, as $\mathbf{e}_3$ is unchanged after a rotation we have that

$$T'_{3-} = e^{-i\psi} T_{3-} \,, \tag{1.67}$$

showing that T_{3-} is helicity 1. Following a similar argument to the independence of the evolution of Fourier modes due to translational invariance, it is possible to determine that as a consequence of rotational invariance modes of different helicities will all evolve independently. Therefore it is possible, for example, to decompose a tensor of rank two T_{ij} into the sum of separate components such as in Eq. (1.66) and Eq. (1.67) which obey independent evolution equations as they have different helicities.

A vector with one index can be decomposed as

$$\beta_i = \beta_i^S + \beta_i^V \,, \tag{1.68}$$

where $\beta_i^S = -ik_i\hat{\beta}$ and β_i^V are the components that cannot be written as the gradient of a scalar, which in Fourier space is equivalent to all those components whose wave vectors are orthogonal to k_i. The circular polarization basis is particularly well suited for this purpose. Under a rotation in this basis we have that $\beta_+^{V\prime} = e^{i\psi}\beta_+^V$. It thus comprises the helicity one components of this vector, which can also be considered as the curl of a vector in three dimensions. In a similar manner, a rank two traceless symmetric tensor can be written as a sum of in helicity zero, one and two components

$$\gamma_{ij} = \gamma_{ij}^S + \gamma_{ij}^V + \gamma_{ij}^T \,. \tag{1.69}$$

with each term representing the decomposition into different helicity states. The helicity zero component must be composed of two derivatives acting on a scalar to preserve the index structure. The tracelessness condition then fixes it to be

$$\gamma_{ij}^S = (-k_i k_j + \frac{1}{3}\delta_{ij}k^2)\gamma \,. \tag{1.70}$$

The symmetric condition on the vector component then fixes this term to be

$$\gamma^V_{ij} = -\frac{i}{2}(k_i\gamma_j + k_j\gamma_i)\,. \tag{1.71}$$

The helicity two component γ^T_{ij} is the transverse and traceless part of the tensor. Couple that with the fact it is also symmetric leaves only two independent components for this part. These components eventually acquire a physical interpretation as the two independent polarizations of gravitational waves.

Now that we have determined that we can separate the perturbation evolution equations into those of different wavenumber and different helicity we shall now apply these ideas to study the form of the perturbation equations for a perturbed FLRW Universe.

Cosmological perturbation theory

A key question which arises in perturbation theory is whether a given perturbation is physical or simply an artefact of a poorly chosen coordinate system. It is a completely analogous situation to the discussion in Sec. 1.2.1 where we discussed fictitious forces which appear as a result of a choice of a non-inertial reference frame. In the same manner, a perturbation may arise from a poor choice of coordinates.[4] Consider a homogeneous FLRW background metric written in conformal time τ as in Eq. (1.37) with the coordinates τ and x^i shifted by a first-order quantity $\zeta^\mu(x,\tau)$ such that $\tilde{x}^\mu = x^\mu + \zeta^\mu(x,\tau)$. If we assume for simplicity that $\zeta^0(x,\tau) = 0$ then after this coordinate shift Eq. (1.37) becomes

$$ds^2 = a^2(\tau)\left[-d\tau^2 + 2\zeta_i' d\tilde{x}^i d\tau + \left(\delta_{ij} + 2\partial_{(i}\zeta_{j)}\right)d\tilde{x}^i d\tilde{x}^j\right]\,, \tag{1.72}$$

where $\partial_{(i}\zeta_{j)}$ is a symmetric sum over the indices i and j. Eq. (1.72) appears to be a FLRW metric with added perturbations. However, we began with a smooth homogeneous background metric and only changed the coordinates, so Eq. (1.72) is still a smooth homogeneous FLRW background metric. It is just written in a poorly chosen coordinate system. As a further example, if $\zeta^0(x,\tau) \neq 0$ then fictitious density perturbations may arise through $\rho(\tau + \zeta^0) = \bar{\rho}(\tau) + \bar{\rho}'\zeta^0$. In this coordinate system it seems there is a density perturbation $\delta\rho = \bar{\rho}'\zeta^0$. This is not a physical density perturbation and is an artefact of the chosen time slicing.

[4] The following discussion follows Daniel Baumann's cosmology notes available from *http://www.damtp.cam.ac.uk/user/db275/Cosmology.pdf.*

This discussion can apply in reverse. With a different choice of time slicing it is possible to work in a coordinate system were the physical matter perturbations are zero by choosing the hypersurface of constant time to be equivalent to the hypersurface of constant energy density.

This freedom to transform between different coordinate systems is often called *gauge freedom*. Although it seems like it may be a problem, it is actually extremely useful as this freedom allows one to perform calculations in particular gauges where there are great simplifications. This is done by fixing the gauge, meaning, choosing a gauge which removes the fictitious perturbations completely. After the gauge has been fixed, one can be rest assured that any remaining perturbations can be treated as physical. Let us see how this is achieved. Consider a general perturbed FLRW metric written in conformal time

$$ds^2 = a^2(\tau)\left[-(1+2A)d\tau^2 + 2B_i dx^i d\tau + (\delta_{ij} + h_{ij})dx^i dx^j\right] . \tag{1.73}$$

For the moment the perturbations A, B_i and h_{ij} may be unphysical as the gauge has not been fixed. Note also that B_i and h_{ij} can be expanded in an SVT decomposition

$$B_i = \hat{B}_i + \partial_i B \,, \tag{1.74}$$

$$h_{ij} = 2C\delta_{ij} + 2\partial_{\langle i}\partial_{j\rangle}E + 2\partial_{(i}\partial_{j)}\hat{E} + 2\hat{E}_{ij} \,, \tag{1.75}$$

where $\partial_{\langle i}\partial_{j\rangle}$ indicate a traceless combination of derivatives and $\partial_{(i}\partial_{j)}$ indicates a symmetric summation of derivatives. Now let us make a gauge transformation by shifting the coordinates $\tilde{x}^\mu = x^\mu + \zeta^\mu(x,\tau)$ with $\zeta^0 = T$ and $\zeta^i = \partial^i L + \hat{L}^i$. Recall that under a coordinate transformation the metric transforms as

$$g_{\mu\nu}(x) = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu}\frac{\partial \tilde{x}^\beta}{\partial x^\nu}\tilde{g}_{\alpha\beta}(\tilde{x}) \,. \tag{1.76}$$

The key quantities which arise from the partial derivatives are given by

$$\frac{\partial \tilde{x}^0}{\partial x^0} = T' + 1 \ , \ \frac{\partial \tilde{x}^0}{\partial x^i} = \partial_i T \ , \ \frac{\partial \tilde{x}^k}{\partial x^0} = \partial_0 L^k \ , \ \frac{\partial \tilde{x}^k}{\partial x^i} = \delta^k_i + \partial_i L^k \,. \tag{1.77}$$

For example, the time-time component transforms as

$$
\begin{aligned}
g_{00} &= a^2(\tau)(1+2A)\,, \\
&= \left(\frac{\partial \tilde{x}^0}{\partial x^0}\right)^2 \tilde{g}_{00}(\tilde{x})\,, \\
&= (1+T')^2 a^2(\tau+T)(1+\tilde{A})\,.
\end{aligned}
\tag{1.78}
$$

Taylor expanding the last line to first-order and solving for $\tilde{A}$ gives

$$\tilde{A} = A - T' - \mathcal{H}T\,, \tag{1.79}$$

where $\mathcal{H} = a'/a$ is the conformal Hubble factor. This defines how the metric perturbation A transforms after a change in coordinates. Similarly we can find that at first-order in the perturbations

$$\tilde{B}_i = B_i + \partial_i T - L_i'\,, \tag{1.80}$$

$$\tilde{h}_{ij} = h_{ij} - 2\partial_{(i}L_{j)} - 2\mathcal{H}T\delta_{ij}\,. \tag{1.81}$$

One way to proceed is to combine the various perturbations in such a way that the combination remains invariant after a gauge transformation. These are called Bardeen variables or gauge invariant variables. Alternatively, by making particular coordinate transformations with $\zeta^\mu(x,\tau)$ the perturbed metric can considerably simplify. For example, by choosing $\zeta^\mu(x,\tau)$ such that $B = E = 0$ the perturbed metric does not contain any mixed time-space components and simplifies to

$$ds^2 = a^2(\tau)\left[-(1+2\Psi)\,d\tau^2 + (1+2\Phi)\,\delta_{ij}dx^i dx^j\right]\,, \tag{1.82}$$

where the new perturbations are $\Phi \equiv C$ and $\Psi \equiv A$. This particular choice of coordinates is called the Newtonian gauge. As we have used up all the coordinate freedom in ζ^μ in order to set $B = E = 0$ and thus obtain equation (1.82) the perturbations Φ and Ψ can be considered physical perturbations, not an artefact of a nonlinear coordinate choice. There are of course other gauges but the Newtonian gauge is the most common and shall be adopted throughout this thesis.

Of course, the perturbations to the metric are only one half of the Einstein equations. See Ref. [41] for a more detailed discussion of the following calculation. It is also necessary to consider perturbations in the matter sector. Taking the four-velocity of a comoving fluid element with a perturbation around the background flow $u^\mu = \overline{u}^\mu + \delta u^\mu$, and applying the condition $u_\mu u^\mu = -1$ it can be shown that

$$\delta u^\mu = a^{-1}\left(-A, v^i\right)\ ,\ \delta u_\mu = a\left(-A, v_i + B_i\right)\ , \tag{1.83}$$

where $\delta u^i \equiv v^i/a$ is the spatial component of the four-velocity perturbation and we have used Eq. (1.73). To make the connection with the Einstein equation we must derive the form of the perturbed energy-momentum tensor for a perfect fluid. This can be obtained by taking $\rho = \overline{\rho} + \delta\rho$ and $p = \overline{p} + \delta p$ along with the expressions above for δu^μ and δu_μ and perturbing Eq. (1.43)

$$\delta T_{\mu\nu} = \left(\delta\rho + \delta p\right)\overline{u}_\mu\overline{u}_\nu + \delta p\overline{g}_{\mu\nu} + 2(\overline{\rho} + \overline{p})\overline{u}_{(\mu}\delta u_{\nu)} + p\delta g_{\mu\nu} + a^2 p\pi_{\mu\nu}\,, \tag{1.84}$$

where $\overline{g}_{\mu\nu}$ is the background FLRW metric and $\pi_{\mu\nu}$ is any transverse traceless component of the stress energy perturbations called anisotropic stress. The contribution from $\pi_{\mu\nu}$ is generally small enough that it shall be neglected from now on. The non-zero components of the stress-energy tensor are given by

$$\delta T_{00} = \overline{\rho}a^2\left(\delta + 2A\right)\,, \tag{1.85}$$

$$\delta T_{0i} = -\overline{\rho}a^2\left[\left(1 + w\right)v_i + B_i\right]\,, \tag{1.86}$$

$$\delta T_{ij} = \delta p a^2 \delta_{ij}\,, \tag{1.87}$$

where $\delta \equiv \delta\rho/\overline{\rho}$ is the matter density perturbation. As with any perturbed quantity in GR it is necessary to ensure these are really physical perturbations. In the same manner as with metric perturbations it is possible to define a set of gauge invariant variables for the stress-energy tensor and treat those as the real physical perturbations. On the other hand one can fix a gauge and work in a simplified coordinate system where all of the gauge freedom has been exploited. A particular choice of gauge invariant variables for the matter perturbations is given by

$$\delta^N = \delta + \frac{\overline{\rho}'}{\overline{\rho}}\left(B - E'\right)\,. \tag{1.88}$$

This is a useful because if we choose to work in the Newtonian gauge with $E = B = 0$ then the matter density perturbation δ is now gauge invariant. It is also the case that in this gauge the velocity perturbation v^i is gauge invariant. For this reason the Newtonian gauge is commonly used to study the dynamics of cosmological perturbations.

We shall now derive the form of the perturbed Einstein equations which enable the computation of the evolution of the perturbations in different cosmologies. For the following we perform an SVT decomposition of the velocity three-vector $v_i = \partial_i v + \tilde{v}_i$. There are four independent Einstein equations and two energy-momentum equations. They determine the evolution of the six perturbation variables Φ, Ψ, δ, v, $\tilde{v}$ and δp. It is therefore a solvable closed system of differential equations. The relevant Einstein equations arise from the (00), $(0i)$ the trace-free (ij) and the trace (ii) components of Eq. (1.24). In the absence of spatial curvature and anisotropic stress the (00) component gives

$$\nabla^2 \Psi = 4\pi G a^2 \bar{\rho} \Delta_m \,, \tag{1.89}$$

where $\Delta_m \equiv \delta - 3\mathcal{H}v$ is called the comoving curvature perturbation. The trace-free (ij) component of the Einstein equation leads to the relation between the metric perturbations $\Phi + \Psi = 0$. Note that in a theory of modified gravity the right hand side of this equation may no longer vanish (see Sec. 4.3.4). The remaining Einstein equations are dynamical and can be solved to obtain the time evolution of the perturbations

$$\Phi' - \mathcal{H}\Psi = 4\pi a^2 G \bar{\rho} (1+w) v \,, \tag{1.90}$$

$$c_s^2 \nabla^2 \Phi - \Phi'' - 3\mathcal{H}\left(1 + c_s^2\right)\Phi' - \left[2\mathcal{H}' + \mathcal{H}^2\left(1 + 3c_s^2\right)\right]\Phi = 0 \,, \tag{1.91}$$

where $c_s^2 = \delta p / \delta \rho$ is the sound-speed. There are two more equations needed to close the system. These arise from the perturbed part of the stress-energy conservation equation $\delta\left(\nabla_\mu T^{\mu\nu}\right) = 0$. They give the evolution of the density perturbation and the scalar component of the velocity perturbation. In the Newtonian gauge they are given by

$$\delta' + 3\mathcal{H}\left(c_s^2 - w\right)\delta = -\left(1+w\right)\left(\nabla^2 v + 3\Phi'\right) \,, \tag{1.92}$$

$$v' + \mathcal{H}\left(1 - 3c_s^2\right) v = \Phi - \frac{c_s^2}{1+w}\delta \,. \tag{1.93}$$

This analysis can also be carried out on the vector and tensor perturbations but as these do not contribute to the growth of structure in the Universe they shall not be discussed in detail.

Of course, it is not possible to precisely predict the underlying density field with perturbation theory. Statistical tools are necessary to compare the theoretical predictions with the observations. In particular a commonly employed observable is the 2-point function, called the power spectrum in Fourier space or the correlation function in real space. There are a number of definitions of the 2-point function. We shall discuss here the principal interpretations. The first is to consider the correlation function as a convolution between two density fields.

$$\zeta(r) = \langle \delta(x)\delta(x+r) \rangle = \int d^3x d^3x' \delta(x)\delta(x')\delta_D(r - |x - x'|) \,. \tag{1.94}$$

where $\delta(x)$ could be an over-density of galaxies, dark matter, temperature anisotropies or so forth and $\delta_D(x)$ is the Dirac δ-function. In this interpretation there is a fixed underlying density field and for a given r, the correlation function samples every point in this density field to obtain the correlation function. A different interpretation of the correlation function is to take an average over an ensemble of different density fields for a fixed r. This case corresponds to a functional integral over the space of different density fields

$$\langle \delta(x_1)\delta(x_2) \rangle_{ensemble} = \int \mathcal{D}[\delta]\delta(x_1)\delta(x_2)P[\delta] \,. \tag{1.95}$$

There is of course only one Universe with one density field and so in practice this is not a particularly useful way of defining the correlation function. It is not too important however, as the *ergodic hypothesis* states that these two definitions are equivalent in the limit of an infinite number of distributions. In other words the ensemble average is equal to the sample average. Of course this is an assumption which cannot test with only the one Universe. It is more common to express the correlation function in Fourier space where it is called the power spectrum $P(k)$ defined by

$$\langle \tilde{\delta}(k)\tilde{\delta}(k') \rangle = (2\pi)^3 P(k)\delta^{(3)}(k - k') \,. \tag{1.96}$$

The power spectrum gives a measure on the amount of matter clustering on each scale. A flat power spectrum indicates that the amplitude of matter clustering is the same no matter what scale is considered. In general, the amount of clustering that occurs is dependent on the particular theory of gravity which is considered. A theory of modified gravity, an extra dark energy component or an exotic dark matter particle may alter the shape and amplitude of the power spectrum. It can therefore provide a useful probe to test different models of gravity. For most models it is assumed that the density field is Gaussian meaning that the power spectrum provides a complete statistical description of the properties of the density field. However it may be the case that in order to distinguish between different theoretical models it is necessary to examine clustering on smaller length scales. Nonlinear effects become increasingly relevant at smaller scales and the underlying density field becomes less and less Gaussian. Higher-order statistics such as the 3-point function, also called the bispectrum, are then utilised in order to study the effects of various models on nonlinear structure [42].

1.3.3 The ΛCDM model

In this section we shall go over some of the key observations which culminated in the current cosmological standard model Λ Cold Dark Matter (ΛCDM). The three key pieces of evidence are the age of the Universe, supernovae observations and the cosmic microwave background (CMB). We shall then conclude this opening chapter which a discussion of photon trajectories in a perturbed Universe, which is a key observational probe into the properties of dark energy.

Age of the Universe

In this section we shall discuss why, simply by considering the age of the Universe, it is clear that we do not live in a Universe dominated by matter with $w = 0$. Recall the Friedmann equation (1.52)

$$\frac{H^2}{H_0^2} = \Omega_{r0}\left(1+z\right)^4 + \Omega_{m0}\left(1+z\right)^3 + \Omega_{k0}\left(1+z\right)^2 + \Omega_{\Lambda 0}\,, \tag{1.97}$$

where the scale factor a has now been written in terms of redshift z. Using the relation between coordinate time t and redshift z given by $dt = -dz/(1+z)H$ we can integrate Eq. (1.97) to obtain the age of a Universe, neglecting radiation which is composed of matter, a cosmological constant and spatial curvature

$$t = \frac{1}{H_0}\int_1^\infty \frac{dx}{x\left(\Omega_{m0}x^3 + \Omega_{\Lambda 0} + \Omega_{k0}x^2\right)}\,, \tag{1.98}$$

where $x \equiv 1 + z$. For a flat matter dominated Universe with $\Omega_{k0} = \Omega_{\Lambda 0} = 0$ this can be evaluated to be

$$t_0 = \frac{2}{3H_0} \,. \tag{1.99}$$

With the definition of $H_0 = 100\,h\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$ and $h \approx 0.72$ the age of a flat matter dominated Universe cannot be more than 10 billion years. The presence of stars in the Universe that are much older than this [43–45] indicates that the composition of the Universe cannot be so simple. There must therefore be some contribution from the cosmological constant and/or spatial curvature to the energy budget of the Universe in order to accommodate these astrophysical constraints.

Supernovae

A key piece of observational evidence for dark energy was provided by observations of Type Ia Supernovae [46, 47]. They are certainly a useful cosmological probe, but as we shall see, are not sufficient to precisely determine the composition of the Universe (see Fig. 1.2). It is only in combination with other observables such as the cosmic microwave background radiation and baryonic acoustic oscillations that the composition of the Universe can be tightly constrained. In this section we shall discuss how supernovae provide a complimentary probe of the stress-energy content of the Universe. Recall the form of an FLRW metric Eq. (1.36) in a spacetime with non-zero spatial curvature

$$ds^2 = -dt^2 + a^2(t)\left(d\chi^2 + S_k^2(\chi)d\Omega^2\right) \,, \tag{1.100}$$

where χ is the comoving distance, the distance that light has travelled from the big bang to today and sets the causal horizon of the Universe given by

$$\chi = \int_0^z \frac{dz'}{H(z')} \,. \tag{1.101}$$

The combination $4\pi S_k(\chi)$ is the surface area of a sphere in a closed, flat or open Universe at comoving distance χ where

$$S_k(\chi) = \begin{cases} \sin\chi & k = -1 \\ \chi & k = 0 \\ \sinh\chi & k = +1 \,. \end{cases} \tag{1.102}$$

The key observable which Supernovae probe is the *luminosity distance*, defined as

$$d_L^2 = \frac{L_s}{4\pi F}\,. \tag{1.103}$$

L_s is the intrinsic luminosity of the source, i.e. the energy emitted per unit time, and the observed flux is $F = L_0/4\pi S_k^2(\chi)$ where L_0 is the observed luminosity of the source. The luminosity distance therefore depends on the comoving distance through $S_k(\chi)$ and is therefore dependent on $H(z)$ through Eq. (1.101). The luminosity distance is then a probe the stress-energy content of the Universe. Observationally, the luminosity distance d_L can be related to the absolute magnitude M of a Supernovae which is defined to be its apparent magnitude m at a distance of 10pc

$$m - M = 5\log_{10}\left(\frac{d_L}{10\text{pc}}\right)\,. \tag{1.104}$$

With an object of known absolute magnitude, measuring the apparent magnitude of the same object allows one to infer the luminosity distance directly. For this technique to work it is necessary to observe an astronomical object with a fixed absolute luminosity called a *standard candle*. Type Ia supernovae are very useful in this regard. They are believed to be the end result of a white dwarf star in a binary companion accreting matter. Once the mass of the white dwarf reaches the Chandrasekhar limit where the electron degeneracy pressure in the core of the white dwarf is insufficient to prevent further gravitational collapse, the result is one of the most energetic thermonuclear explosions known to exist in the Universe, often being brighter than an entire galaxy. Due to the specific conditions needed to generate a type Ia supernovae explosion the absolute magnitude of the peak luminosity are all very close to $M \approx -19$. Unfortunately, type Ia supernovae cannot be considered as precise standard candles but rather, as *standardizeable* candles. There is some spread in the intrinsic luminosity of each type Ia supernovae. However, it was found that there is a useful correlation between the width of the light curve and the absolute magnitude [48]. This enables one to correct the estimate of the absolute magnitude to obtain a measure of the luminosity distance. By measuring the wavelength of some of the key absorption lines the redshift can be determined and hence the relationship between luminosity distance and redshift. Refs. [46, 47] used this relationship to disfavour a matter dominated Universe, instead favouring models dominated by a cosmological constant. Fig. 1.2 demonstrates that the constraints arising from type Ia supernovae are orthogonal to spatial curvature. A determination

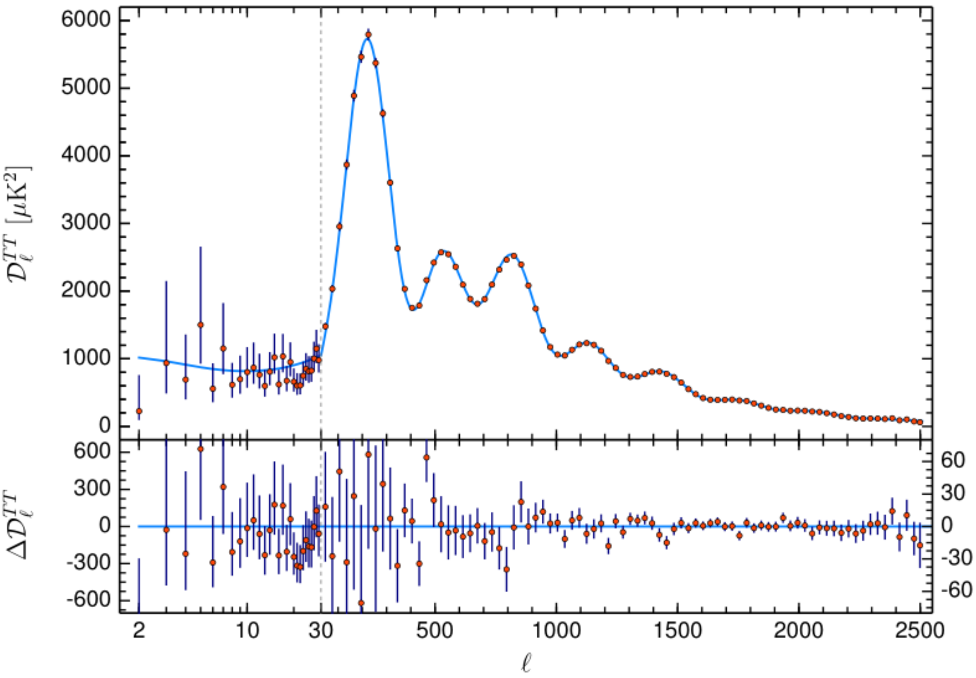

Figure 1.3 *The spectrum of the temperature fluctuations in the cosmic microwave background from the Planck 2018 results [4]*

of the spatial curvature of the Universe could then highly constrain the stress-energy composition. It was fortuitous then that this constraint had already been provided by the CMB.

Cosmic Microwave Background and Baryonic Acoustic Oscillations

The Cosmic Microwave Background (CMB) radiation provides the most stringent constraints that we have on the properties of dark energy to-date. The principal reason for high constraining power of the CMB is its capability of measuring the spatial curvature of the Universe, with the latest Planck 2018 constraint on the contribution of spatial curvature to the stress-energy budget being $\Omega_k = 0.001 \pm 0.002$ [4]. It is the radiation that is left over from the Big Bang, bathing the Universe in photons at a temperature of around $2.73K$. Detailed observations of the properties of this radiation revealed small fluctuations away from this smooth background at the order of 10^{-5}. These anisotropies in the temperature distribution of the CMB were initially sourced by quantum fluctuations generated during the inflationary era. The statistical distribution of the amplitude of the anisotropies at different angular scales is characterised by an angular wavenumber ℓ. The wave pattern that can be observed in figure 1.3 is a signature of the sound waves in the photon-baryon plasma which filled the Universe prior to recombination.

The CMB has many different sources of anisotropy arising from different physical effects occurring on different length scales which all alter the shape of the spectrum in Fig. 1.3. A primary source of anisotropy originates from the Sachs-Wolfe effect, occurring on angular scales of $\theta > 1°$. This is a large-scale super-horizon effect were gravity is dominating. The hot plasma falls into dark matter potential wells and heats up. The temperature anisotropy that occurs on these scales is roughly given by [38, 41]

$$\frac{\Delta T}{T} = \frac{1}{3}\Psi \,. \tag{1.105}$$

This implies that the temperature anisotropies on these scales is determined directly from the dark matter potential. The potential is nearly scale invariant and so the CMB spectrum is essentially flat for small ℓ. On smaller angular scales the dominating effects are baryon acoustic oscillations (BAO). These dominate on scales of $\theta \approx 1°$. The radiation pressure of the plasma prevents complete collapse, and an oscillatory behaviour is observed with a soundspeed of $c_s \approx c/\sqrt{3}$. At recombination, these waves of plasma are frozen into the matter distribution at a characteristic scale that corresponds to the sound horizon of the plasma when the baryons decoupled from the photons. In an interval dt of coordinate time light can travel a comoving distance $d\chi = cdt/a(t)$. After a time t from the Big Bang, this can be integrated to obtain the total comoving distance that light has travelled at redshift z since the Big Bang. Integrating over the scalar factor out to redshift z the comoving distance is given by [38, 39]

$$r_{H,\mathrm{com}}(z) = \int_0^{(1+z)^{-1}} \frac{cda}{a^2 H(a)} \,. \tag{1.106}$$

After matter-radiation equality, the integral is dominated by dust. Approximating $H(a) \approx H_0\sqrt{\Omega_m} a^{-\frac{3}{2}}$ and plugging it into the integral gives

$$r_{H,\mathrm{com}}(z) \approx \frac{2c}{H_0\sqrt{(1+z)\Omega_m}} \,. \tag{1.107}$$

The proper distance is obtained by dividing this expression by $a = 1/(1+z)$.

The angular size on the sky of this length scale at recombination is given by

$$\theta_{H,\mathrm{rec}} = \frac{r_{H,\mathrm{com}}(z_{\mathrm{rec}})}{D_A(z_{\mathrm{rec}})} \,, \tag{1.108}$$

where $D_A(z)$ is the angular diameter distance. As the size of the horizon is related to the horizon of the plasma, $\theta_{H,\mathrm{rec}}$ is directly obtainable once one measures the redshift of recombination. The scale of the first acoustic peak of the CMB is $\propto r_{H,\mathrm{prop}}(z_{\mathrm{rec}})$. By measuring the scale of the acoustic peak from the CMB, one is measuring $\theta_{H,\mathrm{rec}}$. As $r_{H,\mathrm{prop}}(z_{\mathrm{rec}})$ is known, one now has to determine what $D_A(z_{rec})$ should be in order to give the observed max acoustic peak at 1°. It turns out that the best fit for $D_A(z)$ comes from a flat universe. Therefore the CMB very tightly constrains the spatial curvature of the Universe to be flat, meaning that it lies along the line of $\Omega_m + \Omega_\Lambda = 1$ in Fig. 1.2. With the previous constraints on Ω_m this implies that $\Omega_\Lambda \approx 0.7$. The Universe is then dominated by a constant energy density with a negative equation of state.

Due to the tight coupling of the baryons and the photons prior to recombination, the oscillations in the plasma should also leave an imprint in the baryonic matter distribution after recombination. This imprint takes the form of an enhanced clustering of matter at the scale of the sound horizon at recombination, giving a peak in the galaxy correlation function which can be measured at different redshifts, which can be seen in Fig. 1.4. These are the baryonic acoustic oscillations. They provide a standard ruler which can be used as a complementary probe of H_0 which can then be used to constrain the composition of the Universe.

1.3.4 Matter clustering

The power spectrum of matter can provide a measurement of Ω_m, and can therefore be used to determine that the majority of matter in the Universe is dark with the remaining energy budget composed of dark energy. This can be seen by considering the growth of the density perturbations in the early Universe. The growth of perturbations in the early Universe can be obtained by solving the background and perturbation equations in a radiation dominated and then, following matter-radiation equality, a dust-dominated Universe. The growth of a perturbation depends on its scale in relation to the size of the horizon in Eq. (1.106). If the Universe is dominated by radiation free streaming prevents the growth of perturbations within the horizon. Super-horizon perturbations do grow

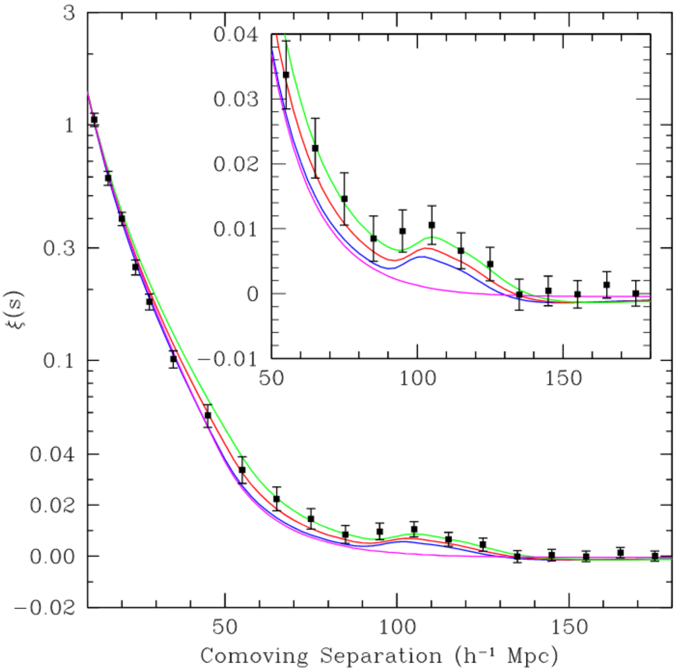

Figure 1.4 *The enhancement in the clustering of galaxies at a comoving separation of around 100 h^{-1}Mpc detected in Ref. [5], demonstrating the imprint of oscillations in the baryon-photon plasma at early times on the large-scale structure at late times.*

as $\delta \propto a^2$ in a radiation dominated Universe, which can be obtained by solving the perturbation equation in the limit $k \ll r_H^{-1}$. There are no physical interactions which can be propagated on super-horizon scales leaving them unaffected by free streaming. When the Universe becomes dominated by matter after matter-radiation equality at redshift z_{eq} the sub-horizon perturbations begin to grow and the super-horizon perturbations grow at a different rate $\propto a$. The change in the rate of growth of the perturbations leads to a singling out of a specific scale in the matter power spectrum corresponding to the size of the comoving horizon at matter-radiation equality. By evaluated in the integral (1.106) at this redshift it can be shown that this scale goes as Ω_m^{-1} (see Ref. [39]). A measurement of this scale in the matter power spectrum where the slope changes then gives another measurement on the matter content of the Universe.

We shall conclude this introductory chapter with a discussion of how a light ray propagates through a perturbed Universe, principally following the discussion in Ref. [49]. In particular, we shall cover the ideas of weak lensing and the integrated Sachs-Wolfe (ISW) effect which are two key observational probes of theories which go beyond ΛCDM.

The two key equations that determine how a photon propagates through the Universe are the geodesic equation and the null ray condition. Defining the affine parameter along the trajectory of the photon to be λ the four-momentum of the photon is given by $p^\mu = dx^\mu/d\lambda$. The null condition is given by

$$p^\mu p_\mu = 0\,, \tag{1.109}$$

and the geodesic equation by

$$\frac{dp^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta = 0\,. \tag{1.110}$$

We shall restrict ourselves to outlining the key elements of this calculation and refer the reader to Refs. [38, 49] for a comprehensive discussion. By perturbing the photon four-vector $p^\mu = \hat{p}^\mu + \delta p^\mu$ and computing the propagation of δp^μ in a perturbed FLRW metric in the Newtonian gauge the following two expressions can be derived from the time component and the spatial component of the geodesic equations

$$\frac{d}{d\tau}\left(\frac{\delta p^0}{p^0}\right) = -\left(\frac{\partial \Psi}{\partial \tau} + \frac{\partial \Phi}{\partial \tau} + 2\frac{\partial \Phi}{\partial r}\right)\,, \tag{1.111}$$

$$\frac{d^2 x^i}{d\lambda^2} + 2\mathcal{H}\frac{d\tau}{d\lambda}\frac{dx^i}{d\lambda} = \left(\frac{d\tau}{d\lambda}\right)^2 \frac{\partial}{\partial x^i}\left(\Psi + \Phi\right)\,. \tag{1.112}$$

Each index i in Eq. (1.112) corresponds to a coordinate transverse to a radial line of sight r. Note also that we have not used the Einstein equation $\Psi = -\Phi$ which would set the right hand side of equation (1.112) to zero, as it may not hold in more general models to be discussed in chapter 2.

Let us examine Eq. (1.111). Along a null trajectory $dr = d\tau$. This then implies that the total derivative $d\Phi/d\tau$ becomes

$$\frac{d\Phi}{d\tau} = \frac{\partial \Phi}{\partial \tau} + \frac{\partial \Phi}{\partial r}\,. \tag{1.113}$$

We can therefore eliminate $\partial\Phi/\partial r$ in equation (1.111) in favour of derivatives with respect to τ. This implies it can be integrated along the line of sight to obtain the change in the perturbed energy of the photon between emission E and observation O to obtain

$$\frac{\delta p^0}{p^0} = -2\Phi\Big|_E^O - \int_E^O \left(\frac{\partial \Psi}{\partial \tau} + \frac{\partial \Phi}{\partial \tau}\right) d\tau\,. \tag{1.114}$$

The average temperature of a black-body distribution, such as a thermal bath of photons, is proportional to the average frequency of the photons $\bar{\nu}$. For an observer moving with comoving velocity u^{μ} this average frequency is given by $\bar{\nu} = -p_{\mu}u^{\mu}$, which can simply be taken to be in the rest frame. The left hand side of Eq. (1.114) can therefore be interpreted as $\delta T/T$ and relates the temperature anisotropies observed in the CMB to the metric perturbations at emission and observation, known as the Sachs-Wolfe effect. We have already encountered this in Eq. (1.105) which can be derived from Eq. (1.114) by restricting to large scales with adiabatic initial conditions and neglecting the time derivatives of the potentials. The second term on the right-hand side of Eq. (1.114) describes the effect on the energy of the photon from metric potentials which change in time. This is the ISW effect, and it can be a particularly powerful observational probe of alternative models to ΛCDM (see for example, Ref. [50]).

All light which travels through the Universe undergoes some degree of gravitational lensing due to the inhomogeneous matter that makes up the Universe. The larger the amplitude of the matter perturbations, the greater the deflection angle which lead to larger deformations in the observed images of galaxies. The distortions of high redshift galaxies are extremely hard to detect individually. The presence of structure such as dark matter halos, can then only be inferred on a statistical basis by correlating the distortion of a larger number of galaxies. This technique is called *weak lensing* (see for example Refs. [51**?** –53]). It is particularly useful as it provides a measure of the *total* matter distribution, not just the baryons, and so acts as a valuable probe of the properties of dark matter. The clustering properties of dark matter are in turn, affecting by the underlying theory of gravity as well as any additional contributions to the stress-energy content of the Universe. Weak lensing observations can therefore provide powerful constraints on models beyond concordance cosmology.

Taking the affine parameter to be the conformal time $\lambda \equiv \tau$ in Eq. (1.112) and using $d\tau = dr$ the spatial component of the geodesic equation reduces to

$$\frac{d^2x^i}{dr^2} = \frac{\partial}{\partial x^i}\left(\Psi + \Phi\right) . \tag{1.115}$$

As x^i is the transverse coordinate to the line of sight it can be written in terms of a deflection angle θ^i defined by by $x^i = r\theta^i$. Eq. (1.115) can then be integrated twice to obtain

$$\theta^i = \theta^i_o + \frac{1}{r}\int_0^r dr'' \int_0^{r'} dr' \frac{\partial}{\partial x^i}\left(\Psi + \Phi\right) . \tag{1.116}$$

By reversing the order of integration in the double integral it can be re-written as a single integral

$$\theta^i = \theta^i_o + \int_0^r dr' \left(1 - \frac{r'}{r}\right) \frac{\partial}{\partial x^i} \left(\Psi + \Phi\right) . \tag{1.117}$$

The deflection angle along the line of sight is therefore dependent of the derivatives of the two metric potentials transverse to the line of sight. For our purposes, the dependence on *both* the metric potentials makes gravitational lensing a particularly powerful probe of theories in which they may not be equal (see Sec. 4.3.4).

This chapter has discussed some of the key elements of General Relativity and Cosmology. It has been by no means a comprehensive overview and the reader is encouraged to follow the references for more details. The following chapter will begin to delve into the deeper theoretical issues that this thesis is concerned with, namely, the problems surrounding ΛCDM as well as theortical ideas that have been developed to overcome them.

Chapter 2

Beyond Λ

All you really need to know for the moment is that the Universe is a lot more complicated than you might think, even if you start from a position of thinking it is pretty damn complicated in the first place.

Douglas Adams

2.1 The cosmological constant problem

By far the simplest physical mechanism to obtain accelerated background expansion is to introduce a constant into the Einstein-Hilbert action. This *cosmological constant* may seem like a rather ad-hoc solution, but its inclusion is perfectly allowed by the symmetries of the Einstein-Hilbert action. Einstein himself famously added such a constant in order to obtain a static Universe with $\dot{a} = 0$. Later on he retracted the idea, describing it as the greatest blunder of his life when Hubble discovered the expansion of the Universe. This section opens the chapter with a discussion of the theoretical issues which arise when introducing this constant. The remainder of the chapter is then concerned with alternative theoretical models which don't necessarily need a cosmological constant to accelerate the Universe. There are many excellent reviews on the cosmological constant which we refer the reader to for more details [54–59].

The action of General Relativity with a classical cosmological constant is given by

$$\frac{M_*^2}{2}\int d^4x\sqrt{-g}\,[R-2\Lambda_{GR}]+S_M\,[g^{\mu\nu},\psi]\ , \tag{2.1}$$

where Λ_{GR} is the bare classical cosmological constant with no quantum effects included and S_M is the matter action coupling the matter fields ψ to the metric $g^{\mu\nu}$. By re-deriving the background Friedmann equation (1.46) it is possible to show that $\ddot{a}\propto\Lambda_{GR}$. In short, a positive cosmological constant acts to accelerate the background expansion. In order for the classical cosmological constant Λ_{GR} to be responsible for the observed rate of accelerated expansion of the Universe it should be of the order of the Hubble parameter H_0. Written in units of the Planck mass we encounter the unfortunate situation that the observed value of the cosmological constant is unnaturally small

$$\Lambda_{obs}\sim H_0^2\sim 10^{-120}M_*^2. \tag{2.2}$$

For a parameter in a theory to be "natural" it should be of the same order as the other parameters in the theory. In this case, the measured value of the cosmological constant is many orders of magnitude smaller than the Planck mass which can be considered the natural scale in the Einstein-Hilbert action. On the face of it this is not in itself a fundamental problem, albeit an aesthetically unpleasing one. It does however hint that more issues may arise further down to the line.

In fact, the real problems associated with the cosmological constant arise when quantum corrections are included. It is necessary to re-tune the classical cosmological constant to the observed value each time higher-order corrections in the quantum perturbative expansion are computed. The cosmological constant is highly sensitive to high energy, or Ultra-Violet (UV), physics. A small change in the UV has significant consequences for the low energy, or Infrared (IR), physics. This usually is not a problem in other areas of quantum field theory. For example, the quantum corrections to the electron mass are proportional to the electron mass itself which keep them under control. The reason this occurs for the electron mass and not the cosmological constant is that in the limit of the electron mass going to zero the Lagrangian possesses an extra symmetry, namely chiral symmetry, which acts on the fermion fields. In general, a parameter in a theory is *technically natural* if the theory possesses an extra symmetry in the limit of the parameter going to zero. As quantum corrections respect the symmetries of the underlying Lagrangian, this means that the quantum corrections to the electron mass must be proportional to the electron mass. It could very well be the case that there is a UV theory which posses an extra symmetry in the limit where the cosmological constant goes to zero. This would provide a natural reason for why the cosmological constant is small relative to the Planck scale but, so far, such a theory remains elusive. Note there is an analogous situation in the case of the Higgs mass known as the hierarchy problem. The Higgs mass should receive quantum corrections from physics in the UV which drive it towards a higher scale. In order to prevent the Higgs mass from being of the order of the cutoff scale of the Standard Model of particle physics (see Sec. 2.3) there must be some mechanism to ensure the quantum corrections to the Higgs mass remain stable. This may arise via a hidden symmetry in a UV-completion of the Standard Model.

We shall now examine more quantitative arguments. To begin, we will follow the discussion in Ref. [58] and show that the vacuum expectation value (VEV) of any field placed in the vacuum state must have a constant energy density. The VEV of the matter field must then contribute to the bare value of the cosmological constant. The only invariant tensor in Minkowski spacetime is $\eta_{\mu\nu}$ which implies that, as the vacuum state must be the same for all observers in a flat spacetime $\langle T_{\mu\nu}\rangle \propto \eta_{\mu\nu}$. This is a local approximation to what we should expect on a curved spacetime through the equivalence principle, and so on a curved background it follows $\langle T_{\mu\nu}\rangle = -\rho_{\rm vac}(x,t) g_{\mu\nu}$ where $\rho_{\rm vac}(x,t)$ is a free function of space and time and the negative sign ensures the 00 component is positive. Using stress-energy conservation and metric compatibility it immediately follows that $\rho_{\rm vac}(x,t)$ must be a constant $\rho_{\rm vac}$. In which case we can write

$$\langle T_{\mu\nu}\rangle = \langle 0|T_{\mu\nu}|0\rangle = -\rho_{\rm vac} g_{\mu\nu}\,, \tag{2.3}$$

where $\rho_{\rm vac}$ is the constant energy density of the vacuum state $|0\rangle$. Because the vacuum state has a none-zero constant energy density, and according to GR energy gravitates, the vacuum gravitates. Adding in the contribution from the matter fields to the Einstein equation the effective cosmological constant takes the form

$$\Lambda_{eff} = \Lambda_{GR} + \rho_{\rm vac}\,. \tag{2.4}$$

It is this effective cosmological constant which appears to drive the accelerated expansion of the Universe. Now it remains to estimate the amplitude of $\rho_{\rm vac}$.

As this was derived from the VEV of the matter fields its value is inherently a quantum field theory prediction. A commonly employed approach in the literature is to impose a sharp cutoff such as the Planck scale M_* and sum up the contributions of the zero point energy modes up to this cutoff scale. For a scalar field of mass m this would contribute as

$$\rho_{\rm vac} = \frac{1}{4\pi^2}\int_0^{M_*} dk k^2 \sqrt{k^2+m^2} \propto M_*^4 \,. \tag{2.5}$$

In other words the cosmological constant should go as the fourth power of the cut-off scale of theory. Note however that, as pointed out in Refs. [58, 60] that if the above calculation was taken seriously then the vacuum energy should behave like radiation with an equation of state $w = 1/3$. The reason for this is that by imposing a sharp cutoff Lorentz invariance is not maintained. In order to obtain physical predictions using regularisation it is necessary that the adopted scheme respects the underlying symmetries of the theory. If a Lorentz invariant scheme is used, such as dimensional regularisation, then a different result is obtained [58, 60] which goes as

$$\rho_{\rm vac} \sim \sum_i \mathcal{O}_i(1) m_i^4 \,. \tag{2.6}$$

The sum is over all of the particles that exist in the Standard Model of particle physics with mass m_i raised to the fourth power multiplied by order-one constants. Unfortunately this value of $\rho_{\rm vac}$ is still far too large by many orders of magnitude, even when the sum is dominated by the top quark mass, to adequately account for the accelerated expansion without a fine-tuned choice of Λ_{GR} to match Λ_{obs}. It is debatable whether this one off fine tuning is a fundamental issue. The real problem is that it is not stable against quantum corrections.

We shall not discuss any loop diagrams here to demonstrate this radiative instability but rather draw on an elegant argument outlined in Refs. [59, 61] using effective actions. The Wilsonian effective action assumes that the path integral for a field theory can be split into light modes ϕ_ℓ and heavy modes ϕ_h between some cutoff scale μ. The low energy effective action $\mathcal{S}_{eff}\left[\phi_\ell\right]$ is defined by integrating out the heavy modes ϕ_h of the full action $\mathcal{S}\left[\phi_\ell, \phi_h\right]$

$$e^{i\mathcal{S}_{eff}[\phi_\ell]} = \int \mathcal{D}\phi_h e^{i\mathcal{S}[\phi_\ell,\phi_h]} \,. \tag{2.7}$$

As the cutoff μ is the largest mass scale that appears in the effective action $\mathcal{S}_{eff}\left[\phi_\ell\right]$ the vacuum energy for this theory should scale as $\rho_{\rm vac} \sim \mu^4$. We then require the classical cosmological constant to be tuned such that the combination of Λ_{GR} and the vacuum energy matches the observed value

$$\Lambda_{obs} = \Lambda_{GR} + \mathcal{O}(1)\mu^4 \, . \tag{2.8}$$

So far no problem. Let us now move the scale μ which separates the light modes from the heavy modes to some new scale μ'. In exactly the same manner, the cosmological constant predicted from this new effective action should now scale as $\rho_{\text{vac}} \sim \mu'^4$. The bare cosmological constant has already been fixed by the requirement of cancelling the contribution of the vacuum energy with the cutoff scale at μ. We could re-tune it to cancel the contribution from this new scale, but that is precisely the point. The effective description of a healthy theory in the IR should not be dependent on the choice of the cutoff scale. This constant re-tuning to match the IR physics is the essence of the cosmological constant problem.

These theoretical issues associated with the cosmological constant motivated a great deal of work in going beyond it. Perhaps the accelerated expansion of the Universe is not driven by a cosmological constant, but a modification to the laws of General Relativity which apply on cosmological scales? The following sections will look at how theorists have attempted in recent years to go beyond the cosmological constant. However, it should be stressed that many of these ideas do not directly address the cosmological constant problem. Even if one of these theories end up being favoured by data there would still be the need to address the radiative instability of the vacuum from a purely theoretical standpoint, without considering cosmic acceleration.

2.2 Beyond ΛCDM: Dark Energy and Modified Gravity

Having assessed various issues related to the cosmological constant problem we shall now examine models which go beyond Λ. The majority of these models do not tackle the cosmological constant problem directly. They seek instead to provide an alternate explanation for the accelerated expansion of the Universe which does not rely on a cosmological constant. A natural extension is to replace the constant with a scalar field field. The existence of such a cosmological scalar field is particularly motivated by the discovery of the Higgs boson, which is the particle excitation of a quantum scalar field. It is natural to suppose that the observed value of the cosmological constant is simply the value the scalar field takes at the minimum of its potential. The question is then how to determine the form of the scalar field potential. This rather natural solution is unfortunately not viable. Weinberg showed [54] that one would have to tune the minimum of the scalar field potential just as much as one would have had to tune the cosmological constant. This is a famous no-go result which we shall now review. It is an important concept that determines that scalar fields cannot provide a simple solution to the cosmological constant problem, and that, even if they are included to drive cosmic acceleration the issues discussed in Sec. 2.1 remain.

2.2.1 Weinberg's no-go theorem

The cosmological constant problem was an issue from a purely theoretical perspective in the decades preceding the discovery of cosmic acceleration. This subsection reviews the result of Weinberg [54] that shows the cosmological constant cannot be interpreted as the minimum value of the potential of a scalar field. The potential of the field would have to be fine-tuned just as much as the cosmological constant. A more detailed discussion of the following proof can be found in Refs. [54, 59, 62]. The assumptions that go into Weinberg's result are quite general but we shall explicitly state them here.

- The theory consists of a local Lorentz four-dimensional field theory including a metric field $g_{\mu\nu}$ and a collection of scalar fields ϕ_i.
- We further assume that the fields are transitionally invariant on-shell, such that $g_{\mu\nu} = const$ and $\phi_i = const$.
- The Lagrangian is built out of invariant combinations of these quantities.

The residual symmetry that is left over once one assumes translational invariance is the rotational symmetry of four dimensional constant matrices. In other words, the four dimensional general linear group $GL(4)$. If these matrices are denoted by $M_{\mu\nu}$, the coordinates change by $x^\mu = M^\mu{}_\nu x^\nu$ so that the metric transforms as

$$g^{\mu\nu} = M^\mu{}_\alpha M^\nu{}_\beta g^{\alpha\beta} \,. \tag{2.9}$$

The term $\sqrt{-g}$ implies that the Lagrangian also transforms as $\mathcal{L} \to \det(M)\mathcal{L}$. Now assuming that the transformation is infinitesimal $M^\mu{}_\nu = \delta^\mu{}_\nu + \delta M^\mu{}_\nu$ the change in the metric is simply

$$\delta g_{\mu\nu} = \delta M_{\mu\nu} + \delta M_{\nu\mu} \,. \tag{2.10}$$

As we are neglecting higher powers of $\delta M_{\mu\nu}$ the change in the Lagrangian reduces to

$$\delta\mathcal{L} = \mathrm{Tr}\,(\delta M)\mathcal{L} \,, \tag{2.11}$$

which can be more clearly determined by working in a basis where $M_{\mu\nu}$ is diagonal. The determinant is then the product over all of the diagonal elements, which at first-order is the trace. The variation of the Lagrangian then reads

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial g_{\mu\nu}}\delta g_{\mu\nu} \,. \tag{2.12}$$

By setting $\delta\mathcal{L} = 0$ the field equations read

$$\frac{\partial\mathcal{L}}{\partial\phi_i} = 0 \ , \ \frac{\partial\mathcal{L}}{\partial g_{\mu\nu}} = 0 \,. \tag{2.13}$$

There are two distinct scenarios to consider. The first is that Eqs. (2.13) hold independently of one another. Assuming that $\partial\mathcal{L}/\partial\phi_i = 0$ we then have that

$$\frac{\partial\mathcal{L}}{\partial g_{\mu\nu}}\left(\delta M_{\mu\nu} + \delta M_{\nu\mu}\right) = \mathrm{Tr}\,(\delta M)\mathcal{L} \,, \tag{2.14}$$

where have used Eq. (2.10). This equation must hold for any matrix $M \in GL(4)$. This then implies that

$$\frac{\partial\mathcal{L}}{\partial g_{\mu\nu}} = \frac{1}{2}g^{\mu\nu}\mathcal{L} \,, \tag{2.15}$$

which can be seen by contracting both sides with $\delta g_{\mu\nu}$ and noting that $g^{\mu\nu}\delta g_{\mu\nu} = 2g^{\mu\nu}\delta M_{\mu\nu}$. By applying the relation

$$\frac{\partial}{\partial g_{\mu\nu}}\sqrt{-g} = \frac{1}{2}g^{\mu\nu}\sqrt{-g} \,, \tag{2.16}$$

we see that, if the Lagrangian is to satisfy Eq. (2.15), the solution must be of the form

$$\mathcal{L} = \sqrt{-g}V(\phi_i) \,. \tag{2.17}$$

Taking the second field equation $\partial\mathcal{L}/\partial g_{\mu\nu} = 0$ we see that this is only satisfied for $V(\phi_i) = 0$. We shall now consider the case where the field equations do not hold independently of one another. This then implies that there should be some relation between the two sets of field equations. The most general expression that one may write down is of the form

$$g_{\mu\nu}\frac{\partial\mathcal{L}}{\partial g_{\mu\nu}} = \sum_i f_i(\phi)\frac{\partial\mathcal{L}}{\partial\phi_i} \,. \tag{2.18}$$

If there exists a symmetry such that

$$\delta g_{\mu\nu} = 2\epsilon g_{\mu\nu} \ , \ \delta\phi_i = -\epsilon f_i \,, \tag{2.19}$$

for some small parameter ϵ then the degeneracy condition in Eq. (2.18) immediately implies that Eqs. (2.13) also hold. By rotating the set of scalar fields such that only one of them transforms under the $GL(4)$ transformation so that now the symmetry becomes

$$\delta g_{\mu\nu} = 2\epsilon g_{\mu\nu} \ , \ \delta\tilde{\phi}_0 = -\epsilon \ , \ \delta\tilde{\phi}_{i\neq 0} = 0 \,, \tag{2.20}$$

we can construct an invariant quantity $e^{2\tilde{\phi}_0}g_{\mu\nu}$ as

$$\delta(e^{2\tilde{\phi}_0}g_{\mu\nu}) = 2\delta\tilde{\phi}_0 e^{2\tilde{\phi}_0}g_{\mu\nu} + e^{2\tilde{\phi}_0}\delta g_{\mu\nu} = 0 \,, \tag{2.21}$$

where the last equality uses Eq. (2.20). As the Lagrangian is constructed from invariant quantities and we require the scalar field and metric degeneracy condition in Eq. (2.18) to hold, we conclude that this is equivalent to assuming no degeneracy condition but with the metric replaced by $g_{\mu\nu} \to \tilde{g}_{\mu\nu} = e^{2\tilde{\phi}}g_{\mu\nu}$ and $\tilde{\phi}_i \to \phi_{i\neq 0}$. The latter transformation follows from the fact that $\tilde{\phi}_{i=0}$ is a scalar and thus doesn't transform under the $GL(4)$ symmetry group. We conclude that when applying the degeneracy condition and assuming the Lagrangian must take the form

$$\mathcal{L} = \sqrt{-g}e^{4\tilde{\phi}_0}V(\tilde{\phi}_{i\neq 0}) \,, \tag{2.22}$$

this then implies that either $V(\tilde{\phi}_{i\neq 0}) = 0$, corresponding to fine tuning, or $e^{4\tilde{\phi}_0} = 0$. The trajectories of massive particles satisfy $e^{2\tilde{\phi}_0}g_{\mu\nu}u^\mu u^\nu = -m^2 e^{2\tilde{\phi}_0}$ and so every massive particle now comes with an extra factor of $e^{2\tilde{\phi}_0}$. Setting this to zero implies the theory cannot contain massive particles. In other words, the theory must respect conformal symmetry and this simply is not the Universe we live in.

2.2.2 Scalar fields and the Ostrogradsky Theorem

In this section we review another important theorem which forms the background of the work in this thesis. Even if the scalar field is not used in the theory to tackle the cosmological constant, it is nevertheless important to take care when adding a scalar field for the purposes of cosmic acceleration. The following theorem is purely classical and makes no assumptions about the physical system under consideration other than it is described in a Lagrangian formulation. The Ostrogradsky theorem states that it is necessary to restrict the equations of motion of a dynamical system to have at most two time derivatives. If this is not the case then it will lead to an unbounded Hamiltonian producing an Ostrogradsky ghost. In the context in which we are interested, namely a scalar field coupled to GR, this theorem has important consequences. In 1974 Horndeski [63] wrote down the most general local Lorentz invariant Lagrangian in four dimensions describing a scalar field coupled to the Einstein-Hilbert action. The detailed form of the theory will be discussed in Sec. 2.2.5, but for now it is sufficient to note that the structure of each of term in the theory is such that Ostrogradsky ghosts do not appear in the field equations. Due to the relevance this theorem had in the construction of Horndeski scalar-tensor theory it is worth discussing how it works in a simple example. More detailed discussions can be found in Refs. [62, 64, 65] which we follow.

Let us assume that a Lagrangian describing the motion of a particle with position $q(t)$ depends explicitly on the first and second time derivatives $\mathcal{L}(q, \dot{q}, \ddot{q})$. After varying $q(t) \to q(t) + \delta q(t)$ the equation of motion can be determined to be

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} + \frac{d^2}{dt^2}\frac{\partial \mathcal{L}}{\partial \ddot{q}} = 0\,. \tag{2.23}$$

If $\partial^2\mathcal{L}/\partial\ddot{q}^2 \neq 0$ then the equation of motion is a fourth order differential equation. It therefore needs four pieces of initial data in order to obtain a complete solution $\{q_0, \dot{q}_0, \ddot{q}_0, \dddot{q}_0\}$. This condition is known as non-degeneracy. Note that it may be that the higher derivatives satisfy some constraint equations so that they do not introduce new degrees of freedom. This is a degenerate system, and Degenerate Higher Order Scalar-Tensor (DHOST) theories have recently been a subject of some interest [66].

The four initial data correspond to four canonical coordinates Q_1, Q_2, P_1, P_2 given by

$$Q_1 = q\,,\; Q_2 = \dot{q}\,, \tag{2.24}$$

$$P_1 = \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \ddot{q}}\,,\; P_2 = \frac{\partial \mathcal{L}}{\partial \ddot{q}}\,, \tag{2.25}$$

which can be used to construct the Hamiltonian

$$H(Q_1, Q_2, P_1, P_2) = \sum_{i=1}^{2} P_i q^{(i)} - \mathcal{L}(Q_1, Q_2, F(Q_1, Q_2, P_2))\,, \tag{2.26}$$

where $q^{(1)} \equiv \dot{q}$, $q^{(2)} \equiv \ddot{q}$, and $\ddot{q} = F(Q_1, Q_2, P_2)$ is the inversion allowed by the non-degeneracy condition. This choice of canonical coordinates reproduces the standard Hamiltonian equations of motion

$$\frac{\partial H}{\partial P_i} = \dot{Q}_i\,,\; \frac{\partial H}{\partial Q_i} = -\dot{P}_i\,. \tag{2.27}$$

We can now write out the Hamiltonian for this theory as

$$H = P_1\dot{Q}_1 + P_2\dot{Q}_2 - \mathcal{L}\left(Q_1, Q_2, F(Q_1, Q_2, P_2)\right) \tag{2.28}$$

$$= P_1 Q_2 + P_2 F(Q_1, Q_2, P_2) - \mathcal{L}\left(Q_1, Q_2, F(Q_1, Q_2, P_2)\right)\,, \tag{2.29}$$

where we have replaced $\ddot{q}$ in the Lagrangian with $\ddot{q} = F(Q_1, Q_2, P_2)$. The first term in the Hamiltonian is *linear* in the momentum variable P_1. This is a disaster for constructing a healthy theory. Physical states will cascade down towards infinitely low energy with this Hamiltonian which is unbounded from below. This is an Ostrogradsky instability. The presence of non-degenerate higher order derivatives in the Lagrangian implies a Hamiltonian which will produce negative energy states. In order to construct a theory that has higher derivative terms in it is necessary to do so in such a way as to remove them at the level of the equations of motion.

2.2.3 Quintessence and k-essence models

In this section we begin to examine models of cosmic acceleration which go beyond a cosmological constant. The first and perhaps simplest extension is to add a dynamical scalar field to the stress-energy component of GR. We do not assume at this stage any direct coupling to the metric (see Sec. 2.2.4). A model which includes a standard scalar field kinetic term and a potential is called *quintessence*

[67]. The form of the theory is

$$S = \int d^4x\sqrt{-g}\left[\frac{M_*^2}{2}R - \frac{1}{2}X - V(\phi)\right] . \tag{2.30}$$

where $X \equiv \partial_\mu\phi\partial^\mu\phi$ is the standard kinetic term and $V(\phi)$ is the potential of the scalar field. The energy-momentum tensor for the scalar field is given by

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\varphi - g_{\mu\nu}\left[\frac{1}{2}(\partial\phi)^2 - V(\phi)\right] . \tag{2.31}$$

We can use this expression to assign a pressure from the spatial components and energy density from the time component to the scalar field and thus derive an equation of state parameter. If the scalar field is completely homogeneous and isotropic, a reasonable assumption on large scales, then it can only be a function of time. The equation of state then takes the form

$$w_\phi = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi} + V(\phi)} . \tag{2.32}$$

In the limit where the scalar field is slowly rolling with $\dot{\phi} \ll 1$ in units of the Planck mass, the potential dominates the energy budget and $w \approx -1$. By constructing models such that this slow role condition holds, just as in inflation [40], it is possible to construct models which mimic the expansion history of ΛCDM with a scalar field instead of a cosmological constant.

It is possible to generalise quintessence to models with non-canonical kinetic terms. These are called *k-essence* models [68, 69]. By adding a general function of the derivatives of the scalar field and kinetic terms $K(\phi, X)$ one can construct more exotic models. A k-essence model takes the form

$$S = \int d^4x\sqrt{-g}\left[\frac{M_*^2}{2}R + K(\phi, X)\right] . \tag{2.33}$$

The precise phenomenology of the theory entirely depends on the functional form of $K(\phi, X)$. The equation of state w_K is now more general. For a model with no explicit ϕ dependence such that $K(\phi, X) \equiv K(X)$ it is given by [70]

$$w_K = \frac{K}{2XK_X - K} . \tag{2.34}$$

These models will automatically avoid the Ostrogradsky instability as they only contain products of first derivatives ensuring that the resulting equations of motion remain no higher than second order.

2.2.4 Non-minimal couplings

It is also possible to directly couple the metric to the scalar field via a *non-minimal* coupling. This will allow a natural distinction to be drawn between variations of the equivalence principle following Ref. [71]. The *weak equivalence principle* states that all test particles follow the geodesics of a universal metric. It can be elevated to a *strong equivalence principle* by extending it to bodies which self-gravitate. For example, in a theory which violates the strong equivalence principle black holes may follow different trajectories to ordinary matter species. Formally these ideas can be expressed by writing the action for a scalar field theory where each matter species ψ_i follows the geodesics set by the *Jordan frame* metric $\tilde{g}^{\mu\nu} = A_i^2(\phi)g^{\mu\nu}$. The subscript i indicates that different matter species may follow different trajectories depending on the form of the non-minimal coupling function $A_i(\phi)$ for each ψ_i. The action is given by

$$S = \int d^4x\sqrt{-g}\left[\frac{M_*^2}{2}R - \frac{1}{2}X - V(\phi)\right] + S_M\left[A_i^2(\phi)g^{\mu\nu}, \psi_i\right] . \tag{2.35}$$

If each matter species couples to the Jordan frame metric in the same way such that $A_i(\phi) \equiv A(\phi)$ then this theory satisfies the weak equivalence principle, where each particle follows the trajectory set by a universal Jordan frame metric. On the other hand, the no-hair theorem guarantees that a black hole cannot possess any physical quantities other than charge, mass and angular momentum [72, 73]. This excludes the possibility that the black hole can have a scalar charge. Non-minimal couplings imply that the trajectory of black holes will be different to that of ordinary matter. Another way to see this is to recognise that a black hole is a property of the first term on the right hand side of Eq. (2.35) and knows nothing about the form of $A(\phi)$. As black holes are self-gravitating bodies the presence of $A(\phi)$ implies that the strong equivalence principle has also been broken. The modification to the geodesic equation in this case takes the form

$$\frac{d^2x^i}{dt^2} = -\partial_i\left[\Phi + \ln A(\phi)\right] . \tag{2.36}$$

It is always possible to perform a conformal transformation such that the matter fields are universally coupled to the Jordan frame metric. This will alter the gravitational sector of the theory such that, although matter particles follow the geodesics of a universal metric, this metric is now no longer determined from the GR field equations but from a modified field equation involving the scalar field. We stress that physics in the Einstein and the Jordan frame are equivalent and can be related [74, 75]. Under a conformal transformation in four dimensions the Ricci scalar in the Jordan frame $\tilde{R}$ is related to the Ricci scalar in the Einstein frame R through [76]

$$\tilde{R} = \frac{R}{A^2(\phi)} + \frac{6}{A^3(\phi)} \Box A(\phi) \tag{2.37}$$

where $\Box \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu$. In the Jordan frame there is therefore a function of ϕ which is directly coupled to the Ricci scalar along with additional interaction terms of the scalar and the metric through the $\Box$ operator. Such a theory which involves this non-minimal coupling function is called a scalar-tensor theory. How these theories connect with cosmological observables forms the basis for much of the work in this thesis.

2.2.5 Generalised scalar-tensor theories: Galileons and Horndeski theory

When adding a scalar field into the Einstein equations, with or without the non-minimal coupling function $A^2(\phi)$, it is important to ensure that the Ostrogradsky theorem is respected. As discussed in Sec. 2.2.2 higher derivatives of the scalar field appearing in the action may lead to an unbounded Hamiltonian. There exists a class of theories which exist on flat space, possess higher derivatives and yet retain second order equations of motion which are called Galileons [77]. They possess a Galilean shift symmetry $\phi \to \phi + a + b_\mu x^\mu$ from which they obtained their name. The original motivation for these theories came from examining theories of massive gravity [78] in various limits, but they also possess a number of nice theoretical properties such as a non-renormalisation theorem as well as possessing a finite number of terms in d-dimensions [79, 80].

Generalising Galileons beyond flat space leads to an interesting class of scalar-tensor theories which can be applied on a cosmological background while avoiding the Ostrogradsky ghost [81]. These generalised Galileons were then found to be equivalent to a theory written down by Horndeski many years earlier [63]. Horndeski theory is the most general way of incorporating a scalar field into the Einstein field equations in four dimensions with at most second-order equations of motion. Because of its generality it has been applied in many areas of theoretical physics such as black holes to cosmology [82–85]. The freedom in the theory is determined by a choice of five free functions of both the scalar field ϕ and its kinetic term $X \equiv \partial_\mu \phi \partial^\mu \phi$.

$$S = \sum_{i=2}^{5} \int d^4x \sqrt{-g}\, \mathcal{L}_i \,, \tag{2.38}$$

where the four Lagrangian densities are defined as

$$\mathcal{L}_2 \equiv G_2(\phi, X)\,, \tag{2.39}$$

$$\mathcal{L}_3 \equiv G_3(\phi, X)\Box\phi\,, \tag{2.40}$$

$$\mathcal{L}_4 \equiv G_4(\phi, X)R - 2G_{4X}(\phi, X)\left[(\Box\phi)^2 - (\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi)\right]\,, \tag{2.41}$$

$$\mathcal{L}_5 \equiv G_5(\phi, X)G_{\mu\nu}\nabla^\mu\nabla^\nu\phi + \frac{1}{3}G_{5X}(\phi, X)\left[(\Box\phi)^3 - 3(\Box\phi)(\nabla_\mu\nabla_\nu\phi)(\nabla^\mu\nabla^\nu\phi) + 2(\nabla_\mu\nabla_\nu\phi)(\nabla^\sigma\nabla^\nu\phi)(\nabla_\sigma\nabla^\mu\phi)\right]\,, \tag{2.42}$$

where $G_{iX} \equiv \partial G_i / \partial X$. The non-minimal coupling function has now been generalised to include derivative interactions of the scalar field. Quintessence and k-essence models are subsets of Horndeski theory with $G_4 = 1$ and G_2 set by the specific model. GR is also a subset of Horndeski theory with $G_4 = 1$ and $G_2 = G_3 = G_5 = 0$. It is theoretically interesting to study in its own right, but there are deeper theoretical reasons to test this theory which we now touch on.

2.2.6 Motivation for Scalar-Tensor Theories: Brane World and Kaluza-Klein theories

Before exploring the phenomenology of scalar-tensor theories it is worth examining their theoretical motivation beyond simply extending the ΛCDM model. Even if it is discovered that a scalar-tensor theory is preferred over standard GR in cosmological observations it should not be considered to be fundamental. There may well be many potential UV-completions which give rise to equivalent scalar-tensor theories and cosmological observations. In this section we follow an example from Ref. [86] which demonstrates how a scalar-tensor theory can arise from a more fundamental theory.

Let us assume that there is an D-dimensional metric $g^{(D)}_{\bar{\mu}\bar{\nu}}$ which can be separated into a four-dimensional metric $g_{\mu\nu}$ and an n-dimensional metric $\tilde{g}_{\alpha\beta}$ where $n = (D-4)$ such that

$$ds^2 = g^{(D)}_{\bar{\mu}\bar{\nu}} dx^{\bar{\mu}} dx^{\bar{\nu}} \,, \tag{2.43}$$

$$= g_{\mu\nu} dx^{\mu} dx^{\nu} + \Omega^2(x) \tilde{g}_{\alpha\beta} d\theta^{\alpha} d\theta^{\beta} \,. \tag{2.44}$$

The indices μ, ν run from 0 to 4 and the indices α, β run over the remaining n indices. We label the coordinates on the $(D-n)$-dimensional space as θ_α. Note also that we neglect any mixed components such $g_{\mu\alpha}$ for simplicity.

The Lagrangian for the full theory is given by

$$\mathcal{L} = \frac{1}{2} \mathcal{C} \sqrt{-g^{(D)}} R^{(D)} \,, \tag{2.45}$$

where $\mathcal{C}$ is a constant to ensure the dimensions match and $R^{(D)}$ is the Ricci scalar in the full D-dimensional theory. The determinant can be decomposed as

$$\sqrt{-g^{(D)}} = \sqrt{-g} \Omega^n \sqrt{\tilde{g}} \,. \tag{2.46}$$

We can obtain an effective Lagrangian of the four-dimensional theory by integrating the full theory only over the θ_α coordinates

$$\tilde{\mathcal{L}}_4 = \tilde{V}_n^{-1} \int \mathcal{L} d^n\theta \,, \tag{2.47}$$

where

$$\tilde{V}_n = \sqrt{\tilde{g}} d^n\theta \,. \tag{2.48}$$

This can then be related to $\mathcal{L}_4 = \sqrt{-g} L_4$ where

$$L_4 = \frac{1}{2} \Omega^n \tilde{V}_n^{-1} \int \sqrt{\tilde{g}} R d^n\theta \,. \tag{2.49}$$

We have omitted the relation between the $R^{(D)}$ and R which includes extra factors of Ω. See Appendix A in Ref. [86] for the full relation. For $n > 1$ and by making the field redefinition

$$\phi = 2\sqrt{\frac{n-1}{n}} \Omega^{\frac{n}{2}} \,, \tag{2.50}$$

L_4 can be written as

$$L_4 = \frac{1}{8} \frac{n}{n-1} \phi^2 R + \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + \frac{1}{2} \left(\frac{1}{4} \frac{n}{n-1} \phi^2 \right)^{1-2/n} \,. \tag{2.51}$$

One can see that the effective Lagrangian in four-dimensions reduced from the full D-dimensional metric behaves like a scalar-tensor theory with a non-minimal coupling function in front of the Ricci scalar, a kinetic term and a potential. Further motivation for the existence of the scalar field comes from the dilaton in string theory and other brane-world models. By testing theories which go beyond ΛCDM one is also testing more fundamental theories of gravity. It is possible therefore that such analyses will eventually shed light on the connections between quantum field theory and general relativity which has been a highly active field of research during recent decades [25, 26].

2.2.7 Screening mechanisms

The Chameleon mechanism

Einstein's theory of General Relativity has to date been extremely well tested on Solar System scales [87]. Therefore it is necessary have some form of *screening mechanism* which ensures that the effects of a fifth force which may act on large scales disappear in the regime where GR has been well tested. We follow the distinction drawn in Refs. [62, 71] and classify different screening mechanisms as to whether they screen through a local field value, via the first derivative of the scalar field or the second derivative of the scalar field. The first example we shall give is one which exhibits screening through the local field value, called the chameleon mechanism [88, 89].

We shall begin again with the action of a general scalar-tensor theory written in the Einstein frame

$$S = \int d^4x\sqrt{-g}\left\{\frac{M_*^2}{2}R - \frac{1}{2}X - V(\phi)\right\} + S_M\left[A^2(\phi)g_{\mu\nu}, \psi_i\right] . \tag{2.52}$$

It is necessary to derive the equation of motion for the field ϕ. The first term on the right hand side of equation (2.52) results in the standard Klein-Gordon equation. However it is necessary to also compute the additional contribution from the non-minimal coupling in the matter action through the Jordan frame metric $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$. After a variation of ϕ this becomes

$$\frac{\delta S_M\left[A^2(\phi)g_{\mu\nu}, \psi_i\right]}{\delta\phi} = \frac{\delta S_M\left[A^2(\phi)g_{\mu\nu}, \psi_i\right]}{\delta\tilde{g}^{\mu\nu}}\frac{\partial\tilde{g}^{\mu\nu}}{\partial\phi} . \tag{2.53}$$

The second term on the right hand side above becomes

$$\begin{aligned}\frac{\partial \tilde{g}^{\mu\nu}}{\partial\phi} &= \frac{\partial}{\partial\phi}(A^2(\phi)g^{\mu\nu}), \\ &= 2A(\phi)g^{\mu\nu}\frac{\partial A}{\partial\phi}, \\ &= 2A^{-1}(\phi)\tilde{g}^{\mu\nu}\frac{\partial A}{\partial\phi}\,.\end{aligned} \tag{2.54}$$

Using the definition of the Jordan frame stress-energy tensor

$$\frac{\delta S_M\left[A^2(\phi)g_{\mu\nu},\psi_i\right]}{\delta\tilde{g}^{\mu\nu}} = \frac{\sqrt{-\tilde{g}}}{2}\tilde{T}_{\mu\nu}\,, \tag{2.55}$$

and the transformation of the determinant of the metric

$$\sqrt{-\tilde{g}} = A^4(\phi)\sqrt{-g}\,, \tag{2.56}$$

the variation of the matter action with respect to ϕ becomes

$$\frac{\delta S_M\left[A^2(\phi)g_{\mu\nu},\psi_i\right]}{\delta\phi} = A^3(\phi)\tilde{T}\frac{\partial A(\phi)}{\partial\phi}\sqrt{-g}\,, \tag{2.57}$$

where $\tilde{T} = \tilde{g}^{\mu\nu}\tilde{T}_{\mu\nu}$ is the trace of the Jordan frame stress energy tensor. The equation of motion for ϕ is then

$$\Box\phi = \frac{\partial V}{\partial\phi} - A^3(\phi)\tilde{T}\frac{\partial A(\phi)}{\partial\phi}. \tag{2.58}$$

In the Jordan frame matter fields are universally coupled the Jordan frame stress-energy tensor $\tilde{T}_{\mu\nu}$ which is conserved $\tilde{\nabla}_\mu\tilde{T}^{\mu\nu} = 0$. After using the relation $\tilde{T} = A^{-4}T$ the divergence of the stress-energy tensor becomes

$$\nabla_\mu T^{\mu\nu} = \frac{T}{A(\phi)}\frac{\partial A(\phi)}{\partial\phi}\partial^\nu\phi\,. \tag{2.59}$$

After defining an energy density $\rho \equiv T/A$ it can be shown that Eq. (2.59) is equivalent to the standard continuity equation in an expanding FLRW Universe in Eq. (1.47). The equation of motion for the scalar field is then given by the standard Klein-Gordan equation with an effective potential which is dependent on the local matter density

$$\Box\phi = \frac{\partial V_{eff}}{\partial\phi}\,, \tag{2.60}$$

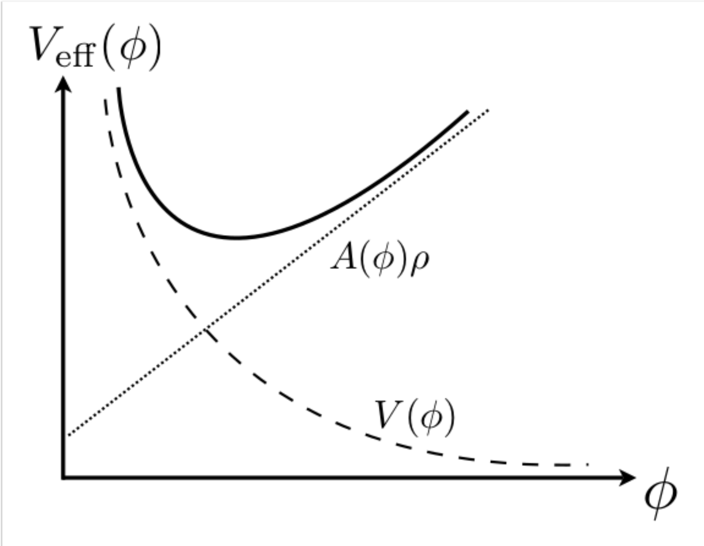

Figure 2.1 *The shape of the effective density dependent potential which gives rise to the chameleon mechanism in high density regions. Reproduced from Ref.[6].*

with

$$V_{\rm eff}(\phi) = V(\phi) + A(\phi)\rho\,. \tag{2.61}$$

The shape of this potential is sketched in figure 2.1. The effective mass of the scalar particle is given by the second derivative of the effective potential with respect to ϕ. As the mass of the particle determines the range of propagation of the fifth force through a Yukawa type potential, the larger the mass of the scalar, the lower the propagation range. In high density regions the scalar cannot propagate much at all and the fifth force is suppressed.

Derivative screening

In this section we shall examine a screening mechanism which arises through the derivatives of the scalar field. In the first instance, let us consider a model which only includes first derivatives of the scalar field, such as k-essence models, using an example from Ref.[62]. Consider the Lagrangian

$$\mathcal{L} = \frac{1}{2}X + \frac{\alpha}{4\Lambda^4}X^2 + \frac{g}{M_*}\phi T\,, \tag{2.62}$$

where in this subsection Λ is a mass scale which sets the regime where the derivative interactions become relevant. Computing the equation of motion gives

$$\Box\phi - \frac{\alpha}{\Lambda^4}\nabla_\mu\left(\partial^\mu\phi X\right) = -\frac{g}{M_*}T\,. \tag{2.63}$$

It is not obvious by simply examining this equation of motion that there is a screening mechanism at work to suppress the fifth force propagated by ϕ. Let us examine the radial field profile around a point source to study how kinetic screening works. With $T = -M\delta^{(3)}(x)$, restricting to radial coordinates r and integrating each side of the equation over a sphere of radius r the scalar field equation becomes

$$\phi' - \frac{\alpha}{\Lambda^4}\phi'^3 = \frac{gM}{4\pi r^2 M_*}\,, \tag{2.64}$$

where a prime indicates a derivative with respect to r. It is possible to solve this equation analytically for $\phi'(r)$, but for our purposes it is sufficient to examine the solution in the region close to the source. This scale is characterised by a crossover distance $r_* \equiv \left(gM/M_*\Lambda^2\right)^{1/2}$ where $r \ll r_*$ defines the region where the screening mechanism operates. In this region the solution to equation (2.64) is

$$\phi'(r)\Big|_{r\ll r_*} = \Lambda^2\left(\frac{r_*}{r}\right)^{2/3}\,. \tag{2.65}$$

The fifth force force goes as $\phi'(r) \sim r^{-2/3}$ and the standard gravitational force scales as r^{-2}. Therefore in regions close to a massive object the scalar force is suppressed relative to the gravitational force. We have worked through this example using a particularly simple k-essence model, however it can be shown to hold in more general models with higher powers of the first derivative of the scalar field appearing in the action [90, 91].

It is also possible to use higher derivatives to obtain a screening mechanism. We shall briefly review an example of higher derivative screening, also called Vainshtein screening [92], which works in a similar way to kinetic screening. Following the example in Ref. [62] we begin with a theory which contains higher derivatives of the scalar field

$$\mathcal{L} = -3X - \frac{1}{\Lambda^3}X\Box\phi + \frac{g}{M_*}\phi T\,. \tag{2.66}$$

Despite the higher derivatives which appear in the Lagrangian the equation of motion remains second order

$$6\Box\phi + \frac{2}{\Lambda^3}\left[(\Box\phi)^2 - (\nabla_\mu\nabla_\nu\phi)^2\right] = -\frac{g}{M_*}T\,. \tag{2.67}$$

Considering a point mass and examining the radial profile of the field around this source one finds that the equation of motion is given by

$$6\phi'(r) + \frac{4\phi'^2(r)}{\Lambda^3 r} = \frac{gM}{4\pi r^2 M_*} \,. \tag{2.68}$$

Defining the Vainshtein radius to be $r_V = (gM/M_*\Lambda^3)^{1/3}$ it is possible to examine the form of the solution for $\phi'(r)$ with Eq. (2.68) in the regime $r \gg r_V$ and $r \ll r_V$. In the first regime of $r \gg r_V$ the solution scales as $\phi' \sim r^{-2}$ which has the same behaviour as the standard gravitational force up to $\mathcal{O}(1)$ numerical factors. However, in the region close to the source the solution goes as

$$\phi'(r)\Big|_{r \ll r_V} \sim r^{-1/2} \,. \tag{2.69}$$

It is therefore suppressed relative to the r^{-2} scaling of the gravitational force, allowing standard GR to be recovered in the vicinity of the source.

2.3 Effective Field Theory

2.3.1 What is EFT?

It may not be too far fetched to state that a synonym for effective field theory is simply "physics". Any physical theory is only applicable within a certain energy range or range of length scales. Calculating the trajectory of a tennis ball thrown in the air does not necessitate the full mechanics of GR. Newtonian mechanics is sufficient. Similarly, when the ball hits the ground one does not need to use quantum mechanics to compute the electromagnetic force between every atom in the ball and the ground which stops it from falling through the earth. We care even less about the interactions between the quarks in the nuclei of the atoms when it comes to tennis ball throwing. In the same way, GR is incapable of describing the physical state at the singularity of a black hole. GR must be a low energy description of a theory with new degrees of freedom at a higher energies. Furthermore, the Standard Model of particle physics has currently been well verified up to technologically feasible energy scales of $\sim$ TeV. This cannot be the full story. The Higgs hierarchy problem and the presence of neutrino masses are only two issues which the Standard Model cannot account for. One must treat it as an effective field theory, valid only up to a certain energy scale, and introduce new operators which become relevant at higher energies which may help resolve some of these issues.

No theory is valid at all length scales. The infinities which arise in the conventional renormalisation process in standard QFT [93, 94] are a consequence of the assumption that the theory is valid on all length scales. These infinities can be removed by absorbing them into a redefinition of a finite number of coupling parameters whose value changes with energy scale such that one can predict measurable quantities. By insisting that the theory is renormalizable in the first place, the theory is automatically restricted to a finite set of operators. EFT accepts that it is not possible to write down a theory which is valid at all energy scales and so there is no problem including non-renormalizable operators in the theory. The drawback is that there are now infinitely many operators. In principle, this means an infinite set of couplings which would require an infinite set of measurements. Naively, the predictivity of the theory is destroyed. We are saved by the fact that EFT works within a finite range of energy scales within which not all of the operators are relevant. There is some cutoff energy scale in the theory Λ_c. The theory is capable of making definite physical predictions at energies below Λ_c because there are only a finite number of operators which are not suppressed below this cutoff.

Let us examine how this works in practice. We shall present a simple example of a scalar field theory in four dimensions which demonstrates the key ideas of EFT without getting lost in the technical details. We require the theory to respect Lorentz invariance and $\phi \to -\phi$ symmetry. The action is then

$$\mathcal{S} = \int d^4x \, \mathcal{L}(\phi, \partial_\mu \phi) \,, \tag{2.70}$$

where the Lagrangian is a Lorentz invariant function of the scalar field and its derivatives. For example, there could be terms that involve ϕ^2, X^2 or $X(\Box\phi)^2$ and so on, where $X \equiv \partial_\mu \phi \partial^\mu \phi$ and $\Box \equiv \nabla_\mu \nabla^\mu$. Of particular importance in this theory is the standard kinetic term X. This sets the dynamical scale of the theory. If we are working in a regime where the kinetic term is suppressed then the field equation will not contain any derivative terms and so ϕ is non-propagating. Therefore we assume that the theory in Eq. (2.70) is valid near the scale set by $X \sim M_*^4$. We shall now compare the relevance of all the other operators relative to this kinetic term.

As an example Lagrangian consider

$$\mathcal{L} = g_0 + g_2\phi^2 + g_4\phi^4 + (\partial\phi)^2 + g_6\phi^6 + \dots \,, \tag{2.71}$$

where the dots indicate the addition of a potentially infinite number of operators involving higher powers and derivatives of ϕ. Each coupling g_i has a mass dimension to ensure that the action itself is dimensionless. The mass dimension of a scalar field in four dimensions is $[\phi] = 1$ which can be derived by examining the kinetic term and noting that $[\partial_\mu] = 1$. The relevance of each term in the action can then be studied by introducing another mass scale Λ_c such that each g_i can be rendered dimensionless, their original dimension being accounted for by factors of Λ_c. The Lagrangian can then be written as

$$\mathcal{L} = \tilde{g}_0\Lambda_c^4 + \tilde{g}_2\Lambda_c^2\phi^2 + \tilde{g}_4\phi^4 + (\partial\phi)^2 + \frac{\tilde{g}_6}{\Lambda_c^2}\phi^6 + \dots \,, \tag{2.72}$$

where each $\tilde{g}_i$ is dimensionless. The scale Λ_c is the cutoff of the theory. Every higher dimensional operator involving higher powers and more derivatives of the scalar field are suppressed by powers of this cutoff. In order to obtain concrete predictions from the theory this cutoff needs to be larger than the dynamical scale of the field ϕ, i.e. $\Lambda_c >> M_*$. As $\phi \sim M_*$, the ratio of the mass term to the kinetic term goes as Λ_c^2/M_*^2. The mass term therefore becomes more relevant when the theory is applied at low energies where the cutoff is large relative to M_*. This is an EFT approach to understanding the Higgs hierarchy problem.

2.3.2 Broken symmetries and the unitary gauge

Postulating the existence of a scalar degree of freedom to drive the accelerated expansion of the Universe in both the inflationary era or the late-time dark energy dominated era may seem ad-hoc. A particularly powerful argument exists to suggest in fact that this can be rather natural. A light scalar degree of freedom can arise from any theory that possesses broken time translational symmetry. This section reviews this concept. Accelerated expansion in the background indicates broken time translational symmetry, which is then associated with a pseudo-Nambu Goldstone boson. *Pseudo* here means the symmetry is only approximately broken and the Goldstone boson acquires a small mass as a consequence.

A symmetry is said to be *spontaneously broken* if the ground state is no longer invariant under the symmetries of the full theory. In other words, if there is a collection of scalar fields Φ_i whose theory is globally invariant under a symmetry group represented by a matrix M, then the vacuum field configuration $\tilde{\Phi}_0$ changes under the global transformation such that $M\tilde{\Phi}_0 \neq 0$. If the broken generators of the transformation matrix are labelled τ_i the full transformation between degenerate vacua can be expressed as $M = \exp\left(i \sum_i \pi_i \tau_i\right)$. After promoting π_i to a field $\pi_i(x)$, it can now be interpreted as one field per broken group generator which are called *Goldstone bosons.* Let us consider a concrete example, namely the Higgs mechanism. The $SU(2) \times U(1)$ symmetry is broken and the Higgs doublet is given by excitations $h(x)$ around the minimum ν

$$\phi = (\nu + h(x), 0) \ . \tag{2.73}$$

The *unitary gauge* is defined such that all the Goldstone fields are zero. In this gauge one has explicitly broken the gauge symmetry and is left only with the relevant physical degrees of freedom that the theory contains. The Lagrangian is no longer invariant under the broken group generators but will still be invariant under the unbroken generators. In the case of the Higgs mechanism, after breaking the $SU(2) \times U(1)$ symmetry one still has a remaining $U(1)$ symmetry corresponding to the massless photon. However the other three Goldstone modes are absorbed by the gauge fields which become massive, thus breaking the gauge symmetry. It is possible to reintroduce the gauge symmetry by "undoing" the gauge transformation at the expense of reintroducing the Goldstones. This is called the Stückelberg method. Starting from a theory which has a broken symmetry, applying such a symmetry transformation will involve new terms appearing involving the Goldstones which now realise the symmetry in combination. For example, starting with a theory describing a free massive vector field A_μ with the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} m^2 A_\mu A^\mu \,, \tag{2.74}$$

where the field strength tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ we can perform a gauge transformation $A_\mu \to A_\mu - \frac{1}{q}\partial_\mu \pi$ to obtain

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} m^2 \left(A_\mu - \frac{1}{q}\partial_\mu \pi \right) \left(A^\mu - \frac{1}{q}\partial^\mu \pi \right) \,. \tag{2.75}$$

If we now treat the field π as the Goldstone field and make a further gauge transformation with a new field χ this will take $A_\mu \to A_\mu + \partial_\mu \chi$ and $\pi \to \pi + q\chi$. It is straightforward to check that this transformation leaves the theory in equation

(2.75) invariant.

How is this discussion relevant for cosmology? Throughout the inflationary and dark energy dominated epochs time translational symmetry is broken. The expansion of the Universe in both periods has a preferred time direction. Therefore it is natural to suppose that there is an associated Goldstone boson introduced through this breaking of time translations. A theory constructed in the unitary gauge involving a scalar field and gravity does not necessarily have to include the scalar field explicitly in the action. The dynamics of the scalar field can be absorbed by the metric with a specific choice of time foliation such that constant ϕ hypersurfaces correspond to constant t hypersurfaces (See also the discussion in Sec. 1.3.2). As we are constructing a theory which explicitly breaks time transitional symmetry, we shall now discuss a formulation of General Relativity which does precisely this. This provides the necessary formalism to write down an effective field theory of a scalar field and gravity.

2.3.3 The Arnowitt-Deser-Misner (ADM) formalism

Let us begin this section by noting some facts about the structure of the Einstein equations. The Einstein equation (1.24) is really ten partial differential equations which determine the dynamical evolution of the metric tensor $g_{\mu\nu}$ in the presence of a source of stress-energy $T_{\mu\nu}$. Four of these equations are constraint equations and only six are dynamical. To define a dynamical system there must be two time derivatives acting on the dynamical quantity. From the structure of the Riemann tensor in Eq. (1.23) only R_{i0j0} can contain two time derivatives acting on the spatial components of the metric g_{ij}. The time derivatives of g_{0i} and g_{00} do not appear in any of the equations, and no *second* time derivatives appear on g_{ij} in the space-time or time-time equations. Is it therefore possible to perform a split such that the generally covariant structure of GR is maintained but the relevant dynamical variables are made explicit? This is the basis the Arnowitt-Deser-Misner (ADM) formulation, or $(1+3)$ formulation, of General Relativity [21, 95]. The following discussion follows that in Ref. [76].

In this formalism four dimensional general covariance is broken by foliating the spacetime with a series of spacelike hypersurfaces. We can describe these hypersurfaces with a scalar field $t(x^\mu)$, such that $t =$ constant define the family of hypersurfaces $\Sigma(t)$. A congruence of curves can be set up to intersect the hypersurface at one point per hypersurface. They are not necessarily geodesics

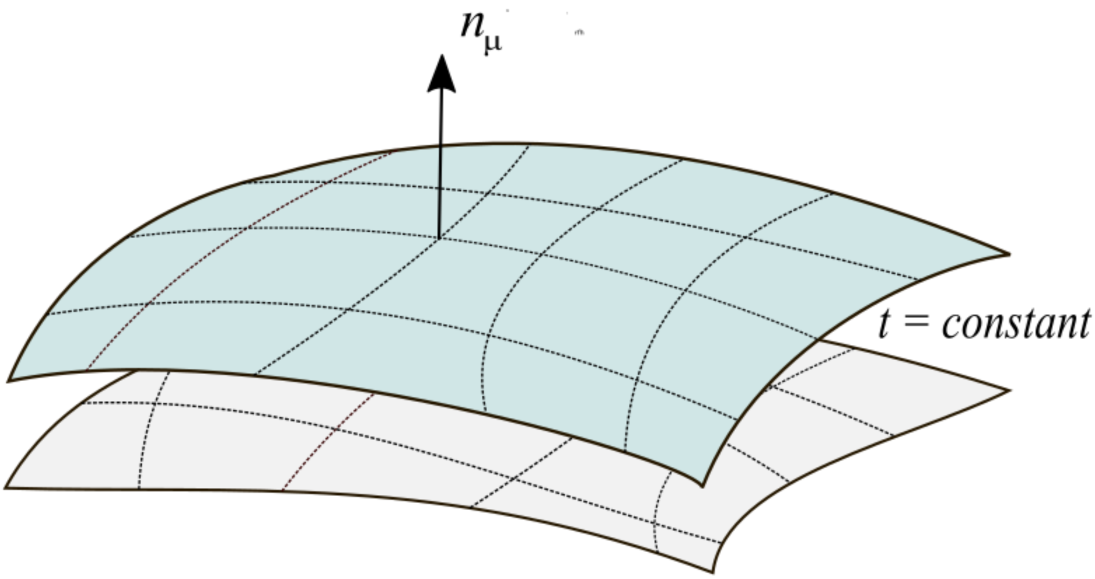

Figure 2.2 *Pictorial representation of the ADM decomposition of spacetime into constant time hypersurfaces.*

or orthogonal to the hypersurface. We shall denote the parameter along the curve as t. The unit normal vector to a hypersurface $\Sigma(t_*)$ where t has been set to a particular value t_*, is defined as

$$n_\mu = -N\partial_\mu t_*. \tag{2.76}$$

Here N is a normalisation, chosen such that $n_\mu n^\mu = -1$, and the negative sign ensures that it is timelike. Each $\Sigma(t)$ is given coordinates y^i. The smoothness of the spacetime is maintained with the congruence of curves that flow between the hypersurfaces. Each curve defines the trajectory through the spacetime that keeps y^i constant for successive values of t. This defines the coordinate system on the entire spacetime $x^\mu = (t, y^i)$. The *projection tetrad* is defined as

$$e^\mu_i = \frac{\partial x^\mu}{\partial y^i}\,, \tag{2.77}$$

which takes any vector and projects it along the i^{th} coordinate direction. A tangent vector belonging to the congruence of curves in the spacetime can then be decomposed in terms its components perpendicular to the hypersurface and orthogonal to the hypersurface such that

$$t^\mu = Nn^\mu + N^i e^\mu_i\,, \tag{2.78}$$

where the set of three functions N^i form a vector called the *shift*. From these definitions one can write the infinitesimal coordinate shift as

$$\begin{aligned} dx^\mu &= t^\mu dt + e^\mu_i dy^i \,, \\ &= N n^\mu dt + \left(N^i dt + dy^i\right) e^\mu_i \,. \end{aligned} \tag{2.79}$$

The spacetime line element is then

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \,, \\ &= -N^2 dt^2 + g_{\mu\nu} e^\mu_i e^\nu_j \left(N^i dt + dy^i\right)\left(N^j dt + dy^j\right) \,, \\ &= -N^2 dt^2 + h_{ij} \left(N^i dt + dy^i\right)\left(N^j dt + dy^j\right) \,, \end{aligned} \tag{2.80}$$

and the induced spatial metric is the projection of the full four dimensional metric onto the hypersurface $h_{ij} = g_{\mu\nu} e^\mu_i e^\nu_j$. By defining $h_{00} = N^i N_i$ and $h_{0i} = N_i$ we can write this in four dimensional notation as

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \,. \tag{2.81}$$

It is possible to define the notion of a covariant derivative that acts on vectors tangential to the spacelike hypersurface. The natural definition for such a derivative operator takes the usual covariant derivative operator acting on a vector V_μ satisfying $V^\mu n_\mu = 0$, and uses the induced metric to project onto the hypersurface. We then have that

$$D_\mu V_\nu = h^\sigma{}_\mu h^\rho{}_\nu \nabla_\sigma V_\rho \,. \tag{2.82}$$

The derivative operator acts on scalars as $D_\mu \phi = h^\sigma{}_\mu \nabla_\sigma \phi$.

The question of how $\Sigma(t)$ is precisely embedded in the spacetime manifold is addressed with idea of the extrinsic curvature tensor. Intuitively the more n_μ changes direction as it is moved around one hypersurface the more curved the hypersurface is within the bulk spacetime. The normal vector will not change if the slicing is flat leading to no extrinsic curvature. Quantitatively, the extrinsic curvature tensor $K_{\mu\nu}$ is defined as the covariant derivative of the normal vector, projected onto the hypersurface with the induced metric

$$K_{\mu\nu} = h^\sigma{}_\mu \nabla_\sigma n_\nu \,. \tag{2.83}$$

We shall now demonstrate that $K_{\mu\nu}$ is symmetric. The following proof requires the Frobenius theorem which states that

$$n_{[\mu}\nabla_{\lambda}n_{\nu]} = 0\,, \tag{2.84}$$

where the brackets indicate an antisymmetric sum over all the indices. Starting from the definition of $K_{\mu\nu}$ and expanding it out with the definition of $h_{\mu\nu}$ we have that

$$K_{\mu\nu} = \nabla_{\mu}n_{\nu} + n^{\lambda}n_{\mu}\nabla_{\lambda}n_{\nu}\,. \tag{2.85}$$

From the Frobenius theorem we can derive the fact that

$$n_{\mu}\nabla_{\lambda}n_{\nu} = n_{\lambda}\nabla_{\mu}n_{\nu} + n_{\mu}\nabla_{\nu}n_{\lambda} + n_{\nu}\nabla_{\lambda}n_{\mu} - n_{\lambda}\nabla_{\nu}n_{\mu} - n_{\nu}\nabla_{\mu}n_{\lambda}\,. \tag{2.86}$$

Inserting this identity in equation (2.85) and using $n_{\lambda}\nabla_{\mu}n^{\lambda} = 0$ and $n_{\mu}n^{\mu} = -1$ we arrive at the result

$$K_{\mu\nu} = \nabla_{\nu}n_{\mu} + n_{\nu}n^{\lambda}\nabla_{\lambda}n_{\mu} = K_{\nu\mu}\,. \tag{2.87}$$

The symmetry of the extrinsic curvature tensor is a useful property for future calculations. By examining the spatial components in equation (2.87) we find that $K_{ij} = \nabla_i n_j = N\Gamma^{0}{}_{ij}$. After expanding out the Christoffel symbol in terms of the induced metric it is possible to show that

$$K_{ij} = \frac{1}{2N}\left(\dot{h}_{ij} - D_i N_j - D_j N_i\right)\,, \tag{2.88}$$

where recall the covariant derivative D_i acts on tensors with spatial indices in the same way as ∇_{μ} acts on tensors with spacetime indices. Now we set the congruence of curves that parameterize the time coordinate to align exactly with the normal vector to each time slice. In this coordinate system one has that $N_i = 0$ and the extrinsic curvature becomes

$$K_{ij} = \frac{1}{2N}\dot{h}_{ij}\,. \tag{2.89}$$

This gives a natural geometrical interpretation of the extrinsic curvature tensor as the time derivative of the induced spatial metric.

Of course, we have previously encountered another form of curvature that does not rely on how the surface is embedded in a higher dimensional space, namely the Riemann curvature tensor in Eq. (1.23). This is an example of intrinsic curvature. It is natural to suppose that each hypersurface has its own intrinsic as well as extrinsic curvature. This is quantified with $R^{(3)}$, defined in the same way as the full Ricci scalar but using the induced metric h_{ij} in place of the full metric $g_{\mu\nu}$. It can be related to the extrinsic curvature and the full four-dimensional Ricci scalar through the *Gauss-Codazzi* relation

$$R^{(3)} = R - K_{\mu\nu}K^{\mu\nu} + K^2 - 2\nabla_\nu(n^\nu\nabla_\mu n^\mu - n^\mu\nabla_\mu n^\nu)\,, \tag{2.90}$$

which can be obtained by projecting out the full Riemann tensor onto the hypersurface using the induced metric (see Sec 12.2 of Ref. [76] for the full derivation).

The ADM formulation has proved to be very useful in the study of dark energy and modified gravity models. It can be used to define the unitary gauge for scalar-tensor theories by associating the scalar field with the uniform time hypersurface which absorbs the scalar field perturbations into the metric, greatly simplifying the computation of cosmological perturbations.

2.3.4 Effective Field Theory of Dark Energy

This section discusses the application of effective field theory to dark energy. The formalism originally was applied to inflation [96, 97] before being applied later to dark energy [7–9, 85, 97–103]. It provides an efficient and generalised description of the evolution of the cosmological perturbations in a large range of scalar-tensor theories. Following the spirit of effective field theory every operator that satisfies the symmetries we impose can be included in the theory. In this construction we use the ADM decomposition in order to break time diffeomorphism invariance. We construct the theory in the unitary gauge by associating ADM spacelike hypersurfaces to correspond to uniform scalar field hypersurfaces. The scalar field is therefore "hidden" in the choice of time coordinate. The EFT operators are the cosmological perturbations. Operators which respect spatial diffeomorphism invariance but which could change under a shift in the time coordinate can be included in the action. At the level of the background the action is given by

$$S^{(0,1)} = \frac{M_*^2}{2}\int d^4x\sqrt{-g}\left[\Omega(t)R - 2\Lambda(t) - \Gamma(t)\delta g^{00}\right]\,, \tag{2.91}$$

where the index $(0, 1)$ indicates that this action contains only background terms and a term depending on a first-order perturbation δg^{00}. The function $\Omega(t)$

introduces a non-minimal coupling between the scalar field and the metric. $\Lambda(t)$ can be added as a free function of time as it is consistent with the symmetries. The next term involves a first order perturbation δg^{00} which vanishes when computing the equations of motion so that $\Gamma(t)$ affects the background. One may wonder why there are no perturbations linear in δK. The reason is that $K = 3H$ on an FLRW background and so $\delta K \equiv K - 3H = \nabla_\mu n^\mu - 3H$. These terms can be absorbed into re-definitions of the terms that already appear in equation (2.91). Taking the variation of Eq. (2.91) with respect to the metric $g^{\mu\nu} \to g^{\mu\nu} + \delta g^{\mu\nu}$ we can obtain the background equations of motion which correspond to the Friedmann equations in a general scalar-tensor theory of gravity. See Eqs. (3.16) and (3.17) for their explicit form.

The utility of the EFT approach is that now we can write down an action which can describe the first order perturbations in a unified way. The action that is sufficient to describe the perturbations in Horndeski theory will be presented in chapter 3 (see Eq.(3.15)). Every perturbation that is compatible with the symmetries of broken time diffeomorphisms, such as δg^{00}, δK and $R^{(3)}$, are included as operators in the EFT expansion.

We shall now follow an example given in Ref. [7] which nicely demonstrates the principles involved in deriving the equations of motion from an EFT action. This is important if the theory is to be connected with observable parameters. We shall see that the EFT expansion incorporates the effect of a modified gravitational slip and Poisson equation. In the following we shall work on a perturbed FRW background metric in the Newtonian gauge keeping the two metric potentials independent

$$ds^2 = -(1+2\Psi)\,dt^2 + a^2(t)\,(1+2\Phi)\,d\mathbf{x}^2\,. \tag{2.92}$$

The Stückelberg transformation reintroduces the explicit dependence of the EFT action on the scalar field. By transforming the time coordinate by an infinitesimal shift such that

$$t \to t + \pi(x,t)\,, \tag{2.93}$$

the operators will change non-trivially. Gauge invariance is then restored, albeit realised in a nonlinear manner. Furthermore any function that explicitly depends on time will introduce terms that depend on π after Taylor expanding around the new time coordinate. For example $\Lambda(t)$ transforms as

$$\Lambda(t+\pi) \approx \Lambda(t) + \dot{\Lambda}(t)\pi + \frac{1}{2}\ddot{\Lambda}(t)\pi^2\,. \tag{2.94}$$

The perturbation operators also transform non-trivially under a time diffeomorphism. By using the tensor transformation law to transform the metric into the new coordinate system with $t' = t + \pi$ in equation (1.10) the time components of the metric and the extrinsic curvature transforms as [7]

$$g^{00} \to g^{00} + 2g^{0\mu}\partial_\mu \pi + g^{\mu\nu}\partial_\mu \pi \partial_\nu \pi \,, \tag{2.95}$$

$$g^{0i} \to g^{\mu i} + g^{\mu i}\partial_\mu \pi \,, \tag{2.96}$$

$$\delta K_{ij} \to \delta K_{ij} - \dot{H}\pi h_{ij} - \partial_i \partial_j \pi \,, \tag{2.97}$$

$$\delta K \to \delta K - 2\dot{H}\pi - \frac{1}{a^2}\nabla^2 \pi \,. \tag{2.98}$$

The scalar field equation of motion from the EFT is then obtained in the usual way by taking $\pi \to \pi + \delta\pi$ and requiring the variation of the action to be zero. With the perturbed FLRW metric in equation (2.92) we can calculate the Ricci scalar. If we restrict to scales much smaller than the horizon by sending $H \to 0$ in the resulting expression R becomes

$$\frac{1}{2}\sqrt{-g}R = -3\dot{\Psi}^2 + (\nabla\Phi)^2 - 2\nabla\Phi\nabla\Phi \,. \tag{2.99}$$

By combining equation (2.99) with equations (2.94) to (2.98) we can obtain an action describing the dynamics of the gravitational and scalar field perturbations from which the modified equations of motion are derived. Taking an action of the form

$$\frac{1}{2}\int \sqrt{-g}M_*^2\Omega(t)R + c(\partial\pi)^2 \,, \tag{2.100}$$

where c is a constant and then applying the Stückelberg method we obtain a series of terms involving powers of the perturbations. Varying this action with respect to Φ, Ψ and π gives an equation of motion for each. The first is a modified Poisson equation which takes the form

$$\left(1 - \frac{M_*^2\dot{\Omega}^2/\Omega}{4\left(c + M_*^2\dot{\Omega}^2/\Omega\right)}\right)\nabla^2\Psi = \frac{\delta\rho_m}{2M_*^2\Omega} \,. \tag{2.101}$$

Another quantifies the difference between the two metric potentials, called the gravitational slip $\eta \equiv -\Phi/\Psi$, which is given by

$$\eta = 1 + \frac{M_*^2 \dot{\Omega}^2/\Omega}{2\left(c + M_*^2\dot{\Omega}^2/\Omega\right)} \,. \tag{2.102}$$

In general the metric potentials will not be equal in a modified gravity or exotic dark energy model. This is why weak gravitational lensing can provide a powerful probe of such models (see Eq. (1.117)).

2.3.5 $\mathcal{L}_3$ in the unitary gauge

In this section we derive in detail how the Horndeski Lagrangian $\mathcal{L}_3$ in Eq. (2.40) can be expressed in the unitary gauge. This will therefore enable an effective field theory description of models with a nonzero $\mathcal{L}_3$, i.e. terms of the form $X^m \Box\phi$. As $K = \nabla^\mu n_\mu$ with $n_\mu = -\partial_\mu\phi/\sqrt{-X}$ we can express $\Box\phi$ as

$$\Box\phi = -\nabla^\mu\left(n_\mu\sqrt{-X}\right) \,. \tag{2.103}$$

The result depends on whether the m is even or odd as we shall soon see. After expanding out each derivative term $X^m\Box\phi$ becomes

$$X^m\Box\phi = -X^m\sqrt{-X}K + \frac{X^m}{2\sqrt{-X}}n^\mu\partial_\mu X \,. \tag{2.104}$$

The first term can readily be expanded in cosmological perturbations with $K = 3H + \delta K$ and $X = (1-\delta g^{00})\dot{\phi}^2$. The second term is not so trivial but we shall now derive another form. We now focus on the second term

$$\frac{X^m}{2\sqrt{-X}}n^\mu\partial_\mu X = -X\nabla_\mu\left\{\frac{X^m n^\mu}{2\sqrt{-X}}\right\} \,, \tag{2.105}$$

$$= -\frac{X^{m+1}}{2\sqrt{-X}}K - \frac{Xn^\mu}{2}\partial_\mu\left(\frac{X^m}{\sqrt{-X}}\right) + \text{b.t} \,, \tag{2.106}$$

$$= \pm\left\{\frac{(-X)^{m+1}}{2\sqrt{-X}}K - \frac{Xn^\mu}{2}\partial_\mu\left(\frac{(-X)^m}{\sqrt{-X}}\right)\right\} \,, \tag{2.107}$$

where the plus sign is for m even and and minus sign for m odd and b.t stands for a boundary term which can be neglected. As the only difference now is an overall minus sign we consider case of even m. The first term simplifies to

$$\frac{(-X)^{m+1}}{2\sqrt{-X}}K = \frac{1}{2}(-X)^{m+\frac{1}{2}}K. \tag{2.108}$$

After simplification and taking into account an integration by parts, the second term becomes

$$-\frac{Xn^{\mu}}{2}\partial_{\mu}\left(\frac{(-X)^{m}}{\sqrt{-X}}\right) = -\frac{Xn^{\mu}}{2}\partial_{\mu}(-X)^{m-\frac{1}{2}}\,, \tag{2.109}$$

$$= \frac{1}{2}(-X)^{m-\frac{1}{2}}n^{\mu}\partial_{\mu}X - \frac{1}{2}(-X)^{m+\frac{1}{2}}K + \text{b.t}\,, \tag{2.110}$$

$$= -\frac{n^{\mu}\partial_{\mu}(-X)^{m+\frac{1}{2}}}{2m+1} - \frac{1}{2}(-X)^{m+\frac{1}{2}}K\,, \tag{2.111}$$

$$= \frac{K(-X)^{m+\frac{1}{2}}}{2m+1} - \frac{1}{2}(-X)^{m+\frac{1}{2}}K + \text{b.t.} \tag{2.112}$$

Combining all of the expressions above we arrive at the final expression for $X^m\Box\phi$ we see that the second term cancels with (2.108) to give

$$X^m\Box\phi = \mp(-X)^{m+\frac{1}{2}}K \pm \frac{1}{(2m+1)}(-X)^{m+\frac{1}{2}}K \tag{2.113}$$

There will be additional contributions from the successive integrations by parts from any function $\xi(\phi)$ which multiplies $X^m\Box\phi$. Here we shall repeat the calculation above with fewer details in order to account for these extra terms. Taking the case of m being even we have as before

$$\xi(\phi)X^m\Box\phi = \xi(\phi)(-X)^{m+\frac{1}{2}}K + \frac{\xi(\phi)X}{2\sqrt{-X}}n^{\mu}\partial_{\mu}X\,. \tag{2.114}$$

After the integration by parts this becomes

$$\xi(\phi)X^m\Box\phi = -\frac{\xi(\phi)X^{m+1}}{2\sqrt{-X}}K - \frac{\xi(\phi)Xn^{\mu}}{2}\partial_{\mu}\left(\frac{X^m}{\sqrt{-X}}\right) + \frac{1}{2}\xi'(\phi)(-X)^{m+1}\,, \tag{2.115}$$

where the prime indicates a derivative with respect to ϕ. An extra two boundary terms arise from the integration by parts of the second term above. In other words

$$-\frac{\xi(\phi)Xn^{\mu}}{2}\partial_{\mu}\left(\frac{X^m}{\sqrt{-X}}\right) = \frac{\xi(\phi)}{2}(-X)^{m-\frac{1}{2}}n^{\mu}\partial_{\mu}X - \frac{\xi(\phi)}{2}(-X)^{m+\frac{1}{2}}K - \frac{1}{2}\xi'(\phi)(-X)^{m+1}, \tag{2.116}$$

An extra term arises from

$$-\frac{\xi(\phi)n^{\mu}\partial_{\mu}(-X)^{m+\frac{1}{2}}}{2m+1} = \frac{\xi(\phi)K(-X)^{m+\frac{1}{2}}}{2m+1} + \frac{\xi'(\phi)(-X)^{m+1}}{2m+1} \tag{2.117}$$

Putting these results together we have that

$$\xi(\phi)X^m\Box\phi = \mp\frac{2m}{2m+1}\xi(\phi)(-X)^{m+\frac{1}{2}}K \pm \frac{1}{2m+1}\xi'(\phi)(-X)^{m+1}, \qquad (2.118)$$

where the sign on top indicates the solution for m being even. This expression is now in a useful form to obtain the cosmological perturbations of any term that appears in $\mathcal{L}_3$.

2.3.6 $\mathcal{L}_4$ in the unitary gauge

Here we shall present the full calculation of transforming $\mathcal{L}_4$ in Eq. (2.41) into the unitary gauge, following Ref. [100]. an important term to analyse is the second spacetime derivative of ϕ.

$$\nabla_\mu\nabla_\nu\phi\,. \qquad (2.119)$$

as partial derivatives commute $\partial_\mu\partial_\nu$ and $\Gamma^\sigma_{\mu\nu}$ is symmetric in μ and ν Eq. (2.119) is symmetric. In order to simplify the notation we follow Ref. [100] and define $\gamma \equiv 1/\sqrt{-X}$ for this section only. The acceleration vector is defined as

$$\dot{n}_\mu = n^\sigma\nabla_\sigma n_\mu\,. \qquad (2.120)$$

It can be related to the extrinsic curvature via

$$\nabla_\mu n_\nu = K_{\mu\nu} - n_\mu\dot{n}_\nu\,. \qquad (2.121)$$

Using these relations one finds that

$$\begin{aligned}\nabla_\mu\nabla_\nu\phi &= -\nabla_\mu\left(\frac{n_\nu}{\gamma}\right)\,,\\ &= -\gamma^{-1}\nabla_\mu n_\nu - n_\nu\partial_\mu\sqrt{-X}\,,\\ &= -\gamma^{-1}\left(K_{\mu\nu} - n_\mu\dot{n}_\nu\right) + \frac{\gamma}{2}n_\nu\partial_\mu X\,.\end{aligned} \qquad (2.122)$$

where in the final line we have used the expression in equation (2.121). The last term in the expression above can be decomposed in the following way

$$\begin{aligned}\frac{\gamma}{2}n_\nu\partial_\mu X &= \frac{\gamma}{2}n_\nu g_{\mu\sigma}\partial^\sigma X \\ &= \frac{\gamma}{2}n_\nu h_{\mu\sigma}\partial^\sigma X - \frac{\gamma}{2}n_\nu n_\mu n_\sigma\partial^\sigma X \\ &= \frac{\gamma}{2}n_\nu h_{\mu\sigma}\partial^\sigma X + \frac{\gamma^2}{2}n_\nu n_\mu \partial_\sigma\phi\partial^\sigma X\,.\end{aligned} \tag{2.123}$$

The final step is to show that the first term above can be written as $\gamma^{-1}n_\nu\dot{n}_\mu$. Staring from the definition of $\dot{n}_\mu$ we have that

$$\begin{aligned}\dot{n}_\mu &= -n^\lambda\nabla_\lambda(\gamma\nabla_\mu\phi)\,, \\ &= -n^\lambda\nabla_\mu\phi\nabla_\lambda\gamma - \gamma n^\lambda\nabla_\lambda\nabla_\mu\phi\,, \\ &= \gamma^{-1}n^\lambda n_\mu\nabla_\lambda\gamma - \gamma n^\lambda\nabla_\lambda\nabla_\mu\phi\,, \\ &= \gamma^{-1}\left(h_\mu^\lambda - \delta_\mu^\lambda\right)\nabla_\lambda\gamma - \gamma n^\lambda\nabla_\lambda\nabla_\mu\phi\,, \\ &= \gamma^{-1}h_\mu^\lambda\nabla_\lambda\gamma - \gamma^{-1}\nabla_\mu\gamma - \gamma n^\lambda\nabla_\lambda\nabla_\mu\phi\,.\end{aligned} \tag{2.124}$$

Let us first take a look at the final term. Commuting the derivatives acting on ϕ

$$\begin{aligned}\gamma n^\lambda\nabla_\lambda\nabla_\mu\phi &= \gamma n^\lambda\nabla_\mu\nabla_\lambda\phi\,, \\ &= -\gamma n^\lambda\nabla_\mu(\gamma^{-1}n_\lambda)\,, \\ &= \gamma\nabla_\mu\gamma^{-1}\,, \\ &= -\gamma^{-1}\nabla_\mu\gamma\,,\end{aligned} \tag{2.125}$$

where in the third line we used the fact that $n^\lambda\nabla_\mu n_\lambda = 0$. This then brings about a cancellation such that

$$\begin{aligned}\dot{n}_\mu &= \gamma^{-1}h_\mu^\lambda\nabla_\lambda\gamma\,, \\ &= \gamma^{-1}h_\mu^\lambda\nabla_\lambda(-X)^{-\frac{1}{2}}\,, \\ &= \frac{\gamma^2}{2}h_\mu{}^\lambda\nabla_\lambda X\,.\end{aligned} \tag{2.126}$$

Plugging this back into Eq. (2.119) and using its trace, it is possible to show that [100]

$$\mathcal{L}_4 = G_4R^{(3)} + (2XG_{4X} - G_4)\left(K^2 - K_{\mu\nu}K^{\mu\nu}\right) - 2\sqrt{-X}G_{4\phi}K\,. \tag{2.127}$$

This expression forms the basis for studying the cosmological perturbations of $\mathcal{L}_4$. We do not discuss $\mathcal{L}_5$ as it does not introduce new free functions at the level of the background or linear pertrbations.

Chapter 3

Reconstructing Horndeski theories from the effective field theory of dark energy

You could find out most things, if you knew the right questions to ask. Even if you didn't, you could still find out a lot.

Ian M. Banks

Identifying the nature of the observed late-time accelerated expansion of the Universe [46, 47] is one of the major outstanding problems in physics. The cosmological constant provides the simplest explanation but, as discussed in Sec. 2.1, it is associated with a range of theoretical challenges [58]. We therefore discussed in the previous chapter the approach of including an additional dark energy component in the matter sector or modifying General Relativity on cosmological scales [23, 62, 71, 104], to address the observed cosmic acceleration without necessarily including a cosmological constant. Large-scale modifications of gravity may be motivated by low-energy extra degrees of freedom that could arise as effective remnants of a more fundamental theory of gravity and couple to the metric non-minimally (see Sec. 2.2.6). Moreover, non-standard gravitational effects can also be of interest to address problems in the cosmological small-scale structure [105]. Cosmological observations provide a new laboratory for tests of gravity that differ by about fifteen orders of magnitude in length scale to the more conventional tests in the Solar System [87]. Therefore it is well worth studying the range of possible large-scale modifications that can arise and the independent constraints on them that can be inferred from cosmology.

In the simplest case the modification is introduced by a universally non-minimally coupled scalar field. This is the scenario considered throughout this thesis. Recall from Sec. 2.2.5 that the most general scalar-tensor theory introducing at most second-order equations of motion to evade Ostrogradsky instabilities is described by the Horndeski action [63, 81, 106]. Despite providing restrictions on the space of possible scalar-tensor models, there remains considerable freedom within Horndeski theory. As a result, testing any observational consequences of the free functions in the Horndeski action directly is inefficient. It is necessary to solve the equations of motion for each model that one wishes to test in turn, and then compare it with observations.

Fortunately the formalism of effective field theory (EFT), introduced in Sec. 2.3, addresses these issues. One starts from the bottom up, with minimal assumptions about the underlying theory, and then constrains a smaller set of functions that parametrize a much larger class of covariant theories. It has proved to be a fruitful approach. For example, it was shown using EFT that Horndeski theories cannot yield an observationally compatible self-acceleration that is genuinely due to modified gravity, unless the speed of gravitational waves significantly differs from the speed of light [107, 108], now known to be incompatible with observations [109] (see Sec. 4.3.2). The same techniques used in EFT were also utilized in the discovery that there exists a class of scalar-tensor theories that contain higher order time derivatives, yet still avoid ghost-like instabilities [110] (also see Ref. [111]). Further applications can be found in Ref. [112–119].

Despite the utility of EFT, some issues remain to be addressed. For instance, it is not clear whether the chosen parametrization of the EFT functions arises naturally in modified gravity models [120–122]. Moreover, constraints on parametrized EFT functions describing the cosmological background and perturbations around it, cannot be connected to the non-perturbative nonlinear regime or to different backgrounds than the cosmological setting. This omits, for instance, constraints arising from the requirement of screening effects [62] in high-density regions. Hence, in order to connect the observational constraints and interpret them in terms of the allowed forms of the Horndeski functions, one requires a covariant description of the phenomenological modifications adopted.

In this chapter we present the reconstruction of a baseline covariant scalar-tensor action from the EFT functions of a second-order unitary gauge action, defined in Sec. 3.1, that shares the same cosmological background and linear perturbations around it. Variations can then be applied to this action to move to another covariant theory that is equivalent at the background and linear perturbation level. This reconstruction enables measurements of parametrized EFT functions to be related to a range of sources from the covariant Horndeski terms, which can then be used to address the theoretical motivation of the phenomenological parameterizations, a topic discussed in chapter 4. It can also be employed to extend predictions to the nonlinear sector or to non-cosmological environments and implement screening conditions on the theoretical parameter space. This shall be explored in chapter 5.

The chapter is organised as follows. In Sec. 3.1, for convenience we briefly review Horndeski scalar-tensor theory and the unitary gauge formalism that provides the tools for an EFT approach to the cosmological perturbations. More details can be found in chapter 2. We then present in Sec. 3.2 the covariant action that is constructed to reproduce the unitary gauge action up to second order in the perturbations and hence yield the equivalent cosmological background dynamics and the linear perturbations around it. In Sec. 3.3, the derivation of the reconstructed action is discussed, before applying it to a few simple example models in Sec. 3.4. Finally, we present the conclusion of this chapter in Sec. 3.5.

3.1 Horndeski gravity and effective field theory

Horndeski gravity [63, 81, 106] describes the most general local, Lorentz-covariant, four-dimensional theory of a single scalar field interacting with the metric that yields at most second-order equations of motion and hence avoids Ostrogradsky instabilities. We have encountered this action in Sec. 2.2.5 but we shall repeat it here for convenience. It is given by

$$S = \sum_{i=2}^{5} \int d^4x \sqrt{-g}\, \mathcal{L}_i \,, \tag{3.1}$$

where the four Lagrangian densities are defined as

$$\begin{aligned}
\mathcal{L}_2 &\equiv G_2(\phi, X)\,, &(3.2)\\
\mathcal{L}_3 &\equiv G_3(\phi, X)\Box\phi\,, &(3.3)\\
\mathcal{L}_4 &\equiv G_4(\phi, X)R \\
&\quad -2G_{4X}(\phi, X)\left[(\Box\phi)^2 - (\nabla^\mu\nabla^\nu\phi)(\nabla_\mu\nabla_\nu\phi)\right]\,, &(3.4)\\
\mathcal{L}_5 &\equiv G_5(\phi, X)G_{\mu\nu}\nabla^\mu\nabla^\nu\phi \\
&\quad +\frac{1}{3}G_{5X}(\phi, X)\left[(\Box\phi)^3 - 3(\Box\phi)(\nabla_\mu\nabla_\nu\phi)(\nabla^\mu\nabla^\nu\phi)\right. \\
&\quad \left.+2(\nabla_\mu\nabla_\nu\phi)(\nabla^\sigma\nabla^\nu\phi)(\nabla_\sigma\nabla^\mu\phi)\right]\,, &(3.5)
\end{aligned}$$

were $X \equiv g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. These Lagrangians have been studied in a variety of different systems including black holes [123, 124], neutron stars [125, 126] and inflationary models [127, 128]. For cosmological purposes, at the background and linear level, it has proven useful to adopt a unitary gauge description of Eq. (3.1) [7, 85, 100, 101]. In this EFT formalism the freedom in the cosmological background metric and each $G_i(\phi, X)$ reduces to five free time-dependent functions. One describes the background dynamics while the other four functions encompass the linear perturbations around it.

In the following, we shall briefly discuss the principles that go into building this EFT for the cosmological dynamics in the unitary gauge (see Refs. [7, 85] for more details). The Friedmann-Lemaître-Robertson-Walker (FLRW) background metric is defined as

$$ds^2 = -dt^2 + a^2(t)d\mathbf{x}^2\,, \tag{3.6}$$

where $a(t)$ is the scale factor. The general procedure then invokes the Arnowitt-Deser-Misner (ADM) formalism of General Relativity (see Sec. 2.3.3 for more details) on an FLRW background to foliate the spacetime with spacelike hypersurfaces. The ADM line element is given by [95]

$$ds^2 = -N^2dt^2 + h_{ij}\left(dx^i + N^i dt\right)\left(dx^j + N^j dt\right)\,, \tag{3.7}$$

where N is the lapse, N^i is the shift and h_{ij} is the induced metric on the spacelike hypersurface. The induced metric can also be written in four-dimensional notation as

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu\,, \tag{3.8}$$

by identifying $h_{00} = N^iN_i$ and $h_{0i} = N_i$. This framework provides a natural motivation for the introduction of the scalar field by treating it as the pseudo-Nambu-Goldstone boson of spontaneously broken time translational symmetry [85, 129]. By associating the time coordinate with the scalar field, the scalar perturbations are absorbed into the metric. One is free to choose the functional form of the spacetime foliation, as long as the scalar field is a smooth function with a time-like gradient. We can then simplify the calculations by setting

$$\phi = t\, M_*^2\,, \tag{3.9}$$

where M_* is a mass scale to match the dimensions. It can be thought of as a bare Planck mass related to the physical Planck mass through corrections from the EFT parameters [7]. Note that as the coordinate time is related to the scale factor in the FLRW background metric $a(t)$, and this in turn is related to the matter content of the universe through the Friedmann equations, the gravitational action and the matter action are now no longer independent after this identification has been made.

In this unitary gauge, we furthermore have

$$X = g^{00}\dot{\phi}^2 = (-1+\delta g^{00})M_*^4\,, \tag{3.10}$$

where g^{00} is related to the lapse via $g^{00} = -N^{-2}$. Here and throughout this chapter dots denote time derivatives and primes will represent derivatives with respect to the scalar field ϕ. Another geometrical quantity that will be used in the EFT action is the extrinsic curvature $K_{\mu\nu}$ defined as

$$K_{\mu\nu} = h_{\mu\sigma}\nabla^{\sigma} n_{\nu}\,, \tag{3.11}$$

where n_μ is the normal vector on the uniform time hypersurface,

$$n_\mu = -\frac{\delta^0_\mu}{\sqrt{-g^{00}}}\,. \tag{3.12}$$

On a spatially flat FLRW background $K_{\mu\nu} = H h_{\mu\nu}$, where $H \equiv \dot{a}/a$ is the Hubble parameter, and hence the perturbation of the extrinsic curvature becomes $\delta K_{\mu\nu} = K_{\mu\nu} - H h_{\mu\nu}$. The final geometrical quantity that will be used is the three dimensional Ricci scalar $R^{(3)}$, defined in the usual way but with the metric $h_{\mu\nu}$.

The full unitary gauge action that describes the background and linear dynamics of Horndeski gravity is then given by [7, 85, 100, 101]

$$S = S^{(0,1)} + S^{(2)} + S_M[g_{\mu\nu}, \psi]\,, \tag{3.13}$$

where

$$S^{(0,1)} = \frac{M_*^2}{2} \int d^4x \sqrt{-g}\left[\Omega(t)R - 2\Lambda(t) - \Gamma(t)\delta g^{00}\right]\,, \tag{3.14}$$

and

$$\begin{aligned} S^{(2)} = \int d^4x \sqrt{-g}\, \bigg[& \frac{1}{2}M_2^4(t)(\delta g^{00})^2 - \frac{1}{2}\bar{M}_1^3(t)\delta K \delta g^{00} \\ & -\bar{M}_2^2(t)\left(\delta K^2 - \delta K^{\mu\nu}\delta K_{\mu\nu} - \frac{1}{2}\delta R^{(3)}\delta g^{00}\right)\bigg]\,. \end{aligned} \tag{3.15}$$

For the zeroth and first-order action $S^{(0,1)}$ we have adopted the notation of Ref. [113]. $S^{(2)}$ is the action at second order and S_M is the matter action with minimal coupling between metric and matter fields. Note here that $R^{(3)}$ is itself a perturbation on flat FLRW. Although everything in this work assumes flat space we keep the above notation of $\delta R^{(3)}$ to emphasize that it is a first- order quantity throughout.

The EFT action (3.13) separates out the background dynamics and the perturbations around it in a systematic way. We have six free functions of time, where a seventh free function of time enters through the FLRW metric with the scale factor $a(t)$ or equivalently $H(t)$. Four free functions are introduced at the background level, while another three enter the dynamics of the linear perturbations. Note, however, that two of the background EFT functions in Eq. (3.14), including $H(t)$, will be fixed by the Friedmann equations with a specified matter content. Given $H(t)$, this leaves a degenerate background function which is only fixed at the level of the linear perturbations. The separation of the background and linear perturbations is more manifest in the notation introduced in Ref. [8], in which there is one free function $H(t)$ that determines the background evolution and four free functions describing the perturbations. More specifically, the background equations that follow from the EFT action, providing the two constraints, are given by [7, 113]

$$\Gamma \quad + \Lambda = 3(\Omega H^2 + \dot{\Omega} H) - \frac{\rho_m}{M_*^2}\,, \tag{3.16}$$

$$\Lambda \quad = 2\Omega \dot{H} + 3\Omega H^2 + 2\dot{\Omega} H + \ddot{\Omega}\,, \tag{3.17}$$

where we assumed a matter-only universe with pressureless dust.

Finally, an important aspect of the unitary gauge action (3.13) for the discussion in Secs. 3.2 and 3.3 is that, at the level of linear theory, no new EFT functions appear in the description of $\mathcal{L}_5$ in addition to those introduced for $\mathcal{L}_{1-4}$ [100, 101]. Hence, it will be sufficient to consider the reconstruction of a baseline covariant action for $\mathcal{L}_{1-4}$ only.

3.2 Reconstructed Horndeski action

So far, much work has been devoted to representing specific theories in terms of the unitary gauge EFT parameters and devising parametrizations of the time-dependent EFT functions (see, e.g. Ref. [7–9, 85, 100, 101, 130]). Here we are interested in the inverse procedure. That is, the class of covariant theories that a set of EFT functions corresponds to. While a previous reconstruction was presented in Ref. [7], the resulting general covariant action is not of the Horndeski type. Therefore it is not guaranteed to be theoretically stable. We shall now present a covariant formulation of a scalar-tensor theory that is embedded in the Horndeski action (3.1) and is reconstructed from the free EFT functions of the second-order unitary gauge action (3.13) such that they share the same cosmological background and linear dynamics. Given that it is not possible to specify a unique covariant theory based on its background and linear theory only, the reconstructed action will serve as a foundation upon which variations can then be applied to move between different covariant theories that are equivalent at the background and linear perturbation level. The basis of this reconstruction is the correspondence between the covariant formalism and the particular unitary gauge adopted, specified by Eq. (3.10).

The covariant Horndeski action that reproduces the same dynamics of the cosmological background and linear perturbations as the EFT action (3.13) is given by (see Sec. 3.3 for a derivation)

$$\begin{aligned} G_2(\phi, X) = &- M_*^2 U(\phi) - \frac{1}{2} M_*^2 Z(\phi) X + a_2(\phi) X^2 \\ &+ \Delta G_2 \,, \end{aligned} \tag{3.18}$$

$$G_3(\phi, X) = b_0(\phi) + b_1(\phi) X + \Delta G_3 \,, \tag{3.19}$$

$$G_4(\phi, X) = \frac{1}{2} M_*^2 F(\phi) + c_1(\phi) X + \Delta G_4 \,, \tag{3.20}$$

$$G_5(\phi, X) = \Delta G_5 \,, \tag{3.21}$$

$U(\phi) = \Lambda + \frac{\Gamma}{2} - \frac{M_2^4}{2M_*^2} - \frac{9H\bar{M}_1^3}{8M_*^2} - \frac{(\bar{M}_1^3)'}{8} + \frac{M_*^2(\bar{M}_2^2)''}{4} + \frac{7(\bar{M}_2^2)'H}{4} + \bar{M}_2^2 H' + \frac{9H^2\bar{M}_2^2}{2M_*^2}$	
$Z(\phi) = \frac{\Gamma}{M_*^4} - \frac{2M_2^4}{M_*^6} - \frac{3H\bar{M}_1^3}{2M_*^6} + \frac{(\bar{M}_1^3)'}{2M_*^4} - \frac{(\bar{M}_2^2)''}{M_*^2} - \frac{H(\bar{M}_2^2)'}{M_*^4} - \frac{4H'\bar{M}_2^2}{M_*^4}$	
$a_2(\phi) = \frac{M_2^4}{2M_*^8} + \frac{(\bar{M}_1^3)'}{8M_*^6} - \frac{3H\bar{M}_1^3}{8M_*^8} - \frac{(\bar{M}_2^2)''}{4M_*^4} + \frac{H(\bar{M}_2^2)'}{4M_*^6} + \frac{H'\bar{M}_2^2}{M_*^6} - \frac{3H^2\bar{M}_2^2}{2M_*^8}$	
$b_0(\phi) = 0$	$b_1(\phi) = \frac{2H\bar{M}_2^2}{M_*^6} - \frac{(\bar{M}_2^2)'}{M_*^4} + \frac{\bar{M}_1^3}{2M_*^6}$
$F(\phi) = \Omega + \frac{\bar{M}_2^2}{M_*^2}$	$c_1(\phi) = \frac{\bar{M}_2^2}{2M_*^4}$

Table 3.1 *The coefficients of powers of X in the Horndeski functions $G_i(\phi, X)$, Eqs.* (3.18) *through* (3.20)*, reconstructed from the EFT functions of the unitary gauge action* (3.13) *(Sec. 3.3).*

where the functional forms of the coefficients of X^n are presented in Table 3.1. The notation in Eqs. (3.18) through (3.20) is motivated such that Eq. (3.13) reduces to the scalar-tensor action of Ref. [131] in the limit that $a_2 = b_{0,1} = c_1 = 0$. The variations ΔG_i characterize the changes that can be performed on the baseline action ($\Delta G_i = 0$) to move between different covariant actions that are degenerate at the level of background and linear cosmology. For example, one may add terms to G_2 which are $\mathcal{O}\left[(1 + X/M_*^4)^3\right]$. In the unitary gauge these terms will be at least of order $(\delta g^{00})^3$ and hence do not affect linear theory. Similarly, after one takes into account an integration by parts relating terms in $b_0(\phi)$ and $Z(\phi)$ the variations ΔG_3 are $\mathcal{O}\left[(1 + X/M_*^4)^3\right]$. In fact, any non-zero contribution in $b_0(\phi)$ can be absorbed into $Z(\phi)$ in this way. Given this freedom, we have set b_0 to zero by default. The ΔG_4 term must be $\mathcal{O}\left[(1 + X/M_*^4)^4\right]$, which is due to the presence of G_{4X} in Eq. (3.35), changing anything of $\mathcal{O}\left[(1 + X/M_*^4)^4\right]$ to $\mathcal{O}\left[(1 + X/M_*^4)^3\right]$ with the variation having no effect on linear theory. Finally, as emphasized in Sec. 3.1, at the linear level contributions from G_5 can be absorbed into G_2, G_3, and G_4, and so the first term that appears in G_5 only affects nonlinear scales. As $\mathcal{L}_5$ in the unitary gauge has at most one X derivative acting on G_5 [100], as with ΔG_4, ΔG_5 starts at $\mathcal{O}\left[(1 + X/M_*^4)^4\right]$.

Importantly, note that the coefficients in Eqs. (3.18) through (3.20) are not independent since there are only five free independent EFT functions in Eq. (3.13). Hence, this leads to constraint equations between the coefficients. Another aspect worth noting is that due to the variations of the form $(1 + X/M_*^4)^n$ around the baseline covariant theory expressed in orders of X^n, the variations introduce well defined changes to all orders of each G_i in Eqs. (3.18) through (3.20). The functional form of each ΔG_i is specified by

$$\Delta G_{2,3} = \sum_{n>2} \xi_n^{(2,3)}(\phi) \left(1 + \frac{X}{M_*^4}\right)^n , \tag{3.22}$$

$$\Delta G_{4,5} = \sum_{n>3} \xi_n^{(4,5)}(\phi) \left(1 + \frac{X}{M_*^4}\right)^n , \tag{3.23}$$

where $\xi_n^{(i)}(\phi)$ are a set of n free functions for each ΔG_i. Note that, using the reconstruction, one can build a model with a non-zero constant EFT function Λ and all the other EFT functions set to zero. As the addition of a ΔG_i term does not affect linear theory, by adding these extra terms one can construct a theory that can only be discriminated from ΛCDM on nonlinear scales.

Given a set of unitary gauge EFT functions $\Omega, \Gamma, \Lambda, M_2^4, \bar{M}_1^3, \bar{M}_2^2$ and H, one can plug them into the relations given in Table 3.1 and Eqs. (3.18) through (3.20) and derive the corresponding baseline covariant action. However, it is important to stress again that the action obtained in the process is not unique. Indeed, it may require the addition of specific ΔG_i as well as several field redefinitions to recover a recognisable form for a given theory. Examples of this are given in Sec. 3.4.

Finally, for ease of use, we present in Table 3.2 the relation of the EFT functions we have adopted to different parameterizations that are frequently used in the literature. These expressions can be thought of as consistency relations. For example, we have the relationship between the background conformal factor Ω, the mass scale M and the speed of gravitational waves c_T^2

$$\Omega(t) = \frac{M^2}{M_*^2} c_T^2 \,. \tag{3.24}$$

As discussed in Ref. [107], a cosmological self-acceleration that is genuinely due to modified gravity implies a significant evolution in Ω departing from the value $\Omega = 1$ of General Relativity. The relation (3.24) makes it explicit that this requires a deviation of the Planck mass from its bare value M_*, or a speed of gravitational waves that differs from that of light. It hence tests the consistency of a self-acceleration effect between the cosmological background, the large-scale structure, and the propagation of gravitational waves.

EFT functions	Notation in Ref. [7]	α-parametrization
$\Omega(t)$	$f(t)$	$\frac{M^2}{M_*^2}c_T^2$
$\Gamma(t)$	$\frac{2c(t)}{M_*^2}$	$-\frac{\rho_m}{M_*^2}-\frac{M^2}{M_*^2}\beta(t)$
$\Lambda(t)$	$\frac{\Lambda(t)-c(t)}{M_*^2}$	$\frac{M^2}{M_*^2}\left[3H^2c_T^2(1+\alpha_M)+\beta(t)+3H\dot{\alpha}_T\right]$
$M_2^4(t)$	$M_2^4(t)$	$\frac{1}{4}\rho_m+\frac{M^2}{4}\left[H^2\alpha_K+\beta(t)\right]$
$\bar{M}_1^3(t)$	$m_3^3(t)$	$M^2\left[H\alpha_M c_T^2+\dot{\alpha}_T-2H\alpha_B\right]$
$\bar{M}_2^2(t)$	$m_4^2(t)$	$-\frac{1}{2}M^2\alpha_T$

Table 3.2 *Relationship of the EFT functions adopted in this thesis to the notation used in Ref. [7]. We have also derived here the expressions of the EFT functions in terms of the α-parametrization of Ref. [8] (with conventions of Ref. [9]). Dots denote derivatives with respect to physical time t, $c_T^2=1+\alpha_T$ is the tensor sound speed squared, and we have defined here $\beta(t)\equiv c_T^2\left[2\dot{H}+H\dot{\alpha}_M+\alpha_M\left(\dot{H}-H^2+H^2\alpha_M\right)\right]+H\dot{\alpha}_T(2\alpha_M-1)+\ddot{\alpha}_T$ for reasons of compactness.*

3.3 Reconstruction Method

We shall now provide a derivation of the reconstructed covariant Horndeski action presented in Sec. 3.2. The general approach to this reconstruction is as follows. We consider the sequence of terms of the unitary gauge action (3.13) contributing at zeroth, first, and second- order. We contrast those with the different $\mathcal{L}_i$ contributions to the covariant Horndeski Lagrangian, Eqs. (3.2) through (3.4). For this, we put them into the unitary gauge, which is a well defined procedure that has been dealt with in previous work [100, 101]. This will identify the Lagrangians that include the required terms in the unitary gauge action, but those will also give rise to extra terms. Using Eq. (3.10) it is possible to make these extra terms covariant and subtract them from the Horndeski Lagrangian that one originally started with. By construction, one is left with a covariant action that reduces to the required terms in the unitary gauge action after making that transformation. This procedure is only necessary for $\mathcal{L}_3$ and $\mathcal{L}_4$, where for $\mathcal{L}_2$ the reconstruction is straightforward. As discussed in Sec. 3.2, $\mathcal{L}_5$ does not introduce terms in the unitary gauge additional to the contributions arising from $\mathcal{L}_{2-4}$ and can thus be omitted. With this procedure we obtain a self consistent and well

defined reconstruction of a baseline covariant theory from the unitary gauge action that shares the same cosmological background and linear perturbations around it, and to which variations can be applied to move to another covariant theory that is equivalent at the background and linear perturbation level (Sec. 3.2). For the discussion of reconstructing a covariant action from the terms in $S^{(2)}$, we introduce the notation $S_i^{(2)}$ with $i = 1, 2, 3$ referring to $S^{(2)}$ with all EFT parameters set to zero apart from M_2^4, $\bar{M}_1^3$, and $\bar{M}_2^2$, respectively.
In Sec. 3.3.1, we discuss the quadratic contribution to Eq. (3.2) arising from the zeroth and first-order EFT action (3.14). The derivation of the first cubic contribution to Eq. (3.3) from second-order perturbations in the EFT action is discussed in Sec. 3.3.2. Finally, the quartic term, Eq. (3.4), is derived in Sec. 3.3.3.

3.3.1 Quadratic term $\mathcal{L}_2$

To start, consider the unitary gauge action up to first order in the perturbations,

$$S_{\Omega=1}^{(0,1)} = \frac{M_*^2}{2} \int d^4x \sqrt{-g} \left\{ R - 2\Lambda(t) - \Gamma(t)\delta g^{00} \right\} , \tag{3.25}$$

where we have set $\Omega = 1$ ($\Omega \neq 1$ will be considered in Sec. 3.3.3). The corresponding covariant action can be obtained through Eq. (3.10), which yields

$$S_{\Omega=1}^{(0,1)} = \int d^4x \sqrt{-g} \left\{ \frac{M_*^2}{2} R - M_*^2 \Lambda(\phi) - \frac{M_*^2}{2} \Gamma(\phi) - \frac{\Gamma(\phi)}{2M_*^2} X \right\} . \tag{3.26}$$

This is simply the action of a quintessence model with a non-canonical kinetic term (see Sec. 3.4.1).

The contribution of the first second-order perturbation in the unitary gauge action (3.15) is

$$S_1^{(2)} = \int d^4x \sqrt{-g} \left\{ \frac{1}{2} M_2^4(t) (\delta g^{00})^2 \right\} . \tag{3.27}$$

Putting this into covariant form, one obtains the action

$$S_1^{(2)} = \int d^4x \sqrt{-g} \left\{ \frac{M_2^4(\phi)}{2} + \frac{M_2^4(\phi)}{M_*^4} X + \frac{M_2^4(\phi)}{2M_*^8} X^2 \right\} . \tag{3.28}$$

Eq. (3.28) is the contribution that a k-essence model [132] makes to Eq. (3.26) at second order in X. The covariant or unitary gauge combinations $S_{\Omega=1}^{(0,1)} + S_1^{(2)}$

describe the same cosmological background and linear theory of any function $G_2(\phi, X)$ in Eq. (3.2) with $G_3 = G_5 = 0$ and $G_4 = M_*^2/2$.

3.3.2 Cubic term $\mathcal{L}_3$

Next, we consider a non-vanishing $\bar{M}_1^3$ coefficient, which is the first term to give rise to a contribution to the cubic Lagrangian $\mathcal{L}_3$. It appears in the EFT action as

$$S_2^{(2)} = \int d^4x\sqrt{-g}\left\{-\frac{1}{2}\bar{M}_1^3(t)\delta g^{00}\delta K\right\}. \tag{3.29}$$

We now reconstruct a covariant action that reduces to Eq. (3.29) to second-order perturbations in the unitary gauge. For this purpose, it is sufficient to consider the special case of $G_3(\phi, X) = \ell_3(\phi)X$ where ℓ_3 is a smooth function of ϕ only. One could do an alternative derivation by making $G_3(\phi, X)$ a function of an arbitrary power of X. Although the reconstructed covariant action would be different, the linear theory would be the same. After a few integrations by parts, in the unitary gauge adopting Eq. (3.9) this term becomes [7, 100, 101]

$$\begin{aligned} M_*^{-6}\ell_3(\phi)X\Box\phi = &\left[\dot{\ell}_3(t) - 3\ell_3(t)H\right]g^{00} - \ell_3(t)\delta g^{00}\delta K \\ &- 3\ell_3(t)H + \frac{3H}{4}\ell_3(t)(\delta g^{00})^2 \\ &- \frac{1}{4}\dot{\ell}(t)(\delta g^{00})^2\,. \end{aligned} \tag{3.30}$$

We take all the terms apart from that involving $\delta g^{00}\delta K$ to the left-hand side of the equation and use Eq. (3.10) to write δg^{00} in covariant form. Comparing Eqs. (3.30) and (3.29), we also make the identification

$$\ell_3(t) \equiv \frac{1}{2}\bar{M}_1^3(t)M_*^{-6}\,. \tag{3.31}$$

Hence, the covariant action that follows is given by

$$\begin{aligned} S_2^{(2)} = \int d^4x\sqrt{-g}&\left\{\frac{9H\bar{M}_1^3}{8} + \frac{M_*^2(\bar{M}_1^3)'}{8} + \frac{\bar{M}_1^3}{2M_*^6}X\Box\phi\right. \\ &\left. + \left[\frac{3H\bar{M}_1^3}{4M_*^4} - \frac{(\bar{M}_1^3)'}{4M_*^2}\right]X + \left[\frac{(\bar{M}_1^3)'}{8M_*^6} - \frac{3H\bar{M}_1^3}{8M_*^8}\right]X^2\right\}, \end{aligned} \tag{3.32}$$

which reduces to Eq. (3.29) at second order in the unitary gauge. Note that after making the replacement (3.10), there are also extra factors of M_* appearing from the replacement of the time derivative with a derivative with respect to the scalar field via $\dot{\bar{M}}_1^3 = M_*^2(\bar{M}_1^3)'$.

3.3.3 Quartic term $\mathcal{L}_4$

Finally, we reconstruct the quartic Lagrangian density $\mathcal{L}_4$. The first contribution arises from the background term $\Omega(t)$,

$$S^{(0,1)}_{\Lambda=\Gamma=0} = \frac{M_*^2}{2}\int d^4x\sqrt{-g}\left\{\Omega(t)R\right\} , \tag{3.33}$$

which, after using equation (3.9), yields the quartic contribution $G_4 = M_*^2\Omega(\phi)/2$.

We now proceed to the reconstruction of a covariant action that reduces to the second order unitary gauge action (3.15) with all the EFT coefficients set to zero apart from $\bar{M}_2^2$,

$$\begin{aligned} S_3^{(2)} = \int d^4x\sqrt{-g}\Big\{&-\bar{M}_2^2(t)\left(\delta K^2 - \delta K^{\mu\nu}\delta K_{\mu\nu}\right) \\ &+\frac{1}{2}\bar{M}_2^2(t)\delta R^{(3)}\delta g^{00}\Big\} . \end{aligned} \tag{3.34}$$

For this purpose, consider the quartic Horndeski Lagrangian (3.4) and transform it into the unitary gauge. This results in [100]

$$\begin{aligned} \mathcal{L}_4 &= G_4 R^{(3)} + (2g^{00}M_*^4 G_{4X} - G_4)(K^2 - K_{\mu\nu}K^{\mu\nu}) \\ &\quad -2M_*^2\sqrt{-g^{00}}G_{4\phi}K. \end{aligned} \tag{3.35}$$

In order to carry out the reconstruction it is necessary to identify G_4 in terms of the EFT parameters. To do this one has to compare the coefficient of R in the covariant Lagrangian with that of $R^{(3)}$ in the unitary gauge Lagrangian. To compare each term consistently, we will make use of the Gauss-Codazzi relation

$$R^{(3)} = R - K_{\mu\nu}K^{\mu\nu} + K^2 - 2\nabla_\nu(n^\nu\nabla_\mu n^\mu - n^\mu\nabla_\mu n^\nu) , \tag{3.36}$$

which relates R to $R^{(3)}$. Hence, the contribution to the quartic term is

$$G_4(\phi, X) = \frac{\bar{M}_2^2(\phi)}{2}\left(1 + \frac{X}{M_*^4}\right) , \tag{3.37}$$

and the covariant Horndeski Lagrangian therefore is

$$\begin{aligned}\mathcal{L}_4 =& \frac{\bar{M}_2^2(\phi)}{2}\left(1+\frac{X}{M_*^4}\right)R \\ &- \frac{\bar{M}_2^2(\phi)}{M_*^4}\left[(\Box\phi)^2 - \nabla_\mu\nabla_\nu\phi\nabla^\mu\nabla^\nu\phi\right] . \end{aligned} \tag{3.38}$$

Note that we have used that $\delta R^{(3)} = R^{(3)}$ in flat space. Putting this into the unitary gauge gives

$$\begin{aligned}\mathcal{L}_4 =& -\bar{M}_2^2\left(\delta K^2 - \delta K^{\mu\nu}\delta K_{\mu\nu}\right) + \frac{1}{2}\bar{M}_2^2\delta R^{(3)}\delta g^{00} \\ &+ 6\bar{M}_2^2H^2 - 4\bar{M}_2^2HK - 3H^2\bar{M}_2^2\delta g^{00} \\ &+ 2\bar{M}_2^2HK\delta g^{00} - \dot{\bar{M}}_2^2\delta g^{00}K + \frac{1}{2}\dot{\bar{M}}_2^2K(\delta g^{00})^2 . \end{aligned} \tag{3.39}$$

To obtain a covariant action that yields the second-order unitary gauge action (3.35), we take the last two lines of Eq. (3.39) and move it to the left-hand side. Care must be taken in the transformation of the term $-4\bar{M}_2^2HK$. In making it covariant one first has to do an integration by parts to take the derivative in $K = \nabla_\mu n^\mu$ onto the other coefficients. Using then the definition of n^μ in Eq. (3.12) one obtains an expansion in powers of δg^{00} that up to second order goes as

$$-4\bar{M}_2^2HK = \frac{d}{dt}(\bar{M}_2^2H)\left\{4 - 2\delta g^{00} - \frac{1}{2}(\delta g^{00})^2\right\} . \tag{3.40}$$

One can then make the usual replacement for δg^{00} in Eq. (3.10) and use the result from Sec. 3.3.2 to transform all the terms involving a $\delta g^{00}\delta K$. This yields the covariant action

$$\begin{aligned}S_3^{(2)} = \int d^4x\sqrt{-g}\Bigg\{& \left[\frac{1}{2}\bar{M}_2^2 + \frac{\bar{M}_2^2}{M_*^4}X\right]R - \frac{\bar{M}_2^2}{M_*^4}\left[(\Box\phi)^2 - \nabla^\mu\nabla^\nu\phi\nabla_\mu\nabla_\nu\phi\right] \\ &- \frac{M_*^4(\bar{M}_2^2)''}{4} - \frac{7M_*^2(\bar{M}_2^2)'H}{4} - M_*^2H'\bar{M}_2^2 - \frac{9H^2\bar{M}_2^2}{2} + \left[\frac{(\bar{M}_2^2)''}{2} + \frac{H(\bar{M}_2^2)'}{2M_*^2} + \frac{2H'\bar{M}_2^2}{M_*^2}\right]X \\ &- \left[\frac{(\bar{M}_2^2)''}{4M_*^4} - \frac{H(\bar{M}_2^2)'}{4M_*^6} - \frac{H'\bar{M}_2^2}{M_*^6} + \frac{3H^2\bar{M}_2^2}{2M_*^8}\right]X^2 + \left[\frac{2H\bar{M}_2^2}{M_*^6} - \frac{(\bar{M}_2^2)'}{M_*^4}\right]X\Box\phi\Bigg\} . \end{aligned} \tag{3.41}$$

After putting action (3.41) back into the unitary gauge, at second order in the perturbations one obtains action (3.34). Note that a different reconstruction of $\delta g^{00}\delta R^{(3)}$ that is not contained within the Horndeski action was recently presented in Ref. [133].

Combining the actions $S^{(0,1)}_{\Lambda=\Gamma=0}$, $S^{(0,1)}_{\Omega=1}$, $S^{(2)}_1$, $S^{(2)}_2$, and $S^{(2)}_3$ in Eqs. (3.33), (3.26), (3.28), (3.32), and (3.41), respectively, we obtain the expressions given for G_i in Eqs. (3.18) through (3.20), which thus are constructed to produce the same cosmological background and linear perturbations as the EFT action (3.13). Note that, as discussed in Sec. 3.2, the quintic term G_5 does not introduce additional EFT functions in $S^{(0-2)}$ and thus its phenomenology at the background and linear perturbation level can be captured by G_{2-4}. For simplicity, we have therefore adopted a baseline reconstruction with $G_5 = 0$ but allow for variations around this solution in Eq. (3.21).

3.4 Simple Examples

For illustration, we provide here a brief discussion of the application of the reconstruction for three simple examples. In Sec. 3.4.1, we show how a quintessence model can be reconstructed and discuss some subtleties about the canonical form of the scalar field action. We then apply the reconstruction to $f(R)$ gravity, cubic galileon gravity and a quartic model in Secs. 3.4.2, 3.4.3 and 3.4.4 respectively.

3.4.1 Quintessence

Let us assume a measurement of $\Omega(t) = 1$, non-vanishing $\Lambda(t)$ and $\Gamma(t)$, and vanishing values for the other EFT functions. Applying this to the reconstructed action defined by Eqs. (3.18) through (3.21), one finds the action (3.26). Note that the kinetic contribution is not in its canonical form. To find the canonical form of the action, we perform the field redefinition

$$\frac{\partial \chi}{\partial \phi} = \frac{1}{M_*}\sqrt{\Gamma(\phi)} \tag{3.42}$$

such that in terms of the new scalar field χ, we obtain

$$S = \int d^4x\sqrt{-g}\left\{\frac{1}{2}M_*^2 R - V(\chi) - \frac{1}{2}(\partial\chi)^2\right\}\,, \tag{3.43}$$

where

$$V(\chi) = M_*^2 \left(\Lambda(\chi) + \frac{1}{2}\Gamma(\chi) \right) . \quad (3.44)$$

Including a non-minimal coupling term $\Omega(t)$ in front of R further allows one to reconstruct a Brans-Dicke action in a similar way by choosing a suitable $\Gamma(t)$ associated with the Brans-Dicke function $\omega(\phi)$.

3.4.2 $f(R)$ gravity

Next, we assume a measurement of varying $\Omega(t)$ and $\Lambda(t)$ while all other EFT functions vanish. This is the scenario that would be expected for a $f(R)$ model. $f(R)$ gravity can be written as a Brans-Dicke type scalar-tensor theory with a scalar field potential and $\omega = 0$ (hence, vanishing Γ). The scalar field in this case can be associated with $f_R \equiv df(R)/dR$, where the potential has a particular dependence on f_R, specified by $f(R)$ and R.

While we can therefore follow the same procedure as in Sec. 3.4.1 for the reconstruction, we also consider here a slightly different approach (cf. [7]). In this case, instead of identifying the time coordinate with the scalar field, one identifies it with the Ricci scalar, adopting a gauge where its perturbations vanish, $\delta R = 0$. Hence, in this case, we directly find

$$\begin{aligned} \mathcal{L}_{f(R)} &= \Omega(R)R - 2\Lambda(R) = R + [\Omega(R)R - R - 2\Lambda(R)] \\ &\equiv R + f(R) . \end{aligned} \quad (3.45)$$

3.4.3 Cubic Galileon

Let us assume a measurement of

$$M_*^2 \Gamma = 4M_2^4 = 3H\bar{M}_1^3 = -\lambda H , \quad (3.46)$$

and $\Omega(t) = \exp(-2M_* t)$ with a positive constant λ and all other EFT functions vanishing. Applying this to the reconstructed action, defined by Eqs. (3.18) through (3.20), and setting $\lambda = 6M_*^5 r_c^2$, defining a crossover distance r_c, we obtain

$$\mathcal{L} = \frac{M_*^2}{2} e^{-2\phi/M_*} R - \frac{r_c^2}{M_*} X \Box \phi + \mathcal{L}_M , \quad (3.47)$$

which is the Lagrangian density of a cubic galileon model [7, 134].

3.4.4 Quartic Lagrangian

To give a simple example of a reconstruction of a model involving G_4, let us assume that the relation $\bar{M}_2^2 = \lambda$ holds for some constant λ. In addition, assume that the other EFT functions are related to H in the following way

$$\begin{aligned} \bar{M}_1^3 = -4\lambda H \, , \, M_2^4 = -\lambda \dot{H} \, , \\ \Gamma + \Lambda = -12 H^2 \, , \, \Gamma - \Lambda = 8\dot{H} \, . \end{aligned} \tag{3.48}$$

Using these relations in the reconstructed action in Eqs. (3.18) to (3.20) it is found that, upon identifying $\lambda = M_*^2$, one recovers the following quartic Horndeski Lagrangian

$$\mathcal{L} = \left(\frac{M_*^2}{2} + \frac{1}{2M_*^2} X \right) R - \frac{1}{M_*^2} \left[(\Box\phi)^2 - \nabla_\mu \nabla_\nu \phi \nabla^\mu \nabla^\nu \phi \right] . \tag{3.49}$$

3.5 Conclusions

This chapter has presented a reconstruction method which maps from parameters in the effective field theory of dark energy back to a fully covariant Horndeski action. The EFT of dark energy and modified gravity does enable a generalised and efficient examination of a large class of theories and there has been much work examining how to express a variety of given covariant theories in terms of the EFT functions. To connect observables with theoretical Horndeski models the inverse mapping is also required. Starting from a given EFT unitary gauge action, for instance provided by measurement, one can derive a covariant Horndeski Lagrangian that shares the same dynamics of the cosmological background and linear fluctuations around it. As the reconstruction cannot be unique, we have focused on the recovery of a baseline covariant Horndeski action that reproduces the desired equivalent background and linear dynamics. We have furthermore characterised the variations of this action that can be performed to move between the covariant theories degenerate at the background and linear level. For illustration, we have applied the reconstruction method to a few simple example models embedded in the Horndeski action: quintessence, $f(R)$ gravity, a cubic galileon model and a quartic model.

The following chapter will present a number of applications of the reconstruction. Of particular interest is the construction of a covariant realisation of the linear shielding mechanism shown to be present in Horndeski theories by analysis of its unitary gauge action [113] (also see Ref. [135]). This mechanism operates in a large class of theories that can become degenerate with ΛCDM in the expansion history and linear perturbations. However, the degeneracy can be broken by the measurement of the speed of cosmological propagation of gravitational waves [107]. This is now known to be equal to the speed of light [136] but as we shall discuss this should not completely eliminate interest in these models. With the reconstruction, we shall also examine the question of how well motivated the frequently adopted parametrizations of the EFT functions in observational studies are [120–122, 137]. Furthermore, the reconstruction will enable one to directly employ measurements of the EFT functions to impose constraints on the covariant Horndeski terms, which will be of particular interest to future surveys such as Euclid [27, 28] or the Large Synoptic Survey Telescope (LSST) [29]. Finally, the covariant reconstruction disentangles the cosmological dependence of the Horndeski modifications in the EFT functions that is due to the spacetime foliation adopted in the unitary gauge. Hence, a reconstructed action from phenomenological EFT functions can be applied to non-perturbative regimes (see e.g. Ref. [138]) or non-cosmological backgrounds and used to connect further observational constraints, for instance, arising from the requirement of screening effects in high-density regimes. This shall be explored in chapter 5. This list of applications of the reconstructed Horndeski action is far from exhaustive.

Chapter 4

Applications of the reconstruction

You can know the name of a bird in all the languages of the world, but when you're finished, you'll know absolutely nothing whatever about the bird. So let's look at the bird and see what it's doing. That's what counts.

Richard P. Feynman

4.1 Introduction

This chapter presents a number of applications of the reconstruction from the EFT of dark energy back to manifestly covariant Horndeski theories introduced in the previous chapter. In particular, it will be concerned with the connection of observational parameters with underlying theories, providing a vital link between the parameter constraints from future cosmological data sets and their theoretical interpretation. Eventually it is hoped that through such techniques it will be possible to significantly rule out or favour Horndeski theory as a whole with various cosmological observations. Recently, the LIGO/Virgo measurement of the gravitational wave GW170817 [109] emitted by a binary neutron star merger with the simultaneous observations of electromagnetic counterparts [136, 139] has already led to a significant reduction of the available theory space within

Horndeski at late times, as was first anticipated in Refs. [107, 108]. The GW170817 event occurred in the NGC 4993 galaxy of the Hydra cluster at a distance of about 40 Mpc and enabled a constraint on the relative deviation of the speed of gravity c_T from the speed of light ($c = 1$) at $\mathcal{O}(10^{-15})$ for $z \lesssim 0.01$ [136]. This agrees with forecasts [107, 140] inferred from the increased likelihood with increasing volume at the largest distances resolved by the detectors, expecting a few candidate events per year, and emission time uncertainties. It was anticipated that the measurement would imply that a genuine cosmic self-acceleration from Horndeski scalar-tensor theory and its degenerate higher-order extensions, including the Galileon theories, can no longer arise from an evolving speed of gravity and must instead be attributed to a running effective Planck mass [107]. The minimal evolution of the Planck mass required for self-acceleration with $c_T = 1$ was derived in Ref. [108] and was shown to provide a 3σ worse fit to cosmological data than a cosmological constant. Strictly speaking, this only applies to Horndeski theories, where $c_T = 1$ breaks a fundamental degeneracy in the large-scale structure produced by the theory space [107, 113]. Generalizations of the Horndeski action reintroduce this degeneracy [113] but self-acceleration in general scalar-tensor theories is expected to be conclusively testable at the 5σ level with Standard Sirens [107] (also see Refs. [141–144]), eventually allowing an extension of this No-Go result. The minimal model serves as a null-test for self-acceleration from modified gravity. It is therefore worth examining whether future observational probes of the large-scale structure are capable of tightening the constraint beyond the 3σ-level. Finally, the measurement of $c_T \simeq 1$ with GW170817 in particular implies that the quintic and kinetically coupled quartic Horndeski Lagrangians must be negligible at late times [145, 146] (also see e.g., Refs. [107, 117, 147–155] for more recent discussions). The measurement also led to a range of further astrophysical and cosmological implications (see, e.g., Ref. [156]).

Despite giving strong restrictions on the set of scalar-tensor theories that could explain the accelerated expansion, there remains a great deal of freedom in the model space after the GW170817 observation and the phenomenological study of the quintic and kinetically coupled quartic Horndeski Lagrangians should not be dismissed so soon. There are two important aspects to be considered in this argument. On the one hand, the speeds of gravity and light are only constrained to be effectively equal at the low redshifts of $z \lesssim 0.01$. This certainly applies to the regime of cosmic acceleration but not to the early Universe, where a decaying deviation in c_T could still lead to observable signatures without invoking fine-tuning. Moreover, for more general scalar-tensor theories, the linear shielding mechanism [113] may be extended to a modified gravitational wave propagation, where the Horndeski terms could cause cosmic self-acceleration while other terms may come to dominate for the wavelengths relevant to GW170817. These wavelengths differ by those associated with cosmic acceleration by $\mathcal{O}(10^{19})$ [157]. Hence, in this chapter we will not exclusively restrict to the models satisfying the GW170817 constraint, envisaging more general applications of the methods presented.

As we have discussed in Sec. 2.3.4 the EFT of dark energy provides a useful approach to studying the cosmological effects of a variety of Horndeski models in a unified way. This generality comes at a cost. Currently, one either has to start from a given fully covariant theory and compute the EFT coefficients in terms of the functions defined in the covariant action, or take a phenomenologically motivated parameterization for the EFT functions. In the first instance, one is essentially left with the original problem of having a large range of theories to compare with observations. Following the second approach gives a general indication of the effects of modified gravity on different observational probes, but it is generally unclear what physical theories are being tested when a particular parameterization is adopted.

In the previous chapter we developed a mapping from the EFT coefficients to the family of Horndeski models which give rise to the same background evolution and linear perturbations. This mapping provides a method to study the form of the Horndeski functions determined from observations on large scales. One can furthermore address the question of what theories various phenomenological parameterizations of the EFT functions correspond to. We shall examine the form of the underlying theories corresponding to two commonly used EFT parameterizations for late-time modifications motivated by cosmic acceleration. Reconstructed actions that exhibit minimal self-acceleration and linear shielding are also presented. We furthermore apply the reconstruction to phenomenological parameterizations such as a modified Poisson equation and gravitational slip [158–162] as well as the growth-index parametrization [163–165]. These are the primary parameters that the next generation of galaxy-redshift surveys will target [27–29]. With the reconstruction it is possible to connect these parameterizations with viable covariant theories, and explore the region of the theory space being sampled when a particular parameterization is adopted. The reconstruction is also applied to a phenomenological model that exhibits a weakening of the growth of structure relative to ΛCDM today, which may be of interest to address potential observational tensions [166, 167]. Finally, in every analysis of the EFT model space it is necessary to ensure that the chosen model parameters do not lead to ghost or gradient instabilities. When comparing models with observations this can, for instance, lead to a highly inefficient sampling of the model space and misleading statistical constraints due to complicated stability priors. To avoid those issues, we propose an alternative parameterization of the EFT function space, which uses the stability parameters directly as the basis set such that every sample drawn from that space is inherently stable.

The chapter is organised as follows. In Sec. 4.2 we briefly review the EFT formalism, in particular reviewing a commonly used alternative basis for the EFT functions, before specifying the stability criteria imposed on the model space. We then propose a new inherently stable EFT basis that we argue is most suitable for statistical comparisons of the available theory space to observations. Sec. 4.3 covers a number of different reconstructions, ranging from commonly adopted parameterizations encountered in the literature (Sec. 4.3.1) to models for minimal self-acceleration (Sec. 4.3.2), linear shielding (Sec. 4.3.3), phenomenological modifications of the Poisson equation and gravitational slip (Sec. 4.3.4), the growth-index parametrization (Sec. 4.3.5), and weak gravity (Sec. 4.3.6). In Sec. 4.3.7 we provide an example of a reconstruction from the inherently stable parameter space. before inspecting the impact of the choice of EFT parametrization on the reconstructed theories in Sec. 4.4. Finally, we discuss conclusions in Sec. 4.5.

4.2 Alternate bases for EFT of dark energy parameters

At second order, Horndeski gravity corresponds to the EFT action in Eqs.(3.14) and (3.15). In general, various subsets of Horndeski theory lead to separate contributions from the EFT coefficients. In particular, theories compatible with the GW1710817 observation must satisfy $\bar{M}_2^2(t) \simeq 0$ at $z \lesssim 0.01$. Taking into account the Hubble expansion $H(t) \equiv \dot{a}/a$ and the two constraints from the Friedmann equations, Eqs. (3.13)–(3.15) contain five independent functions capable of describing the background and linear perturbations of Horndeski theory.

An alternative basis for these EFT functions with a more direct physical interpretation [8]. See Table I of Ref. [9] and Table 3.2 for the connection between the two descriptions, although bear in mind the different conventions. This basis is defined via

$$\alpha_M \equiv \frac{M_*^2\Omega' + 2(\bar{M}_2^2)'}{M_*^2\Omega + 2\bar{M}_2^2}\,, \tag{4.1}$$

$$\alpha_B \equiv \frac{M_*^2 H\Omega' + \bar{M}_1^3}{2H\left(M_*^2\Omega + 2\bar{M}_2^2\right)}\,, \tag{4.2}$$

$$\alpha_K \equiv \frac{M_*^2\Gamma + 4M_2^4}{H^2\left(M_*^2\Omega + 2\bar{M}_2^2\right)}\,, \tag{4.3}$$

$$\alpha_T \equiv -\frac{2\bar{M}_2^2}{M_*^2\Omega + 2\bar{M}_2^2}\,, \tag{4.4}$$

where throughout this section primes denote derivatives with respect to $\ln a$. Here α_M describes the running of the effective Planck mass $M = \sqrt{M_*^2\Omega + 2\bar{M}_2^2}$ defined through $\alpha_M = d\ln M^2/d\ln a$, allowing for some variation in the strength of the gravitational coupling over time. The function α_B describes a braiding or mixing between the kinetic contributions of the scalar and tensor fields. The function α_K enters through the kinetic term of the scalar field and only becomes relevant on scales comparable to the horizon. Finally, α_T describes the deviation of the speed of gravitational waves from the speed of light with $c_T^2 = 1 + \alpha_T$.

4.2.1 Stability Criteria

To ensure the absence of ghost and gradient instabilities it is necessary to impose certain constraints on the EFT functions. For instance, in order to avoid a kinetic term with the wrong sign or an imaginary sound speed for the scalar modes one must have [8]

$$\alpha \equiv \alpha_K + 6\alpha_B^2 > 0\,, \quad c_s^2 > 0\,, \tag{4.5}$$

where the soundspeed c_s is given by

$$c_s^2 = -\frac{2}{\alpha}\Big[\alpha_B' + (1+\alpha_T)(1+\alpha_B)^2 - \left(1 + \alpha_M - \frac{H'}{H}\right)(1+\alpha_B) + \frac{\rho_m}{2H^2M^2}\Big]\,. \tag{4.6}$$

Furthermore, the stability of the background to tensor modes requires

$$c_T^2 > 0\,, \quad M^2 > 0\,. \tag{4.7}$$

One must be careful when using parametrizations of the EFT functions to reconstruct covariant theories that the stability conditions are satisfied. A way to achieve this that we adopt in Secs. 4.3.1–4.3.5 is to set the soundspeed equal to unity and use this as a constraint on the EFT coefficients. It then remains to check that the other conditions are also satisfied by hand. This is somewhat restrictive as there are many viable stable scalar-tensor theories that do not have $c_s^2 = 1$. An alternative approach is to directly parameterize the stability conditions as a new set of EFT functions (Secs. 4.2.2 and 4.3.7).

4.2.2 A New Inherently Stable Parameterization

For generic tests of modified gravity and dark energy, a range of different time parametrizations (see Sec. 4.3.1) are commonly adopted for the EFT coefficients in $S^{(0,1)}$ and $S^{(2)}$ or for the α_i functions. These parameterizations do not a priori satisfy the stability criteria in Eqs. (4.5) and (4.7). As a consequence the sampling in this parametrization, for example when conducting a Markov Chain Monte Carlo (MCMC) analysis to constrain the EFT parameter space with observations, can be highly inefficient. Only a very small fraction of the samples will hit a stable region of parameter space. Moreover, the stability criteria can yield contours on the parameter space that are statistically difficult to interpret. For instance, ΛCDM can be confined to a narrow corner of two intersecting edges produced by the stability requirements. This corner may only be sparsely sampled and could lead to spurious evidence against concordance cosmology.

To avoid those issues, we propose here a new basis for the parametrization of modified gravity and dark energy models in the effective field theory formalism. We will make use of the GW170817 constraint $\alpha_T \simeq 0$ at $z \lesssim 0.01$ and assume that it applies throughout the late-time domain of interest here. We define a function B through

$$1 + \alpha_B \equiv \frac{B'}{B}\,. \tag{4.8}$$

Eq. (4.6) can then be expressed as a linear homogeneous second-order differential equation for B with

$$B'' - \left(1 + \alpha_M - \frac{H'}{H}\right) B' + \left(\frac{\rho_m}{2H^2M^2} + \frac{\alpha\, c_s^2}{2}\right) B = 0\,. \qquad (4.9)$$

By the existence and uniqueness theorem for ordinary differential equations a real solution exists for real boundary conditions on B and B'. Alternatively, we may provide an initial or present value α_{Bi} or α_{B0}, respectively.

Hence an inherently stable parametrization of the EFT function space for modified gravity and dark energy can be defined by parametrizations of the basis

$$M^2 > 0\,, \quad c_s^2 > 0\,, \quad \alpha > 0\,, \quad \alpha_{B0} = \text{const.}\,, \qquad (4.10)$$

along with the Hubble parameter H. The braiding function α_B can be determined from the integration of Eq. (4.9) and α_K from α_B and α.

We advocate that this basis should be used for observational constraints on the EFT function space to avoid the problems described earlier. It also provides a direct physical interpretation of the observational constraints. While parametrizations in H classify quintessence dark energy models where $\alpha > 0$, c_s^2 describes more exotic dark energy models with $\alpha_{B0} \neq 0$ adding an imperfection to the fluid and $M \neq M_*$ modifying gravity. In ΛCDM, $M = M_*$, c_s drops out and the remaining parameters vanish. This parameterization furthermore addresses the measure problem on the parameter space. While it is difficult to know a priori what is a reasonable prior range to place on the α_i parameters, it is much clearer in this physical parameterization. In addition, if measurements of these physical parameters seem to approach a fixed value it becomes easier to place bounds on the desired accuracy. We shall apply the reconstruction to a model defined in this basis in Sec. 4.3.7. Note that one can easily add the beyond-Horndeski parameter α_H to this basis, which will introduce a modification in c_s^2. Various examples of how to map this parameterization to different dark energy and modified gravity models can be found in Ref. [168].

4.3 Reconstructing covariant theories

We now present a series of applications of the mapping relations derived in the previous chapter. We begin with a reconstruction of common parameterizations of the EFT functions used in the literature (Sec. 4.3.1) and then examine the

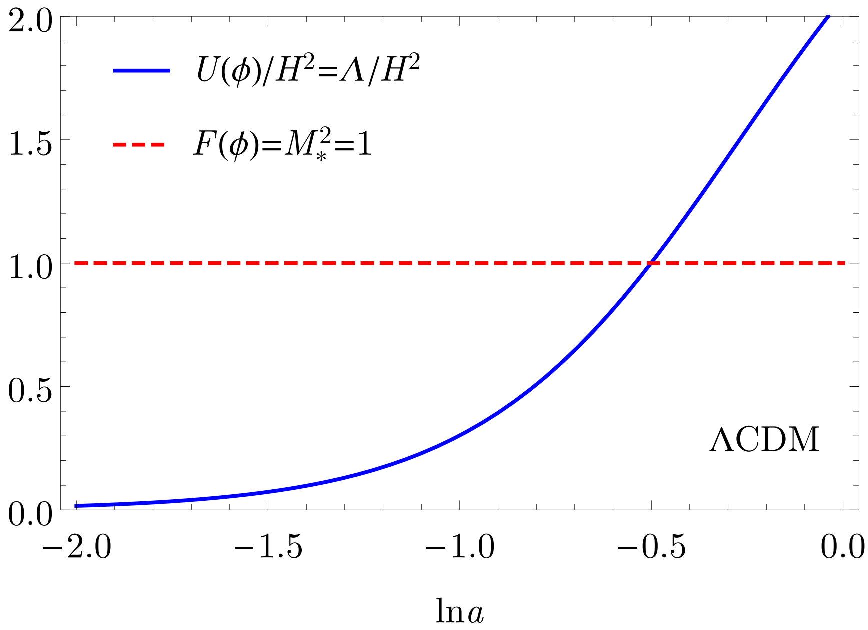

Figure 4.1 *Reconstructed contributions to the Horndeski action for ΛCDM, normalized with powers of H received in the reconstruction (See Table. 3.1). The curves serve as reference for the comparison of the reconstructed modifications in Secs. 4.3.1–4.3.7. Due to the normalization with H^2, the cosmological constant appears to decay at high redshift.*

form of the underlying theory of the minimal self-acceleration model (Sec. 4.3.2) and theories that exhibit linear shielding (Sec. 4.3.3). Following this, we discuss reconstructions from more phenomenological modifications of gravity with a modified Poisson equation and a gravitational slip (Sec. 4.3.4) as well as the growth-index parametrization (Sec. 4.3.5). We then present a reconstruction of a model which has a weakened growth of structure relative to ΛCDM (Sec. 4.3.6) before concluding with an example of a reconstruction from the inherently stable parameterization introduced in Sec. 4.2.2 (Sec. 4.3.7).

Recall the reconstructed Horndeski action is defined such that when expanded up to second order in unitary gauge one recovers Eq. (3.13) with the Horndeski functions given in Eqs. (3.18) through (3.21). It is worth noting that taking $c_T \simeq 1$ as a linear constraint sets $c_1 = 0$ in Eq. (3.20) but does not directly make a statement about $\Delta G_{4X/5}$. However, excluding the highly unlikely cancellation of c_1 and $\Delta G_{4X/5}$, and assuming approximately linear theory from the outskirts of the Milky Way with $c_1 = 0$, the nonlinear contributions $\Delta G_{4/5}$ are still constrained by $|c_T - 1| \lesssim 10^{-13}$.

In illustrations of the reconstructed Horndeski functions G_i, each contributing term is divided by the powers of H it receives multiplying the EFT functions in the reconstruction (see Table 3.1). This ensures a meaningful comparison of the effective modifications from ΛCDM rather than providing illustrations for deviations that are suppressed and do not propagate to the cosmological background evolution and linear perturbations. For instance, we have $U(\phi)/H^2 \sim b_1(\phi)/H \sim \bar{M}_2^2$. As a reference, we show in Fig. 4.1 the Horndeski functions G_i that correspond to ΛCDM, where $G_4 = 1$, $G_2 = -2\Lambda$ and $G_3 = G_5 = 0$, i.e., in particular the term Λ/H^2. The Planck 2015 value $\Omega_m = 0.308$ [169] for the matter density parameter is adopted throughout the thesis. We also work in units where the bare Planck mass $M_* = 1$, such that the vertical axis on each plot indicates the deviation from this value. Because the choice of how the scalar field is defined is arbitrary, we present the reconstructed terms as functions of $\ln a$ rather than ϕ, except for the examples given in Secs. 4.3.1 and 4.3.2. The colour scheme is set such that the terms in blue correspond to terms that can be identified in the matter sector, whereas the red terms couple to the metric and so in that sense are a "modification" of gravity. These modified gravity terms are $F(\phi)$ and $c_1(\phi)$, the latter being non-zero when the $\alpha_T = 0$ constraint is dropped.

It is worth noting that one always has the freedom to redefine the scalar field ϕ in the action. We shall briefly discuss how one can recast the reconstructed coefficients of the covariant theory from functions of $\ln a$ to a more standard description. For this purpose, we choose the Brans-Dicke representation, where $F(\phi) \equiv \psi$, and then re-express all of the terms in the reconstructed action as a function of the new scalar field ψ. This choice implies $\phi = F^{-1}(\psi)$ and

$$\partial_\mu \phi = f(\psi)\partial_\mu \psi \,, \tag{4.11}$$

where for simplicity we have defined the function $f(\psi) \equiv d(F^{-1})/d\psi$. After this field re-definition the reconstructed action written in terms of ϕ is transformed into a scalar-tensor action for ψ with $(\partial\phi)^2 = f^2(\psi)(\partial\psi)^2$ and $\Box\phi = f(\psi)\Box\psi + df/d\psi\,(\partial\psi)^2$. The new representation of the theory then involves the terms

$$\tilde{U}(\psi) = U(\psi)\,, \tag{4.12}$$

$$\tilde{Z}(\psi) = f^2(\psi)Z(\psi)\,, \tag{4.13}$$

$$\tilde{b}_1(\psi) = f^3(\psi)b_1(\psi)\,, \tag{4.14}$$

$$\tilde{a}_2(\psi) = a_2(\psi)f^4(\psi) + b_1(\psi)f^2(\psi)\frac{df}{d\psi}\,. \tag{4.15}$$

Depending on the functional form of $f(\psi)$ higher-derivative terms may be enhanced or suppressed in this description. For consistency, in this representation we also transform the Hubble parameter to be a function of ψ such that $H \to \tilde{H}$. We will show examples of this transformation in Secs. 4.3.1 and 4.3.2.

4.3.1 Reconstruction of common EFT parameterizations

A common choice of phenomenologically motivated functional forms of the EFT modifications is to parameterize them in such a way that they only become relevant at late times. Typically their evolution is tied to the scale factor $a(t)$ or to the dark energy density $\Omega_\Lambda(a) \equiv H_0^2\Omega_\Lambda/H^2$ raised to some power q. Note that now, with the GW170817 constraint, self-acceleration from modified gravity is strongly challenged as a direct explanation for the late-time accelerated expansion [108] and it can be questioned whether the functional form of such parameterizations continues to be well motivated. On the other hand, a dark energy model may still introduce a related modification of gravity, for instance, as a means to remedy the old cosmological constant problem of a non-gravitating vacuum. We set this issue aside for now and adopt the two parametrizations

$$\mathcal{A} \quad : \quad \alpha_i = \alpha_{i0} a(t)^{q_i} \,, \tag{4.16}$$

$$\mathcal{B} \quad : \quad \alpha_i = \alpha_{i0} \left(\frac{\Omega_\Lambda(a)}{\Omega_{\Lambda,0}} \right)^{q_i} \,. \tag{4.17}$$

Here the label i runs over the set of functions $\{i \in M, T, K, B\}$ in Eqs. (4.1)–(4.4). The two parametrizations can be used to study the effect of small deviations from ΛCDM in the linear late-time fluctuations resulting from a set of non-vanishing α_i.

In principle, there are many alternative parameterizations that could be used beyond these simple ones. For the purposes of this chapter we shall however restrict to these two examples which have been frequently used in the literature (see e.g. Ref. [137]). It was recently suggested that those are sufficiently general to encompass the linear effects of the different time dependencies in a variety of modified gravity theories [122] (however, also see Ref. [120]). The reconstruction from EFT back to manifestly covariant theories provides a method to examine how the underlying covariant theory changes with a different choice of parameterization. One can thus begin to address the question of what scalar-tensor theory is actually being constrained when a particular parameterization is adopted.

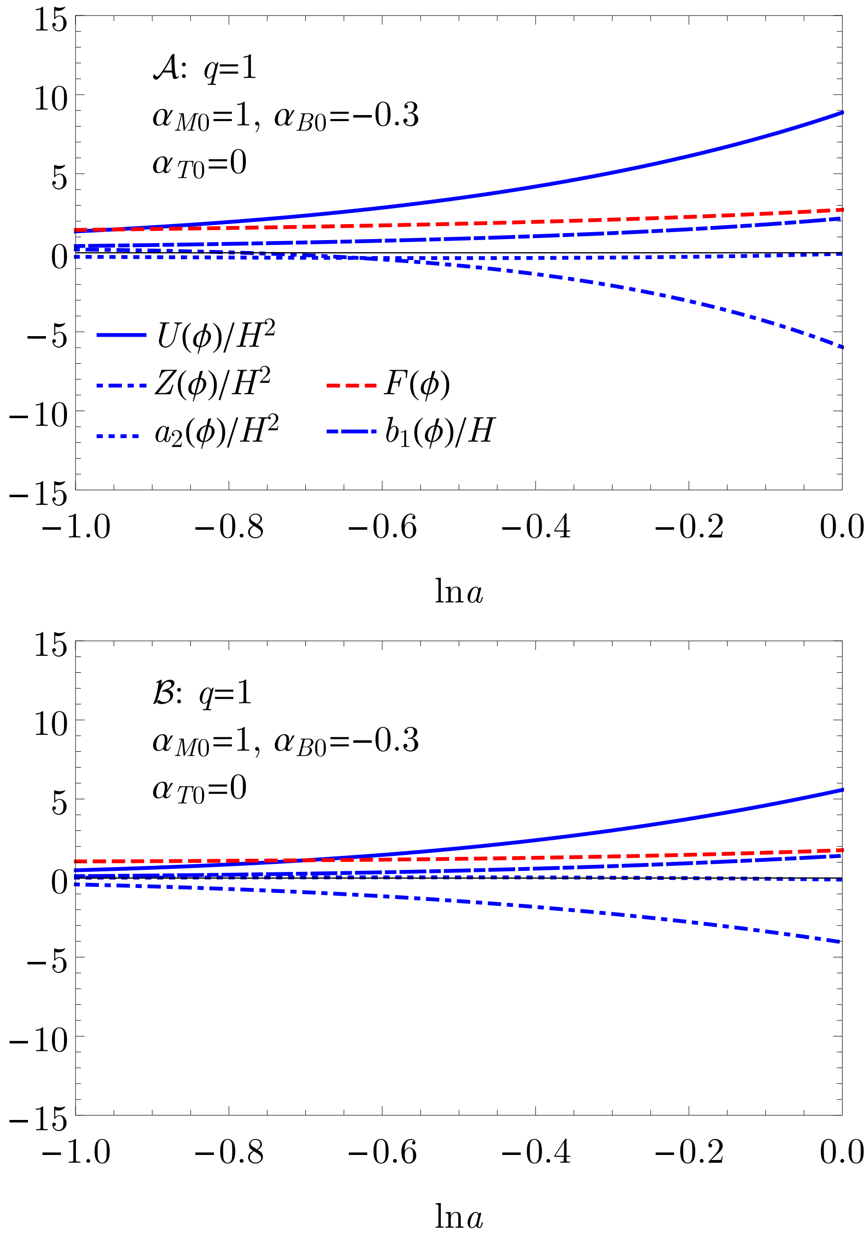

Figure 4.2 *Examples of reconstructed actions arising from two different parameterizations of the EFT functions $\mathcal{A}$ and $\mathcal{B}$ specified in Eqs. (4.16) and (4.17). We chose equal amplitudes for the comparison. The general evolution of the modifications is unaffected by the particular choice of time parametrization, although the magnitude of the various terms is enhanced when using parameterization $\mathcal{A}$. This can be attributed to the convergence to constant α_i at late times in $\mathcal{B}$. The reconstructed terms of the scalar-tensor action can be converted into functions of a scalar field ψ, for instance, by adopting a Brans-Dicke representation and casting the functions in terms of $F \to \psi$ (see Fig. 4.3). However, as the choice of scalar field is arbitrary, the reconstructions shall generally be illustrated as functions of* $\ln a$.

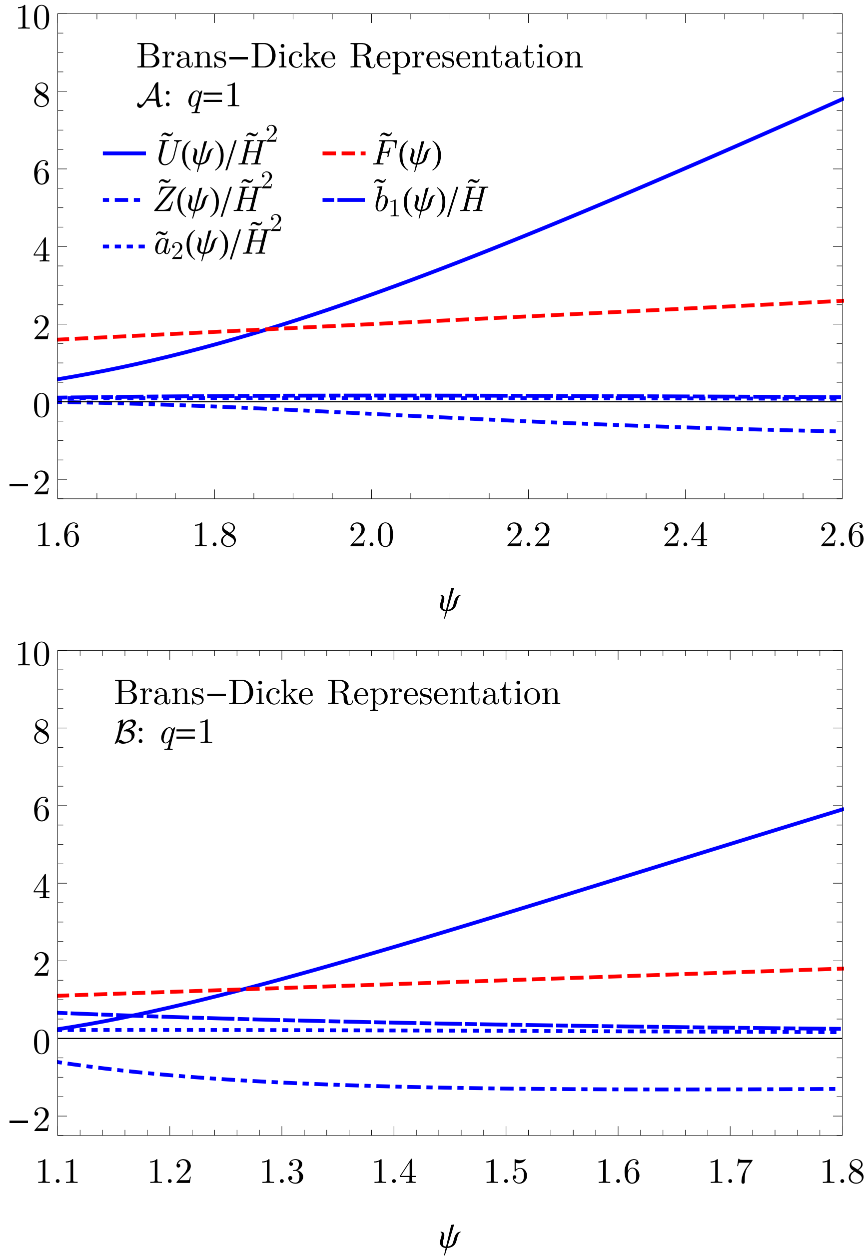

Figure 4.3 *Brans-Dicke representation, with $F(\phi) \equiv \psi$, of the reconstructed scalar-tensor theories illustrated in Fig. 4.2. We have transformed the Hubble rate $H \to \tilde{H}$ such that it is also a function of ψ and divided each term in the action by appropriate powers of $\tilde{H}$ (see Sec. 4.3).*

To provide concrete examples for the models that are reconstructed from Eqs. (4.16) and (4.17), we parametrize α_M, α_B and α_T with $\mathcal{A}$ or $\mathcal{B}$ and then set α_K such that $c_s^2 = 1$ (see discussion in Sec. 4.2.1). Note that strictly speaking this deviates from adopting Eqs. (4.16) and (4.17) for all α_i but it simplifies the stability treatment of the model. Furthermore, the deviation is only relevant on near-horizon scales. Numerical values for α_{i0} are then chosen to ensure that the stability condition $\alpha > 0$ in Eq. (4.5) is satisfied. For illustration, we set $\alpha_{M0} = 1$, $\alpha_{B0} = -0.3$ and $\alpha_{T0} = 0$ with $q_i = q = 1$. This yields a stable scalar-tensor theory for both parameterizations $\mathcal{A}$ and $\mathcal{B}$. The corresponding terms of the Horndeski functions are shown in Fig 4.2. The behavior of the reconstructed theories is tied to the functional form of the parameterization used, with the Horndeski modifications becoming more relevant at later times. We note that the general form of these modifications is independent of the particular parametrization adopted between $\mathcal{A}$ and $\mathcal{B}$. However, one can identify minor differences. For instance, the magnitude of the reconstructed modifications for $\mathcal{A}$ are larger. This is due to saturation of the modifications in $\mathcal{B}$ at late times. This particular choice for each α_{i0} leads to a model with an enhanced potential relative to ΛCDM and the standard kinetic term $Z(\phi)$ dominating the action at late times. There is a small contribution from the cubic term $b_1(\phi)$ but the k-essence term $a_2(\phi)$ is negligible. We present a number of examples which examine the sensitivity of the reconstruction to changes in α_{i0} and q_i in Sec. 4.4. For instance, by increasing the powers q_i in each of the parameterizations, the effects of modified gravity and dark energy are delayed to later times. Changing the amplitude of each α_{i0} on the contrary has a larger effect on the underlying theory. For example, when α_B dominates over α_M the cubic Galileon term $b_1(\phi)$ dominates over the potential $U(\phi)$ at late times, whereas when α_M dominates over α_B the potential and kinetic term $Z(\phi)$ are enhanced with smaller contributions from the k-essence and cubic Galileon terms. However, we find that the mapping is relatively robust, with small deviations in the α_i parameters around some fixed values not significantly altering the underlying theory. While we have checked this for a number of examples, further work is necessary to investigate this aspect more thoroughly.

Finally, in Fig. 4.3 we illustrate the corresponding Brans-Dicke representations of the reconstructed theories for $\mathcal{A}$ and $\mathcal{B}$ that are presented in Fig. 4.2. In this description the behavior of each term in the reconstruction is now dependent on the evolution of $F(\phi)$. It is clear that the functional form of each term in the theory remains broadly similar whether parameterization $\mathcal{A}$ or $\mathcal{B}$ is chosen.

4.3.2 Minimal self-acceleration

The LIGO/Virgo constraint of $|c_T - 1| \lesssim \mathcal{O}(10^{-15})$ and its implication that a genuinely self-accelerated Universe in scalar-tensor gravity must be attributed to a significant evolution in M^2 was first anticipated in Ref. [107]. This trivially excludes acceleration arising from an evolving speed of gravity c_T and the according class of gravitational models such as genuinely self-accelerated quartic and quintic Galileons and their Horndeski and higher-order generalizations with $\alpha_T \neq 0$, i.e., $G_{4X}, G_5 \neq 0$ for Horndeski gravity (see e.g. Ref. [145, 146]). With this expectation, Ref. [108] devised the minimal surviving modification of gravity that can yield cosmic self-acceleration consistent with an event like GW170817. We briefly review this model, before presenting a corresponding reconstructed covariant scalar-tensor theory.

While self-acceleration may generally be defined as cosmic acceleration without a cosmological constant or a scalar field potential, this definition includes exotic dark energy models like k-essence [132] or cubic Galileon and Kinetic Gravity Braiding (KGB) [170] models. Hence, a more precise definition is required if cosmic acceleration is genuinely to be attributed to an intrinsic modification of gravity. This definition also needs to distinguish between models where dark energy or a cosmological constant drives cosmic acceleration but where a modification of gravity may still be present. As a definition of a genuinely self-accelerated modification of gravity in chameleon gravity models, Ref. [171] argued that while cosmic acceleration should be present in the Jordan frame with metric $g_{\mu\nu}$, it should not occur in the conformally transformed Einstein frame $\tilde{g}_{\mu\nu} = \Omega g_{\mu\nu}$ with the conformal factor Ω. Otherwise, the acceleration should be attributed to an exotic matter contribution. In Ref. [107] this argument was generalized to include an evolving speed of gravity c_T in addition to an evolving strength of gravity M^{-2} as the cause of self-acceleration. This encompasses the quartic and quintic Galileon models as well as their generalizations in the full Horndeski action and beyond. These effects can be described by an effective conformal factor in the cosmological background that absorbs the contributions from conformal and disformal couplings in the Einstein frame. An Einstein-Friedmann frame can then be defined from the effective conformal (or pseudo-conformal) transformation of the cosmological background. Alternatively, this can be viewed as assigning genuine cosmic self-acceleration to the magnitude of the breaking of the strong equivalence principle [71]. Note that self-acceleration arising from a dark sector interaction would correspondingly be attributed to the

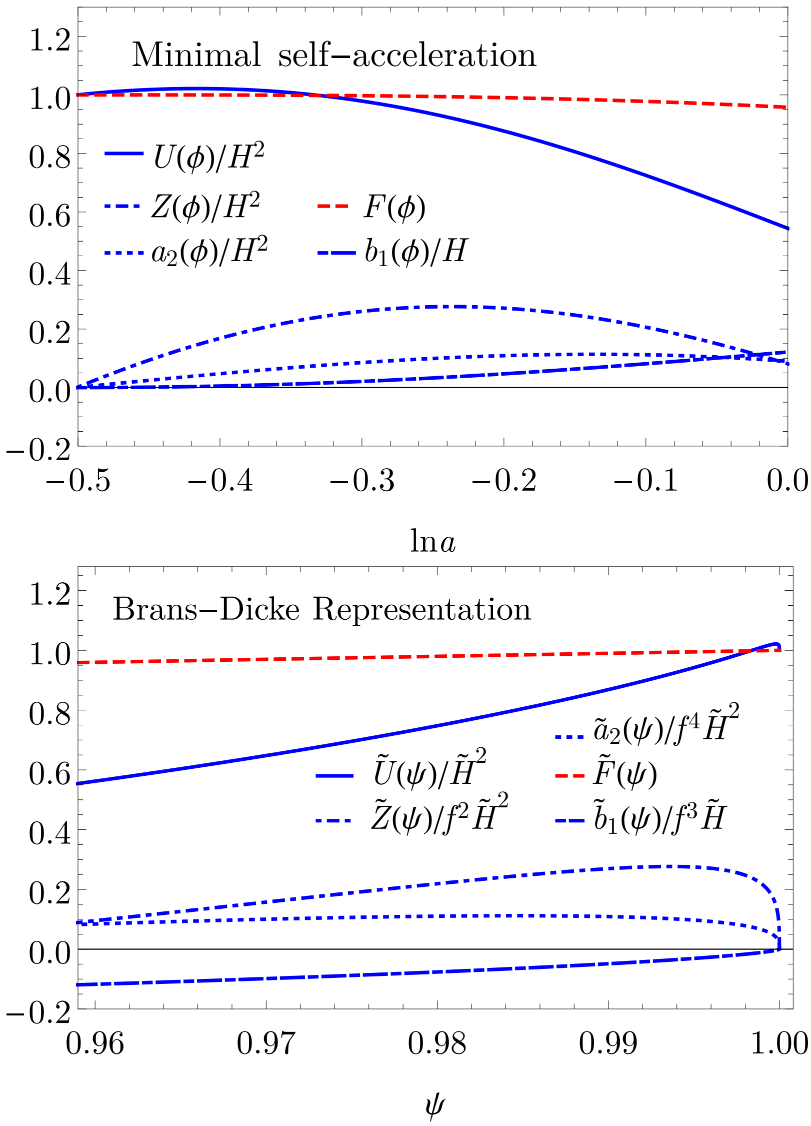

Figure 4.4 Top*: The scalar-tensor theory yielding the minimal modification of gravity required for self-acceleration with $c_T = 1$. Note that the scalar field potential at early times ensures a recovery of the decelerating phase of ΛCDM and decays in the accelerating phase $H^2 < \Lambda$ to barely prevent positive acceleration in Einstein frame.* Bottom*: The minimal self-acceleration model expressed using the Brans-Dicke representation in terms of ψ. We have divided each term by the corresponding factors of $f(\psi)$ for a clearer comparison to the left-hand panel. Note that as $F(\phi)$ is decreasing, the forward direction in time corresponds to decreasing values of ψ.*

breaking of the weak equivalence principle.

With this definition, genuine self-acceleration implies that in the Einstein-Friedmann frame

$$\frac{d^2\tilde{a}}{d\tilde{t}^2} \leq 0\,, \tag{4.18}$$

with the minimal modification obtained at equality. From inspection of the transformed Friedmann equations, it follows that this condition can hold only if the EFT function Ω satisfies [107]

$$-\frac{d\ln\Omega}{d\ln a} \gtrsim \mathcal{O}(1)\,. \tag{4.19}$$

Note that

$$\Omega = \frac{M^2}{M_*^2} c_T^2\,, \tag{4.20}$$

implying that self-acceleration requires a significant deviation in the speed of gravitational waves or an evolving Planck mass. Since GW170817 strongly constrains the deviations of c_T at low redshifts, i.e., in the same regime of cosmic acceleration, one can set $c_T = 1$ ($\alpha_T = 0$) in Eq. (4.20), so that self-acceleration must solely arise from the effect of M^2 (or α_M) [107]. The minimal modification of gravity for genuine cosmic self-acceleration can then be derived by minimizing the impact of a running M^2 on the large-scale structure. For Horndeski gravity, this implies $\alpha_B = \alpha_M$ with $c_s^2 = 1$ setting α_K [108]. The EFT functions of the model are then fully specified by a given expansion history $H(z)$, which for a minimal departure from standard cosmology can be set to match ΛCDM. We present the reconstructed scalar-tensor action for minimal genuine self-acceleration in Fig. 4.4.

Note that for a ΛCDM expansion history, cosmic acceleration in Jordan frame occurs when $H^2 < \Lambda$. Hence, a *minimal* self-acceleration must recover $U/H^2 = 1$ at the transition from a decelerating to an accelerating cosmos. There is therefore still a scalar field potential or cosmological constant that contributes to reproduce the ΛCDM expansion history in the decelerating phase where there are no modifications of gravity but then it decays at a rate so as not to introduce any positive acceleration in the Einstein-Friedmann frame, keeping the Universe at a constant expansion velocity. The cosmic acceleration in Jordan frame is then solely driven by the decaying Planck mass, commencing at the threshold $H^2 < \Lambda$. It is in this sense a model with the minimal gravitational modification required for positive acceleration. Alternatively, the scalar field potential could be removed by hand, but this would lead to a loss of generality and the conservative character of the inferred conclusions.

The reconstructed scalar-tensor terms $F(\phi)$ and $U(\phi)$ for minimal self-acceleration in Fig. 4.4 are decaying functions as expected, with the behavior of the other terms acting to minimize the impact on scalar perturbations and the large-scale structure. At redshift $z = 0$, we find comparable contributions from the quintessence $Z(\phi)$, k-essence $a_2(\phi)$ and cubic Galileon $b_1(\phi)$ terms indicating that they are all required to ensure a minimal self-acceleration. Ref. [108] performed a MCMC analysis of the model with recent cosmological data, finding a 3σ worse fit than ΛCDM and hence strong evidence for a cosmological constant over the minimal modification of gravity required in Horndeski scalar-tensor theories for self-acceleration and consistent with the expectation of the GW170817 result. The constraints are driven by the cross correlation of the integrated Sachs-Wolfe effect with foreground galaxies. It is worth noting that the minimal self-acceleration derived for M^2 also applies to beyond-Horndeski [110, 111] theories or Degenerate Higher-Order Scalar-Tensor (DHOST) theories [66]. Due to the additional free EFT functions introduced in those models, however, the measurement of $\alpha_T \simeq 0$ is not sufficient to break the dark degeneracy and linear shielding is still feasible [113]. However, it was pointed out in Ref. [107] that Standard Sirens tests of the evolution of M^2 are not affected by this degeneracy and may provide a 5σ result on minimal self-acceleration for Horndeksi gravity and its generalizations over the next decade. Independently of future gravitational wave measurements, minimal self-acceleration provides a benchmark model which can quantify to what extent galaxy-redshift surveys like Euclid [27, 28] or LSST [29] can exclude cosmic self-acceleration from modified gravity, precluding dark degeneracies (or linear shielding) in higher-order gravity.

4.3.3 Covariant model with Linear Shielding

A number of classes of scalar-tensor theories that cannot be distinguished from concordance cosmology via observations of the large-scale structure and background evolution alone were presented in Ref. [113]. This phenomena arises through a linear shielding mechanism. It was then shown in Ref. [107] that for Horndeski theories the measurement of $\alpha_T = 0$ breaks this degeneracy. However, linear shielding still remains viable in more general scalar-tensor theories and its extension to the modified gravitational wave propagation may even provide

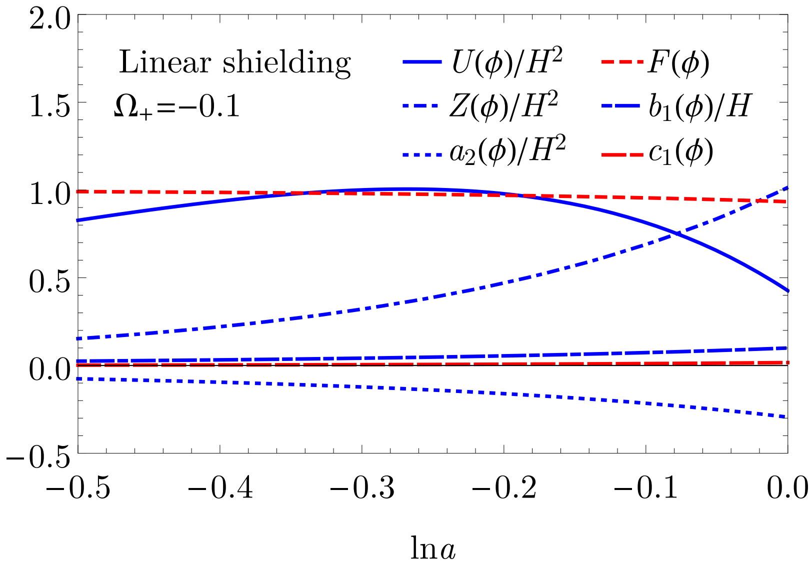

Figure 4.5 *The scalar-tensor theory that exhibits linear shielding for the parameterization in Eq.* (4.23).

a means to evade the GW1701817 constraint for self-acceleration from c_T [157]. It is furthermore worth considering that the $\alpha_T \simeq 0$ constraint only applies at late times and it may remain of interest to examine Horndeski models with non-vanishing α_T at higher redshifts that may also undergo linear shielding. It is therefore worthwhile to examine some basic forms of the scalar-tensor theories that give rise to linear shielding.

In order to recover ΛCDM in the linear cosmological small-scale limit, for models belonging to the $\mathcal{M}_{\rm II}$ class of linear shielding, the EFT functions must satisfy the conditions [107, 113]

$$\alpha_M M^2 = \alpha_B \kappa^2 M^4 - \frac{1-\kappa^2 M^2}{\alpha_B} \times \left\{ \frac{\rho_m}{2H^2} + \left[\alpha_B' + \alpha_B + (1+\alpha_B)\frac{H'}{H} \right] M^2 \right\}, \tag{4.21}$$

$$\alpha_T = \frac{\kappa^2 M^2 - 1}{(1+\alpha_B)\kappa^2 M^2 - 1}\alpha_M \,. \tag{4.22}$$

Applying these constraints, setting the background expansion to match ΛCDM and fixing $c_s^2 = 1$ leaves one free EFT function. With a parameterization of this function and applying the reconstruction, one can then find a scalar-tensor theory that exhibits linear shielding.

Here we adopt the same parameterization as Ref. [113] and choose

$$\Omega(a) = 1 + \Omega_+ a^n \,, \tag{4.23}$$

with $\Omega_+ = -0.1$ and $n = 4$. The general behavior of all the terms in the reconstruction of this linear shielding model is fairly insensitive to changing the magnitude of Ω_+, the one free parameter in the model. The action does differ under a change in the sign of Ω_+, but this acts to decelerate the expansion.

We illustrate the reconstructed scalar-tensor action for the choice of parameters in Fig. 4.5. $U(\phi)$ is dominated by the EFT function $\Lambda(t)$ which behaves in a similar way to the minimal self-acceleration model, acting as a cosmological constant at early times before decaying away at late times. The late-time decay of $\Lambda(t)$ is compensated by the other terms in the reconstruction to ensure that the linear perturbations are not affected in their ΛCDM behavior. $F(\phi)$ also decays which is a consequence of the choice of a negative Ω_+, required for self-acceleration. The linearly shielded Horndeski model requires a decrease in the speed of gravitational waves over time which leads to $c_1(\phi)$ growing in time. The kinetic terms become more dominant at late times, predominantly being driven by Γ and M_2^4 with the form of $a_2(\phi)$ essentially mimicking that of M_2^4. In $b_1(\phi)$ the contributions $\bar{M}_1^3$ and $\bar{M}_2^2$ compete and suppress it relative to the other terms in the action.

Although the conditions for linear shielding may seem contrived when expressed in terms of the EFT parameters, we find that it is nevertheless the case that there is a generic scalar-tensor theory which gives rise to this mechanism for the particular parameterization we adopt. It is also worth bearing in mind that observational large-scale structure constraints allow for a broad variation around the strict conditions in Eqs. (4.21) and (4.22) in which the model space remains observationally degenerate with ΛCDM.

4.3.4 μ and η reconstruction

The effects of modified gravity and dark energy on the large-scale structure can be described phenomenologically by the behaviour of two functions of time and scale that parameterize a deviation in the Poisson equation $\mu(a, k)$ and introduce a gravitational slip $\eta(a, k)$ [158–162]. We shall work with a perturbed FLRW metric in the Newtonian gauge with $\Psi \equiv \delta g_{00}/2g_{00}$ and $\Phi \equiv \delta g_{ii}/2g_{ii}$ and matter density perturbations Δ_m in the comoving gauge. The effects of modified gravity

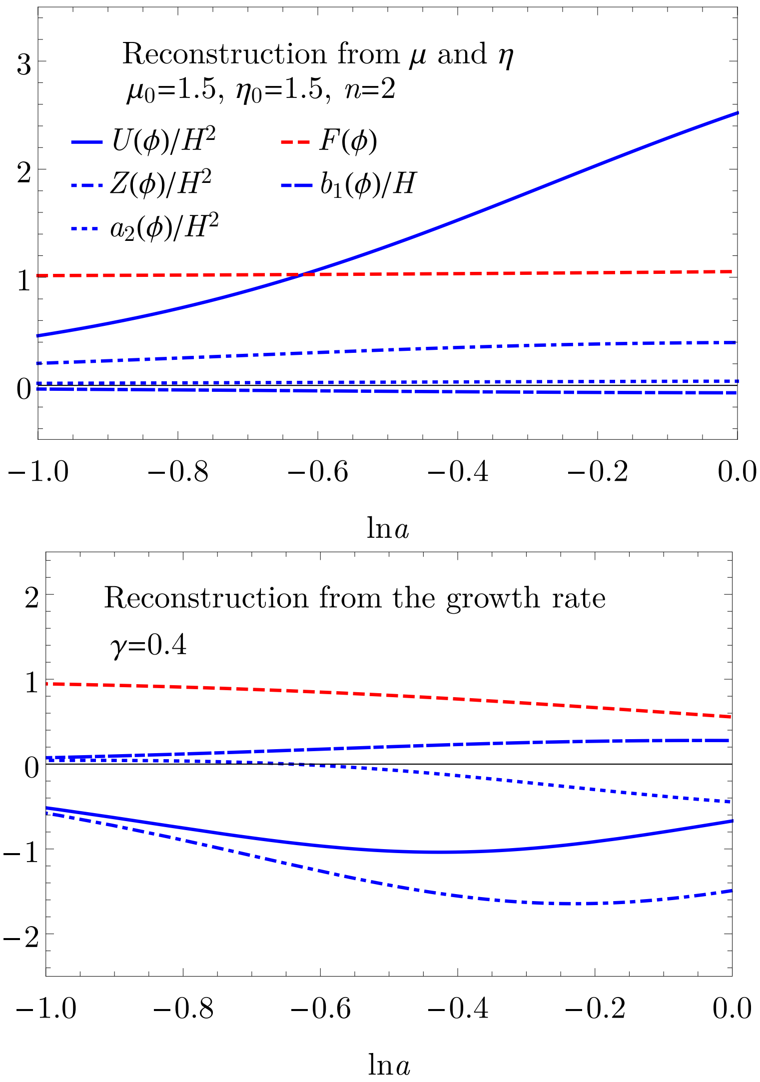

Figure 4.6 Left*: Reconstructed action from a direct parameterization of the modified Poisson equation and the gravitational slip.* Right*: Reconstructed action from the growth-index parametrization.*

and dark energy on the perturbations can be described via the relations

$$k_H^2 \Psi = -\frac{\kappa^2 \rho_m}{2H^2} \mu(a,k) \Delta_m \,, \tag{4.24}$$

$$\Phi = -\eta(a,k) \Psi \,, \tag{4.25}$$

where $k_H \equiv k/(aH)$. Energy and momentum conservation then closes the system of differential equations and one can solve for the evolution of the linear perturbations.

The modifications $\mu(a,k)$ and $\eta(a,k)$ are more general than the EFT formalism but the two can be linked in the domain covered by the EFT functions. Specifically, in the formal linear theory limit of $k_H \to \infty$ the functions μ and η can be treated as only functions of time. In this limit, they can be related to the EFT functions via

$$\mu_\infty = \frac{2\left[\alpha_B(1+\alpha_T) - \alpha_M + \alpha_T\right]^2 + \alpha(1+\alpha_T)c_s^2}{\alpha c_s^2 \kappa^2 M^2} \,, \tag{4.26}$$

$$\eta_\infty = \frac{2\alpha_B\left[\alpha_B(1+\alpha_T) - \alpha_M + \alpha_T\right] + \alpha c_s^2}{2\left[\alpha_B(1+\alpha_T) - \alpha_M + \alpha_T\right]^2 + \alpha(1+\alpha_T)c_s^2} \,. \tag{4.27}$$

For the purposes of this chapter we shall remain in this small-scale regime and parameterize the time-dependent modifications as

$$\mu(a) = 1 + (\mu_0 - 1)a^n \,, \tag{4.28}$$

$$\eta(a) = 1 + (\eta_0 - 1)a^n \,, \tag{4.29}$$

with $n = 2$. For simplicity, we furthermore consider a background evolution $H(t)$ that matches that of ΛCDM and we adopt $\alpha_T = 0$ at all times to break the degeneracy in parameter space. The kineticity function α_K is set by the choice $c_s^2 = 1$. The set of EFT functions is then closed by Eqs. (4.26) and (4.27), determining the evolution of α_B and α_M. Given a choice of parameters μ_0, η_0 we can now reconstruct a corresponding Horndeski scalar-tensor theory. For this example we choose a model that exhibits both a non-zero gravitational slip and an enhanced growth of structure today by setting $\mu_0 = \eta_0 = 3/2$.

The reconstructed scalar-tensor action is illustrated in Fig. 4.6. The dominant term at redshift zero is $U(\phi)$. It behaves as a cosmological constant which is enhanced relative to its ΛCDM value. $F(\phi)$ is determined through the evolution of M^2. Despite the enhanced growth with this parameterization of μ and η the Planck mass increases from its GR value today. The enhanced growth is therefore coming from the clustering effect of α_B. This can be seen more clearly by writing

$$\mu = \frac{M_*^2}{M^2}\left(1 + \frac{2(\alpha_B - \alpha_M)^2}{\alpha c_s^2}\right) . \tag{4.30}$$

Although the Planck mass is increasing, α_B also increases to dominate over α_M and gives rise to the pre-defined evolution in $\mu(a)$. The domination of α_B over α_M also causes $b_1(\phi)$ to be negative. This is because $b_1 \sim \bar{M}_1^3 \sim (\alpha_M - 2\alpha_B)$ up to numerical factors and positive background terms. In this model $\alpha_K \approx 0$. The background terms that contribute to M_2^4 compete to cancel each other out. The dominant term in $a_2(\phi)$ is from $-\bar{M}_1^3$ or α_B, which is small and positive.

4.3.5 Ω_m^γ reconstruction

One of the most commonly used formalisms for testing departures from GR with the large-scale structure is the growth-index parametrization [163–165]. It involves a direct parameterization of a modification of the growth rate

$$f \equiv \frac{d\ln\Delta_m(a,k)}{d\ln a} = \Omega_m(a)^\gamma \tag{4.31}$$

with the growth-index parameter γ, which is generally considered a *trigger* or *consistency* parameter. Any observational deviation from its GR value $\gamma \approx 6/11$ [163] will indicate a breakdown of GR.

On sub-horizon scales ($k \gg aH$) the modified growth equation for the matter density contrast is given by

$$\Delta_m'' + \left(2 + \frac{H'}{H}\right)\Delta_m' - \frac{3}{2}\Omega_m(a)\mu_\infty(a)\Delta_m = 0\,, \tag{4.32}$$

which follows from the modified Poisson equation (4.24) and momentum conservation. Inserting Eq. (4.31) into (4.32), one obtains a relation between $\mu_\infty(a)$ and γ,

$$\mu_\infty = \frac{2}{3}\Omega_m^{\gamma-1}\left[\Omega_m^\gamma + 2 + \frac{H'}{H} + \gamma\frac{\Omega_m'}{\Omega_m} + \gamma'\ln\left(\Omega_m\right)\right] , \tag{4.33}$$

where we allowed γ to be time dependent for generality. Given a particular choice of γ, the functional form of μ_∞ can then be obtained from Eq. (4.33). However, as γ can only be used to determine μ_∞, one must separately parameterize the gravitational slip η_∞ (and some additional specifications are required for a relativistic completion [116, 172, 173]). One can then reconstruct a covariant theory that gives rise to the particular choices of γ and η_∞. This allows to directly examine what kind of theories can be associated with an observational departure from GR in γ.

In this example, we set for simplicity $\eta_\infty = 1$ as in GR. This implies that, with $\alpha_T = 0$, $\alpha_M = 0$ or $\alpha_B = \alpha_M$. We choose the second condition. With this choice we have that $M^2 = 1/\mu_\infty$ and we fix α_K such that $c_s^2 = 1$. We shall reconstruct a theory which gives rise to a constant deviation in the growth index from the GR value of $\gamma \approx 0.55$. The value for γ needs to be chosen such that the stability condition $\alpha > 0$ is satisfied and so we choose $\gamma = 0.4$ for this purpose. In fact, the theoretical stability of the theory requires $0.35 \lesssim \gamma \lesssim 0.55$, preferring enhanced growth of structure, with any value chosen outside this range leading to $\alpha < 0$. As long as the theoretical conditions are satisfied then it is straightforward to apply the reconstruction and obtain a covariant theory for any numerical value for γ.

The corresponding model is illustrated in Fig. 4.6. As we have chosen a rather large departure from ΛCDM the reconstructed theory displays a somewhat unnatural behavior with a potential that is negative and substantial contributions from the kinetic and Galileon terms in order to maintain the background expansion history. Therefore, even with this seemingly simple parameter it is quite possible that exotic regions of the space of theories are being explored when it deviates from its concordance value.

4.3.6 Weak gravity

Typically scalar-tensor theories exhibit an enhanced growth of the matter density fluctuations relative to ΛCDM, with Brans-Dicke gravity being a simple example [174]. More precisely, they lead to a modification such that $\mu > 1$ in Eq. (4.24). However, it is possible that modifications arise such that one obtains a weaker growth of structure, or *weaker gravity*, with $\mu < 1$. This scenario has recently received some attention [112, 114, 155, 175, 176], particularly in the context of potential tensions in the cosmological data [166, 167].

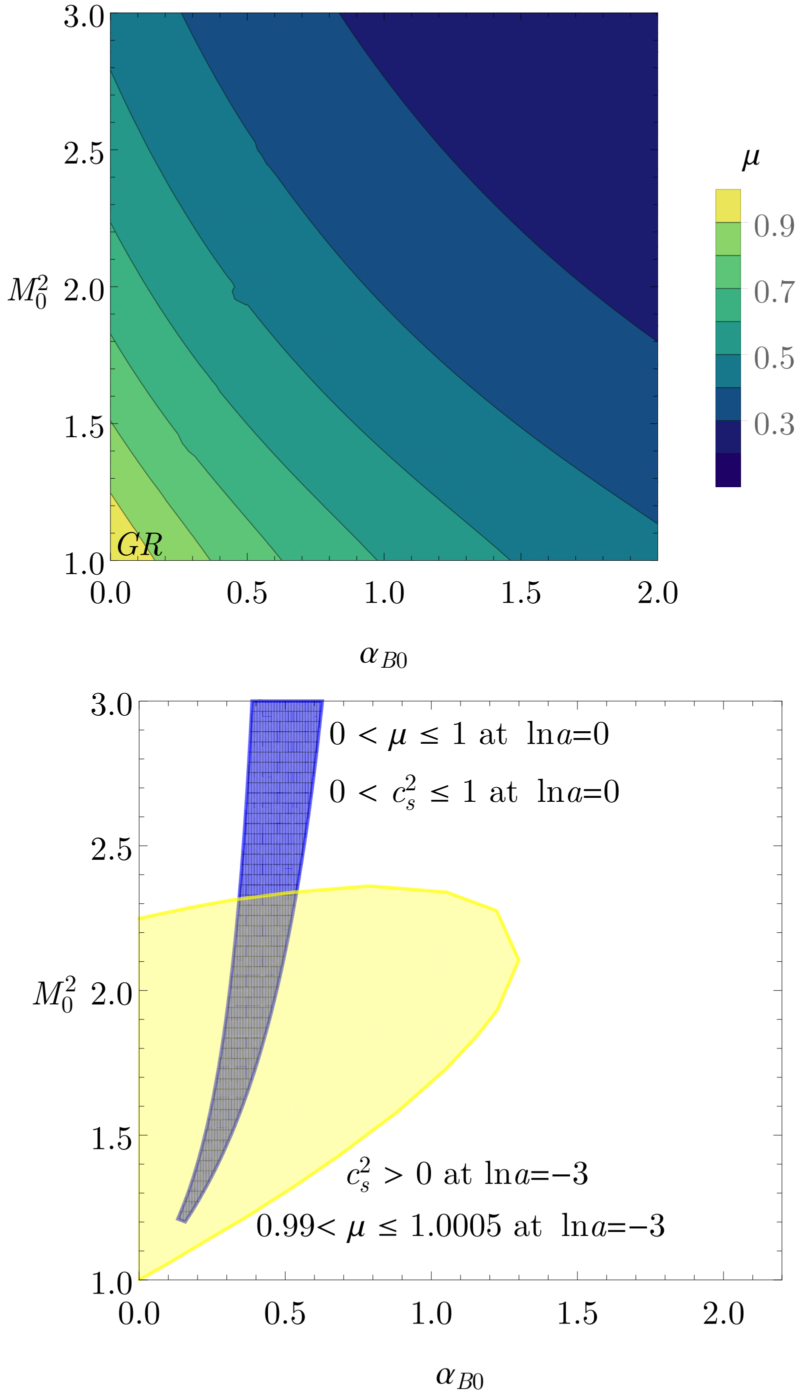

Figure 4.7 Top*: Contour plot in the space of M_0^2 and α_{B0} displaying the regions that allow for a weakened growth of structure with $0 < \mu_0 < 1$ today.* Bottom*: The dark strip indicates the region of EFT parameter space that allows for a weakening of growth with a positive, sub-luminal soundspeed at redshift zero. After imposing the past boundary conditions $\mu = 1$ and $c_s^2 > 0$ at $\ln a = -3$ indicated by the lighter yellow region it is possible to reconstruct a viable covariant model from any point in the intersecting region. We have ensured that the chosen point used for the reconstruction in Fig. 4.8 satisfies $c_s^2 > 0$ for all time.*

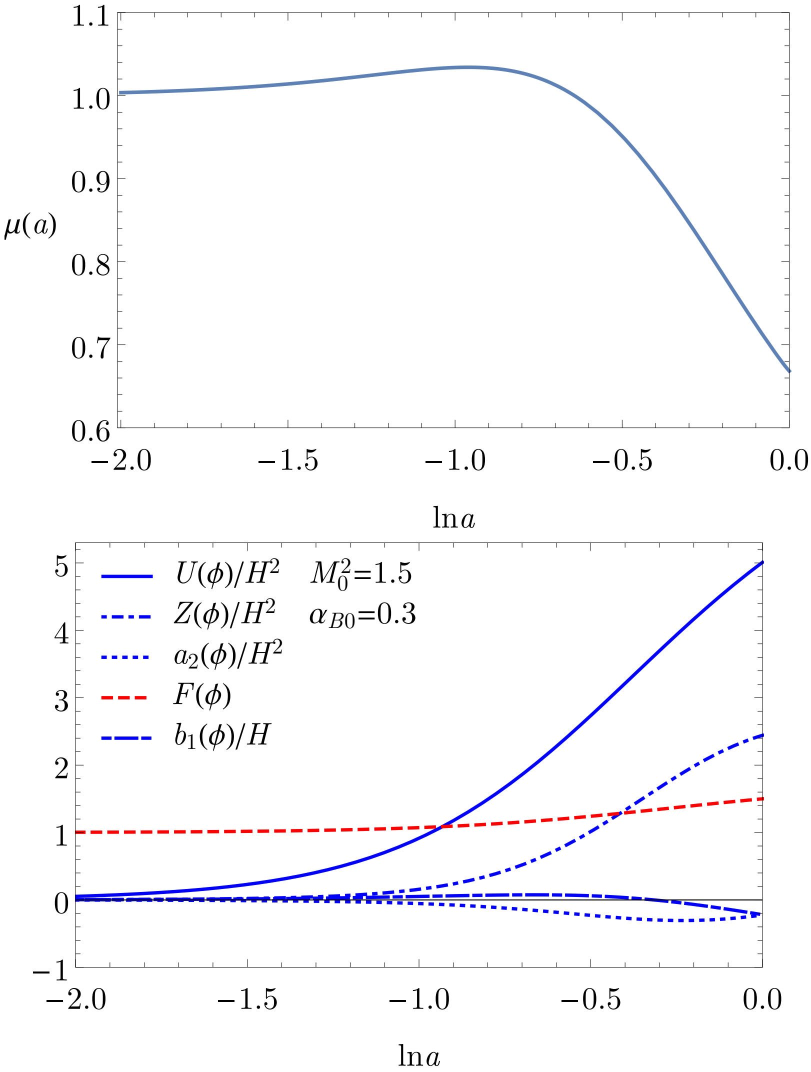

Figure 4.8 Top: *The behaviour of the deviation from the Poisson's equation over time for the model in Sec. 4.3.6, where one may identify a dynamical* $G_{\text{eff}} \equiv \mu$. *There is a characteristic period of enhanced growth at* $\ln a \approx -0.96$ *before entering an epoch of weakening of the growth persisting today.* Bottom: *A reconstructed scalar-tensor theory that exhibits a weakening of the growth of structure ("weak gravity") with* $\alpha_T = 0$*, which satisfies the stability requirements and past boundary conditions. It is essentially a Brans-Dicke type model with a potential and standard kinetic term along with small contributions from the k-essence and cubic terms.*

In this section we demonstrate how one may use the reconstruction to derive a stable scalar-tensor theory of weak gravity for a particular parameterization of the EFT functions with $\alpha_T = 0$.

We begin by choosing the parameterization of the Planck mass M^2 as

$$M^2 = 1 + (M_0^2 - 1)\frac{\Omega_\Lambda(a)}{\Omega_{\Lambda 0}}\,, \tag{4.34}$$

where M_0^2 is the value of the Planck mass today. The particular choice of Planck mass evolution when $M_0^2 > 1$ is a priori suggestive of weak gravity as M^2 appears in the denominator of Eq. (4.26) such that the increasing Planck mass with time leads to a decreasing μ if fixing the other EFT parameters. However, there is still a great deal of freedom in choosing numerical values for M_0^2 and the evolution of the remaining α_i. For instance, it may be the case that the evolution in α_B is enough to compensate for the weakened growth effect and give rise to an enhancement instead. For this example, we adopt the functional form of $\mathcal{B}$ in Sec. 4.3.1 with $q_i = q = 1$ for the parameterization of the α_B function and we set $\alpha_K = 0$ for simplicity and to easily guarantee that the stability condition $\alpha > 0$ is satisfied. As previously mentioned, α_K only becomes relevant on scales comparable to the horizon and so the requirement that $\mu < 1$ is independent of the choice of α_K. Parameter values for M_0^2 and α_{B0} are then chosen to ensure that the condition $c_s^2 > 0$ is satisfied.

We explore the viable regions of parameter space producing a given $\mu_0 \equiv \mu(z = 0)$ in the left-hand panel of Fig. 4.7. One can easily identify a large region that allows for weak gravity with $0 < \mu_0 < 1$ when $M_0^2 > 1$ while remaining stable and having the Planck mass return to its bare value in the past by construction. All of these requirements severely restrict the allowed model space. In fact, we find that within the particular parameterization adopted here, a period of enhanced growth in the past is required in order for all of these criteria to be satisfied.

We explore this circumstance in more detail in the right-hand panel of Fig. 4.7. For this purpose, we allow for a small period of enhanced growth in the past at $\mathcal{O}(10^{-4})$, which allows one to find an overlap of stable parameter choices that also yield weak gravity at late times. Increasing this value causes the viable parameter regions to overlap at an even greater extent. Restricting parameters to an upper bound of exactly unity instead eliminates any overlap.

A suitable parameter choice that satisfies all of the requirements described here is $M_0^2 = 3/2$ and $\alpha_{B0} = 0.3$ and we checked that for this choice the soundspeed remains positive at all times in the past. The left-hand panel of Fig. 4.8 displays the evolution of the gravitational coupling through time with this choice of EFT parameters. One can clearly identify a period of enhanced growth which peaks around $\ln a \approx -0.96$ with $\mu \approx 1.03$ before decaying and producing weak gravity with $\mu \approx 0.65$ at redshift $z = 0$.

Once given the choice of EFT parameters it is straightforward to implement them in the reconstruction and obtain a stable scalar-tensor theory that exhibits a weakening of growth of structure with $\alpha_T = 0$. The corresponding model is illustrated in the right-hand panel of Fig. 4.8. The evolution of $U(\phi)$ mimics that of a cosmological constant, but as $\Lambda \sim M^2H^2$ it is enhanced relative to its ΛCDM behavior due to the increase of the Planck mass over time. This is similar to the behaviour observed in Sec. 4.3.4. The Planck mass also determines the evolution of $F(\phi)$ which increases over time. The behaviour of $b_1(\phi)$ is determined by the combination $\alpha_M - 2\alpha_B$. The braiding term is sub-dominant at early times, but becomes important at late times, where it contributes to drive $b_1(\phi)$ negative. There is also a small negative k-essence term $a_2(\phi)$ that is comparable in magnitude to $b_1(\phi)$.

Bear in mind that different choices of q_0, a non-zero α_K or a parameterization in terms of α_M rather than M^2 impacts the form of the theory. However, it is primarily sensitive to significant changes in the amplitudes of parameters as discussed in Sec. 4.4, and one does not have much freedom in increasing the amplitude of α_B while keeping the theory stable (Fig. 4.7). Finally, note that this weak gravity model differs from Ref. [155] as $\alpha_M \neq \alpha_B$, thus exhibiting a non-vanishing gravitational slip. More work is necessary to understand what general conditions need to hold in order to obtain a stable scalar-tensor theory the exhibits a weakened growth of structure and $\alpha_T = 0$.

4.3.7 Reconstruction from inherently stable parameterizations

Throughout this work it has been necessary to check that the reconstructed theories obey the stability constraints in Eqs. (4.5) and (4.7). This is due to the function space spanned by the basis of α_i, or equivalently the coefficients in the EFT action in Eqs (3.14) and (3.15), not being a priori stable. As discussed

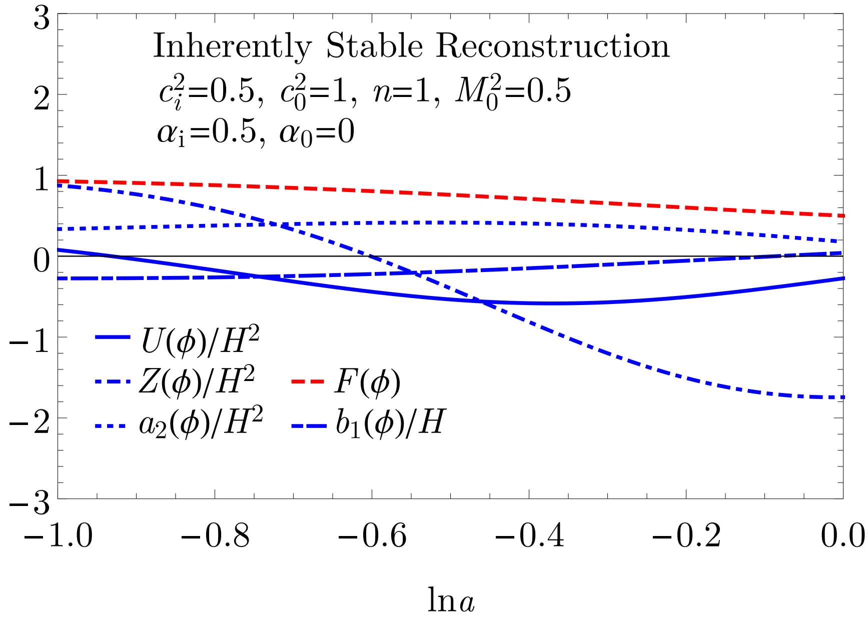

Figure 4.9 *Reconstructed scalar-tensor theory from a direct parametrization of the stability functions $c_s^2 > 0$, $\alpha > 0$ and $M^2 > 0$ with $\alpha_T = 0$.*

in Sec. 4.2.2, rather than cumbersomely checking that these stability criteria are satisfied for a particular parameterization, one may instead consider discarding the α_i functions in favor of another parameterization that automatically satisfies the stability requirements. Therefore, any observational constraints will by definition be restricted to a theory space that obeys the no-ghost and no-gradient instability conditions. We introduced such an inherently stable basis in Sec. 4.2.2.

We shall now briefly present a reconstruction from this basis. For this purpose we adopt the functional forms

$$c_s^2 = c_i^2 + (c_0^2 - c_i^2)a^n \,, \tag{4.35}$$

$$\alpha = \alpha_i + (\alpha_0 - \alpha_i)a^n \,, \tag{4.36}$$

where the constants c_i^2 and α_i are initial conditions for the soundspeed and the kinetic term respectively (defined for the limit $a \to 0$) whereas c_0^2 and α_0 set their values today. Each value should be chosen such that $\alpha, c_s^2 > 0 \; \forall a$. For the Planck mass we adopt the parameterization in Eq. (4.34).

In Fig. 4.9 we illustrate a reconstructed theory with a ΛCDM background, an increasing soundspeed as well as decaying kinetic term and Planck mass. More specifically, we set $c_i^2 = 0.5$, $c_0^2 = 1$, $\alpha_i = 0.5$, $\alpha_0 = 0$, $M_0^2 = 0.5$, and $n = 1$. Although the reconstructed terms seem somewhat exotic, for example

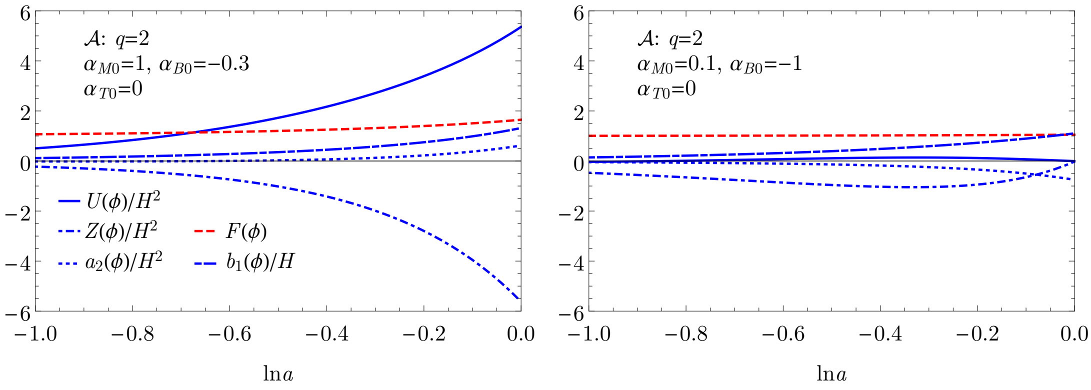

Figure 4.10 *The effect of varying parameter values in a parameterization of EFT functions on the reconstructed scalar-tensor theory for a model with a dominant Planck mass evolution α_M (*left panel*) and a model with a dominant braiding term α_B (*right panel*). Note that α_{M0} and α_{B0} are of opposite sign to satisfy the stability requirements. In the right-hand panel where α_B dominates, the cubic Galileon term b_1 is the most prevalent modification as the potential and quintessence terms decay to zero. There is also a non-negligible contribution from the k-essence term. On the contrary, a dominating α_M leads to a large potential and quintessence kinetic term, with smaller contributions from the cubic and k-essence terms.*

the potential is very different to its ΛCDM behavior despite the concordance background evolution, by construction the model is guaranteed to be stable.

4.4 Effect of varying the parameterization on the underlying theory

Finally, we examine the sensitivity of the reconstructed theories on the variation of parameter values for a given parametrization of the EFT functions. We shall only use the functional form $\mathcal{A}$, discussed in Sec. 4.3.1, which is broadly used in literature. Recall that we have found that the form of the underlying theory is rather insensitive to the choice between functions $\mathcal{A}$ and $\mathcal{B}$ (Fig. 4.2). In all cases we check that the stability condition $\alpha > 0$ is satisfied and with the remaining freedom in α_K we set $c_s^2 = 1$. We furthermore set $\alpha_T = 0$. As a consequence of these choices, the signs of α_B and α_M are opposite.

In Fig. 4.10 we show the effect on the theory when the braiding term α_B dominates

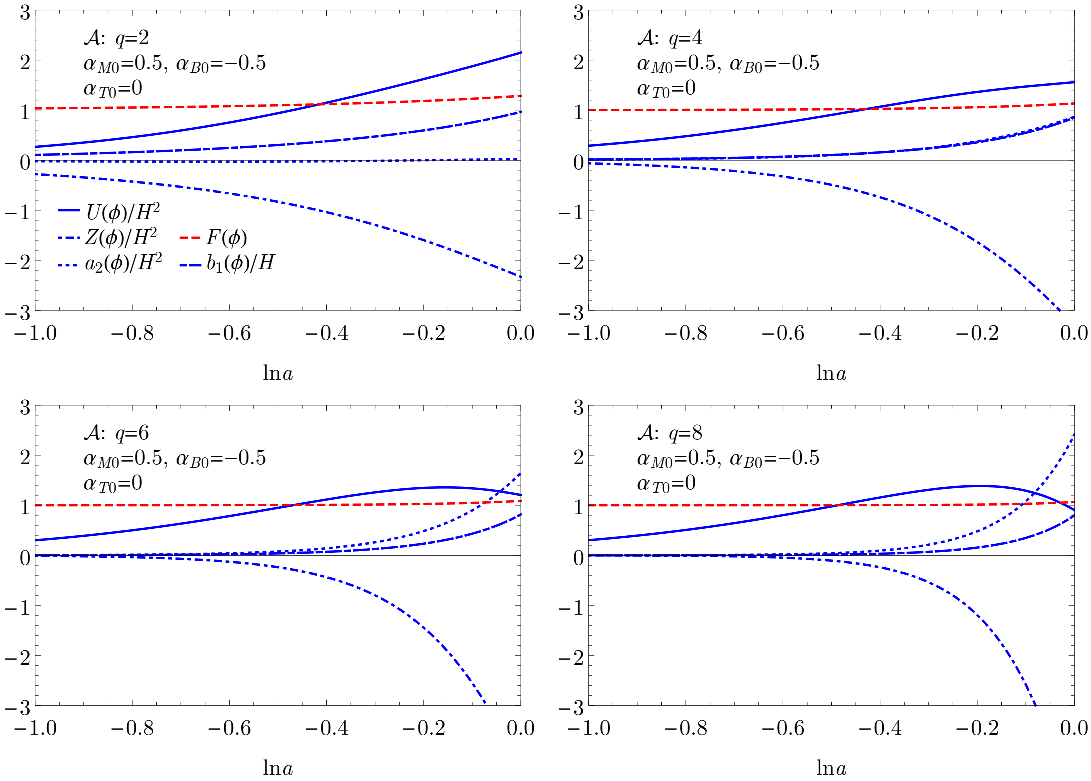

Figure 4.11 *The effect of varying the powers q in the parameterization on the underlying theory. It is apparent that with this choice of α_i functions every term in the reconstruction becomes relevant. Modifications are suppressed at high redshift with increasing power, with a steepening at low redshifts. For this choice of amplitudes, the k-essence term is particularly sensitive, increasing from zero to dominate over the potential for large q. The standard kinetic term and potential become more negative at $z = 0$ for larger powers. This is in contrast to the cubic term $b_1(\phi)$, which remains relatively unaffected by this alteration in the parameterization.*

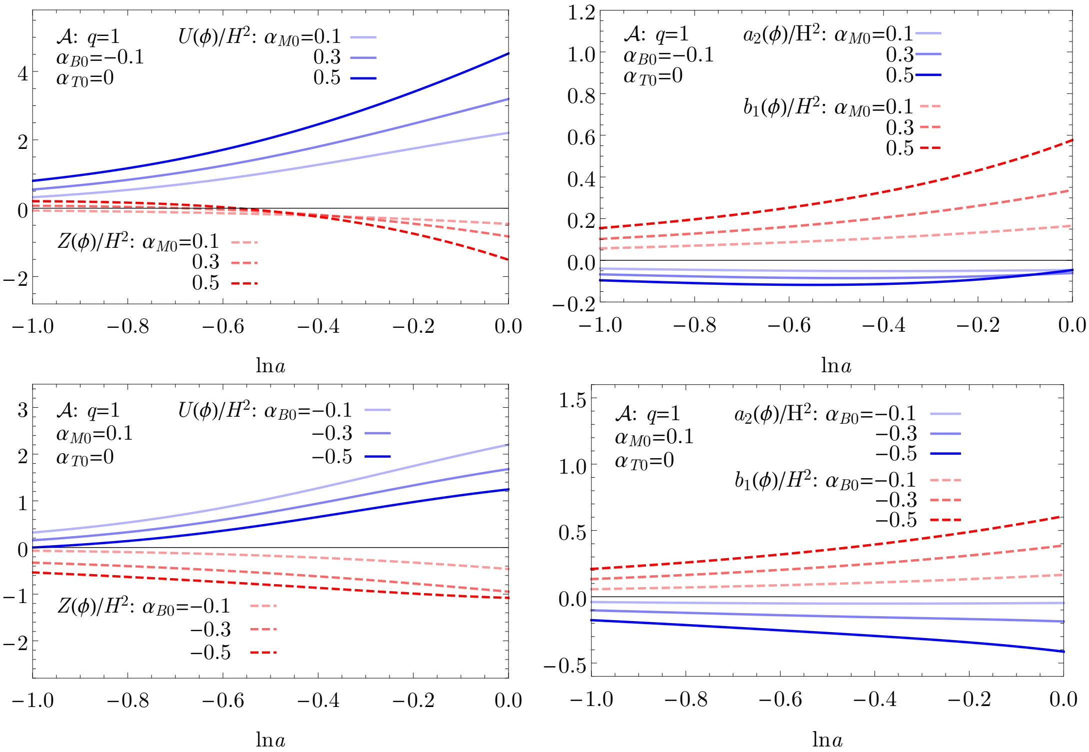

Figure 4.12 *Effects on the reconstructed scalar-tensor theory from incremental changes in the amplitude of α_M for a fixed α_B and vice versa. The general form of the underlying theory is rather insensitive to these changes. Enhancing α_B suppresses the potential and enhances all the other terms whereas enhancing α_M increases every term in the reconstruction other than the k-essence term $a_2(\phi)$. Note that the colour scheme here bears no distinction between dark energy and modified gravity in contrast to all other figures.*

over the variation in the Planck mass α_M and vice versa. In the first instance, the dominant terms are a potential behaving like a cosmological constant and a large kinetic term for the scalar field mimicking a Brans-Dicke theory with small k-essence and cubic Galileon contributions. On the contrary, when α_B dominates over α_M the cubic term $b_1(\phi)$ becomes the most relevant term in the theory with the potential decaying away rapidly towards $z = 0$. In both scenarios the ΛCDM expansion history is maintained by the behaviour of the complementary terms in the reconstruction that compensate for the change in the potential.

Next, we examine the effects of varying the power in the parametrization while retaining consistency in the stability requirements. We fix the magnitude of α_{M0} and α_{B0} to be equal but opposite. The effects of changing the power on the underlying theory are illustrated in Fig. 4.11. When the power of the parameterization is increased the effects of modified gravity become more relevant at later times. The cubic term is generally unaffected by this variation, but the kinetic and k-essence terms are enhanced. When a large power is chosen, the k-essence contribution comes to dominate at late times.

Finally, in Fig. 4.12 we illustrate the effects of changing α_{M0} while keeping α_{B0} fixed and vice versa. We find that the form of the underlying theory is fairly insensitive to small changes in the amplitude, although certain terms may be enhanced or suppressed relative to others with different choices. For example, increasing α_M has the effect of enhancing the potential relative to that of ΛCDM. This is again due to the dependence of $\Lambda \sim M^2$. The kinetic term $Z(\phi)$ is also enhanced although to a lesser degree than the potential whereas the k-essence and cubic Galileon terms $a_2(\phi)$ and $b_1(\phi)$ are rather insensitive to these $\mathcal{O}(10^{-1})$ changes in α_M. The term $a_2(\phi)$ remains least affected with smaller variations restricted to the past. Thus, in general we find that by enhancing α_{M0} for a fixed, small α_{B0}, one is enhancing the potential and the standard kinetic term of the scalar-tensor model. In contrast, for a fixed small value of α_{M0}, enhancing the effects of α_{B0} leads to a suppression of the potential and an enhancement of the cubic Galileon term.

4.5 Conclusions

Finding a natural explanation for the observed late-time accelerated expansion of our Universe continues to be a significant challenge in cosmology. It is therefore important that efficient methods are devised with the aim of connecting cosmological observables with the wealth of proposed theories to obtain a deeper understanding of the underlying physical mechanism driving the expansion. These efforts may furthermore give crucial insights into the persistent issues related to the reconciliation of quantum field theory with general relativity.

As we have emphasised, the effective field theory of dark energy provides a useful tool for studying the dynamics of cosmological perturbations of a large family of scalar-tensor theories in a unified framework. Many of the upcoming surveys of the large-scale structure plan to utilize this formalism to constrain the freedom in modified gravity and dark energy phenomenology [27–29]. It is therefore crucial to be able to connect any observational constraints to the underlying space of scalar-tensor theories, which in turn can be connected to more fundamental theories of gravity.

Chapter 3 developed a reconstruction method that maps from a set of EFT functions to the family of Horndeski theories degenerate at the level of the background and linear perturbations. In this chapter we applied this mapping to a number of examples. These include the comparison of the resulting action when one utilises two frequently adopted phenomenological parameterizations for the EFT functions to study the effects of dark energy and modified gravity at late times. We find that changing between the two parameterizations has a small effect on the general form of the underlying theory, although certain terms can be enhanced relative to others. The underlying theory is instead more sensitive to the amplitudes of the different EFT functions.

Of particular interest is the reconstruction of a model that exhibits minimal self-acceleration. The reconstructed scalar-tensor theory possesses the minimum requirements on the evolution of the Planck mass for self-acceleration from a modification of gravity consistent with a propagation speed of gravitational waves equal to that of light. It is a useful model to test for the next generation of surveys, as it acts as a null-test for self-acceleration from modified gravity.

We also examine models that exhibit a linear shielding mechanism to hide the gravitational modifications in the large-scale structure. Although the simplest models require a non-vanishing α_T, it is worth bearing in mind that the stringent constraint on the speed of gravity with $\alpha_T = 0$ only applies at low redshifts and may also involve scale dependence [157] for more general theories. While the constraints in the space of the EFT functions for linear shielding to operate seem rather complicated, using the reconstruction we find there are generic Horndeski theories that exhibit this effect.

We furthermore provide a direct connection between various parameterizations that exist in the literature and the corresponding underlying theories. For example, we reconstruct theories from a phenomenological parameterization of the modified Poisson equation and gravitational slip as well as from the growth-index parameter. One can use these reconstructions to connect constraints arising from such parameterizations with viable Horndeski models. We also apply the reconstruction to obtain a theory that exhibits a weakening of the present growth of structure relative to ΛCDM, i.e., a weak gravity model, a possibility that may ease potential tensions in the growth rate at low redshift [166, 167].

Finally, we proposed an alternative parameterization basis for studying dark energy and modified gravity models which is manifestly stable. These are the Planck mass, the dark energy soundspeed, the kinetic energy of the scalar field and a braiding amplitude as the new basis of EFT functions. Any constraints placed on these physical parameters are guaranteed to correspond to healthy theories. It is no longer necessary to perform separate and cumbersome stability checks on sampled theories when using this basis.

Chapter 5

Screening and degenerate kinetic self-acceleration from the nonlinear freedom

Mathematical science shows what is. It is the language of unseen relations between things. But to use and apply that language, we must be able fully to appreciate, to feel, to seize the unseen, the unconscious.

Ada Lovelace

5.1 Introduction

The cost of generality in the EFT of dark energy formalism is its restricted applicability to certain length scales, usually just the cosmological background and linear perturbations. Recently however there has been some work in extending the expansion to higher-order perturbations [177, 178]. An alternative approach is to start from the full covariant action. The loss of generality is then traded for the applicability on a much broader range of length scales, allowing nonlinear effects such as screening to be studied. We have presented in chapter

3 a reconstruction from the EFT of dark energy on the level of the background and linear perturbations to the class of Horndeski theories that give rise to the particular set of given EFT functions. With this covariant action it becomes feasible to generally connect the nonlinear regime to that of the background and linear scales. This link shall be the focus of this chapter.

More precisely, within the reconstructed theory there are correction terms defined in Eqs. (3.22) and (3.23) that account for the nonlinear freedom that exists between Horndeski theories that are degenerate at the level of the background and linear perturbations. Specification of these correction terms allows one to move between linearly degenerate theories.

We first discuss the uniqueness of the correction terms in the reconstructed theory. Applying the recent constraint on the equality between the speeds of light and of gravitational waves [136] we show that the number of free functions that are present at higher order in the EFT of dark energy is significantly reduced to two per order in perturbation theory. This then implies that the nonlinear freedom is uniquely specified by the nonlinear correction terms. It is worth noting that out of the four new EFT functions found in Ref. [42] at second order in the cosmological perturbations of Horndeski theory, the two functions dominating in the sub-horizon regime vanish for a luminal speed of gravity, and the impact of the nonlinear correction terms on the weakly nonlinear regime of structure formation remains to be examined in detail.

As an initial demonstration of the implications of the correction terms, we show how this nonlinear freedom can be used to endow a reconstructed theory with a screening mechanism. Due to the tight Solar-System constraints on deviations from GR [179] it is necessary for a large-scale modification of GR to employ a screening mechanism that suppresses the effects of a fifth force on small scales. These screening mechanisms fall into one of three categories [71]: those that screen through deep gravitational potentials such as the chameleon [88] or symmetron mechanisms [180], screening through first derivatives of the potentials such as k-mouflage models [90] or screening through second derivatives as for the Vainshtein mechanism [92]. See Sec. 2.2.7 for an overview.

A simple scaling method was developed in Refs. [146, 181] to determine whether a given theory possesses an Einstein gravity limit or not. We present an application of this scaling method to the reconstructed theory and demonstrate with three examples that there is enough freedom in the nonlinear regime of a reconstructed theory to obtain, in principle, any of these three screening mechanisms.

A further interesting consequence that arises when considering theories built from the correction terms is that it is simple to construct theories that are indistinguishable from ΛCDM to arbitrary level in cosmological perturbations. Only observations in the nonlinear regime can be used to distinguish them from ΛCDM. Such degenerate theories may be built from kinetic terms alone without including a cosmological constant, hence providing a kinetic self-acceleration effect.

Finally, we present a reconstruction from the nonlinear EFT back to the space of manifestly covariant theories. This follows a similar structure to the background and linear reconstruction and in principle provides a method for obtaining a Horndeski theory reconstructed from a range of different length scales from the background to the nonlinear regime.

The chapter is organised as follows. We present in Sec. 5.2 the Horndeski field equations arising from the remaining freedom in Horndeski after the luminal speed of gravity constraint is applied. These will be relevant for the application to screening in Sec. 5.4. The nonlinear correction terms in the reconstructed action are also reviewed for convenience. The uniqueness of the nonlinear correction terms in the reconstructed action is examined in Sec. 5.3. Sec. 5.4 briefly reviews the scaling method and discusses how the nonlinear freedom in the reconstructed scalar-tensor theories can be used to implement screening effects due to large gravitational potentials and large first or second derivatives of the potential. In Sec. 5.5 we discuss how the nonlinear freedom can be used to construct models that accelerate the cosmic expansion without a cosmological constant with a suitable choice of kinetic terms, yet are degenerate with standard cosmology at the background level or even to arbitrary level of perturbations. The derivation of a third-order reconstruction is presented in Sec. 5.6 along with a discussion of the extension to n-th order. Finally, we provide conclusions on the results in the chapter in Sec. 5.7.

5.2 Horndeski field equations and nonlinear freedom

Following the simultaneous detection of gravitational waves with an electromagnetic counterpart [136], the available freedom in Horndeski theory greatly simplified to [146]

$$\mathcal{L}_2 \equiv G_2(\phi, X)\,, \tag{5.1}$$
$$\mathcal{L}_3 \equiv G_3(\phi, X)\Box\phi\,, \tag{5.2}$$
$$\mathcal{L}_4 \equiv G_4(\phi)R\,, \tag{5.3}$$

where now $\mathcal{L}_5$ can be set to zero. We now present the metric and scalar field equations that are obtained from varying $g_{\mu\nu}$ and ϕ in Eqs. (5.1) to (5.3). Although the structure of these equations is complicated the relevance for the application in Sec. 5.4 is simply the number of spacetime derivatives and powers of the scalar field that enter into each of the field equations. The metric field equation is given by [106, 146]

$$\Gamma R_{\mu\nu} = -\sum_{i=2}^{4} \mathcal{T}^{(i)}_{\mu\nu} + \left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right)/M_*^2 \tag{5.4}$$

and the scalar field equation is given by

$$\Gamma \sum_{i=2,3,4} (\nabla^\mu J^{(i)}_\mu - P^{(i)}_\phi) + \Xi \sum_{i=2}^{4} \mathcal{T}^{(i)} = -\frac{T}{M_*^2}\Xi\,, \tag{5.5}$$

where $\Gamma \equiv 2G_4/M_*^2$, $\Xi \equiv 2G_{4\phi}/M_*^2$ and

$$P^{(2)}_\phi = \frac{2}{M_*^2} G_{2\phi}\,, \tag{5.6}$$
$$P^{(3)}_\phi = \frac{2}{M_*^2} \nabla_\mu G_{3\phi} \nabla^\mu \phi\,, \tag{5.7}$$
$$P^{(4)}_\phi = \frac{2}{M_*^2} G_{4\phi} R\,, \tag{5.8}$$
$$J^{(2)}_\mu = -\, G_{2X} \nabla_\mu \phi\,, \tag{5.9}$$

$$\begin{aligned}
J_\mu^{(3)} =& - G_{3X}\Box\phi\nabla_\mu\phi + G_{3X}\nabla_\mu X + 2G_{3\phi}\nabla_\mu\phi\,, && (5.10)\\
\mathcal{T}_{\mu\nu}^{(2)} =& - \frac{1}{M_*^2}G_{2X}\nabla_\mu\phi\nabla_\nu\phi \\
& + \frac{1}{2M_*^2}g_{\mu\nu}\left(XG_{2X}+2G_2\right)\,, && (5.11)\\
\mathcal{T}_{\mu\nu}^{(3)} =& \frac{2}{M_*^2}G_{3X}\mathcal{S}^{(4,3)} + G_{3\phi}\nabla_\mu\phi\nabla_\nu\phi\,, && (5.12)\\
\mathcal{T}_{\mu\nu}^{(4)} =& G_{4\phi}\mathcal{S}^{(2,1)} + G_{4\phi\phi}\mathcal{S}^{(2,2)}\,. && (5.13)
\end{aligned}$$

Note that $J_\mu^{(4)} = 0$. The $\mathcal{S}^{(i,j)}$ notation indicates a term that contains i spacetime derivatives and j powers of the scalar field. As discussed in Sec. 5.4, knowledge of these quantities is sufficient to determine whether a given term will become dominant or sub-dominant in a screened or un-screened limit, not its precise functional form. We refer the reader to the appendix of Ref. [146] for the explicit expressions but note the different definitions of the G_i functions and X.

Under the assumption of luminal speed of gravity [136] we shall show in Sec. 5.3 that the unique nonlinear correction terms in the reconstructed theory are specified only by Eq.(3.22), which we reproduce here for convenience

$$\Delta G_{2,3} = \sum_{n>2}\xi_n^{(2,3)}(\phi)\left(1+\frac{X}{M_*^4}\right)^n\,, \tag{5.14}$$

where $\Delta G_{4,5} = 0$ and $\xi_n^{(i)}(\phi)$ are free functions of the scalar field, reflecting the large degree of freedom that exists on nonlinear scales without affecting linear scales. These terms arise from noting that in the unitary gauge with the foliation $\phi = tM_*^2$ the kinetic term of the scalar field becomes $X = (-1+\delta g^{00})M_*^4$. Eq. (5.14) is therefore an expansion in $(\delta g^{00})^n$.

The freedom in the correction term (5.14) may be exploited to endow the reconstructed theories with some desired nonlinear features without affecting linear theory. In particular, $\xi_n^{(i)}(\phi)$ can be designed to implement a screening mechanism (Sec. 5.4) or even to hide a kinetic self-acceleration effect of the cosmic background expansion to an arbitrary level of nonlinear perturbations (Sec. 5.5).

5.3 Uniqueness of the ΔG_i corrections

Due to the importance of the ΔG_i nonlinear correction terms for the applications of interest in Secs. 5.4, 5.5 and 5.6 we shall first investigate to what extent these terms are the unique corrections to the reconstructed Horndeski action in Eqs. (3.18) to (3.21).

Recall that the correction terms in (5.14) were inferred from the requirement that in covariant language $\delta g^{00} = 1 + X/M_*^4$. Successive powers of $1 + X/M_*^4$ therefore yield corrections that do not affect lower-order perturbations, in particular, the background or linear theory. However, there are of course other operators which can be added to the EFT which will not affect the background and linear dynamics such as δK^3 and $(\delta R^{(3)})^3$. In principle a term such as δK^3 could be added to the EFT action, which would affect the dynamics of the second-order perturbations. Note however that for the same reason that δK^2 only appears in combination with $\delta K_{\mu\nu}\delta K^{\mu\nu}$ after $\mathcal{L}_4$ is written in the unitary gauge and expanded in the perturbations, it is not possible to simply add δK^3 as there are no terms in the Horndeski action that give rise to this term alone. More specifically, on the cosmological background $K_{\mu\nu} = H h_{\mu\nu}$, the perturbation $\delta K = K - 3H$ must appear in the combination

$$K^3 - 3K K_{\mu\nu} K^{\mu\nu} + 2K_{\mu\nu} K^{\mu\sigma} K^{\nu}{}_{\sigma} \,, \tag{5.15}$$

which gives rise to a number of nonlinear operators in the EFT action involving $\delta K_{\mu\nu}$ [177, 178]. The only term in the Horndeski action that gives rise to such a combination is in $\mathcal{L}_5$. Following the spirit of EFT one may add these nonlinear operators because they are consistent with the symmetries that we have imposed, but the theory which is underlying such a combination generally violates the luminal speed of gravity constraint [146] such that we will omit these terms. By use of the Gauss-Codazzi relation

$$R^{(3)} = R - K_{\mu\nu} K^{\mu\nu} + K^2 - 2\nabla_\nu \left(n^\nu \nabla_\mu n^\mu - n^\mu \nabla_\mu n^\nu \right) \,, \tag{5.16}$$

relating the 3-dimensional Ricci scalar $R^{(3)}$ to the 4-dimensional Ricci scalar R and $K_{\mu\nu}$, one can furthermore see that adding on higher powers of $R^{(3)}$ to the EFT in a similar manner will inevitably introduce higher powers of δK, and the previous argument applies. The same logic also requires ΔG_4 and ΔG_5 to vanish and the nonlinear freedom is now completely specified by Eq. (5.14).

An alternative perspective on this argument is to consider a covariant form of the extrinsic curvature tensor, or for simplicity its trace

$$K = -\nabla_\mu \left(\frac{\partial_\mu \phi}{\sqrt{-X}} \right) . \tag{5.17}$$

By expressing the denominator in terms of the metric perturbations, Taylor expanding and performing the replacement of δg^{00} with $1 + X/M_*^2$, one obtains in schematic form

$$K = \Box\phi + F(X, \nabla_\mu \phi, \nabla_\mu X) , \tag{5.18}$$

where $F(X, \nabla_\mu \phi, \nabla_\mu X)$ is some complicated function of the scalar field and derivatives of the scalar field obtained after the expansion, the precise form of which is not relevant to the discussion. Taking higher powers of δK and making use of Eq. (5.18) will lead to terms such as $(\Box\phi)^n$. Such expressions belong either to Horndeski models with non-luminal speed of gravitational waves or beyond-Horndeski theories. Reversing the logic, it is necessary to start from such a model in order to obtain a nonlinear correction involving a higher power of δK. Therefore, any correction terms to the EFT of dark energy that make use of the operators $(\delta K)^n$ with $n \geq 2$ and $R^{(3)}$ will reconstruct a theory that has a non-vanishing G_{4X} or G_5 term or a beyond-Horndeski model.

For Horndeski models with luminal speed of gravity, the only nonlinear operators that appear at n-th order are therefore

$$\left(\delta g^{00}\right)^n \ , \ \left(\delta g^{00}\right)^{n-1} \delta K . \tag{5.19}$$

which adds two new independent EFT functions per order in the perturbations. More explicitly, the n-th order contribution to the EFT action with $n \geq 3$ is given by

$$\delta \mathcal{S}^{(n)} = \int d^4x \sqrt{-g} \sum_{i=3}^{n} \left[\bar{M}_i^4(t) \left(\delta g^{00}\right)^i + \bar{M}_i^3(t) \left(\delta g^{00}\right)^{i-1} \delta K \right] , \tag{5.20}$$

where each $\bar{M}_i^3(t)$ and $\bar{M}_i^4(t)$ are the two free functions that contribute at i-th order in the action. This is a logical extension to n-th order of the first two operators which appear in $\mathcal{S}^{(2)}$ in Eq. (3.15), namely $(\delta g^{00})^2$ and $\delta g^{00} \delta K$.

5.4 Nonlinear freedom for screening

As a first application of the free nonlinear correction term in Eq. (5.14) in the reconstructed scalar-tensor action we shall consider the realization of screening mechanisms that are required to recover GR in the well-tested Solar-System regime [179]. For this purpose, we shall employ the scaling method of Refs. [146, 181] (also see applications in Refs. [138, 182, 183]) that allows an efficient identification of the existence of Einstein gravity regimes for a particular choice of Horndeski functions. We briefly review the method (Sec. 5.4.1) and then apply it for a characterization of the nonlinear correction terms ΔG_i that realize screening by large gravitational potentials $\Phi_N > \Lambda$ for some threshold Λ (Sec. 5.4.2), large first derivatives $\nabla\Phi_N > \Lambda$ (Sec. 5.4.3) or large second derivatives $\nabla^2\Phi_N > \Lambda$ (Sec. 5.4.4) [71].

5.4.1 Scaling method

The scaling method was developed in Refs. [146, 181] to efficiently determine whether a given Horndeski theory possesses an Einstein gravity limit. It proceeds as follows. At the level of the field equations the scalar field ϕ is expanded in terms of a field perturbation ψ as

$$\phi = \phi_0 \left(1 + \alpha^q \psi\right) , \tag{5.21}$$

where ϕ_0 denotes the background value and α is the theoretical parameter relevant to the expansion. For example, it could be the speed of light or the coupling of a Galileon interaction term. After performing this expansion, the scalar field equation of the Horndeski model (see Eq. (5.5)) takes the generic form

$$\alpha^{s+mq} F_1(\psi, \tilde{X}) + \alpha^{t+nq} F_2(\psi, \tilde{X}) = \frac{T}{M_*^2} , \tag{5.22}$$

where $s, m, t, n \in \mathbb{N}$ and $\tilde{X} = \partial_\mu\psi\partial^\mu\psi$. Now consider the limit of $\alpha \to \infty$ or $\alpha \to 0$. As the right-hand side of Eq. (5.22) is independent of α the leading-order term on the left-hand side must also be independent of α to balance the equation. This restricts the possible values of the exponent q. Therefore there must be at least one term which scales as α^0 with every other term involving non-zero powers of α vanishing in the $\alpha \to 0$ or $\alpha \to \infty$ limit. For example, choosing $q = -s/m$

and taking the $\alpha \to \infty$ limit requires $-s/m < -t/q$, so that $t + n(-s/m) < 0$ and the dominating terms in the field equation become

$$F_1(\psi, \tilde{X}) = \frac{T}{M_*^2} \,. \tag{5.23}$$

If in a given α limit the metric field equations reduce to the Einstein equations after performing the expansion (5.21), then the corresponding scalar field equation applies to the screened limit where the fifth force is suppressed. To ensure consistency the value of q chosen to obtain a screened limit must be the same in both the scalar and metric field equations. Note that there may also be terms that involve powers of α that do not depend on q. Depending on whether they are raised to a positive or a negative power they will diverge or vanish in either limit of α. If they vanish then this is not an issue, but if they diverge extra care must be taken. For example, it may be important to use the freedom in the ΔG_i terms to remove any divergences which arise in either limit.

In the following we present the recovery of three distinct screening mechanisms by suitable choices of ΔG_i. Drawing on the distinction discussed in Ref. [71] this will encompass the known screening mechanisms: (i) by large gravitational potentials $\Phi_N > \Lambda$ for some threshold Λ (Sec. 5.4.2), (ii) by large first derivatives $\nabla\Phi_N > \Lambda$ (Sec. 5.4.3) and (iii) by large second derivatives $\nabla^2\Phi_N > \Lambda$ (Sec. 5.4.4). We shall find that there is more than sufficient freedom in the nonlinear sector to, in principle, endow the reconstructed theory with a particular screening mechanism regardless of the constraints of the background and the linear perturbations. Importantly, however, while this generally implies the existence of Einstein gravity limits in the deeply nonlinear regime, this does not guarantee that a given observed region is nonlinear enough for the screening mechanism to be activated. The numerical value of the screening scale needs to be computed separately and ultimately decides whether a theory is compatible with stringent Solar-System tests. It is not surprising that screening mechanisms can be added to linearly reconstructed models as they are inherently nonlinear effects. It is however important to verify this explicitly.

5.4.2 Large field value screening

As a first example we consider the implementation of a screening effect by large field values $\Phi_N > \Lambda$. More specifically, we will focus on the Chameleon

Mechanism [88, 89]. We shall first cast the reconstructed theory into the Brans-Dicke representation with $F(\phi) = \phi/M_*$ (see Sec. 4.3). With this choice we have that $\Gamma = \phi/M_*$ and $\Xi = 1$ in Eqs. (5.4) and (5.5). By making use of the freedom in ΔG_i it is possible to add a term to G_2 that sets the q-value to be arbitrarily positive or negative. To see this let us begin with the full reconstructed Horndeski action in Eqs. (3.18) to (3.20) with a ΔG_2 term that takes the form

$$\Delta G_2 = \xi(\phi) \left(1 + \frac{X}{M_*^4}\right)^n , \tag{5.24}$$

with $n \geq 3$ and $\xi(\phi)$ given by

$$\xi(\phi) = M_*^2 U(\phi) - \frac{\lambda^{-N}}{2} \left(\phi - \phi_{min}\right)^k , \tag{5.25}$$

where λ is a coupling parameter, N and k are both positive integers, $U(\phi)$ is the reconstructed potential in Eq. (3.18) and ϕ_{min} denotes the minimum value of the second contribution to the potential in Eq. (5.25). No other ΔG_i terms are necessary as they all contain derivative terms which vanish in the screened limit. We shall take the scaling parameter α to be the coupling λ.

This choice cancels the potential obtained from the linear reconstruction and replaces it with a power-law potential that takes a similar form to the chameleon screening example in Ref. [146, 184] but with $\alpha \to \alpha^{-N}$. It is with a suitable choice of N that no derivative terms contribute in the screening limit. In this limit we then obtain the Einstein equation

$$\frac{\phi}{M_*} R_{\mu\nu} = -\mathcal{T}^{(2)}_{\mu\nu} + \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\right) / M_*^2 + H_m \left[\nabla_\mu \phi\right] , \tag{5.26}$$

where $\mathcal{T}^{(2)}_{\mu\nu}$ is defined in Eq. (5.11) and $H_m \left[\nabla_\mu \phi\right]$ represents all the terms that involve derivatives of ϕ in the metric field equation, the precise form of which is not relevant as we shall find that they disappear in the $\alpha \to 0$ limit of interest. Taking the trace of Eq. (5.26) leads to $\phi R/M_* = -\mathcal{T}^{(2)}$ which, noting that $\mathcal{T}^{(2)} = 2G_2/M_*^2$, gives a relation between R and G_2. The scalar field equation is given by

$$-\frac{2\phi}{M_*^2} \left(G_{2\phi} + G_{4\phi} R\right) + \mathcal{T}^{(2)} + H_s \left[\nabla_\mu \phi\right] = -T/M_*^2 , \tag{5.27}$$

where $H_s \left[\nabla_\mu \phi\right]$ represents all the terms in the scalar field equation involving derivatives of ϕ which will disappear in the $\alpha \to 0$ limit. With the choice of ΔG_2 in Eq. (5.24) there is no contribution from the reconstructed potential $U(\phi)$ to the scalar field equation. After eliminating R and $\mathcal{T}^{(2)}$ in favour of G_2 the scalar field equation becomes

$$\alpha^{-N} \left(\phi - \phi_{min}\right)^{k-1} \left[\phi k - 2 \left(\phi - \phi_{min}\right)\right] + H_s \left[\nabla_\mu \phi\right] = -T \,. \tag{5.28}$$

Applying the scaling method with the scalar field now expanded in terms of ψ as in Eq. (5.21), we examine the set of q values which leave non-vanishing terms on the left-hand side of Eq. (5.28) in the $\alpha \to 0$ limit. As $\alpha \to 0$ it is necessary to take the largest q value from this set after the scaling in Eq. (5.28). Disregarding the derivative terms in $H_s \left[\nabla_\mu \phi\right]$, we find that q takes one of two possible values

$$q \in \left\{ \frac{N}{k-1}, \frac{N}{k} \right\} . \tag{5.29}$$

We must take $q = N/(k-1)$ as it is the largest in the set of q values from G_2. The integer N can then be chosen in Eq. (5.25) to be arbitrarily large. In the limit of $\alpha \to 0$ this will send all terms involving spacetime derivatives of ϕ to zero, justifying the original choice of $\xi(\phi)$. This is important as in principle the value of n in Eq. (5.24) is only bounded from below by the requirement that it is a nonlinear correction. All the terms involving derivatives of the scalar field scale as $X^m = \phi_0^{2m} \alpha^{2mN/(k-1)} \tilde{X} \to 0$ as $\alpha \to 0$ with $m = \{1, \ldots, n\}$.

Now we expand the scalar field around the minimum of the potential such that $\phi_{min} \approx \phi_0$. This then implies that $\phi - \phi_0 = \phi_0 \alpha^q \psi$. The remaining terms in the scalar field equation for $\alpha \to 0$ relate the local value of the scalar field to the matter density as

$$\psi = \left(\frac{-T}{\phi_0^k k} \right)^{\frac{1}{k-1}} , \tag{5.30}$$

which recovers the chameleon screening effect for $k < 1$. The metric field equation in the same limit reduces to

$$\phi_0 R_{\mu\nu} = \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) / M_* \,, \tag{5.31}$$

recovering the standard Einstein equation with a re-scaled Planck mass set by the background field value ϕ_0. Therefore we have implemented a Chameleon Mechanism in a scalar-tensor action that is reconstructed from an arbitrary cosmological background evolution and linear perturbations by adding a suitable choice of ΔG_2. Whether the screening effect operates in the Solar System to comply with stringent local tests of gravity needs to be checked numerically for a given reconstructed model.

5.4.3 First-derivative screening

Next we examine the implementation of a screening effect that operates through large first derivatives $\nabla\Phi_N > \Lambda$. More specifically, we focus on the k-mouflage screening effect [90, 185]. We may simply choose here the scaling parameter α to be the kineticity function α_K and take the $\alpha \to \infty$ limit. EFT functions such as α_K are typically parameterised as $\alpha_{K0} f(a)$ where $f(a)$ is some function of the scale factor with $f(a = 1) \equiv 1$. Often this is simply a power of the scale factor or the evolution of the dark energy density normalised to the present value $\Omega_{DE}(a)/\Omega_{DE0}$. This ensures that the effects of the modifications only become relevant at late times. We shall take here the scaling parameter to correspond to the value of α_K today $\alpha = \alpha_{K0}$. It is also possible to take α_{B0} or α_{M0} as the scaling parameter but as the reconstruction depends differently on these EFT parameters this will lead to different behaviour in the screened limit (see Sec. 5.4.4). Taking α to be α_{K0}, we see that as the reconstructed action is linear in the EFT functions we have from Table 3.1 that each term scales as $U(\phi) \sim \alpha$, $Z(\phi) \sim \alpha$, $a_2(\phi) \sim \alpha$ and $b_1(\phi) \sim \alpha^0$, which follows from the fact that $\bar{M}_1^3$ is independent of α_K (see Table 3.2 for the full set of relations between the EFT coefficients of the different bases). With this choice we have that the terms in G_2 will scale as α^{1+nq} for some integer n but those in G_3 will scale as α^{nq}.

In order to obtain an Einstein field equation it is necessary to remove the potential to avoid divergences in the $\alpha \to \infty$ limit. This also makes physical sense as the screening mechanism in this case operates via the kinetic terms. We shall also remove all of the dependence on the canonical kinetic term linear in X to ensure that the screening operates through higher powers of X. To this end, we choose $\Delta G_2 = \Delta G_2^{(1)} + \Delta G_2^{(2)}$, where

$$\Delta G_2^{(1)} = \frac{1}{2} M_*^6 Z(\phi) \left(1 + \frac{X}{M_*^4}\right)^4 - \frac{1}{2} M_*^6 Z(\phi) \left(1 + \frac{X}{M_*^4}\right)^3 , \tag{5.32}$$

$$\Delta G_2^{(2)} = 2M_*^2 U(\phi) \left(1 + \frac{X}{M_*^4}\right)^3 - M_*^2 U(\phi) \left(1 + \frac{X}{M_*^4}\right)^6 . \tag{5.33}$$

These nonlinear corrections ensure that every term in G_2 is now at least proportional to X^2 or greater. With this choice the relevant term in the scalar field equation is

$$\nabla^\mu J_\mu^{(2)} = -G_{2XX}\nabla^\mu X \nabla_\mu \phi - XG_{2X\phi}\,, \tag{5.34}$$

where $J_\mu^{(2)}$ is defined in Eq. (5.9). The first term on the right-hand side in Eq. (5.34) scales as α^{1+3q}, which sets the minimum q-value to be $q = -1/3$. As every term in G_3 scales as α^{nq} with $n > 0$ this will send every term involving G_3 to zero in the $\alpha \to \infty$ limit. This particular q-value will also ensure that $\mathcal{T}_{\mu\nu}^{(i)} \to 0$ as $\alpha \to \infty$ so that the metric field equation reduces to the standard Einstein field equation. The resulting scalar field equation corresponds to a k-mouflage model

$$\xi(\phi)\partial^\mu X \partial_\mu \phi = -\frac{T}{M_*^2}\,, \tag{5.35}$$

with

$$\xi(\phi) = a_2(\phi) + \frac{9Z(\phi)}{2M_*^2} - \frac{9U(\phi)}{M_*^8}\,. \tag{5.36}$$

5.4.4 Second-derivative screening

Finally, we consider the implementation of screening through large second derivatives $\nabla^2\Phi_N > \Lambda$, more explicitly the realization of the Vainshtein mechanism in the $\alpha \to \infty$ limit where the scaling parameter α is taken to be α_B only. The procedure is similar to Sec. 5.4.3. In this case $U(\phi) \sim \alpha$, $Z(\phi) \sim \alpha$, $a_2(\phi) \sim \alpha$ as before, but in contrast to Sec. 5.4.3, $b_1(\phi) \sim \alpha$, which follows from the fact that $\bar{M}_1^3 \propto \alpha_B$. We begin by adding on the nonlinear counterterms in Eqs. (5.32) and (5.33) to ensure the X dependence of G_2 is at least X^2.

It turns out that the important term in the scalar field equation which gives rise to a non-trivial equation of motion and Vainshtein screening is $\nabla^\mu J_\mu^{(3)}$ where $J_\mu^{(3)}$ is given in Eq. (5.10). Plugging in the expression in Eq. (3.19) we have that

$$\nabla^\mu J_\mu^{(3)} = b_1(\phi)\mathcal{S}^{(4,2)} + H_s\left[\nabla_\mu\phi\right]\,, \tag{5.37}$$

where again $H_s\left[\nabla_\mu\phi\right]$ represents all of the terms involving derivatives of ϕ that will vanish in the $\alpha \to \infty$ limit. Furthermore $S^{(4,2)}$ is a term that involves four derivative operators and two powers of the scalar field, which is given explicitly by

$$\mathcal{S}^{(4,2)} = \left(\Box\phi\right)^2 + \partial_\mu\phi\partial^\mu\Box\phi + \Box X\,. \tag{5.38}$$

These terms each scale as α^{1+2q} requiring a q-value of $-1/2$ to ensure independence of α on the left-hand side. As we have also ensured that G_2 starts at least at X^2, scaling as α^{4q} with $q=-1/2$, these higher-derivative terms will disappear in the $\alpha\to\infty$ limit. The scalar field equation in this limit then becomes

$$\frac{\phi_0^3}{M_*}b_1(\phi_0)\left[(\Box\psi)^2+\partial_\mu\psi\partial^\mu\Box\psi+\Box\tilde{X}\right]=-\frac{T}{M_*^2}\,, \tag{5.39}$$

where $\tilde{X}\equiv\partial_\mu\psi\partial^\mu\psi$. This is a typical scalar field equation involving higher derivatives of ψ expected for Vainshtein screening. It is necessary to ensure that the standard Einstein equation is obtained in the same limit in the metric field equations so that we can be sure this is the screened limit.

Having already set $q=-1/2$ from the scalar field equation and ensured that G_2 starts at X^2 with $\Delta G_2^{(1)}$ and $\Delta G_2^{(2)}$, every term $\mathcal{T}_{\mu\nu}^{(i)}$ in the metric field equation (5.4) vanishes in the $\alpha\to\infty$ limit. For example the first term in $\mathcal{T}_{\mu\nu}^{(4)}$ scales as

$$G_{4\phi}\mathcal{S}^{(2,1)}\sim\alpha^{2q}\sim\alpha^{-1}\to 0\,, \tag{5.40}$$

and the first one in $\mathcal{T}_{\mu\nu}^{(3)}$ scales as

$$\frac{2}{M_*^2}G_{3X}\mathcal{S}^{(4,3)}\sim\alpha^{1+3q}\sim\alpha^{-\frac{1}{2}}\to 0\,. \tag{5.41}$$

With the choice of the Brans-Dicke representation of $F(\phi)=\phi/M_*$ we have that $\Gamma=\phi_0/M_*$ and $\Xi=1$, and the metric field equation reduces to Eq. (5.31).

To summarize, by choosing α_{B0} as the scaling parameter and removing the constant and linear terms in X from G_2 one can obtain the standard Einstein field equation with a re-scaled Planck mass and a scalar field equation involving second derivatives in ψ as expected in the case of Vainshtein screening.

5.5 Nonlinear freedom for degenerate kinetic self-acceleration

As a further application of the nonlinear freedom in reconstructed scalar-tensor theories, we demonstrate how the correction term in Eq. (5.14) can be configured to construct scalar-tensor theories that are degenerate with standard cosmology to an arbitrary level of cosmological perturbations (Sec. 5.5.1). As a particular interesting example we show how this allows for models that accelerate the Universe without a cosmological constant yet remain dynamically degenerate with ΛCDM through a suitable configuration of the kinetic terms (Sec. 5.5.2).

5.5.1 Perturbative degeneracy with ΛCDM

An important implication of Eq. (5.14) is that it is possible to use the ΔG_i terms to write down a Horndeski theory that possesses a highly non-trivial form for the nonlinear perturbations yet reduces to ΛCDM on the background, where the correction terms vanish. This degeneracy may even be extended to an arbitrary level of perturbations. The existence of such classes of theories is a natural consequence of the reconstruction being an expansion in $(1 + X/M_*^4)^n$ with $n \in \mathbb{N}$. One can therefore construct theories whose physical effects only become relevant at a particular level of higher-order perturbations characterized by the power n.

To see how this works in practice let us choose, for example,

$$G_2 = -M_*^2\Lambda + \xi_n^{(2)}(\phi)\left(1 + \frac{X}{M_*^4}\right)^n , \tag{5.42}$$

with $G_3 = 0$, $G_4 = M_*^2/2$ and $n \geq 3$. After performing an ADM decomposition with $\phi = tM_*^2$ the second term in Eq. (5.42) becomes $\xi_n^{(2)}(t)\,(\delta g^{00})^n$. On the background and linear scales therefore there will be no effects arising from the non-canonical kinetic terms and it will appear to be exactly ΛCDM. Note that this argument does not rely on the specific foliation adopted as we shall verify shortly for a specific example, but for now simply note that any non-zero perturbations that arise from another choice of foliation must be pure gauge. At the nonlinear level Eq. (5.42) departs from ΛCDM and we have discussed the mapping of the $\xi_n^{(2)}(t)$ functions onto nonlinear EFT functions in Sec. 5.6. It is also possible to write a theory with $G_2 = \Lambda$ and

$$G_3 = \xi_n^{(3)}(\phi)\left(1 + \frac{X}{M_*^4}\right)^n . \tag{5.43}$$

In an equivalent manner this corresponds to a Galileon theory that can only be distinguished from ΛCDM on nonlinear scales. Combinations of ΔG_2 and ΔG_3 can also be used to construct more non-trivial theories.

For clarity we shall provide an explicit example of this degeneracy and compute the background equations of motion and check that the expansion is indeed matching that of ΛCDM. A more detailed analysis, including the investigation of possible instabilities and perturbative effects, will be the subject of further analysis. For simplicity, we shall only focus here on the degeneracy at the level of the background and not for the linear perturbations. Hence, we take $n = 2$ in Eq. (5.42) so that

$$\begin{aligned} G_2 &= -M_*^2\Lambda + \xi(\phi)\left(1 + \frac{X}{M_*^4}\right)^2 , \\ &= -M_*^2\Lambda + \xi(\phi) + 2\xi(\phi)X/M_*^4 + \xi(\phi)X^2/M_*^8 , \end{aligned} \tag{5.44}$$

where $\xi(\phi)$ is a free function of ϕ. Not making any assumptions about the space-like foliation we now put this equation into the unitary gauge by setting the scalar field to be just a function of time. With $X = (-1 + \delta g^{00})\,\dot{\phi}^2$ we have at linear order

$$\begin{aligned} G_2 = &- M_*^2\Lambda + \xi(t) + \frac{2\xi(t)X_0}{M_*^4} + \frac{\xi(t)X_0^2}{M_*^8} \\ &- \left[\frac{2\xi(t)X_0}{M_*^4} + \frac{2\xi(t)X_0^2}{M_*^8}\right]\delta g^{00} , \end{aligned} \tag{5.45}$$

where X_0 is the value that X takes on the background, i.e., $X_0 = -\dot{\phi}^2$. This gives an explicit expression for the EFT functions $\Lambda(t)$ and $\Gamma(t)$ in the unitary gauge expansion of G_2 in Eq. (5.45), where first line corresponds to $-M_*^2\Lambda(t)$ and the second line to $-M_*^2\Gamma(t)/2$. Recall that the Friedmann equations in the EFT formulation are given by [7, 9, 113]

$$\Gamma(a) \quad +\Lambda(a) = 3H^2 - \frac{\rho_m}{M_*^2} , \tag{5.46}$$

$$\Lambda(a) \quad = 2HH' + 3H^2 , \tag{5.47}$$

where we have set the non-minimal coupling parameter $\Omega = 1$, we parameterise the time t in terms of the scale factor a. With the expressions for $\Gamma(a)$ and $\Lambda(a)$ obtained from Eq. (5.45) one can take linear combinations of the Friedmann equations (5.46) and (5.47) to eliminate the dependence on the background expansion H and obtain a field equation for the background value of the scalar field. This is determined from the resulting expression

$$\Gamma(a) + \frac{1}{3}\left[\Gamma(a) + \Lambda(a)\right]' = 0 \tag{5.48}$$

to be

$$\left[\frac{4\xi(a)X_0}{M_*^{10}} + \frac{\xi'(a)}{M_*^{10}}\left(X_0 - M_*^4/3\right)\right]\left(X_0 + M_*^4\right) \tag{5.49}$$

$$+ X_0'\left[\frac{2\xi(a)}{3M_*^6} + \frac{2\xi(a)X_0}{M_*^{10}}\right] = 0\,, \tag{5.50}$$

which is the non-trivial Klein-Gordan scalar field equation. It has a trivial solution $X_0 = -M_*^4$. More complicated solutions to the background scalar field equation will be explored in the future. From $X_0 = -M_*^4$, one immediately recognizes in Eq. (5.44) that $G_2(X_0) = -\Lambda$, and hence the recovery of the ΛCDM background expansion. Alternatively, once the solution to the background evolution of the scalar field has been obtained it is possible to derive the equation-of-state parameter for the resulting k-essence model given by [70]

$$w(a) = \frac{-M_*^2\Lambda + \xi(\phi)\left(1 + X/M_*^4\right)^2}{M_*^2\Lambda - \xi(\phi)\left(1 + X/M_*^4\right)\left(1 - 3X/M_*^4\right)}\,. \tag{5.51}$$

After inserting the background solution $X = X_0 = -M_*^4$ one obtains $w = -1$, confirming that the background expansion is indeed matching that of ΛCDM.

5.5.2 Degenerate kinetic self-acceleration

To highlight the implications of the perturbative degeneracy, we will now study a particularly interesting example of Eq. (5.42). Let us consider a class of models specified by $\xi(\phi) = M_*^2\Lambda_\phi$ in Eq. (5.44). The subscript ϕ indicates that Λ_ϕ is a coupling parameter in the higher-order kinetic terms of the scalar field ϕ. Eq. (5.44) then becomes

$$G_2 = -M_*^2\Lambda_{GR} + M_*^2\Lambda_\phi\left(1 + \frac{X}{M_*^4}\right)^2\,, \tag{5.52}$$

where we defined $\Lambda \equiv \Lambda_{GR}$. We also set $G_4 = 1$ and $G_3 = 0$ and stress that any contributions to Λ_{GR} from quantum corrections of matter fields in this discussion are neglected. If we now set $\Lambda_\phi = \Lambda_{GR}$ this model exhibits the particular feature of having no explicit cosmological constant. The model is now simply

$$G_2 = 2\Lambda_\phi X/M_*^2 + \Lambda_\phi X^2/M_*^6\,. \tag{5.53}$$

However, the observed cosmological constant Λ_{obs} in the cosmological background of this model remains $\Lambda_{obs} = \Lambda_\phi = \Lambda_{GR}$. An alternative approach is to start with the model

$$G_2 = 2\Lambda_\phi X/M_*^2 + \Lambda_\phi X^2/M_*^6 - 2M_*^2\Lambda_{GR}\,, \qquad (5.54)$$

and then set $\Lambda_{GR} = 0$. In summary, in one interpretation the coupling Λ_ϕ is tuned to match a non-vanishing Λ_{GR} that corresponds to the observed Λ_{obs} or $\Lambda_{GR} = 0$ and $\Lambda_\phi = \Lambda_{obs}$.

With either interpretation these models generate a *kinetic self-acceleration* effect that is degenerate with the cosmological constant to the $(n-1)$-th order of cosmological perturbations. While this may certainly be viewed as an engineered self-acceleration effect, it also raises more general questions about the genuineness of a kinetic self-acceleration that resembles a cosmological constant for observational compatibility. We note that a similar expansion to Eq. (5.52) can be performed for G_3 with similar implications. For instance, one may consider a kinetic gravity braiding model with nontrivial G_2 and G_3. By combining power series of $(1 + X/M_*^4)^n$ in G_2 and G_3 that only contribute at $(n-1)$-th order in cosmological perturbations, one can choose the coefficients of G_2 and G_3 in an expansion in X to cancel off to just leave a term X^n in G_2 and G_3 for arbitrarily large n. Greater values of n then correspond to models which are more difficult to distinguish from ΛCDM and for which nonlinear data must be used for their discrimination. This may shed some light on the results of Ref. [186], where better agreement with ΛCDM at the linear level was likewise found for kinetic gravity braiding models with $G_3 \propto X^n$ for large n but adopting a canonical G_2 instead, which is not feasible with using ΔG_i corrections only.

It is worth noting however that a further interesting consequence of Λ_{obs} being interpreted as a coupling rather than a bare constant is that it may be possible to render the acceleration effect in Eq. (5.52) technically natural as it can now enter as a coefficient to an irrelevant operator rather than as a non-renormalizable constant [187, 188]. At a more practical level, we emphasise that these models have the interesting property that discriminatory effects of this type of cosmic acceleration are left exclusively to the nonlinear observational regime.

5.6 Higher-order reconstruction

With the higher-order EFT expansion in Eq. (5.20) and the freedom in the nonlinear sector having been significantly reduced by the restriction to a luminal speed of gravity, it becomes straightforward to perform a n-th order reconstruction of the corresponding class of Horndeski theories by fixing the ΔG_i functions order-by-order in terms of the nonlinear EFT functions $\bar{M}_i^{3,4}$. We shall now see how this extra information modifies the reconstruction from the background and linear scales by adding in the new free functions and slightly changing the dependence on the linear EFT functions. We shall elaborate on this explicitly for the case of $i = 3$ before outlining the general n-th order case.

Let us begin by noting that in the unitary gauge a term that takes the form $\xi(\phi)X^m\Box\phi$ becomes

$$\begin{aligned}\xi(\phi)X^m\Box\phi = &\mp \frac{2m}{2m+1}\xi(\phi)(-X)^{m+\frac{1}{2}}K \\ &\pm \frac{1}{2m+1}\xi'(\phi)(-X)^{m+1}\,,\end{aligned} \tag{5.55}$$

where the sign difference on the top and bottom indicate even or odd m respectively and the prime denotes a derivative with respect to ϕ. After expanding Eq. (5.55) in the unitary gauge there will be several terms that contribute and that can be mapped onto the operators in Eq. (5.20).

We shall proceed along the same lines as chapter 3 to obtain a corresponding covariant action. To begin, by using the replacement $\delta g^{00} = 1 + X/M_*^2$ the $(\delta g^{00})^3$ operator becomes

$$\bar{M}_3^4(t)\left(\delta g^{00}\right)^3 = \bar{M}_3^4(\phi)\left(1 + \frac{3X}{M_*^4} + \frac{3X^2}{M_*^8} + \frac{X^3}{M_*^{12}}\right)\,. \tag{5.56}$$

This contributes to $U(\phi)$, $Z(\phi)$, $a_2(\phi)$ along with a new, now necessarily non-vanishing contribution to the coefficient of X^3 that we call $a_3(\phi)$. Let us now derive the covariant action which gives rise to the following expansion in the unitary gauge

$$\bar{M}_1^3(t)\delta g^{00}\delta K + \bar{M}_3^3(t)(\delta g^{00})^2\delta K\,. \tag{5.57}$$

We shall take the case of $m = 1, 2$ in Eq. (5.55) for simplicity and begin with the combination

$$G_3 = b_1(\phi)X\Box\phi + b_2(\phi)X^2\Box\phi + \Delta G_3^{(4)}\,, \tag{5.58}$$

$U(\phi) = \Lambda + \frac{\Gamma}{2} - \frac{M_2^4}{2M_*^2} - \frac{3H\bar{M}_1^3}{2M_*^2} + \frac{3H\bar{M}_3^3}{M_*^2} - \frac{3(\bar{M}_1^3)'}{20} + \frac{(\bar{M}_3^3)'}{5} - \frac{\bar{M}_3^4}{M_*^2}$	
$Z(\phi) = \frac{\Gamma}{M_*^4} - \frac{2M_2^4}{M_*^6} - \frac{3H\bar{M}_1^3}{M_*^6} + \frac{12H\bar{M}_3^3}{M_*^6} + \frac{3(\bar{M}_1^3)'}{5M_*^4} - \frac{4(\bar{M}_3^3)'}{5M_*^4} - \frac{6\bar{M}_3^4}{M_*^6}$	
$a_2(\phi) = \frac{M_2^4}{2M_*^8} - \frac{3H\bar{M}_3^3}{M_*^8} + \frac{(\bar{M}_1^3)'}{5M_*^6} - \frac{3(\bar{M}_3^3)'}{5M_*^6} + \frac{3\bar{M}_3^4}{M_*^8}$	
$a_3(\phi) = \frac{(\bar{M}_1^3)'}{40M_*^{10}} - \frac{(\bar{M}_3^3)'}{5M_*^{10}} + \frac{\bar{M}_3^4}{M_*^{12}}$	$b_1(\phi) = \frac{3\bar{M}_1^3}{4M_*^6} - \frac{2\bar{M}_3^3}{M_*^6}$
$F(\phi) = \Omega$	$b_2(\phi) = \frac{\bar{M}_1^3}{8M_*^{10}} - \frac{\bar{M}_3^3}{M_*^{10}}$

Table 5.1 *Contributions to the reconstructed Horndeski action arising from the nonlinear corrections in the EFT action at third order. The reconstruction can easily be expanded to arbitrary higher order.*

where $\Delta G_i^{(4)}$ indicates that the nonlinear corrections now start at fourth order. We transform Eq. (5.58) into the unitary gauge and then solve for $b_1(\phi)$ and $b_2(\phi)$ in terms of the EFT functions. It is necessary to have two independent functions in the covariant expansion as there are two independent EFT functions. At third order in the perturbations we obtain

$$G_3 \supset -b_1(\phi)M_*^6\delta g^{00}\delta K + \frac{1}{4}b_1(\phi)M_*^6(\delta g^{00})^2\delta K \tag{5.59}$$

$$+ 2b_2(\phi)M_*^{10}\delta g^{00}\delta K - \frac{3}{2}b_2(\phi)M_*^{10}(\delta g^{00})^2\delta K\,, \tag{5.60}$$

where for the sake of clarity we have not shown the terms which are independent of δK. We then require that

$$-b_1(\phi)M_*^6 + 2b_2(\phi)M_*^{10} = \bar{M}_1^3(\phi)\,, \tag{5.61}$$

$$b_1(\phi)M_*^6 - 6b_2(\phi)M_*^{10} = 4\bar{M}_3^3(\phi)\,. \tag{5.62}$$

This system of equations can be straightforwardly solved to obtain $b_1(\phi)$ and $b_2(\phi)$. The results are shown in Table 5.1 along with the contributions to G_2.

Importantly, this method can straightforwardly be extended to higher orders, where at each order it is necessary to invert an $n \times n$ matrix to obtain the corresponding EFT coefficients in terms of covariant functions in G_3. It is then possible to derive a reconstruction from the $\bar{M}_i^4, \bar{M}_i^3$ terms which proceeds in exactly the same manner as discussed for $n = 3$. It is also important to stress that a different combination of the terms in Eq. (5.55) with different choices of m could have been chosen to develop the reconstruction. From the structure of Eq. (5.55) there will always be terms involving $(\delta g^{00})^n \delta K$ to arbitrary order for any m which can be used as the basis for deriving the reconstructed theory. There is therefore a degeneracy in the space of models which go as $X^m \Box\phi$ on the behaviour of the background and perturbations.

The reconstructed Horndeski theory that covers the background, linear- and second-order cosmological perturbations is given by

$$\begin{aligned} G_2(\phi, X) =& - M_*^2 U(\phi) - \frac{1}{2} M_*^2 Z(\phi) X + a_2(\phi) X^2 \\ &+ a_3(\phi) X^3 + \Delta G_2 \,, \end{aligned} \tag{5.63}$$

$$G_3(\phi, X) = b_0(\phi) + b_1(\phi) X + b_2(\phi) X^2 + \Delta G_3 \,, \tag{5.64}$$

$$G_4(\phi, X) = \frac{1}{2} M_*^2 F(\phi) \,. \tag{5.65}$$

The precise form of each term written in terms of the EFT functions is presented in Table 5.1. Note that now that we have extended the reconstruction to nonlinear order it is necessary to include higher powers of X in the reconstruction, both in G_2 and G_3. In the same manner, if we were to extend the reconstruction to $(n-1)$-th order in cosmological perturbations it would introduce terms of the form X^n in G_2 and G_3.

Finally, it is also of interest to examine what effect these higher-order perturbations have on the physical EFT basis introduced in chapter 4, and developed in Ref. [168]. It consists of parameterizing the EFT formalism in terms of inherently stable basis functions: The effective Planck mass squared M^2, the sound-speed squared c_s^2, the kinetic energy of the scalar field α and the background expansion $H(t)$, along with α_{B0}. Any constraints placed on these parameters are guaranteed to satisfy the conditions for avoiding ghost and gradient instabilities, which otherwise must be checked independently for other bases. For higher-order perturbations, note that by shifting the time coordinate infinitesimally such that $t \to t + \pi$ the important operators for our purpose in the EFT action change in accordance with the following Stückelberg transformations [7, 85]

$$g^{00} \to g^{00} + 2g^{0\mu}\partial_\mu \pi + g^{\mu\nu}\partial_\mu\partial_\nu\pi \,, \tag{5.66}$$

$$\delta K \to \delta K - 3\dot{H}\pi - a^{-2}\Box\pi \,, \tag{5.67}$$

where π is interpreted as the extra scalar degree of freedom which was hidden when the action was written in the unitary gauge. An operator of the form $\bar{M}_3^4(t)(\delta g^{00})^3$ will introduce terms in the full Lagrangian such as $\bar{M}_3^4(t)\dot{\pi}^2$ after applying the time diffeomorphism. As the physical basis for the EFT functions is defined through the coefficients of such terms, this implies that these higher-order operators act to correct the lower-order EFT functions. For example, the soundspeed will now depend on these higher-order EFT functions and so the linear stability may be affected by what occurs at the nonlinear level. Physically this makes sense. If one has a second-order perturbation which is unstable, it will produce a runaway effect such that it will grow to affect the linear and background scales. In other words, the perturbations of the perturbations must be kept under control if the theory is to be completely stable. The stability of the full theory can of course be computed at the level of the covariant action. EFT naturally splits up the dynamics of the different length scales, and in order to obtain a theory that is stable, this stability must be kept at all orders in the EFT expansion.

5.7 Conclusions

Constraining models beyond ΛCDM is a worthwhile and promising endeavour of modern cosmology. We are about to see an enormous influx of observational data from surveys such as Euclid [27, 28] and LSST [29], which will provide percent-level constraints on the cosmological parameters. The outcome of these surveys will be twofold. Either the Universe turns out to be consistent with ΛCDM, which will motivate a more directed effort in tackling the cosmological constant problem (see, e.g., Refs. [189–202]). On the other hand, if recent observational tensions [4, 203, 204] persist then that will be strong evidence that the theory describing the Universe on cosmological scales requires revision and potentially will go beyond a cosmological constant. Constraints on deviations from GR are obtained on a broad range of different length scales, and a potential new theory acting on large cosmological scales must also be consistent with observations at nonlinear scales.

In this chapter we have discussed how in generalised scalar-tensor theories observations made at the level of the background and the linear perturbations may be connected with the nonlinear regime and vice-versa. This is made possible through the reconstruction of covariant Horndeski theory from the EFT of dark energy derived in chapter 3. The reconstructed theories are degenerate to linear order in cosmological perturbations and differ only by nonlinear correction terms ΔG_i. We first explored the uniqueness of these correction terms. At n-th order in perturbation theory the number of EFT operators that one can write down which are consistent with the symmetry of broken time diffeomorphisms becomes unmanageable. However, we have argued that by restricting to Horndeski theories that respect the GW170817 constraint of luminal speed of gravity [136, 146] the number of free functions that enter the EFT expansion at each order is limited to two. The two correction terms at n-th order can then be related to the free functions $\xi_n^{(2,3)}(\phi)$ specifying ΔG_2 and ΔG_3.

As a first application of the nonlinear correction terms, we have considered the implementation of screening mechanisms. With the reconstructed covariant theory it is possible to apply techniques that have been developed [146, 181] to identify the existence of Einstein gravity limits within a given Horndeski theory. With the use of these methods we have demonstrated that there is enough freedom on nonlinear scales to employ a particular type of a screening mechanism by a suitable configuration of the correction terms. More specifically, we have provided the examples of realizing a chameleon, k-mouflage and Vanshtein mechanism.

A further consequence of the reconstruction method concerns the identification of a class of models that is degenerate with ΛCDM at the level of the background and linear perturbations but departs from it at arbitrary order of nonlinear perturbations. A subclass of these models further exhibits kinetic self-acceleration, where the background expansion is accelerating exactly like ΛCDM but there is no explicit cosmological constant written in the theory. The acceleration is instead driven by the kinetic terms. An immediate consequence of the existence of such models is that even if the background expansion and linear matter power spectrum is measured to agree with ΛCDM from the next generation of surveys, the degenerate alternatives may not generally be excluded. Moreover, a theoretically appealing aspect of these models is that, with the cosmological constant now acting as a coefficient of kinetic terms rather than a bare constant, it may be possible to render it technically natural. These implications warrant a more detailed study of these models. Finally, the same techniques that were employed in the development of the reconstruction of the Horndeski action to linear order in cosmological perturbations were utilized here to derive a reconstructed theory that includes the nonlinear EFT functions. For given constraints on these functions this enables a reconstruction of the Horndeski theory across a broad range of length scales, which may be supplemented with a restriction of the allowed forms of ΔG_i to those that employ a screening mechanism.

There remain many further applications to be examined for the nonlinear sector of the reconstruction method. For example, obtaining the stability conditions is an important step in understanding the viability of the sampled models in parameter estimation analyses and it is as yet unclear what effect the nonlinear correction terms have on the stability of the theory. There may also be a more physical basis for the correction terms such as that presented in Ref. [168] for linear perturbations, which automatically satisfies the stability constraints at the nonlinear level.

Chapter 6

Conclusion

There are three stages in scientific discovery. First, people deny that it is true, then they deny that it is important; finally they credit the wrong person.

Bill Bryson

Understanding the accelerated expansion of the Universe remains one of the most fascinating problems in modern physics. Cosmic acceleration could provide crucial clues to develop our understanding of the deep connections between quantum field theory and General Relativity. Intimately entwined in this story is the cosmological constant problem (CCP) [54, 58, 59]. Tackling the CCP will almost certainly shed light on the physical mechanism driving cosmic acceleration. Is it being driven solely by a cosmological constant, or is it a hint of new physics appearing on cosmological length scales? Theoretical issues abound with the cosmological constant as discussed in chapter 2. This motivated a concerted effort in the development of a large number of alternative mechanisms for cosmic acceleration [23, 62, 71]. One of the simplest approaches is to generalise the constant in the Einstein-Hilbert action to a scalar field which permeates the Universe, and whose dynamics drives the accelerated expansion as in inflation. The existence of the Higgs boson, the excitation of the Higgs scalar field, and low energy effective theories which arise from a broad range of string theory and higher-dimensional brane-world models can motivate the presence of such a field acting on cosmological scales. Even ignoring these theoretical motivations, the vast difference in length scales between the earth and the observable Universe warrants the examination of whether there are new forces which become relevant on cosmological scales. After all, so far GR has only been rigorously tested within the Solar System [179].

Adding a scalar field to GR does not come without difficulties. It can often lead to the appearance of Ostrogradsky ghosts as we have seen in chapter 2. The most general theory which avoids these theoretical pathologies is Horndeski scalar-tensor theory. It therefore provides a well motivated starting point as a theory to test beyond-ΛCDM cosmology involving scalar fields. Horndeski theory is a fully covariant theory which can be applied to black holes [123, 124], neutron stars [125, 126] and inflation [127, 128]. In order to study the phenomenology of Horndeski theory on cosmological scales the effective field theory (EFT) of dark energy was developed [7–9, 85, 100–102] to provide a generalised description of the dynamics of the cosmological background as well as linear perturbations in Horndeski theory. Only five free functions of time are needed in the EFT of dark energy to completely capture the dynamics of the background and linear perturbations. This may still seem like too much freedom, but compared with the essentially infinite amount of freedom in the full Horndeski theory it is a great reduction in the number of parameters needed to constrain Horndeski using cosmological observations.

The remaining freedom may naïvely suggest that it is potentially impossible to distinguish between a particular Horndeski model from ΛCDM in the data. Indeed, in Ref. [107] it was shown that there were in fact infinitely many Horndeski models with the same background and linear cosmology as ΛCDM, as long as the speed of gravitational waves is allowed to deviate from the speed of light. Counting the number of observables, H for the background and $\mu(a)$ and $\eta(a)$ for the linear perturbations (see Sec. 4.3.4) against the five free functions immediately leads to this degeneracy. One of the EFT functions becomes irrelevant in the sub-horizon regime and so this degeneracy is a consequence of one EFT function, chosen to be the deviation of the speed of gravitational waves from the speed of light. Fortunately, the recent discovery that the speed of gravitational waves is equal to the speed of light [109] immediately broke this degeneracy, and in doing so ruled out a large subclass of Horndeski theory [145, 146]. It is worth noting however that there are some interesting caveats which should be considered [157, 205].

Even in light of the gravitational wave speed result, there does remain a great deal of freedom in Horndeski theory. The important difference post-GW170817 is that this freedom is falsifiable. This thesis has explored the connection between cosmological observables and Horndeski models. The eventual aim will be to use the vast data sets generated from the next generation of cosmological surveys [27–29] to directly constrain Horndeski theory. Many of the constraints which will arise from these surveys will be placed on generic parameters which aim to encompass a large number of beyond-ΛCDM models. It is therefore crucial to provide the link between commonly used parameterisations for dark energy and modified gravity models and the underlying Horndeski theory. Constraints alone do mean anything without theoretical interpretation. This thesis provides such a connection.

Chapter 3 developed a reconstruction from the effective field theory of dark energy back to fully covariant Horndeski theories. In principle, this reconstruction will enable one to bypass the effective field theory of dark energy altogether, and map directly from generic parameterizations back onto Horndeski theory. Chapter 4 explored ways in which this could be done, by taking specific examples of the modified Poisson parameter $\mu(a)$ and the gravitational slip $\eta(a)$ and using them to reconstruct the corresponding covariant theory. Although these examples were idealised, it provides a method to take constraints on these parameters and map them onto a Horndeski theory. Future work will tackle connecting the reconstruction with cosmological data sets. In principle, this provides a method to determine the shape of, for example, the scalar field potential directly from data.

In order to conclusively rule it out it is necessary to determine possible observable signatures of Horndeski theory in the nonlinear regime. This was studied in chapter 5, with a more detailed examination of the nonlinear corrections which are added on to the action reconstructed from background and linear cosmology. An interesting consequence which arose from studying these nonlinear corrections is the existence of a class of models which possess exactly the same phenomenology as ΛCDM at the level of the background and linear perturbations, but differ only in the nonlinear regime. These theories pose an important question to be resolved with upcoming surveys. Even if ΛCDM remains fully consistent with linear and background data it is not possible to determine whether it really is ΛCDM without an analysis of the nonlinear regime. Alternatively, there may be theoretical arguments which rule them out, an area which deserves a more thorough investigation. A specific sub-class of these models can even give rise to ΛCDM-like background and linear cosmology and yet does not include a cosmological constant. They exhibit *kinetic self-acceleration*. As the cosmological constant now appears as a coupling rather than as a bare constant in these kinetically self-accelerating models it may render it radiatively stable. For this reason, and for their interesting cosmological phenomenology, these models will be the subject of a more detailed study in the future.

Throughout the following years it may very well be the case that concordance cosmology holds out well against the observational constraints and, just as the search for new physics at the LHC seems to be unsuccessful, so the search for new physics in cosmology may be unsuccessful. Of course, we may be simply asking the wrong questions and exploring dead-end avenues, but until the arrival of revolutionary insights it is surely a worthy goal. In spite of this, let us not forget the elephant in the room. Many theories that have been proposed to explain cosmic acceleration do not address the cosmological constant problem. There are of course models which do precisely this [59, 193, 206], but they are precisely constructed with this aim in mind. The outcome of the analysis of future cosmological data sets will, with any luck, answer one question: ΛCDM or not ΛCDM? In the first instance, it is vital that there is a more concerted effort in tackling the theoretical issues associated with the cosmological constant. Such research could have a tremendous impact not just on our understanding of cosmology, but also of fundamental physics. If the latter scenario occurs, this thesis provides a signpost towards a theory, or a collection of theories, which is/are more compatible with what the data might reveal.

Nature is full of surprises. At every turn the only law that seems to hold is to expect the unexpected. Cosmology is no different. It is no overstatement to suggest that we are on the cusp of a revolution in our understanding of the Universe in which we live. The data could tell an unforeseen story, transforming the expected into the implausible. In twenty or thirty years we may look up at the sky through a different lens. After all, whenever the impossible has been eliminated whatever remains, however improbable, must be the truth.

Bibliography

[1] J. Kennedy, L. Lombriser and A. Taylor, *Reconstructing Horndeski models from the effective field theory of dark energy*, *Phys. Rev.* **D96** (2017) 084051, [1705.09290].

[2] J. Kennedy, L. Lombriser and A. Taylor, *Reconstructing Horndeski theories from phenomenological modified gravity and dark energy models on cosmological scales*, *Phys. Rev.* **D98** (2018) 044051, [1804.04582].

[3] J. Kennedy, L. Lombriser and A. Taylor, *Screening and degenerate kinetic self-acceleration from the nonlinear freedom of reconstructed Horndeski theories*, 1902.09853.

[4] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, 1807.06209.

[5] SDSS collaboration, D. J. Eisenstein et al., *Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies*, *Astrophys. J.* **633** (2005) 560–574, [astro-ph/0501171].

[6] B. Jain and J. Khoury, *Cosmological tests of gravity*, *Annals of Physics* **325** (July, 2010) 1479–1516, [1004.3294].

[7] G. Gubitosi, F. Piazza and F. Vernizzi, *The Effective Field Theory of Dark Energy*, *JCAP* **1302** (2013) 032, [1210.0201].

[8] E. Bellini and I. Sawicki, *Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity*, *JCAP* **1407** (2014) 050, [1404.3713].

[9] J. Gleyzes, D. Langlois and F. Vernizzi, *A unifying description of dark energy*, *Int. J. Mod. Phys.* **D23** (2015) 1443010, [1411.3712].

[10] J. Warren and C. C. W. Taylor, *The atomists: Leucippus and democritus. fragments: A text and translation with a commentary*, *Journal of Hellenic Studies* **120** (2000) 175.

[11] B. Pullman, *The atom in the history of human thought*. 1998.

[12] C. Rovelli, *Reality is not what it seems: The journey to quantum gravity, Chapter 1.* 2016.

[13] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophysical Journal* **159** (Feb, 1970) 379.

[14] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *Astrophysical Journal* **86** (Oct, 1937) 217.

[15] R. Massey, T. Kitching and J. Richard, *The dark matter of gravitational lensing*, *Rept. Prog. Phys.* **73** (2010) 086901, [1001.1739].

[16] T. Marrodn Undagoitia and L. Rauch, *Dark matter direct-detection experiments*, *J. Phys.* **G43** (2016) 013001, [1509.08767].

[17] G. B. Gelmini, *The Hunt for Dark Matter*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 559–616, 2015.

[18] J. Liu, X. Chen and X. Ji, *Current status of direct dark matter detection experiments*, *Nature Phys.* **13** (2017) 212–216, [1709.00688].

[19] B. Carr, F. Kuhnel and M. Sandstad, *Primordial Black Holes as Dark Matter*, *Phys. Rev.* **D94** (2016) 083504, [1607.06077].

[20] T. Padmanabhan, *Gravitation: Foundations and Frontiers.* 2010.

[21] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation.* 1973.

[22] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, *American Journal of Physics* **41** (Apr., 1973) 598–599.

[23] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1–189, [1106.2476].

[24] LIGO Scientific, Virgo collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102, [1602.03837].

[25] C. Kiefer, *Quantum Gravity.* 2014.

[26] C. Rovelli, *Quantum Gravity.* 2007.

[27] Euclid Theory Working Group collaboration, L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **16** (2013) 6, [1206.1225].

[28] EUCLID collaboration, R. Laureijs et al., *Euclid Definition Study Report*, 1110.3193.

[29] LSST collaboration, Z. Ivezic, J. A. Tyson, R. Allsman, J. Andrew and R. Angel, *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, 0805.2366.

[30] S. M. Carroll and J. Traschen, *Spacetime and Geometry: An Introduction to General Relativity*, *Physics Today* **58** (Jan., 2005) 52.

[31] *The Works of Richard Bentley, pg 212-213*, vol. III. 1836.

[32] R. M. Wald, *General relativity.* 1984.

[33] P. Szekeres, *A Course in Modern Mathematical Physics.* 2005.

[34] J. Michell, *On the Means of Discovering the Distance, Magnitude, c. of the Fixed Stars, in Consequence of the Diminution of the Velocity of Their Light, in Case Such a Diminution Should be Found to Take Place in any of Them, and Such Other Data Should be Procured from Observations, as Would be Farther Necessary for That Purpose. By the Rev. John Michell, B. D. F. R. S. In a Letter to Henry Cavendish, Esq. F. R. S. and A. S.*, *Philosophical Transactions of the Royal Society of London Series I* **74** (1784) 35–57.

[35] J. K. Yadav, J. S. Bagla and N. Khandai, *Fractal dimension as a measure of the scale of homogeneity*, **405** (July, 2010) 2009–2015, [1001.0617].

[36] D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris and J. D. McEwen, *How isotropic is the Universe?*, *Phys. Rev. Lett.* **117** (2016) 131302, [1605.07178].

[37] V. Mukhanov, *Physical Foundations of Cosmology.* 2005.

[38] S. Dodelson, *Modern cosmology.* 2003.

[39] P. Schneider, *Extragalactic Astronomy and Cosmology: An Introduction.* 2015.

[40] D. Baumann, *TASI Lectures on Inflation*, 0907.5424.

[41] J.-P. Uzan and P. Peter, *Primordial Cosmology.* 2005.

[42] E. Bellini, R. Jimenez and L. Verde, *Signatures of Horndeski gravity on the Dark Matter Bispectrum*, *JCAP* **1505** (2015) 057, [1504.04341].

[43] E. Carretta, R. G. Gratton, G. Clementini and F. Fusi Pecci, *Distances, ages and epoch of formation of globular clusters*, *Astrophys. J.* **533** (2000) 215–235, [astro-ph/9902086].

[44] R. Jimenez, P. Thejll, U. Jorgensen, J. MacDonald and B. Pagel, *Ages of globular clusters: a new approach*, *Mon. Not. Roy. Astron. Soc.* **282** (1996) 926–942, [astro-ph/9602132].

[45] B. M. S. Hansen, J. Brewer, G. G. Fahlman, B. K. Gibson, R. Ibata, M. Limongi et al., *The white dwarf cooling sequence of the globular cluster messier 4*, *Astrophys. J.* **574** (2002) L155–L158, [astro-ph/0205087].

[46] Supernova Search Team collaboration, A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009–1038, [astro-ph/9805201].

[47] Supernova Cosmology Project collaboration, S. Perlmutter et al., *Measurements of Omega and Lambda from 42 high redshift supernovae*, *Astrophys. J.* **517** (1999) 565–586, [astro-ph/9812133].

[48] M. Hamuy, M. M. Phillips, R. A. Schommer, N. B. Suntzeff, J. Maza and R. Aviles, *The Absolute luminosities of the Calan/Tololo type IA supernovae*, *Astron. J.* **112** (1996) 2391, [astro-ph/9609059].

[49] L. Amendola and S. Tsujikawa, *Dark Energy*. 2015.

[50] J. Renk, M. Zumalacrregui, F. Montanari and A. Barreira, *Galileon gravity in light of ISW, CMB, BAO and H_0 data*, *JCAP* **1710** (2017) 020, [1707.02263].

[51] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, *Phys. Rept.* **340** (2001) 291–472, [astro-ph/9912508].

[52] M. Kilbinger, *Cosmology with cosmic shear observations: a review*, *Rept. Prog. Phys.* **78** (2015) 086901, [1411.0115].

[53] R. Mandelbaum, *Weak lensing for precision cosmology*, *Ann. Rev. Astron. Astrophys.* **56** (2018) 393–433, [1710.03235].

[54] S. Weinberg, *The cosmological constant problem*, *Rev. Mod. Phys.* **61** (1989) .

[55] S. Nobbenhuis, *Categorizing different approaches to the cosmological constant problem*, *Found. Phys.* **36** (2006) 613–680, [gr-qc/0411093].

[56] J. Polchinski, *The Cosmological Constant and the String Landscape*, in *The Quantum Structure of Space and Time: Proceedings of the 23rd Solvay Conference on Physics. Brussels, Belgium. 1 - 3 December 2005*, pp. 216–236, 2006. hep-th/0603249.

[57] R. Bousso, *TASI Lectures on the Cosmological Constant*, *Gen. Rel. Grav.* **40** (2008) 607–637, [0708.4231].

[58] J. Martin, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*, *Comptes Rendus Physique* **13** (2012) 566–665, [1205.3365].

[59] A. Padilla, *Lectures on the Cosmological Constant Problem*, 1502.05296.

[60] E. K. Akhmedov, *Vacuum energy and relativistic invariance*, hep-th/0204048.

[61] C. P. Burgess, *The Cosmological Constant Problem: Why it's hard to get Dark Energy from Micro-physics*, in *Proceedings, 100th Les Houches Summer School: Post-Planck Cosmology: Les Houches, France, July 8 - August 2, 2013*, pp. 149–197, 2015. 1309.4133.

[62] A. Joyce, B. Jain, J. Khoury and M. Trodden, *Beyond the Cosmological Standard Model*, *Phys. Rept.* **568** (2015) 1–98, [1407.0059].

[63] G. W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363–384.

[64] R. P. Woodard, *Avoiding dark energy with 1/r modifications of gravity*, *Lect. Notes Phys.* **720** (2007) 403–433, [astro-ph/0601672].

[65] R. P. Woodard, *Ostrogradsky's theorem on Hamiltonian instability*, *Scholarpedia* **10** (2015) 32243, [1506.02210].

[66] D. Langlois and K. Noui, *Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability*, *JCAP* **1602** (2016) 034, [1510.06930].

[67] R. R. Caldwell, R. Dave and P. J. Steinhardt, *Cosmological imprint of an energy component with general equation of state*, *Phys. Rev. Lett.* **80** (1998) 1582–1585, [astro-ph/9708069].

[68] T. Chiba, T. Okabe and M. Yamaguchi, *Kinetically driven quintessence*, *Phys. Rev. D* **62** (Jun, 2000) 023511.

[69] C. Armendariz-Picon, V. Mukhanov and P. J. Steinhardt, *Essentials of k-essence*, *Phys. Rev. D* **63** (Apr, 2001) 103510.

[70] R. Cordero, E. L. Gonzalez and A. Queijeiro, *An equation of state for purely kinetic k-essence inspired by cosmic topological defects*, *Eur. Phys. J.* **C77** (2017) 413, [1608.06540].

[71] A. Joyce, L. Lombriser and F. Schmidt, *Dark Energy Versus Modified Gravity*, *Ann. Rev. Nucl. Part. Sci.* **66** (2016) 95–122, [1601.06133].

[72] S. L. Adler and R. B. Pearson, *"no-hair" theorems for the abelian higgs and goldstone models*, *Phys. Rev. D* **18** (Oct, 1978) 2798–2803.

[73] P. T. Chrusciel, J. Lopes Costa and M. Heusler, *Stationary Black Holes: Uniqueness and Beyond*, *Living Rev. Rel.* **15** (2012) 7, [1205.6112].

[74] R. H. Dicke, *Mach's principle and invariance under transformation of units*, *Phys. Rev.* **125** (Mar, 1962) 2163–2167.

[75] E. E. Flanagan, *The Conformal frame freedom in theories of gravitation*, *Class. Quant. Grav.* **21** (2004) 3817, [gr-qc/0403063].

[76] T. Padmanabhan, *Gravitation, Foundations and Frontiers.* Cambridge University Press, Cambridge, 2010.

[77] A. Nicolis, R. Rattazzi and E. Trincherini, *Galileon as a local modification of gravity*, *Phys. Rev. D* **79** (Mar, 2009) 064036.

[78] C. de Rham, *Massive Gravity*, *Living Rev. Rel.* **17** (2014) 7, [1401.4173].

[79] M. A. Luty, M. Porrati and R. Rattazzi, *Strong interactions and stability in the DGP model*, *JHEP* **09** (2003) 029, [hep-th/0303116].

[80] K. Hinterbichler, M. Trodden and D. Wesley, *Multi-field galileons and higher co-dimension branes*, *Phys. Rev.* **D82** (2010) 124018, [1008.1305].

[81] C. Deffayet, X. Gao, D. A. Steer and G. Zahariade, *From k-essence to generalised Galileons*, *Phys. Rev.* **D84** (2011) 064039, [1103.3260].

[82] E. Babichev, C. Charmousis and A. Lehébel, *Black holes and stars in Horndeski theory*, *Classical and Quantum Gravity* **33** (2016) 154002, [1604.06402].

[83] C. Deffayet and D. A. Steer, *A formal introduction to Horndeski and Galileon theories and their generalizations*, *Class. Quant. Grav.* **30** (2013) 214006, [1307.2450].

[84] C. Charmousis, *From Lovelock to Horndeski's Generalized Scalar Tensor Theory*, *Lect. Notes Phys.* **892** (2015) 25–56, [1405.1612].

[85] J. K. Bloomfield, E. E. Flanagan, M. Park and S. Watson, *Dark energy or modified gravity? An effective field theory approach*, *JCAP* **1308** (2013) 010, [1211.7054].

[86] Y. Fujii and K. Maeda, *The scalar-tensor theory of gravitation.* 2007.

[87] C. M. Will, *The Confrontation between general relativity and experiment*, *Living Rev. Rel.* **9** (2006) 3, [gr-qc/0510072].

[88] J. Khoury and A. Weltman, *Chameleon fields: Awaiting surprises for tests of gravity in space*, *Phys. Rev. Lett.* **93** (2004) 171104, [astro-ph/0309300].

[89] J. Khoury and A. Weltman, *Chameleon cosmology*, *Phys. Rev. D* **69** (Feb, 2004) 044026.

[90] E. Babichev, C. Deffayet and R. Ziour, *k-Mouflage gravity*, *Int. J. Mod. Phys.* **D18** (2009) 2147–2154, [0905.2943].

[91] C. de Rham and R. H. Ribeiro, *Riding on irrelevant operators*, *Journal of Cosmology and Astro-Particle Physics* **2014** (Nov., 2014) 016, [1405.5213].

[92] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett.* **39B** (1972) 393–394.

[93] M. Srednicki, *Quantum field theory.* 2007.

[94] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Addison-Wesley, Reading, USA, 1995.

[95] R. L. Arnowitt, S. Deser and C. W. Misner, *The Dynamics of general relativity*, *Gen. Rel. Grav.* **40** (2008) 1997–2027, [gr-qc/0405109].

[96] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan and L. Senatore, *The Effective Field Theory of Inflation*, *JHEP* **03** (2008) 014, [0709.0293].

[97] P. Creminelli, G. D'Amico, J. Norena and F. Vernizzi, *The Effective Theory of Quintessence: the w < -1 Side Unveiled*, *JCAP* **0902** (2009) 018, [0811.0827].

[98] M. Park, K. M. Zurek and S. Watson, *A Unified Approach to Cosmic Acceleration*, *Phys. Rev.* **D81** (2010) 124008, [1003.1722].

[99] J. K. Bloomfield and E. E. Flanagan, *A Class of Effective Field Theory Models of Cosmic Acceleration*, *JCAP* **1210** (2012) 039, [1112.0303].

[100] J. Gleyzes, D. Langlois, F. Piazza and F. Vernizzi, *Essential Building Blocks of Dark Energy*, *JCAP* **1308** (2013) 025, [1304.4840].

[101] J. Bloomfield, *A Simplified Approach to General Scalar-Tensor Theories*, *JCAP* **1312** (2013) 044, [1304.6712].

[102] S. Tsujikawa, *The effective field theory of inflation/dark energy and the Horndeski theory*, *Lect. Notes Phys.* **892** (2015) 97–136, [1404.2684].

[103] M. Lagos, T. Baker, P. G. Ferreira and J. Noller, *A general theory of linear cosmological perturbations: scalar-tensor and vector-tensor theories*, *JCAP* **1608** (2016) 007, [1604.01396].

[104] K. Koyama, *Cosmological Tests of Modified Gravity*, *Rept. Prog. Phys.* **79** (2016) 046902, [1504.04623].

[105] L. Lombriser and J. Peñarrubia, *How chameleons core dwarfs with cusps*, *Phys. Rev.* **D91** (2015) 084022, [1407.7862].

[106] T. Kobayashi, M. Yamaguchi and J. Yokoyama, *Generalized G-Inflation Inflation with the Most General Second-Order Field Equations*, *Progress of Theoretical Physics* **126** (2011) 511–529, [1105.5723].

[107] L. Lombriser and A. Taylor, *Breaking a Dark Degeneracy with Gravitational Waves*, *JCAP* **1603** (2016) 031, [1509.08458].

[108] L. Lombriser and N. A. Lima, *Challenges to Self-Acceleration in Modified Gravity from Gravitational Waves and Large-Scale Structure*, *Phys. Lett.* **B765** (2017) 382–385, [1602.07670].

[109] Virgo, LIGO Scientific collaboration, B. Abbott et al., *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101, [1710.05832].

[110] J. Gleyzes, D. Langlois, F. Piazza and F. Vernizzi, *Healthy theories beyond Horndeski*, *Phys. Rev. Lett.* **114** (2015) 211101, [1404.6495].

[111] M. Zumalacrregui and J. Garca-Bellido, *Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian*, *Phys. Rev.* **D89** (2014) 064046, [1308.4685].

[112] F. Piazza, H. Steigerwald and C. Marinoni, *Phenomenology of dark energy: exploring the space of theories with future redshift surveys*, *JCAP* **1405** (2014) 043, [1312.6111].

[113] L. Lombriser and A. Taylor, *Classifying Linearly Shielded Modified Gravity Models in Effective Field Theory*, *Phys. Rev. Lett.* **114** (2015) 031101, [1405.2896].

[114] S. Tsujikawa, *Possibility of realizing weak gravity in redshift space distortion measurements*, *Phys. Rev.* **D92** (2015) 044029, [1505.02459].

[115] E. Bellini, A. J. Cuesta, R. Jimenez and L. Verde, *Constraints on deviations from CDM within Horndeski gravity*, *JCAP* **1602** (2016) 053, [1509.07816].

[116] L. Lombriser and A. Taylor, *Semi-dynamical perturbations of unified dark energy*, *JCAP* **1511** (2015) 040, [1505.05915].

[117] L. Pogosian and A. Silvestri, *What can cosmology tell us about gravity? Constraining Horndeski gravity with Σ and μ*, *Phys. Rev.* **D94** (2016) 104014, [1606.05339].

[118] L. Perenon, C. Marinoni and F. Piazza, *Diagnostic of Horndeski Theories*, *JCAP* **1701** (2017) 035, [1609.09197].

[119] Y. Cai, Y. Wan, H.-G. Li, T. Qiu and Y.-S. Piao, *The Effective Field Theory of nonsingular cosmology*, *JHEP* **01** (2017) 090, [1610.03400].

[120] E. V. Linder, G. Sengör and S. Watson, *Is the Effective Field Theory of Dark Energy Effective?*, *JCAP* **1605** (2016) 053, [1512.06180].

[121] E. V. Linder, *Challenges in connecting modified gravity theory and observations*, *Phys. Rev.* **D95** (2017) 023518, [1607.03113].

[122] J. Gleyzes, *Parametrizing modified gravity for cosmological surveys*, *Phys. Rev.* **D96** (2017) 063516, [1705.04714].

[123] C. Charmousis, B. Gouteraux and E. Kiritsis, *Higher-derivative scalar-vector-tensor theories: black holes, Galileons, singularity cloaking and holography*, *JHEP* **09** (2012) 011, [1206.1499].

[124] E. Babichev and C. Charmousis, *Dressing a black hole with a time-dependent Galileon*, *JHEP* **08** (2014) 106, [1312.3204].

[125] A. Cisterna, T. Delsate and M. Rinaldi, *Neutron stars in general second order scalar-tensor theory: The case of nonminimal derivative coupling*, *Phys. Rev.* **D92** (2015) 044050, [1504.05189].

[126] A. Cisterna, T. Delsate, L. Ducobu and M. Rinaldi, *Slowly rotating neutron stars in the nonminimal derivative coupling sector of Horndeski gravity*, *Phys. Rev.* **D93** (2016) 084046, [1602.06939].

[127] X. Gao and D. A. Steer, *Inflation and primordial non-Gaussianities of 'generalized Galileons'*, *JCAP* **1112** (2011) 019, [1107.2642].

[128] A. De Felice and S. Tsujikawa, *Shapes of primordial non-Gaussianities in the Horndeski's most general scalar-tensor theories*, *JCAP* **1303** (2013) 030, [1301.5721].

[129] F. Piazza and F. Vernizzi, *Effective Field Theory of Cosmological Perturbations*, *Class. Quant. Grav.* **30** (2013) 214007, [1307.4350].

[130] N. A. Lima, V. Smer-Barreto and L. Lombriser, *Constraints on decaying early modified gravity from cosmological observations*, *Phys. Rev.* **D94** (2016) 083507, [1603.05239].

[131] G. Esposito-Farese and D. Polarski, *Scalar tensor gravity in an accelerating universe*, *Phys. Rev.* **D63** (2001) 063504, [gr-qc/0009034].

[132] C. Armendariz-Picon, T. Damour and V. F. Mukhanov, *k - inflation*, *Phys. Lett.* **B458** (1999) 209–218, [hep-th/9904075].

[133] Y. Cai and Y.-S. Piao, *A covariant Lagrangian for stable nonsingular bounce*, *JHEP* **09** (2017) 027, [1705.03401].

[134] N. Chow and J. Khoury, *Galileon Cosmology*, *Phys. Rev.* **D80** (2009) 024037, [0905.1325].

[135] I. Sawicki, I. D. Saltas, M. Motta, L. Amendola and M. Kunz, *Nonstandard gravitational waves imply gravitational slip: On the difficulty of partially hiding new gravitational degrees of freedom*, *Phys. Rev.* **D95** (2017) 083520, [1612.02002].

[136] Virgo, Fermi-GBM, INTEGRAL, LIGO Scientific collaboration, B. P. Abbott et al., *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J.* **848** (2017) L13, [1710.05834].

[137] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XIV. Dark energy and modified gravity*, *Astron. Astrophys.* **594** (2016) A14, [1502.01590].

[138] L. Lombriser, *A parametrisation of modified gravity on nonlinear cosmological scales*, *JCAP* **1611** (2016) 039, [1608.00522].

[139] GROND, SALT Group, OzGrav, DFN, INTEGRAL, Virgo, Insight-Hxmt, MAXI Team, Fermi-LAT, J-GEM, RATIR, IceCube, CAASTRO, LWA, ePESSTO, GRAWITA, RIMAS, SKA South Africa/MeerKAT, H.E.S.S., 1M2H Team, IKI-GW Follow-up, Fermi GBM, Pi of Sky, DWF (Deeper Wider Faster Program), Dark Energy Survey, MASTER, AstroSat Cadmium Zinc Telluride Imager Team, Swift, Pierre Auger, ASKAP, VINROUGE, JAGWAR, Chandra Team at McGill University, TTU-NRAO, GROWTH, AGILE Team, MWA, ATCA, AST3, TOROS, Pan-STARRS, NuSTAR, ATLAS Telescopes, BOOTES, CaltechNRAO, LIGO Scientific, High Time Resolution Universe Survey, Nordic Optical Telescope, Las Cumbres Observatory Group, TZAC Consortium, LOFAR, IPN, DLT40, Texas Tech University, HAWC, ANTARES, KU, Dark Energy Camera GW-EM, CALET, Euro VLBI Team, ALMA collaboration, B. P. Abbott et al., *Multi-messenger Observations of a Binary Neutron Star Merger*, *Astrophys. J.* **848** (2017) L12, [1710.05833].

[140] A. Nishizawa and T. Nakamura, *Measuring Speed of Gravitational Waves by Observations of Photons and Neutrinos from Compact Binary Mergers and Supernovae*, *Phys. Rev.* **D90** (2014) 044048, [1406.5544].

[141] I. D. Saltas, I. Sawicki, L. Amendola and M. Kunz, *Anisotropic Stress as a Signature of Nonstandard Propagation of Gravitational Waves*, *Phys. Rev. Lett.* **113** (2014) 191101, [1406.7139].

[142] A. Nishizawa, *Generalized framework for testing gravity with gravitational-wave propagation. I. Formulation*, *Phys. Rev.* **D97** (2018) 104037, [1710.04825].

[143] E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore, *Gravitational-wave luminosity distance in modified gravity theories*, *Phys. Rev.* **D97** (2018) 104066, [1712.08108].

[144] L. Amendola, I. Sawicki, M. Kunz and I. D. Saltas, *Direct detection of gravitational waves can measure the time variation of the Planck mass*, *JCAP* **1808** (2018) 030, [1712.08623].

[145] R. Kimura and K. Yamamoto, *Constraints on general second-order scalar-tensor models from gravitational Cherenkov radiation*, *JCAP* **1207** (2012) 050, [1112.4284].

[146] R. McManus, L. Lombriser and J. Pearrubia, *Finding Horndeski theories with Einstein gravity limits*, *JCAP* **1611** (2016) 006, [1606.03282].

[147] J. Beltran Jimenez, F. Piazza and H. Velten, *Evading the Vainshtein Mechanism with Anomalous Gravitational Wave Speed: Constraints on Modified Gravity from Binary Pulsars*, *Phys. Rev. Lett.* **116** (2016) 061101, [1507.05047].

[148] P. Brax, C. Burrage and A.-C. Davis, *The Speed of Galileon Gravity*, *JCAP* **1603** (2016) 004, [1510.03701].

[149] D. Bettoni, J. M. Ezquiaga, K. Hinterbichler and M. Zumalacrregui, *Speed of Gravitational Waves and the Fate of Scalar-Tensor Gravity*, *Phys. Rev.* **D95** (2017) 084029, [1608.01982].

[150] P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017) 251302, [1710.05877].

[151] J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, *Phys. Rev. Lett.* **119** (2017) 251303, [1710.05893].

[152] J. M. Ezquiaga and M. Zumalacrregui, *Dark Energy After GW170817: Dead Ends and the Road Ahead*, *Phys. Rev. Lett.* **119** (2017) 251304, [1710.05901].

[153] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller and I. Sawicki, *Strong constraints on cosmological gravity from GW170817 and GRB 170817A*, *Phys. Rev. Lett.* **119** (2017) 251301, [1710.06394].

[154] M. Crisostomi and K. Koyama, *Self-accelerating universe in scalar-tensor theories after GW170817*, *Phys. Rev.* **D97** (2018) 084004, [1712.06556].

[155] E. V. Linder, *No Slip Gravity*, *JCAP* **1803** (2018) 005, [1801.01503].

[156] H. Wang et al., *The GW170817/GRB 170817A/AT 2017gfo Association: Some Implications for Physics and Astrophysics*, *Astrophys. J.* **851** (2017) L18, [1710.05805].

[157] R. A. Battye, F. Pace and D. Trinh, *Gravitational wave constraints on dark sector models*, *Phys. Rev.* **D98** (2018) 023504, [1802.09447].

[158] J.-P. Uzan, *The acceleration of the universe and the physics behind it*, *Gen. Rel. Grav.* **39** (2007) 307–342, [astro-ph/0605313].

[159] L. Amendola, M. Kunz and D. Sapone, *Measuring the dark side (with weak lensing)*, *JCAP* **0804** (2008) 013, [0704.2421].

[160] R. Caldwell, A. Cooray and A. Melchiorri, *Constraints on a New Post-General Relativity Cosmological Parameter*, *Phys. Rev.* **D76** (2007) 023507, [astro-ph/0703375].

[161] W. Hu and I. Sawicki, *A Parameterized Post-Friedmann Framework for Modified Gravity*, *Phys. Rev.* **D76** (2007) 104043, [0708.1190].

[162] P. Zhang, M. Liguori, R. Bean and S. Dodelson, *Probing Gravity at Cosmological Scales by Measurements which Test the Relationship between Gravitational Lensing and Matter Overdensity*, *Phys. Rev. Lett.* **99** (2007) 141302, [0704.1932].

[163] P. J. E. Peebles, *The large-scale structure of the universe.* Princeton University Press, 1980.

[164] L.-M. Wang and P. J. Steinhardt, *Cluster abundance constraints on quintessence models*, *Astrophys. J.* **508** (1998) 483–490, [astro-ph/9804015].

[165] E. V. Linder, *Cosmic growth history and expansion history*, *Phys. Rev.* **D72** (2005) 043529, [astro-ph/0507263].

[166] DES collaboration, T. M. C. Abbott et al., *Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev.* **D98** (2018) 043526, [1708.01530].

[167] A. Amon et al., *KiDS+2dFLenS+GAMA: Testing the cosmological model with the $E_{\rm G}$ statistic*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 3422–3437, [1711.10999].

[168] L. Lombriser, C. Dalang, J. Kennedy and A. Taylor, *Inherently stable effective field theory for dark energy and modified gravity*, *Journal of Cosmology and Astro-Particle Physics* **2019** (Jan., 2019) 041, [1810.05225].

[169] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13, [1502.01589].

[170] C. Deffayet, O. Pujolas, I. Sawicki and A. Vikman, *Imperfect Dark Energy from Kinetic Gravity Braiding*, *JCAP* **1010** (2010) 026, [1008.0048].

[171] J. Wang, L. Hui and J. Khoury, *No-Go Theorems for Generalized Chameleon Field Theories*, *Phys. Rev. Lett.* **109** (2012) 241301, [1208.4612].

[172] L. Lombriser, *Consistency check of ΛCDM phenomenology*, *Phys. Rev.* **D83** (2011) 063519, [1101.0594].

[173] L. Lombriser, J. Yoo and K. Koyama, *Relativistic effects in galaxy clustering in a parametrized post-Friedmann universe*, *Phys. Rev.* **D87** (2013) 104019, [1301.3132].

[174] C. Brans and R. H. Dicke, *Mach's principle and a relativistic theory of gravitation*, *Phys. Rev.* **124** (Nov, 1961) 925–935.

[175] L. Perenon, F. Piazza, C. Marinoni and L. Hui, *Phenomenology of dark energy: general features of large-scale perturbations*, *JCAP* **1511** (2015) 029, [1506.03047].

[176] G. D'Amico, Z. Huang, M. Mancarella and F. Vernizzi, *Weakening Gravity on Redshift-Survey Scales with Kinetic Matter Mixing*, *JCAP* **1702** (2017) 014, [1609.01272].

[177] G. Cusin, M. Lewandowski and F. Vernizzi, *Nonlinear Effective Theory of Dark Energy*, *JCAP* **1804** (2018) 061, [1712.02782].

[178] N. Frusciante and G. Papadomanolakis, *Tackling non-linearities with the effective field theory of dark energy and modified gravity*, *JCAP* **1712** (2017) 014, [1706.02719].

[179] C. M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4, [1403.7377].

[180] K. Hinterbichler and J. Khoury, *Symmetron Fields: Screening Long-Range Forces Through Local Symmetry Restoration*, *Phys. Rev. Lett.* **104** (2010) 231301, [1001.4525].

[181] R. McManus, L. Lombriser and J. Pearrubia, *Parameterised Post-Newtonian Expansion in Screened Regions*, *JCAP* **1712** (2017) 031, [1705.05324].

[182] L. Lombriser, *Parametrizations for tests of gravity*, *International Journal of Modern Physics D* **27** (2018) 1848002.

[183] N. Bolis, C. Skordis, D. B. Thomas and T. Zlosnik, *The Parameterized Post-Newtonian-Vainshteinian formalism for the Galileon field*, 1810.02725.

[184] L. Lombriser, K. Koyama and B. Li, *Halo modelling in chameleon theories*, *JCAP* **1403** (2014) 021, [1312.1292].

[185] P. Brax, C. Burrage and A.-C. Davis, *Screening fifth forces in k-essence and DBI models*, *JCAP* **1301** (2013) 020, [1209.1293].

[186] R. Kimura, T. Kobayashi and K. Yamamoto, *Observational Constraints on Kinetic Gravity Braiding from the Integrated Sachs-Wolfe Effect*, *Phys. Rev.* **D85** (2012) 123503, [1110.3598].

[187] B. Gripaios, *Lectures on Effective Field Theory*, 1506.05039.

[188] D. Pirtskhalava, L. Santoni, E. Trincherini and F. Vernizzi, *Weakly Broken Galileon Symmetry*, *JCAP* **1509** (2015) 007, [1505.00007].

[189] J. D. Barrow and D. J. Shaw, *A New Solution of The Cosmological Constant Problems*, *Phys. Rev. Lett.* **106** (2011) 101302, [1007.3086].

[190] D. J. Shaw and J. D. Barrow, *A Testable Solution of the Cosmological Constant and Coincidence Problems*, *Phys. Rev.* **D83** (2011) 043518, [1010.4262].

[191] J. D. Barrow and D. J. Shaw, *The Value of the Cosmological Constant*, *Gen. Rel. Grav.* **43** (2011) 2555–2560, [1105.3105].

[192] R. Casadio, *Lorentz invariance without trans-Planckian physics?*, *Phys. Lett.* **B724** (2013) 351–354, [1303.1914].

[193] N. Kaloper and A. Padilla, *Sequestering the Standard Model Vacuum Energy*, *Phys. Rev. Lett.* **112** (2014) 091304, [1309.6562].

[194] N. Kaloper, A. Padilla, D. Stefanyszyn and G. Zahariade, *Manifestly Local Theory of Vacuum Energy Sequestering*, *Phys. Rev. Lett.* **116** (2016) 051302, [1505.01492].

[195] Q. Wang, Z. Zhu and W. G. Unruh, *How the huge energy of quantum vacuum gravitates to drive the slow accelerating expansion of the Universe*, *Phys. Rev.* **D95** (2017) 103504, [1703.00543].

[196] L. Lombriser and V. Smer-Barreto, *Is there another coincidence problem at the reionization epoch?*, *Phys. Rev.* **D96** (2017) 123505, [1707.03388].

[197] S. Appleby and E. V. Linder, *The Well-Tempered Cosmological Constant*, *JCAP* **1807** (2018) 034, [1805.00470].

[198] L. Lombriser, *Late-time acceleration by a residual cosmological constant from sequestering vacuum energy in ultimate collapsed structures*, 1805.05918.

[199] S.-J. Wang, *Electroweak relaxation of cosmological hierarchy*, *Phys. Rev.* **D99** (2019) 023529, [1810.06445].

[200] O. Evnin and K. Nguyen, *Graceful exit for the cosmological constant damping scenario*, *Phys. Rev.* **D98** (2018) 124031, [1810.12336].

[201] F. Canales, B. Koch, C. Laporte and A. Rincon, *Vacuum energy density: deflation during inflation*, 1812.10526.

[202] L. Lombriser, *On the cosmological constant problem*, 1901.08588.

[203] H. Hildebrandt et al., *KiDS-450: Cosmological parameter constraints from tomographic weak gravitational lensing*, *Mon. Not. Roy. Astron. Soc.* **465** (2017) 1454, [1606.05338].

[204] A. G. Riess, S. Casertano, W. Yuan, L. Macri, J. Anderson, J. W. MacKenty et al., *New Parallaxes of Galactic Cepheids from Spatially Scanning the Hubble Space Telescope: Implications for the Hubble Constant*, *Astrophysical Journal* **855** (Mar., 2018) 136, [1801.01120].

[205] E. J. Copeland, M. Kopp, A. Padilla, P. M. Saffin and C. Skordis, *Dark energy after GW170817 revisited*, *Phys. Rev. Lett.* **122** (2019) 061301, [1810.08239].

[206] E. J. Copeland, A. Padilla and P. M. Saffin, *The cosmology of the Fab-Four*, *JCAP* **1212** (2012) 026, [1208.3373].