Abstract: The current standard cosmological model is constructed within the framework of general relativity with a cosmological constant $\Lambda$, which is often associated with dark energy, and phenomenologically explains the accelerated cosmic expansion. Understanding the nature of dark energy is one of the most appealing questions in achieving a self-consistent physical model at cosmological scales. Modification of general relativity could potentially provide a more natural and physical solution to the accelerated expansion. The growth of the cosmic structure is sensitive in constraining gravity models. In this paper, we aim to provide a concise introductory review of modified gravity models from an observational point of view. We will discuss various mainstream cosmological observables, and their potential advantages and limitations as probes of gravity models.

Keywords: modified gravity; large-scale structure; cosmology; observation

1. Introduction

Our current physical description of gravity, Einstein’s general relativity (GR), has profound implications for various astrophysical observables. The theory precisely predicts the orbital decay of binary pulsars [1], gravitational time dilation observed in the spectra of white dwarf Sirius B [2], and gravitational waves (GWs) from merging black holes [3] or other compact objects [4].

General relativity is also one of the cornerstones for our understanding of the Universe over billions of years of evolution. However, there still lacks a satisfying explanation for the observed accelerated expansion of our Universe. The first evidence for cosmic acceleration was found in measurements of type-Ia supernovae [5,6] and was later supported by various observations including the cosmic microwave background (CMB), the 3D distribution of large-scale structures, and GW measurements. To explain the observed acceleration under GR, an exotic form of energy with negative pressure has to be introduced in the field equations, the so-called “dark energy” (DE) (see reviews, e.g., [7,8]). In its simplest form, DE is just a cosmological constant $\Lambda$. This additional source of energy could be explained through vacuum energy predicted by quantum field theory. However, this simplest explanation faces various theoretical problems [9], mainly driven by the fact that
the vacuum energy predicted by quantum field theory disagrees with the observed values by more than a hundred orders of magnitude.

In addition to DE, an exotic form of matter that only interacts gravitationally, “dark matter”, also needs to be introduced to describe, e.g., the observed cosmic structure formation and rotation curves of galaxies [10]. The $\Lambda$CDM model—a model containing DE in the form of a cosmological constant plus cold dark matter—is currently the best fit to most observations. Although the $\Lambda$CDM model can well explain various astrophysical observations, the nature of its dark sectors is not well understood: neither DE nor dark matter could be detected so far in lab experiments. Therefore, theoretical physicists around the world have been trying to come up with solutions for the accelerated expansion of our Universe that do not involve introducing new exotic components but rather modify or extend Einstein’s theory of gravity.

This paper is one of the review series on the cosmological constant problem. The series will discuss the cosmological constant problem from different perspectives. Here, we will be focusing on testing modified theories of GR, “Modified Gravity” (MG) through probes of the growth of structure or the gravitational potential at extragalactic scales $\sim O(1 \, h^{-1}\text{Mpc})$. For each probe, we will present their motivations, their promise to detect modifications of GR, as well as their challenges and limitations.

The review organised as follows. We will start with a brief overview of the modified gravity models in Section 2; we will cover two of the most popular parametrizations of the MG models, the screening mechanisms, and various parameterised frameworks of gravity. We will proceed with astrophysical probes at different scales in Section 3 and tools to distinguish the deviation from GR and list constraints from current surveys. We will discuss the available simulations which assist further tests and our understanding of a few most popular MG parametrizations in Section 4. In the next Section 5, we will outline current and next-stage cosmological surveys, including galaxy, CMB, and radio surveys. Finally, we will conclude the work in Section 6.

2. An Overview of the Modified Gravity Models

In this section, we briefly overview the most popular MG models. As discussed in Section 1, although the cosmological constant provides an excellent explanation for the accelerated cosmic expansion, more natural choices for the cosmological constant can be classified into two categories: (i) modifications of the stress-energy tensor on the right-hand side of Einstein’s field equation, leading to DE models; (ii) modifications of, e.g., the Einstein–Hilbert action on the left-hand side of the field equation, leading to MG models. In practice, there are no clear boundaries between DE or MG models (for a review see, e.g., [11]). Nevertheless, it is possible to make certain distinctions based on the strong equivalence principle and observe whether ordinary matter experiences additional forces beyond gravity. Figure 1 presents a flowchart distinguishing MG vs. DE following [11].

In the conformal Newtonian gauge, the line element is given by

$$ds^2 = a^2 \left[-(1 + 2\Psi)d\tau^2 + (1 - 2\Phi)dx^2\right],$$

with $\tau$ being the conformal time; the gauge invariant Newtonian potential $\Psi$ and curvature potential $\Phi$ are functions of space and time. In the case of GR, the difference between two potentials is negligibly small, and the 00 component of the Einstein equation on sub-horizon scales is given by the Poisson equation

$$\nabla^2 \Psi = 4\pi Ga^2 \delta \rho_m,$$

where $\delta \rho_m \equiv \rho_m - \bar{\rho}_m$ is the fluctuation of the matter density. In the presence of a fifth force, Equation (2) is modified and the structure formed, albeit from the one predicted by GR. There are different ways to “modify gravity”, among which scalar–tensor theories (e.g., Brans and Dicke [12]) are probably the most well-studied models. One of the most
general scalar–tensor theories with second-order field equations in four dimensions is the Horndeski theory [13], whose action is constructed as

\[ S = \int d^4x \sqrt{-g} \left\{ \sum_{i=2}^{5} \mathcal{L}_i [\phi, g_{\mu\nu}] + \mathcal{L}_m (g_{\mu\nu}, \psi) \right\}. \] (3)

Each of the four Lagrangian densities in the summation is a function of four parameters \{\mathcal{L}_2[K], \mathcal{L}_3[G_3], \mathcal{L}_4[G_4], and \mathcal{L}_5[G_5]\}, where \{K, G_3, G_4, G_5\} are arbitrary functions of (\phi, X), with \phi as a scalar field, \(X \equiv -\nabla^\nu \phi \nabla_\nu \phi / 2\). \(\mathcal{L}_m\) as the matter Lagrangian density, and \(\psi\) as the matter field.

Figure 1. A flowchart distinguishing the DE and modified gravity (MG) models (based on Joyce et al. [11]). The weak equivalence principle (WEP) states that there exists a Jordan frame metric to which matter species are minimally coupled, independent of their composition or structure. Taking scalar–tensor theory as an example, all matter species couple universally to the metric in the presence of a scalar field. The strong equivalence principle (SEP) further restricts the WEP such that the equivalence principle also applies to objects that exert a self-gravitational force. If the coupling between matter fields and the metric does not involve any scalar fields, it is termed a DE model. Whereas if the matter fields couple to the metric non-trivially via the scalar field, it is termed a MG model (red box) to be considered in this paper.

In the following, we will only review two MG models, the \(f(R)\) model and the braneworld model, which can be linked to Horndeski theory by properly choosing the four Horndeski functions. These two models are also the most representative and have been widely tested using different astrophysical probes (see Section 3) and implemented in cosmological simulations (see Section 4).

2.1. Conformal Coupling Models: \(f(R)\) Gravity

\(f(R)\) gravity [14,15] is a very popular class of MG models, described by the following gravitational action

\[ S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} [R + f(R)] + S_m (g_{\mu\nu}, \psi), \] (4)

where the cosmological constant \(\Lambda\) is replaced by an algebraic function of the Ricci scalar, \(f(R)\), and \(S_m\) is the action for the matter field \(\psi\). Various functional forms for \(f(R)\) were proposed in early works to explain the cosmic acceleration [16–18]. However, they do not pass solar system tests of gravity [19,20]. The authors of [21] suggested the functional form of \(f(R)\) which is compatible with local gravity tests.
\[ f(R) = -m^2 \frac{c_1 \left( \frac{R}{m^2} \right)^n}{c_2 \left( \frac{R}{m^2} \right)^n + 1} \approx \rho_{\Lambda, \text{eff}} + f_{R0} \left( \frac{R}{R_0} \right)^{-n}. \]  

(5)

This functional form is chosen such that in the small curvature limit \( \lim_{R \to 0} f(R) \to 0 \) we recover \( \Lambda \)CDM at high redshift and in the large curvature limit \( \lim_{R \to \infty} f(R) \to \text{const.} \) it mimics cosmic acceleration at low redshift. \( \bar{R}_0 \) is the background Ricci scalar today. To satisfy cosmological and local observations, the condition \( R/m^2 \ll 1 \) is required, with \( m \) being a free parameter representing a mass scale. The two free parameters \( c_1 \) and \( c_2 \) are adjusted accordingly such that the effective DE density \( \rho_{\Lambda, \text{eff}} \) gives rise to a cosmological constant that matches observations. \( f_{R0} \equiv f_R(z = 0) \) is the present-day value of the background field, with the scalar field being \( f_R \equiv df(R)/dR \). This scalar field \( f_R \) is of particular importance, and the impact of \( f(R) \) gravity can be viewed in terms of the dynamics of \( f_R \). The larger the amplitude of the scalar field \( |f_R| \), the stronger deviation from GR.

In varying the action in Equation (4) with respect to the metric, we can derive the modified Einstein equations, whose trace can be interpreted as the equation of motion for \( f_R \)

\[ \Box f_R = \frac{\partial V_{\text{eff}}}{\partial f_R}, \quad V_{\text{eff}}(f_R) = V(f_R) + A(f_R)\rho_{m}, \]  

(6)

where \( \Box \equiv \nabla_{\mu} \nabla^{\mu} \) is the d’Alembert operator. The effective potential \( V_{\text{eff}} \) consists of two parts; one is a bare function that depends on the scalar field itself; another part knows the external matter field density \( \rho_{m} \) (see Equation (20) in Hu and Sawicki [21] for the exact form of \( V_{\text{eff}} \)). The density-dependent term of the effective potential in Equation (6) is the key that this functional form of \( f(R) \) can pass the local gravity tests as its dynamics associate high-density regions with the high curvature of spacetime (see Section 2.3).

The interaction range of the scalar field is determined by the Compton wavelength \( \lambda_c \):

\[ m_{f_R}^2 \equiv \frac{\partial^2 V_{\text{eff}}}{\partial f_R^2}, \quad \lambda_c \equiv m_{f_R}^{-1}, \]  

(7)

the larger the scalar field, the shorter the Compton wavelength, and the shorter range the fifth force can mediate its interaction. The presence of the fifth force leads to a different structure formation history, embodied in the modified Poisson equation. Under the quasi-static\(^1 \) and weak field approximations\(^2 \), the scalar field \( f_R \) can be rewritten as

\[ \nabla^2 \Psi \approx \frac{16\pi G}{3} a^2 \delta \rho_{m} - \frac{1}{6} a^2 \delta R \]

\[ = 4\pi G a^2 \delta \rho_{m} - \frac{1}{2} \nabla^2 f_R, \]  

(8)

where the first term is the standard Poisson equation, and the second term represents the fifth force, \( F_{5th} \propto \nabla f_R \), generated by the scalar field \( f_R \). Under the static limit, Equation (6) can be rewritten as

\[ \nabla^2 f_R = -\frac{a^2}{3} [\delta R - 8\pi G \delta \rho_{m}], \]  

(9)

where \( \delta R = R(f_R) - \bar{R} \) and \( \delta \rho_{m} = \rho_{m} - \bar{\rho}_{m} \) is the perturbation of non-relativistic matter density. The \( f(R) \) model can be shown to be equivalent to a scalar–tensor theory in which the scalar field has a universal coupling to different matter species by a conformal
transformation \([19,22]\). At the same time, the connection between the \(f(R)\) gravity and Horndeski can be seen by setting \([23]\)

\[
G_4 = \phi = f_R, \quad K = f(R) - Rf_R, \quad G_3 = 0.
\]

Hereafter, we follow the naming convention for the \(f(R)\) model: for a given value of the present-day \(f_{R0}\), we take the absolute value of its logarithm and call it \(F\) \(\log(f_{R0})\). As an example, \(f_{R0} = 10^{-5}\) is called the F5 model.

2.2. Derivative Coupling Models: DGP Gravity

In the DGP model \([24]\), the Universe is a four-dimensional “brane” embedded in a five-dimensional spacetime or bulk. The total action of the model is

\[
S = \int_{\text{brane}} d^4x \sqrt{-g} \frac{K}{16\pi G} + \int d^5x \sqrt{-g^{(5)} \frac{R^{(5)}}{16\pi G^{(5)}}} + S_m(g_{\mu\nu}, \phi_i),
\]

where \(g_{\mu\nu}, g, R \) and \(G\) are the metric tensor, the determinant of the metric, the Ricci scalar and the gravitational constant in the 4D brane, respectively, and \(g^{(5)}, R^{(5)} \) and \(G^{(5)}\) are their equivalents in the 5D bulk. \(S_m\) is the action of the matter fields \(\psi_i\) which are assumed to be confined on the brane. The transition from 4D to 5D is governed by the crossover scale \(r_c \equiv G^{(5)}/(2G)\).

There are two branches of solutions for the DGP model. The self-accelerating branch \([25]\) can lead to cosmic acceleration purely gravitationally without introducing a cosmological constant. However, this branch of the solution is theoretically unstable \([26]\) and the observed expansion history does not seem to align with the predictions of the self-accelerating DGP model \([27]\). The normal branch is theoretically stable, but it cannot lead to an accelerated Hubble expansion; a trivial negative pressure energy stress component \([28]\) or an extra DE component must be added to match the observational data \([29]\). Nevertheless, the DGP model remains attractive as a benchmark model with the so-called Vainshtein screening mechanism that will be described in detail below, and in this paper, we will only discuss the normal branch of the DGP model (nDGP). The structure formation in the nDGP model is governed by the modified Poisson and scalar equations in the quasi-static and weak field limits \([30]\):

\[
\nabla^2 \Psi = 4\pi G a^2 \delta \rho_m + \frac{1}{2} \nabla^2 \phi, \quad (12)
\]

\[
\nabla^2 \phi + \frac{r_c^2}{3\beta_{dgp} a^2 c^2} \left[ (\nabla^2 \phi)^2 - (\nabla i \nabla \phi)^2 \right] = \frac{8\pi G a^2}{3\beta_{dgp}} \delta \rho_m, \quad (13)
\]

where \(\phi\) is a scalar degree of freedom related to the bending modes of the brane, i.e., it describes the position of the brane in the fifth dimension. The total modified gravitational potential \(\Psi\) is given by \(\Psi = \Psi_N + \phi/2\), with \(\Psi_N\) being the standard Newtonian potential. Again, the fifth force is proportional to the field’s gradient and is given by \(F_{5th} = \nabla \phi/2\). The parameter \(\beta_{dgp}(a)\) is a time-dependent function that depends on the crossover scale \(r_c\) (see Equation 2.2.5 in \([30]\)).

The DGP model gives rise to a cubic interaction \(\sim (\partial \phi)^2 \Box \phi\) in its four-dimensional effective theory \([26]\), with \(\phi\) being the Galileon \([31]\). The Galileon is a scalar field with a shift symmetry (in analogy to Galilei transformation in classical mechanics). The generalized Galileon \([32]\) can be mapped to the Horndeski theory \([33]\). The cubic Galileon model can be reduced to a Horndeski model with the mapping

\[
G_4 = 1, \quad K = -c_2 X, \quad G_3 = c_3 X/M^3,
\]

where \(c_3\) and \(M\) are free parameters in the Galileon model. Note that it is common to fix \(c_2 = -1\).
In this paper, we follow the naming convention for the nDGP model and denote the model as \( NH_0 r_c \), where the product of the Hubble parameter at \( z = 0 \) and the crossover scale characterises the departure from GR. As an example, \( H_0 r_c = 5.0 \) is called the N5 model.

2.3. Screening Mechanisms

As discussed in Sections 2.1 and 2.2, MG models lead to a change in the Poisson equation, which inevitably results in an altered equation of motion for the gravitational acceleration. To pass the constraints imposed within the solar system, such as radio frequency shift [34], the lunar laser ranging constraints [35,36], and the earth-based torsion balance experiments [37,38], one needs solutions that suppress the modification of GR within the solar system. In the examples of early \( f(R) \) models, the problem was that those models introduce an extremely light scalar degree of freedom, producing a long-range fifth force and dissociate the space-time’s curvature from the local density. Therefore, the essence of the screening mechanism is to re-associate the fifth force with the local environment. Suppose we focus on the modified Poisson equation; in this case, the mechanisms can be summarized as (i) thin-shell screening: adding a non-linear scalar potential such that the effective potential becomes density-dependent via the dynamics of the scalar field itself; (ii) kinetic screening: non-linear generalization of the Laplacian operator.

2.3.1. Thin-Shell Screening

In the thin-shell models, the effective potential is usually split into two terms: a bare potential that only depends on the scalar field and an environmental-dependent term \( V_{\text{eff}} = V(\phi) + \rho_m A(\phi) \), with \( \rho_m \) being the environmental density (see Equation (6)). By adjusting the functions \( V(\phi) \) and \( A(\phi) \), the scalar field can acquire a large mass in the high-density regions. Given that the Compton wavelength is inversely proportional to the mass \( \lambda_c \propto m^{-1}_{\text{eff}} \), the larger the mass, the shorter the distance the scalar field can mediate the fifth force. There are two widely studied screening mechanisms, the chameleon mechanism and symmetron mechanism. In the following, we will focus on the chameleon mechanism since it is related to the \( f(R) \) model, presented in Section 2.1. More discussions about the symmetron mechanisms can be found in [39,40].

Chameleon screening can be achieved with certain choices of the function \( f(R) \) [41–44]. A typical choice of the potential for the chameleon mechanism is

\[
V(\phi) \propto \phi^{-n}, \quad A(\phi) \propto e^{\beta \phi/M_{\text{pl}}},
\]

with \( M_{\text{pl}} \) as the reduced Planck mass and \( \beta \) as the coupling strength of the scalar field to the matter. The choice of the potential leads to the modified Poisson equation in Equations (8) and (9), where one can have a quick peek into two opposite regimes of solutions. In the large-field limit, when \( |f_R| \) is relatively large (e.g., in the case of large \( |f_{R0}| \)), the perturbation \( \delta f_R \) is small compared to the background field \( |f_R| \), and \( |\delta R| \ll 8 \pi G \rho_m \), so that the Poisson Equation (8) can be approximated as

\[
\nabla^2 \Psi \approx \frac{16 \pi G}{3} \delta \rho_m a^2.
\]

Comparing Equation (16) with the standard Poisson equation in \( \Lambda \)CDM, the enhancement of gravity is \( 4/3 \) the strength of the standard Newtonian force (see also Equation (8) in [25]). In the small-field where \( |f_R| \) takes very small values, the left-hand side of Equation (9) is negligible and therefore we have \( \delta R \approx 8 \pi G \delta \rho_m \). Plugging this into Equation (8) we recover the standard Poisson equation.

2.3.2. Kinetic Screening

Kinetic screening works by modifying the Laplacian operator in the equation of motion. Two examples of kinetic screenings are the Vainshtein mechanism and the K-mouflage mechanism. In the following, we will briefly overview the Vainshtein mechanism since...
the nDGP model (see Section 2.2) is a representative class of modified gravity models that feature the Vainshtein screening mechanism [45]. More details for K-mouflage mechanism can be found in [46–48].

The difference in thin-shell and kinetic screening can already be seen by comparing Equations (8) and (9) and Equations (12) and (13). Instead of adding a scalar potential, the kinetic screening modifies the Laplacian operator of the equation of motion. To further illustrate how the Vainshtein mechanism works, let us for simplicity consider solutions in spherical symmetry, defining excess mass enclosed in radius \( r \) as \( M(r) \equiv 4\pi \int_0^r \delta \rho_m(r') r'^2 \, dr' \).

Solving Equation (13) we find that the fifth force is given by

\[
F_{5\text{th}} = \frac{1}{2} \frac{d\phi}{dr} = \frac{2}{3 \beta_{dgp}} \frac{r^3}{r_V^3} \left[ 1 + \frac{r^3}{r_V^3} - 1 \right] F_{\text{GR}},
\]

where \( r_V \) is the Vainshtein radius

\[
r_V \equiv \left( \frac{8r_s^2}{9\beta_{\text{dgp}}^2} \right)^{1/3} = \left[ \frac{4GM(R)}{9\beta_{\text{dgp}}^2 H_0^2 \Omega \, c^2} \right]^{1/3},
\]

with the Schwarzschild radius \( r_s \equiv 2GM(R)/c^2 \), \( \Omega \, c = 1/(4H_0^2r_s^2) \), and the total mass of the spherical object \( M(R) \equiv 4\pi \int_0^R \delta \rho_m(r') r'^2 dr' \). On scales larger than the Vainshtein radius \( r \gg r_V \), gravity receives an additional contribution \( 1/3\beta_{\text{dgp}} \) (\( \beta_{\text{dgp}} > 0 \) for the nDGP model). On the other hand, for \( r \ll r_V \) the fifth force \( F_{5\text{th}} \approx 2(r/r_V)^{2/3}/(3\beta_{\text{dgp}}) \) is suppressed.

2.4. Parameterised Frameworks of Gravity

Above we give an overview of specific models of DE/MG. When approaching cosmological observations, it might be useful to first assess in which direction the data constrain gravity by adopting a phenomenological approach. Working with parameterised frameworks of gravity is a powerful and efficient way to test GR with no preference for a given model. In the following, we present the most popular parameterised frameworks.

2.4.1. MG Phenomenological Functions

Let us consider a Universe described by the metric of Equation (1), where perturbations are determined by the potentials \( \Psi \) and \( \Phi \). A straightforward approach to MG/DE theories is to cast all the modifications to the perturbation evolution into two parameterised functions. One can refer to these as modified growth or MG phenomenological parameters. Usually, one describes how the coupling between gravity and matter density is modified, i.e., how it changes the Poisson equation. The other accounts for the variation among the potentials \( \Psi \) and \( \Phi \). In the past decade, various choices of parametrisation have been explored. We review some of the most popular ones.

Following the notation of [49], possible MG parameters can be introduced by modifying the standard equations in the sub-horizon quasi-static approximation

\[
\begin{align*}
(i) \quad Q(a,k) & : \quad -k^2 \Phi = 4\pi Ga^2 Q(a,k) \rho \Delta , \\
(ii) \quad \mu(a,k) & : \quad -k^2 \Psi = 4\pi Ga^2 \mu(a,k) \rho \Delta , \\
(iii) \quad \Sigma(a,k) & : \quad -k^2 (\Phi + \Psi) \equiv 8\pi Ga^2 \Sigma(a,k), \\
(iv) \quad \eta(a,k) & : \quad \eta(a,k) \equiv \Phi \div \Psi ,
\end{align*}
\]

with \( \Delta \equiv \delta + 3aH_0 / k \) as the Gauge-invariant co-moving density perturbation, \( \rho \) as the energy density, and \( v \) as the irrotational component of the peculiar velocity. \( Q \) and \( \mu \) modify the Poisson equation for \( \Phi \) and \( \Psi \), respectively. \( \Sigma \) parametrizes the change in the lensing response to the massless particle in a given matter field, and \( \eta \), the so-called gravitational slip parameter, reflects the non-zero anisotropic tensor. It is sufficient to choose two from
the equations (i–iv) given that they are not independent. In Table 1, we summarise the relation between the gravity models and the phenomenological functions. These functions are identical to ones for GR and the standard DE models (e.g., quintessence). For clustering DE models (e.g., $k$-essence [50,51]), there is no gravitational slip, but the Poisson equations can be modified (e.g., see review [52]). In general, MG models can introduce modifications to the Poisson equation and incorporate anisotropic stress terms with slip parameters that deviate from unity (e.g., Brans–Dicke, $f(R)$, and DGP theories). However, it is important to note that there can be exceptions within specific subclasses of models, such as the generalized cubic covariant Galilean model, which exhibits an equal potential $\Psi = \Phi$ [55]. Nevertheless, it is worth mentioning that the opposite signs of $\mu - 1$ and $\Sigma - 1$ would tend to disfavour all Horndeski models [56]. Moving forward, our focus will centre on $(\mu, \Sigma)$ due to their observational significance. Specifically, $\mu$ holds a connection to the Newtonian potential and can be constrained through galaxy clustering via RSD (see Section 3.2.1). On the other hand, $\Sigma$ is linked to the Weyl potential and can be probed by the lensing and ISW effects discussed in Section 9 (1, 2 and 4), respectively, while the slip parameter $\eta$ often possesses a simple functional form in many models, tending to have weaker constraints from the data [57].

Table 1. Connection between well-known models of DE/MG and the EFT formalism (from Bloomfield et al. [58], Frusciante and Perenon [59]). Here, 0 means that the EFT function is identically equal to 0, ✓ indicates that the function is present, – indicates that the function is not present, while ⋆ indicates that the function is related to other EFT functions. In addition, we also show the relationship between different gravity models and the phenomenological functions $\mu$, $\Sigma$, and $\eta$. For each case, it is sufficient to specify two of these functions, as the remaining functions can be derived from the other two.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\Omega$</th>
<th>$\Lambda$</th>
<th>$\alpha$</th>
<th>$M_0^2$</th>
<th>$M_1^2$</th>
<th>$M_2^2$</th>
<th>$M_3^2$</th>
<th>$m_1^2$</th>
<th>$\mu$</th>
<th>$\Sigma$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Lambda$CDM</td>
<td>0</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$f(R)$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
<td>$\leq 1$</td>
</tr>
<tr>
<td>Brans–Dicke</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
<td>$\leq 1$</td>
</tr>
<tr>
<td>Quintessence</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
<td>1</td>
</tr>
<tr>
<td>$k$-essence</td>
<td>0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
<td>1</td>
</tr>
<tr>
<td>DGP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>–</td>
<td>$\neq 1$</td>
<td>$\neq 1$</td>
<td>$\neq 1$</td>
</tr>
<tr>
<td>Horndeski</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>–</td>
<td>$(\mu - 1)(\Sigma - 1) &gt; 0$</td>
<td>$\neq 1$</td>
<td>$\neq 1$</td>
</tr>
</tbody>
</table>

2.4.2. Effective Field Theory of Dark Energy

Effective field theory (EFT) is a general theoretical technique, first employed in a cosmological scenario to describe inflation [60–62]. It was later applied to describe DE by means of a unifying and model-independent framework, referred to as EFTofDE [58,63]. Indeed, the idea behind EFTofDE is to construct the most general, single scalar field action to be effective, i.e., to be easily interfaced with observations, and unifying, in the sense that it aims to include the highest possible number of DE–MG models as special cases.

The recipe to construct the EFTofDE action can be summarized as follows:

1. Usually, the validity of the weak equivalence principle is assumed a priori. This makes the Jordan frame, where the metric is universally coupled to the matter fields, the best-suited framework. We refer to [63] for details on the Jordan frame and the alternative formulation in the Einstein frame;

2. The action is constructed within the unitary gauge [60,61]. In practice, this means that the perturbation of the extra scalar degree of freedom $\phi$ representing the DE–MG framework is vanishing. This corresponds to foliate the 4D spacetime in 3D hypersurfaces by breaking the time-translation symmetry and fixing a preferred time slicing;
3. The chosen foliation is characterised through the unit vector \( n_\mu \) perpendicular to the time slicing

\[
n_\mu = - \frac{\partial_\mu \phi}{\sqrt{-\hat{g}^{\mu\nu} \partial_\mu \phi \partial_\nu \phi}} = - \frac{\delta_\mu^0}{\sqrt{-\hat{g}^{\mu\nu}}},
\]

where \( \hat{g}_{\mu\nu} \) is the Jordan frame metric. From the unit vector, we can define the extrinsic curvature \( K_{\mu\nu} \) as \( K_{\mu\nu} = \nabla_\mu n_\nu \);

4. We construct the action from all the perturbed operators invariant under the residual symmetry of spatial diffeomorphisms, such as the upper time component of the metric \( \delta \hat{g}^{00} \), the Riemann tensor \( \delta R_{\mu\nu\alpha\beta} \), the Ricci tensor \( \delta R_{\mu\nu} \) and scalar \( \delta R \), the extrinsic curvature \( \delta K^\mu_\nu \) and its trace \( \delta K \);

5. Due to the broken time-translation symmetry, the coefficient of the operators in the action are allowed to be time-dependent functions. We call these parameters EFT functions.

The resulting EFTofDE action in conformal time up to the second order in perturbations is

\[
S_{\text{EFT}} = \int \, d^4x \sqrt{-\hat{g}} \left\{ \frac{M_{\text{Pl}}^2}{2} \left[ 1 + \Omega(\tau) \right] R + \Lambda(\tau) - c(\tau) a^2 \delta \hat{g}^{00} \right.
\]

\[
+ \frac{M_2^2(\tau)}{2} \left( a^2 \delta \hat{g}^{00} \right)^2 - \frac{M_1^2(\tau)}{2} a^2 \delta \hat{g}^{00} \delta K - \frac{M_2^2(\tau)}{2} (\delta K)^2
\]

\[
- \frac{M_3^2(\tau)}{2} \delta K^\mu_\nu \delta K^\nu_\mu + \frac{M_4^2(\tau)}{2} a^2 \delta \hat{g}^{00} \delta R(3)
\]

\[
+ m_2^2(\tau)(\delta \hat{g}^{\mu\nu} + n^\mu n^\nu) \partial_\mu \left( a^2 \delta \hat{g}^{00} \right) \partial_\nu \left( a^2 \delta \hat{g}^{00} \right)
\]

\[
+ \ldots \right\} + S_m \left[ \hat{\psi}^{(i)}_m, \hat{g}_{\mu\nu} \right],
\]

where \( S_m \) is the matter action. There are nine time-dependent EFT functions: \( \{ \Omega, \Lambda, c \} \) that multiply first-order (in perturbations) operators and affect both the background and the perturbation evolution, and \( \{ M_2^2, M_1^2, M_3^2, M_4^2, M_5^2, m_2^2 \} \) for second-order operators that only enter in the evolution of the perturbations. We recover the GR limit when all the EFT functions vanish, with the exception of \( \Lambda \), and the EFT action reduces to the standard Einstein–Hilbert one.

The EFTofDE framework allows for exploring a wide range of models. Within the so-called pure EFT approach, one can test the bare effect of each operator in the action, by parameterising the evolution of the EFT function. Otherwise, one can link the EFT parameters to well-known DE–MG models, i.e., the mapping approach. For example, \( f(R) \) theories of gravity (see Section 2.1) are simply connected to the EFT formalism as follows [63]

\[
\Omega = f_{R}, \quad \Lambda = \frac{M_{\text{Pl}}^2}{2} (f - R f_R), \quad c = 0,
\]

with \( M_2^2 = M_1^2 = M_3^2 = M_4^2 = M_5^2 = m_2^2 = 0 \). The corresponding relations can also be found for the other models considered in this paper. In Table 1 we gather a summary of how the EFT formalism is linked to specific models of DE.

The EFT formalism offers a powerful parameterised framework derived from solid theoretical assumptions. It is widely used in the literature and numerical tools to work within the EFT frameworks are the Einstein–Boltzmann solver EFTCAMB [64] and the Monte Carlo Markov chain sampler EFTCosmoMC [65]. For a complete review of EFTofDE and a summary of the most recent constraints, we refer the reader to [59].
2.4.3. The $\alpha$-Basis Parameterisation

Although the EFTofDE framework allows for exploring a very wide parameter space, it is not always straightforward connecting each EFT function to a specific physical interpretation. An alternative parameterisation that offers a clearer physical meaning is the $\alpha$-basis. First introduced by [66], with the $\alpha$-basis all the modifications of gravity described by Horndeski theories are cast into four independent functions of time: $\alpha_K(t), \alpha_B(t), \alpha_M(t)$ and $\alpha_T(t)$. In the following we summarize the physical meaning attached to each of these functions [66]:

- $\alpha_K$, dubbed as kineticity, is connected to the kinetic energy term in the Horndeski Lagrangian (see Equation (3)) and depends on the functions $K, G_3, G_4$ and $G_5$;
- $\alpha_B$ quantifies the braiding, i.e., the mixing between the kinetic terms of the scalar field and the metric. It depends on the $G_3, G_4$ and $G_5$ functions;
- $\alpha_M$ describes the running of the effective Planck mass and can be easily expressed in terms of Horndeski functions as
  \[ H M_{\text{Pl}}^2 \alpha_M \equiv \frac{d M_{\text{Pl}}}{d t}, \tag{26} \]
  with $H$ being the Hubble parameter and $M_{\text{Pl}}^2 \equiv 2 \left( G_4 - 2 X G_{4,X} + X G_{5,\phi} - \phi H X G_{5,X} \right)$;
- $\alpha_T$ is the tensor speed excess, that quantifies the speed excess of GWs with respect to the speed of light. It depends on the Horndeski functions as
  \[ \alpha_T \equiv \frac{2 X}{M_{\text{Pl}}^2} \left[ 2 G_{4,X} - 2 G_{5,\phi} - (\dot{\phi} - \phi H) G_{5,X} \right]. \tag{27} \]

The $\alpha$-basis parametrization is widely used and is implemented in the publicly available Einstein–Boltzmann solver hi-class [67,68]. For further discussion on this parametrization, including its connection to the EFTofDE formalism, we refer the reader to [59,66].

2.5. Why Do We Still Study $f(R)$ and DGP?

Given the discussion in Sections 2.1 and 2.2, we may now ask ourselves this question: why would we further study $f(R)$ and DGP gravity models given that we still need a cosmological constant $\Lambda$ or DE component to explain the cosmic acceleration? One of the motivations for studying the MG theories was to explain accelerated expansion by replacing the cosmological constant with a more natural model. Now that a large class of self-accelerated models is being ruled out by constraints from large-scale structures and GWs [69,70], MG theories are less attractive from a cosmological perspective.

Nevertheless, these two benchmark models remain meaningful in certain ways. First, they are the few survivors from various tests at different astrophysical scales. At small scales, they can successfully incorporate screening mechanisms to recover Newtonian gravity; at large scales, the enhancement in structural growth that they predict can be allowed by astrophysical probes when properly choosing the corresponding parameters. Second, we need to bridge observations and theory. These two models, with well-developed simulations and analytic templates, are illustrative examples of what MG signals could possibly look like for a given observable, especially when MG-induced features are dominated by systematics of instrumental and astrophysical origin. Finally, the cosmological constant is a phenomenological model that empirically describes cosmic acceleration. It is not surprising that there is room to introduce further parameters to this model, and, in the meantime, maintain this model’s validity while numerous attempts teach us what room is left.

As an alternative to the model-specific approach, the parametrization framework is another way of searching for deviations from GR and high complementary, especially given that there are many MG models and it is hard to study all models exhaustively. We briefly mentioned this idea in Section 2.4.1 and will mention some recent developments below.
2.5.1. Survivors from GWs

In GR, the GWs travel at the speed of light. Many MG models alter this prediction. Therefore, GW can put constraints on the MG models. A simple model-independent parametrization of the speed of the GW is

\[ c_T^2 = c^2 (1 + \alpha_T(t)), \]  

(28)

where \( c_T \) is the speed of the tensor modes, \( t \) is the physical time, and \( \alpha_T \) is given in Equation (27). The difference in the arrival time of the photons and GWs results in a bound of the property function \( \alpha_T(t) \). The property function \( \alpha_T \) can be related to the \( G_4 \) and \( G_5 \) functions defined in Equation (3) [33,66]. Applying the condition \( \alpha_T = 0 \) we have the speed of the GW as the speed of light \( c_T = c \), and at the same time, eliminating part of \( L_4 \) and the entire \( L_5 \) in the Horndeski Lagrangian given by Equation (3). Hence, models that involve quartic and quintic Galileans are ruled out as they invoke the full \( L_4 \) and \( L_5 \), while the cubic Galileans can still survive.

In the same way, we can ask ourselves which EFT functions survive the GW170817 event [4]. Being careful with the adopted notation and formalism, it is possible to map \( \alpha_T \) into EFT functions. In the formalism of this paper this relation would be [58,59,66]

\[ \alpha_T = - \frac{\bar{M}_2^2/M_{Pl}^2}{1 + \Omega + \bar{M}_2^2/M_{Pl}^2}. \]  

(29)

If \( \alpha_T \) is suppressed by the GW detection, this implies that the \( \bar{M}_2 \) function is vanishing. This simple condition on the EFT theories allows for a better understanding of the implications of GW detection on the surviving theories. In the \( f(R) \) case, \( \bar{M}_2 \) is zero (see discussions around Equation (25)); therefore, GW does not constrain the \( f(R) \) model.

Another impact of MG is through the running Planck mass \( \alpha_M \), linked to the \( \Omega \) parameter in the EFT formalism. In the tensor sector, this acts as a friction term in the equation of motion of GWs. A running Planck mass can thus change the GW amplitude and lead to a difference in the luminosity distance by a GW and its electromagnetic counterpart [71]; at the same time, it can also move around the amplitude of the primordial peak of the B-mode spectrum [72]. As the tensor/scalar ratio \( r \) also shifts the peak amplitude, there is a degeneracy between the \( r \) and \( \alpha_M \). This degeneracy can be lifted by combining with probes sensitive to the scalar sector (e.g., see Section 3 for lensing or integrated Sachs–Wolfe effect) since a running Planck mass can affect background evolution and structure growth [72].

Ref. [73] studied the GW luminosity distance in \( f(R) \) gravity with simulated mock data for an Einstein telescope (ET) [74]. They found that the ET-like data in the first running decade could only provide constraints for \( |f_{R0}| < 10^{-2} \) and was not helpful in constraining the \( f(R) \) gravity. In the case of DGP models, the GW amplitude could “leak” into higher dimensions. The leakage could cause additional GW damping during its propagation. Therefore, constraints can also be obtained by comparing GW and EM luminosity distances. The authors of Ref. [75] performed a study on the DGP model and found that the model was very poorly constrained by the GW170817 event, where only wavelengths comparable to the cosmic horizon (very low frequency) could leak into extra dimensions.

2.5.2. Further Parametrization

Let us review the idea of parametrizing MG effects. The parametrization scheme in Equations (20)–(22) is built on cosmological perturbation theories and is applicable only at linear regimes. The authors of [76] took the post-Friedmann approach [77,78] and developed a parametrization framework which can be applied to all scales. Going beyond the post-Friedmann equation at the leading order, the authors re-grouped the first-order post-Friedmann equation into a form that separates the large- and small-scale limit. By neglecting some additional terms corresponding to “intermediate scales” (we will clarify this term shortly), the parametrization can be performed analogously to linear parametrization.
To understand what gravity models could be parameterized by this approach, let us take one step back and examine the GR-ΛCDM scenario. In GR-ΛCDM, on large scales one can use the cosmological perturbation theory (PT) and on small scales one can use Newtonian gravity; the potential arises from the cosmological PT (Equation (1)) and coincides with the Newtonian potential (Equation (2)), while there are no “intermediate scales” at which neither of the theories apply. The approach in [76] applies to most GR models that do not have an intermediate scale, including $f(R)$ and cubic Galilean models. In a companion paper [79], the authors implemented this idea into $N$-body simulations and demonstrated the framework for weak lensing (WL) observables.

While the parametrization approach allows one to assess more general classes of models, connections back to physics models can become ambiguous as the phenomenological behaviour of the models can be non-unique. This approach can nevertheless provide a null test of the standard model of cosmology and a general test of gravity.

Lastly, an interesting alternative model-independent approach that does not rely on fixing a parametrization is gravity reconstruction [80–84]. Latest results from [85] show how it is possible to restrict the theory space directly from cosmological observations, reconstructing the redshift evolution of the MG functions via numerical interpolation. The author of [85] developed a sophisticated machinery, prompting its use with future high-precision observations.

3. Observational Probes of Gravity Models

Modifying GR can impact the expansion history of the Universe and the process of structure formation, while the expansion history inferred via geometric probes, such as standard rulers from baryon acoustic oscillation [86,87] and distances of type-Ia supernovae [5,6], can be indistinguishable from that of standard GR, the physics behind structure formation can be substantially modified.

At mildly non-linear to linear scales ($O(10) - O(100) h^{-1}\text{Mpc}$), the theory of gravity can be described by the following quantities: the Weyl potential $\Phi_\gamma(k,t)$, the density field $\rho(k,t)$, and the velocity field $v(k,t)$. As shown in Figure 2, these three quantities are not fully independent: the potential and density contrast field are connected via the Poisson equation, the potential and velocity field are linked through the Euler equation, and the time derivative of the density contrast field is connected to the velocity field through the continuity equation. Modifications to the gravitational potential can affect the photon trajectories emitted from the CMB or galaxies, create secondary temperature anisotropies in the CMB photons, and induce new galaxy-clustering patterns. At non-linear scales ($\sim O(10^{-2}) - O(1) h^{-1}\text{Mpc}$), the screening mechanisms have further influence on the Poisson equations at the characteristic screening scales, and essentially modify the potentials, matter distribution, and velocity field. Hence, the screening effects can lead to modified density profiles, environmental-dependent density distributions, and cluster abundance and probability density distributions.

Therefore, we divide the astrophysical probes into the following three categories based on basic physical quantities: potential, density, and velocity. There are derived quantities from each of these categories, such as convergence, density contrast field, and temperature anisotropies. Further, one can construct summary statistics to extract the information. In this section, we will start with the relevant observables for the various astrophysics probes derived from the basic physical quantities described above. Then we will further discuss how the different datasets can be summarized to detect modifications of gravity.
3.1. Derived Observables

3.1.1. Potential-Based Derived Observables

From the current constraints, we expect the cosmological potential perturbations in Equation (1) to be very small, implying that linear approximations tend to be valid. Here are some simple interpretations of the potentials: \( \Psi \) governs the dynamics of non-relativistic objects, and therefore determines the structural growth. The Weyl potential, \( \Phi_\gamma = (\Phi + \Psi)/2 \), determines the null geodesics (light propagation), since \( \Phi_\gamma \) is conformally invariant at linear order.\(^7\)

(i) **Temperature fluctuation.** As the CMB photons travel between the last scattering surface and the observers, their wavelengths are altered when they travel through the time-varying gravitational potentials. The integrated Sachs–Wolfe(ISW)\(^8\) effect measures the decay of gravitational potential due to cosmic expansion. The integrated temperature anisotropies can be expressed as\(^3\)

\[
\frac{\Delta T}{T_{\text{CMB}} \text{ ISW}}(\hat{\mathbf{n}}) \approx - \int \frac{d}{d\tau} (\Phi + \Psi) \, d\tau,
\]

where \( T_{\text{CMB}} \) is the average temperature of the CMB background. Here, the two potentials \( \Phi \) and \( \Psi \) are allowed to be different, with \( \tau \) as the conformal time.

(ii) **Deflection of photon trajectories.** In the presence of MG, the modified potential can source different inhomogeneous density distributions between the photons sources and the observer compared to the ones given by GR. The additional inhomogeneity of the density field thus deflects the trajectories of photons. When the detected photons are emitted by the last scattering surface, the derived observables are CMB lensing(see reviews\([89,90]\)). Alternatively, when the photons are emitted from distant source galaxies, the galaxy shapes can be subtly distorted by intervening mass distribution and are hence named cosmic shear or galaxy WL (see review\([91]\)).

---

**Figure 2.** Connections between the basic physical quantities (boxed in grey) and derived observables (boxed in red). The derived observables can form different combinations as observational probes in terms of summary statistics, topological properties, and environmental dependence (boxed in dotted blue). We also highlight the probes sensitive to the screening mechanism in blue. They are useful tools in probing deviations from GR. More details see Section 3.
The deflection of photons can be expressed in terms of the lensing potential $\psi$, the integral of the Weyl potential $\Phi$, along the photon trajectory from its source to the observer

$$\psi(\hat{n}) = \frac{2}{c^2} \int_0^{\chi_s} d\chi \frac{\chi_s - \chi}{\chi_s} \Phi(\chi_0, \chi),$$

(31)

where $\chi_s$ is the co-moving radial distance to the photon source. A spatially flat universe and the Born approximation are assumed. Observables, such as the deflection angle and the isotropic and anisotropic distortions, can be expressed in terms of derivatives of the lensing potential. For example, the isotropic component of the distortion matrix, the convergence $\kappa$, is given by

$$\kappa(\hat{n}) = \frac{1}{2} \nabla^2 \psi(\hat{n}) = \frac{1}{c^2} \int_0^{\chi_s} d\chi \frac{\chi(\chi_s - \chi)}{\chi_s} \nabla^2 \Phi(\chi_0, \chi),$$

(32)

where $\nabla_{\hat{n}} \equiv \chi \nabla_{\perp}$ are the gradients with respect to the angular position and the co-moving transverse distance, respectively. If the perturbations in the potential $\Phi$ are varying on scales much smaller than $\chi_s$, the longitudinal component of the Laplacian $\nabla^2_{\parallel}$ does not contribute to the integral in Equation (32). This thin-lens approximation allows substitution of the full Laplacian $\nabla^2$ for its transversal part $\nabla^2_{\perp}$ in Equation (32), thus applying the Poisson equation to cast the convergence as an integral over the mass density contrast, weighted by the lensing efficiency $\chi(\chi_s - \chi)/\chi_s$ (see review [92]).

3.1.2. Density-Based Derived Observables

(i) Density contrast field. As can be seen from the Poisson equation, the metric perturbations directly source the density contrast field given by

$$\delta(x) \equiv \frac{\rho(x) - \bar{\rho}}{\rho}.$$

(33)

The density contrast field is a Gaussian random field at the leading order. Hence, its two-point statistics capture most of the information in the field. Late-time gravitational interaction induces non-linearities, and information in the field thus spreads beyond two-point statistics. However, the matter field is not a direct observable. Instead, it is traced by luminous tracers such as individual galaxies or diffuse gases. As galaxies or large-scale structures are only formed when the density crosses a certain threshold, the contrast of the field of galaxies and the matter are related via galaxy bias (see review by [93]). Furthermore, due to the galaxies’ peculiar velocities, the observed galaxy positions are distorted by the velocity fields, known as redshift space distortions (see Section 3.1.3).

In addition to overdensities, underdense regions of the Universe, $\delta_u$, contain complementary information (see Figure 2). Due to the screening mechanisms, MG effects are only suppressed in high-density regions, while density profiles of underdense regions are modified with respect to GR: the centre will become emptier as more mass outflows towards their high-density surroundings [94,95]. In addition, the lensing signal is also modified due to the modified lensing potential [96,97]. Moreover, underdense regions have an additional advantage, where properties are largely insensitive to the baryonic and galaxy formation physics [98]. Three-dimensional spherical underdense regions are usually recognized as cosmic voids. One typical way of identifying voids is using a rectangular-grid-based spherical void finder [99]. Alternatively, there are other approaches to quantify underdense regions, e.g., watershed or Voronoi-based algorithms [100,101]. In 2D, the spherical void finder also has its projected version over the line-of-sight (LoS). When underdense regions are identified via circumspheres of triangular Delaunay tessellation cells, they are “tunnels” [102]; alternatively, if underdense regions are identified by placing random circles of identical radii in the sky plane, they are “troughs” [103].
3.1.3. Velocity-Based Derived Observables

(i) **Redshift space distortions.** MG forces leave strong imprints on the matter velocity fields. On sufficiently large scales, galaxies trace matter velocity fields and there is no velocity bias between the galaxies and matter distribution. The galaxies’ peculiar velocities add an additional component to the galaxies’ redshifts and lead to an anisotropic clustering pattern, which is known as redshift space distortion (RSD). The observed density contrast field \( \delta_g(s) \) is in redshift space, thus receiving a velocity correction that is parallel to the LoS direction. In the linear regime and under the distant observer approximation, the density contrast field in redshift space is given by

\[
\delta_g(s) \approx \delta_g(r) - H^{-1} \frac{\partial v_\parallel}{\partial r_\parallel},
\]

with the co-moving Hubble scale \( H = aH \), the LoS-parallel velocity \( v_\parallel = v \cdot \hat{n} \), and the real and redshift space positions are related as \( s = r + v_\parallel \cdot \hat{n} / H \).

(ii) **Temperature anisotropies in CMB Photons.** As CMB photons travel through the Universe, they can be inverse Thompson scattered by hot ionized gas. The induced shifts in the photon temperature are referred to as Sunyaev–Zeldovich effects. In the case of the kinetic Sunyaev–Zeldovich (kSZ) effect [104], the shift in photon temperature is caused by the bulk motion of the ionised gas in clusters. The shift temperature can be expressed as

\[
\left( \frac{\Delta T}{T_{\text{CMB}} \Delta z} \right)_{\text{kSZ}}(\hat{n}) = \left( \frac{\sigma_T \rho_{g0}}{\mu_e m_p} \right) \int_0^{z_*} dz \left( \frac{1+z}{H(z)} \right)^2 e^{-\tau(z)} q \cdot \hat{n},
\]

where \( \sigma_T \) is the electron Thomson cross-section, \( \rho_{g0} \) is the mean gas density at redshift \( z = 0 \), \( \mu_e m_p \) is the mean mass per electron, \( z_* \) is the redshift at which reionization ends, \( \tau \) is the Thomson optical depth, \( \Theta \) is the electron ionization fraction which depends on the primordial helium abundance and the number of ionized helium electrons [105,106], and \( q \equiv (1 + \delta) v \) is the density-weighted peculiar velocity of the electron (bulk peculiar velocity of ionized regions or clouds). The dot product \( q \cdot \hat{n} \) in Equation (35) implies that only the LoS-parallel component velocity contributes to the temperature anisotropies. Given that the velocity is proportional to the Fourier modes \( k \) of matter density contrast \( \tilde{\psi} \propto k \) (we use \( \sim \) to denote Fourier space quantities) in linear theory, only \( k \) modes parallel to the LoS can contribute to the anisotropy. However, since Fourier modes parallel to the LoS direction cancel with each other after the projection, the angular correlation at smaller scales mainly receives contributions from components perpendicular to the LoS direction \( q_{\perp} \).

(iii) **Direct peculiar velocity measurements.** The radial component of the peculiar velocity of a galaxy can be directly estimated when both the redshift and redshift-independent distance estimates are available. Distance estimates can be obtained via well-known correlations of galaxy properties such as the Tully–Fisher relation for spiral galaxies [107,108] and the fundamental plane for ellipticals [109]. These relations provide distances to galaxies with 20% uncertainties for redshifts reaching \( z \sim 0.1 \). Another alternative to obtain distances is using type-Ia supernovae, which have smaller intrinsic scatters in luminosity after standardisation and can provide distances with 7% uncertainties. The statistics of the peculiar velocities and their cross-correlation with the galaxy density field can yield tighter constraints on the growth rate at \( z < 0.1 \), where cosmic variance limits the constraints from RSD.

3.2. Two-Point Statistics

The lowest-order summary statistic is the two-point statistics. They can capture most information in a mildly non-Gaussian field. In principle, one can make any combinations
\(2C_N \equiv N!/(N-2)!/2!\) from the derived quantities discussed in Section 3.1 to form a
two-point statistics in Fourier space
\[
\langle \mathcal{X}(\mathbf{k}_1), \mathcal{Y}(\mathbf{k}_2) \rangle \equiv (2\pi)^3 \delta_D(\mathbf{k}_1 - \mathbf{k}_2) P_{xy}(k_1),
\]
or in configuration space
\[
\langle X(x_1), Y(x_2) \rangle \equiv \zeta_{xy}(x_1 - x_2)
\]
where \(\langle \ldots \rangle\) denotes a statistical ensemble average. One can compute auto-correlation when
\(X\) and \(Y\) are the same tracers and cross-correlation when \(X\) and \(Y\) are different tracers. In
the following, we will discuss both cases.

3.2.1. Redshift Space Power Spectrum

From Equation (34) we can construct the galaxy power spectrum in redshift space. At linear order at large scales, the galaxy power spectrum receives an LoS-dependent boost due to the velocity-dependant term and becomes anisotropic [110]

\[
P^g_s(k) = (b_1 + f\mu^2)^2 P_m(k),
\]
with \(\mu = \hat{k} \cdot \hat{n}, b_1\) as the linear bias and the linear growth rate as \(f = d\ln D/d\ln a\). In GR, the growth rate can be accurately described by \(f(z) = \Omega_m^0\) with \(\gamma \approx 0.55\) (with a weak dependence on other parameters) [111,112]. The overall amplitude of the matter power spectrum in the later epoch depends on the amplitude of primordial scalar power spectrum \(A_s\). The combination \(f(z)\sigma_8(z)\), with \(\sigma_8\) being the root-mean-square of the linear density field smoothed with a top-hat filter of \(8\ h^{-1}\)Mpc, has also been one of the most common methodologies applied in current surveys to constrain deviation from GR (see Section 5.5 and Figure 16), particularly at large scales (\(\sim O(20) - O(150)\)h\(^{-1}\)Mpc). Alternatively, this amplitude-induced degeneracy can be re-cast as the ratio between \(f\) and the galaxy bias and expressed as \(\beta \equiv f/b_1\), with \(b_1\) as the galaxy linear bias.

The power spectrum modelling is often performed following a perturbative approach (see Section 3.5), where most of the perturbation models are developed for GR-only scenarios. It is thus important to validate the GR-based galaxy-clustering models on the mock catalogues. The authors of [113] studied the power spectrum of \(N\)-body simulation with \(f(R)\) gravity. The authors of [114] applied the BOSS wedge analysis pipeline to the \(N\)-body simulation with nDGP. The authors of [115] further extended the study to the \(f(R)\) and nDGP models.

However, the large-scale RSD test requires relatively high statistical precision to detect deviations from GR. For example, the authors of [115] studied mock simulations for a BOSS-like survey and found significant deviation for DGP models while the potential for \(f(R)\) is less promising. Figure 3 shows the impact of different gravity models on the measured galaxy two-point correlation when decomposed into Legendre polynomial basis. The quadrupole moment, which encodes the leading-order anisotropic information, is most affected by the MG effects due to the peculiar velocities. There is an obvious deviation of the N1 model from GR, while the rest of the models are all within the 1\(\sigma\) standard deviation. Future surveys, such as DESI, will improve statistical precision, but models with smaller deviations from GR would still remain challenging for large-scale RSD [116].

Given that MG models, such as \(f(R)\), can produce large deviations at small non-linear scales, RSD effects in these regimes are important. The authors of Ref. [117] studied the imprints of \(f(R)\) gravity for small-scale RSD (< \(O(10)\)h\(^{-1}\)Mpc) and found no evidence of \(f(R)\) gravity in BOSS data. However, the analysis needs to assume certain relations between the circular velocity profile and the effective halo mass, while the time-dependent screening mechanism could evade this assumption. Nevertheless, the small-scale information is valuable and helps to distinguish different gravity models.
Figure 3. Impact of MG on the galaxy two-point correlation function multipoles at a snapshot $z = 0.5$. For visibility, the multipole moments $\ell = 0, 2, 4$ are shifted by a constant, respectively. GR is in black with error bars inferred from the standard deviation of five simulations. Coloured curves denote different MG models. A clear deviation from GR can be seen for the quadrupole moments $\zeta_2$, while other models are consistent within the 1σ error bar of GR. (Credit: Hernández-Aguayo et al. [115]).

In addition to the galaxy auto-correlation, cosmic voids as large underdense regions of the Universe are also affected by the RSD. Since the underdense regions are less screened, they are expected to be more sensitive to the fifth force [118–120]. One may further expect that the modelling of the galaxy–void cross-correlation is simpler as the dynamics around voids are more linear. Linear RSD models were studied in [121–123] developing a linear RSD model for galaxy–void cross-correlation. The cross-correlation has been used to infer the growth rate of cosmic structure [121,124–127]. Yet, the modelling relies on the following assumptions when mapping between real and redshift space: the void number conservation and the void centre position invariance. Additionally, one also needs to assume that the real space galaxy density and velocity fields around the void centres are isotropic [128]. The potential breakdown of these assumptions boils down to the unknown void centre in real space for realistic applications. Based on these concerns, the authors of [129] applied the reconstruction methods to restore the galaxies’ real space positions before applying the void finders, followed by an extended modelling [128] as a complementary approach to the non-reconstructive analysis.

3.2.2. Angular Power Spectrum for LoS-Integrated Observables

The angular power spectrum is practical in studying projected quantities. The angular power spectrum for LoS-integrated observables at linear order with Limber approximation [130] can be written as (see, e.g., Equation (45) in [11])

$$C_{XY}^{\ell} = \int \frac{dz}{\chi^2(z)} \frac{H(z)}{c} \left[ F_X(k,z)F_Y(k,z)P(k) \right] \bigg|_{k=\ell/\chi(z)},$$  

(38)

where the kernels $F_i(k,z)$ for tracers $i = X$ or $Y$ consist of the scale-dependant growth factor $D(k,z) = \delta(k,z)/\delta(k,0)$. In this paper, we will show the expressions under linear approximations and specify the power spectrum $P(k)$ for each tracer separately in the following. In practice, various non-linear effects need to be accounted for in the modelling.
The power spectrum at large scales is sensitive to MG effects, as screening mechanisms work less efficiently at large scales with low density. For potential-induced observables, it is also practical to introduce a scale-dependent variable $G$ that quantifies deviation from GR:

$$G(k, z) \equiv \frac{1}{2} |1 + \eta(k, z)| \mu(k, z)(1 + z) D(k, z),$$

with $\eta$ and $\mu$ defined in Equations (21) and (22). In the following, we will discuss four types of angular power spectra for LoS-integrated observables.

(i) **Galaxy lensing.** One advantage of galaxy WL is that it can access smaller-scale information relative to galaxy-clustering analysis. In contemporary galaxy WL analyses measure two-point statistics of the observed shear field, which is the anisotropic distortion induced by the lensing potential introduced in Equation (31) (e.g., [131–137]). In photometric surveys, source galaxies are binned into tomographic redshift bins, and the correlation of the shear field within these tomographic bins, as well as their cross-correlations, are estimated with a selection of two-point statistics. Common examples of employed two-point statistics are correlation functions, (pseudo) angular power spectra, or complete orthogonal sets of E-/B-mode integrals (COSEBs) [138]. All these statistics can be related to the angular power spectrum of the shear field through linear transformations. For first-order, the angular power spectrum of the shear field is equivalent to that of the convergence given in Equation (32) (see, e.g., Equation (6.32) in [92]. The power spectrum is the matter power spectrum at redshift today $P(k) \equiv P_{\text{mm}}(k, z = 0)$, and the kernel $F_\kappa(k, z)$ that enters the angular power spectrum Equation (38) can be written as

$$F_\kappa(k, z) = G(k, z) \frac{3H_0^2 \Omega_m}{2a(z)} \chi(z) \int_z dz' n_i(z') \frac{\chi(z') - \chi(z)}{\chi(z')},$$

where $n_i(z)$ is the distribution of source galaxies in tomographic redshift bin $i$.

As both a higher matter fraction and more clustered matter cause an increase in the lensing signal, there is a strong degeneracy between the constraints on the matter fraction $\Omega_m$ and the mass fluctuation $\sigma_8$ derived from WL data. For galaxy WL, the parameter combination

$$S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}$$

quantifies the well-constrained direction in the $\Omega_m - \sigma_8$ space.

Figure 4 shows the lensing convergence power spectrum at two source redshifts $z_{\text{source}} = 0.5$ and 1.0 for the F5 model and in comparison to GR, where all MG models have more power compared to the GR case [139]. There is a power enhancement of 10–30% for $10 < \ell < 10^4$ (corresponds to $O(10^{-1}) - O(10^2)h^{-1}\text{Mpc}$) due to MG, with an enhancement increasing towards smaller scales.

Current stage 3 surveys measure cosmic shear with high precision: the detection significances are in the range of 20–30$\sigma$, depending on the analysis and employed scale cuts [135–137]. These high-precision measurements require careful mitigation of astrophysical and observational systematics to avoid biases during parameter inference.

One challenge in galaxy lensing analysis is the contamination by intrinsic alignment of galaxies. The observed quantity in galaxy WL is the ellipticities of source galaxies. If the orientations of these galaxies are uniformly distributed, then the observed ellipticity of a galaxy, despite being very noisy, is an unbiased estimate of the shear imparted by gravitational lensing. Alignment of galaxies with the large-scale structures they reside in, as well as the structures that cause the gravitational lensing, can induce a coherent correlation of galaxy ellipticities that mimics the effect of gravitational lensing. This intrinsic alignment of galaxies is one of the leading astrophysical systematics in analysing galaxy WL data; modelling the intrinsic alignment ultimately depends on details of galaxy formation,
while direct observation of the effect in the faint population of galaxies used in WL surveys is very challenging.

Figure 4. Upper: Convergence power spectrum (see Equation (38) with kernel Equation (40)) for GR (black) and F5 model \( |f(R)| = 10^{-5} \) at two source redshifts \( z_{\text{source}} = 0.5 \) (dashed) and \( z_{\text{source}} = 1.0 \) (solid). Lower: fractional difference in the convergence power spectrum for GR and F5 model (Credit: [139]).

Another concern is modelling the non-linear scales and baryon feedback. The LoS kernels in Equation (40) have very broad supports, reaching from the observer up to the source galaxy. While the kernel peaks roughly halfway between the source and the observer, this broad support means that in principle, any angular scale receives contributions from all physical scales from the matter power spectrum. In practice, these contributions can be small but they still need to be accounted for to avoid biases in the modelling. This sensitivity to smaller scales, the historically small areas covered by galaxy lensing surveys, as well as the observational challenges in analysing WL at larger scales, cause WL analyses to push deep into the non-linear regime. At these scales, N-body simulations, or phenomenological models based thereon, are required to model the non-linear structure formation. While a range of models and emulators are available for the matter power spectrum in GR, this is generally not the case for MG. Furthermore, galaxy formation processes, such as the feedback from active galactic nuclei, can redistribute large amounts of gas, significantly impacting the matter power spectrum (e.g., [142]). Modelling these feedback processes relies again on the details of galaxy formation, while observational data is challenging for the redshift and mass ranges relevant for galaxy WL.

On the observational side, photometric redshift uncertainties are one of the biggest concerns in galaxy WL analyses. The galaxy samples from which the shear is estimated are typically faint. This makes it challenging to find representative spectroscopic samples from which to estimate the redshift distribution \( n_i(z) \) of the \( i \)th tomographic bin of the photometric source sample, while current surveys use a combination of sophisticated methodologies to address this problem (e.g., [143,144]), the greater depth and statistical power of future surveys (see Section 5.1) still poses significant challenges to the accurate estimation and calibration of their photometric redshifts.
Shape estimation is another challenge. Galaxy WL relies on accurate estimation of galaxies’ ellipticities or shears. However, most galaxies used for WL are faint, noisy, pixelated, and—in the case of ground-based surveys—poorly resolved due to the large atmospheric point spread function (e.g., [145,146]). As the depth of surveys increases, galaxies increasingly overlap with their neighbours. This blending further complicates the accurate estimate of the shape of single galaxies.

(ii) CMB lensing. As CMB photons travel along the LoS, they are deflected by the gravitational potential gradients associated with the large-scale structure in the late-time epoch. The CMB lensing thus indirectly traces the underlying matter distribution. Although CMB lensing and galaxy lensing share common features, the broad CMB lensing kernel guarantees its sensitivity to higher redshifts. Unlike galaxy lensing, the source redshifts of CMB lensing are almost exactly known; CMB lensing is thus free from photometric redshift uncertainties.

In the case of CMB lensing, one also inputs the power spectrum of matter (the same as for galaxy lensing), and the source distribution $n(z)$ in Equation (40) is given by a Dirac delta function at the redshift of the last scattering $z_*$. The resulting kernel $F_{\kappa_{\text{CMB}}}(k,z)$ peaks at a redshift of $\sim 2$, compared to galaxy WL with $z \sim 0.5$, and is thus less sensitive to the (non-linear) growth of structure at lower redshifts [147]. This difference in redshift sensitivity results in a different degeneracy direction in the $\Omega_m - \sigma_8$-plane, and can be approximated by $\sigma_8(\Omega_m/0.3)^{0.25}$ [148]. While the CMB lensing measurements are noise-dominated at smaller scales, the large sky area covered by CMB experiments results in a high detection significance of 40$\sigma$ [148].

One main observational challenge in CMB lensing analysis is foreground contaminations. There are contaminations by various sources: thermal Sunyaev–Zeldovich effect (tSZ [149]), cosmic infrared background (CIB), kSZ effect, and galactic dust or extinction. These foreground contaminations, which can bias cross-correlation measurements, are particular important for CMB experiments with high angular resolution (e.g., ACT [150]). Among these foregrounds, tSZ is of particular importance. It originates from hot electron gases in galaxy clusters that scatter off the CMB photons and can correlate with galaxies at low redshifts. CIB could be another source of contamination since it is also peaked at $z \sim 2$ (e.g., [151]). Various foreground mitigation schemes are adopted in the CMB analysis pipelines (e.g., [148,150]).

On the modelling side, CMB and galaxy lensing share similar concerns in non-linear gravitational evolution [152] and baryonic effects. In particular, there are simulations and mitigation schemes for baryonic effects studied in [153–156]. In addition to the observational and theoretical challenges, quadratic estimators (QE [157,158], which are widely used to recover CMB lensing potential, could be sub-optimal themselves. QEs are numerically efficient but are only statistically optimal when the lensing effect on the primary CMB power spectrum is small. This is the case at large scales ($\ell < 3500$) but breaks down for smaller scales ($\ell \gg 3500$), and can be sub-optimal for high-sensitivity experiments [159]. The reader can find more details and alternatives in the works [159–163].

(iii) Combined probes analyses. The two-point statistics of various combinations of different fields can be combined into $N \times 2\text{pt}$ analyses. A common joint analysis consists of the combination of the two-point statistics of the lensing fields, the galaxy density fields, and their cross-correlations. These $3 \times 2\text{pt}$ analyses form the backbone of current and future photometric surveys (e.g., [164–166]). Jointly analysing the different probes helps to break parameter degeneracies due to different sensitivities of the respective probes to both cosmological parameters, as well as astrophysical and observational parameters. For example, in the linear regime, the two-point statistics of the galaxy density field scale as $\propto b^4\sigma_8^2$, while the cross-correlation of galaxy density and lensing—called galaxy–galaxy lensing—and cosmic shear scale as $\propto b\sigma_8^2$ and $\propto \sigma_8^2$, respectively, since lensing does not depend on the galaxy bias $b$. This combination of probes can therefore break the degeneracy between galaxy bias and clustering.
amplitude. Similarly, observational systematics in lensing, such as multiplicative shear biases, affect different probe combinations to varying degrees and therefore can be handled more robustly in joint analyses than in a single-probe analysis. Combined with the high signal-to-noise of these measurements (e.g., \( \sim 90\sigma \) for DES-Y3 [166]), this allows for tight parameter constraints.

Such joint 3 \times 2\text{pt} analyses can be extended with other combinations of fields. For example, joint analyses of the two-point statistics of galaxy lensing, galaxy density, CMB lensing, and their cross-correlations are often referred to as 6 \times 2\text{pt} analyses (e.g., [167]). Cross-correlating CMB lensing (a projected quantity, tracing matter directly via the potential) and late-time galaxy (3D quantity, tracing matter indirectly via density fluctuation up to a galaxy bias) is very powerful in probing structure growth [151,168–171]. The cross-correlation between CMB lensing and galaxies of different types also helps to break parameter degeneracy between bias and amplitude fluctuations [172]. Further, the cross-correlation of two independent experiments is also free from any correlation in instrumental noise or systematics.

An example of combining probes designed to be sensitive to deviations from GR is the \( E_G \) statistic [173]:

\[
E_G \equiv \frac{P_{\delta \nabla^2 \Phi \gamma}}{P_{\delta \theta}} \approx \frac{1}{2} (1 + \eta) \mu \frac{\Omega_{m,0}}{f(z)} = \frac{\Omega_{m,0}}{f(z)}, \tag{42}
\]

where \( P_{\delta \nabla^2 \Phi \gamma} \) is the cross-correlation between the galaxy positions and the lensing map, \( P_{\delta \theta} \) is the cross-correlation between the galaxy and velocity divergence field \( \theta \equiv \nabla \cdot v \). At linear order, galaxy bias drops out by taking the ratio of the cross-power spectrum. The \( E_G \) probe is interesting because it is sensitive to the lensing potential and the RSDs simultaneously at larger scales; the relative lensing amplitude and the assembly rate of large-scale structures will differ in the presence of MG. In the case of GR, the effective parameters \( \mu = \eta = 1 \) (see Equations (21) and (22)).

The \( E_G \) statistic was first applied to galaxy WL [173–177]. The authors of [178] proposed to apply \( E_G \) to CMB lensing and forecasted the cross-correlation with the BOSS sample. The authors of [179] applied \( E_G \) statistics to the cross-correlation between the Planck CMB lensing map [148] and eBOSS DR16 QSO sample [180,181]. Figure 5 shows the consistency of the \( E_G \) statistic with the scale-independent \( \Lambda \text{CDM} \) prediction across the five scale bins.

![Figure 5. \( E_G \) statistic by cross-correlating the Planck CMB lensing map and eBOSS DR16 QSO sample for the Northern Galactic Cap (NGC, blue), Southern Galactic Cap (SGC, orange), and the combined sky (NS, red). The green solid line denotes the \( \Lambda \text{CDM} \) using Planck [182] with the shaded area being 1\( \sigma \) uncertainty from the simulation. The \( E_G \) probe is consistent with the scale-independent \( \Lambda \text{CDM} \) prediction across the five scale bins. (Credit: [179]).](image-url)
ISW and cross-correlation with galaxy. The ISW effect is sensitive to the time-dependent Weyl potential and can thus be used to constrain $\Sigma$ in Equation (22). During the expansion of the Universe, gravitational potential decays. At the same time, the $f(R)$ gravity enhances structure growth below the Compton wavelength, and photons thus become colder (decrease in angular power spectrum). Therefore, the ISW effect affects the larger scales by suppressing the power for low-multipole moments. The sensitivity at larger scales also implies that the ISW effect is cosmic variance-limited due to the limited number of modes [183]. While a direct measurement of the ISW effect from the CMB spectrum is very small [184], expected to be measurable through the cross-correlation with the large-scale structure. The authors of [185] proposed the correlation of the X-ray survey and the CMB anisotropy measurements, e.g., the authors of [186] performed a combined analysis for galaxy surveys, radio survey, and hard X-ray counts and found a 4.5$\sigma$ detection significance of the ISW signal. In the ISW–galaxy cross-correlation case, the power spectrum is given by

$$F_{\text{ISW}}(k,z) = 3T_{\text{CMB}} H_0^2 \Omega_{m,0} k^{-2} \frac{\partial G(k,z)}{\partial z},$$

while the galaxy kernel is

$$F_g(k,z) = b(z)\Pi(z) D(k,z),$$

with $b(z)$ as the galaxy bias and $\Pi(z)$ as a normalized galaxy selection function. Alternatively, expressing the ISW kernel in terms of the ratio between the Newtonian and curvature potential can be used to test the $\eta$ function (see Equation (22)). The authors of [184,186] used the ISW–galaxy cross-correlation to constrain $\Sigma$.

The ISW–galaxy cross-correlation has been used to test different gravity and DE models [187–191], and was forecasted to have constraining power for sources selected from radio continuum surveys [192]. Figure 6b shows the impact of the $f(R)$ model on the ISW–galaxy cross-correlation power spectrum for different Compton wavelength parameters $B_0 \equiv B|_{z=0}$ (see Equation (17) in [21]). The data points are from one galaxy subsample of SDSS (see Section 5 for more details on the survey). Comparing the $\Lambda$CDM prediction (blue), and two parametrizations of the $f(R)$ model (red), the difference is most visible at larger scales with low-multipole moments $\ell \lesssim 100$. The authors of [193] used the ISW effect to constrain models which attempt to replace DE with local-inhomogeneity-sourced backreactions [194,195].

While the ISW amplitude can be used to probe the MG effects, the amplitude degenerates with the galaxy bias. A robust measurement thus relies on careful calibration of the galaxy bias and contamination fraction of the sources [184]. When cross-correlating the ISW effect with galaxies, one can still be affected by galaxy shot noise and imperfect overlap between the galaxy and ISW kernels in redshift [196]. Additionally, ISW is a linear probe at larger scales and is not sensitive to screening effects.

kSZ power spectrum. The amplitude of the kSZ angular power spectrum strongly depends on the normalization of matter perturbations and is thus sensitive to the growth of structure. As mentioned in Section 3.1.3, the photon temperature is shifted by the bulk motion of the ionized gas, and thus the kSZ effect is also sensitive to the velocity field. The kSZ angular power is given by taking the expectation value of Equation (35). As discussed previously, the angular power spectrum only receives contribution components perpendicular to the LoS direction as the LoS-component averages out due to the integral at linear scales. The power spectrum $P$ is given by the Vishniac power spectrum [197], which requires modelling the power spectrum of the density field, velocity field, as well as the cross-correlation of the density–velocity fields [198–200]. The kernel is given by
\[
F_{\text{KSZ}} = \frac{\sigma_T p_0 \Theta}{\mu_e m_p} (1 + z)^2 e^{-\tau(z)} \frac{DD}{D_0},
\]

with dots denoting the derivative with respect to the proper time. Compared to other secondary anisotropies of the CMB, the kSZ effect is weak, with detections of the kSZ power spectrum and cross-correlations with galaxies only in the 3–5\(\sigma\) range \[201–205\].

![Image 1](https://via.placeholder.com/150)

\[\text{(a)}\]

**Figure 6.** (a) Ratio of the kSZ angular power spectrum in the presence of MG compared to GR. The angular power spectrum computed using MGCAMB \[206\] and MGHaloFit \[207\] for \(|f_0^R| \in [5 \times 10^{-6}, 5 \times 10^{-5}]\) (crosses), while a fitted model is the solid curve (see Equations (16) and (17) and Figure 4 in \[208\]). (Credit: \[208\]). (b) Impact of \(f(R)\) gravity on the ISW–galaxy cross-correlation function. The \(\Lambda\)CDM prediction is given in blue, and the modified gravity models with different Compton wavelength parameters \(B_0 = 0.1\) and \(B_0 = 0.5\) are in red dashed curves. The grey data points are from the SDSS Luminous Red Galaxy (LRG) sample \[209\] (see \[189\] for data points from other SDSS samples). The ISW effect in CMB is suppressed in the presence of MG (reproduced from \[189\]).

The authors of \[210\] pointed out that galaxy cluster velocities can be used as gravity tests via the kSZ effect. Figure 6a shows the angular power spectrum model for different \(f(R)\) gravities, where the linear growth rate is from MGCAMB \[206\] and the matter power spectrum is computed using MGHaloFit \[207\]. A fitted model is shown as the solid curve on top of the model curve (shown as crosses). As pointed out in \[208\], the amplitude of the kSZ power spectrum depends on the redshift of the reionization epoch. Further, the power spectrum scales as the square of the ionization fraction (see \(\Theta\) in Equation (35)), and underestimating the helium ionization fraction can lead to underestimation of the power spectrum. Moreover, the baryon physics can reduce the gas density in halos and counteract the MG effects \[105\]. The authors of \[211\] studied the kSZ effect for both the effect of the nDGP and \(f(R)\) models using simulations with full physics, such as star formation, cooling, stellar and black hole feedback, and realistic galaxy populations. Using these simulations, they found that both \(f(R)\) and nDGP models can enhance power by boosting the abundance and peculiar velocity of large-scale structures. At the same time, the sub-grid baryonic physics can suppress the electron transverse momentum (see Figures 6–8 in \[211\]).

### 3.2.3. Power Spectrum of Line Intensity Mapping

Line intensity mapping (LIM or IM) \[212\] has recently become a competitive cosmological probe. It has the potential to fill the gap between CMB and galaxy-clustering observations. LIM measures the integrated emission of spectral lines from multiple galaxies and the intergalactic medium in exceptionally large 3D volumes. A strong point of LIM is its tomographic nature, which allows probing the matter distribution from the present time up to the epoch of reionization (EoR, \(z \sim 10\)). Therefore, LIM observations may be able to explore the Universe at unprecedented redshifts not accessible by other LSS probes.
The LIM signal is a biased tracer of the underlying matter distribution at different redshifts. The observable is the power spectrum of the brightness temperature ($T_b$) perturbations. At a given redshift, this can be expressed as [212]

$$P_{\text{LIM}}(k, z) = \bar{T}_b(z) b(z) P_m(k, z) + P_{\text{SN}}(z), \quad (46)$$

where $\bar{T}_b$ is the mean brightness temperature of the line (see, e.g., Equation (10) in [213]), $b$ is the bias of the considered probe, and $P_m$ and $P_{\text{SN}}$ are the matter and shot noise power spectra, respectively. Both the bias and the shot-noise term are dependent on the line luminosity function. In the literature, both the angular power spectrum and power spectrum multipoles are widely used. We refer to Bernal et al. [214] for a review on how to model the LIM signal.

The most studied line is the 21 cm signal arising from the spin–flip transition in the neutral hydrogen ground state [213,215–218]. During the reionization epoch at high redshifts, the Compton wavelength of the scalar-field perturbations can still be considered in the linear regime. Thus, the fifth force is yet unsuppressed before the screening effects undertake the extra structure growth, where the modified growth rate results in modified brightness temperature [219,220]. Hence, 21 cm cosmology offers a possibility of probing MG effects at high redshifts. At lower redshifts, neutral hydrogen HI-dominated galaxies also emit 21 cm (e.g., [221]). Figure 7 shows theory-predicted cross power spectra between CMB temperature and 21 cm in response to different gravity models at $z = 0.35$, demonstrating changes of 10–25% [222]. Rotational lines of carbon monoxide (CO) [223–225], fine structure lines of ionized carbon (CII) [226,227], as well as Lyman-α, Hα, and Hβ lines [228–231] are also interesting spectral features that complement the 21 cm line.

Observationally, LIM suffers from strong continuum foreground contamination, including diffuse galactic synchrotron emission, bright point sources, and atmospheric turbulence, which can potentially degrade the predicted constraining power. Several foreground removal techniques are currently being proposed and tested [232–239]. Overall, LIM is suited to probe gravity at higher and lower redshifts. It additionally allows to constrain BAO feature evolution through a wide range of time, thus also constraining the background evolution [212]. The LIM power spectrum is expected to be sensitive to variations in the background equation of state [240]. Although only a few detections are available so far [227,241–246], there is a plethora of ongoing and planned experiments (see Section 5.3).

![Figure 7](image_url)

**Figure 7.** Cross-correlation between the 21 cm and CMB temperature for different gravity models relative to the $\Lambda$CDM model. The ratio between the $\Lambda$CDM model to itself is the solid line, while the ratio of $f(R)$ models are shown in dashed curves (reproduced from: [222]).
3.3. Higher-Order Statistics and Morphology of the Structure

In this section, we focus on the gravitational non-linearities induced by MG. In general, higher-order statistics are expected to be more sensitive to the MG effects than two-point statistics as they are more sensitive to non-linear effects. Higher-order statistics are also helpful when combined with two-point statistics in breaking galaxy bias degeneracies in models beyond GR.

3.3.1. Galaxy NPCFs/Polyspectra

Non-linear gravitational evolution induces coupling between different modes and leads to non-Gaussianity. Information is thus spread beyond the two-point statistics. Higher-order statistics thus help to break the parameter degeneracies and can be more sensitive to non-Gaussianity when gravity is enhanced. As a generalization of two-point statistics, polyspectra or $N$-point correlation functions (NPCFs) measure the probability excess of finding a galaxy $N$-tuplet over a uniform distribution (see, e.g., [247]). The joint probability $dP_N$ of finding a galaxy in both volume elements $dV_i$ for $i = 1, \ldots, N$ is given as

$$dP_N = \bar{\rho}^N (1 + \bar{\zeta}^{(N)}) dV_1 \ldots dV_N. \quad (47)$$

Taking $N = 3$ as an example, the total 3PCF $\bar{\zeta}^{(3)}$ is composed of the reducible 2PCF contribution and a connected contribution, where the connected contribution (reduced 3PCF) is given by

$$\bar{\zeta}(r_{12}, r_{13}, r_{23}) = \langle \delta(r_1) \delta(r_2) \delta(r_3) \rangle. \quad (48)$$

Despite the numerous studies that have presented 3PCF measurements [248–253], a brute force measurement of the NPCF poses computational challenges. For $N_g$ galaxies, while a brute force approach has the complexity that scales as $O(N_g^N)$, an algorithm with the complexity of $O(N_g^2)$ was proposed for the 3PCF [254] and was extended to the arbitrary integer $N \geq 4$ [255–257]. The algorithm relies on the radial-angular decomposition and can either be applied to discrete data points or gridded data. In the latter case, the computation can alternatively be carried out by a fast Fourier transform (FFT) algorithm for the 3PCF [258], which has complexity $O(N_m \log N_m)$, with $N_m$ being the number of mesh grids. An extended FFT version was implemented in [259], including the projected 3- and 4PCF. These improvements drastically accelerated the speed of the NPCF estimators in configuration space and also opened up new avenues for probing fundamental physics [260–262].

Polyspectra are Fourier counterparts of NPCFs. When applied to surveys with realistic geometry and observational effects, they encode complementary information to each other. The bispectrum as the Fourier counterpart of the connected 3PCF is defined as

$$\langle \delta(k_1) \delta(k_2) \delta(k_3) \rangle = (2\pi)^3 \delta_D(k_{123}) B(k_1, k_2, k_3), \quad (49)$$

with $k_{123} \equiv k_1 + k_2 + k_3$. In Fourier space, most early studies on bispectrum [263–267] were limited to certain choices of triangular configurations due to computational challenge. Efficient FFT-based algorithms were developed to make use of all $N$-tuplet configurations [268,269].

The authors of [270] studied the matter power spectrum and bispectrum for $f(R)$ simulations. They found that the MG-sensitivity of the bispectrum depends on the setup of the initial power spectrum. Nevertheless, the bispectrum helps break degeneracies in bias and MG models. The authors of [98] measured three-point statistics of HOD catalogues built from MG simulations run with $N$-body codes (see Section 4). Figure 8 shows the impact of $f(R)$ and nDGP models on the 3PCF and real-space bispectrum. The left panel shows the difference of 3PCF between GR (black dashed curve) and the MG models by fixing $r_1 = r_2 = 3 h^{-1} \text{Mpc}$ only. Apart from F6 and N5, all other models show significant deviations from
GR at these smaller scales. Moreover, the results are insensitive to the HOD parameters. However, the deviations decrease towards larger scales $\sim 5 \, h^{-1}\text{Mpc}$. The right panel shows the comparison of GR and the $f(R)$ model for the bispectrum monopole. The difference between the MG models and GR is quantified by $\Delta B \equiv (B_{\text{MC}} - B_{\text{GR}}) / \sqrt{\sigma_{\text{MC}}^2 + \sigma_{\text{GR}}^2}$. For F4 and F6, $\Delta B \sim 3$ and 1.6, respectively. Interestingly, the difference between F5 and GR is even smaller than F6, which is likely driven by the HODs.

Higher-order statistics are shown as a promising tool for probing MG effects. They are more sensitive to MG signals compared to the two-point statistics towards smaller scales. However, modelling non-linear scales poses challenges (see Section 3.5). At larger scales, higher-order statistics usually have a lower signal-to-noise ratio compared to the two-point statistics. Observational effects such as treatment of non-trivial survey geometry are also more complicated compared to the two-point statistics.

Figure 8. Impact of MG on three-point statistics. Left: difference in the 3PCF between GR (black dashed curve) and different MG models (coloured dashed curve) by fixing $r_1 = r_2 = 3 \, h^{-1}\text{Mpc}$ only. The grey-shaded region represents the standard deviation of the five HOD GR catalogues. Right: upper panel shows the real-space bispectrum for GR (black) and the $f(R)$ model (coloured). Vertical lines correspond to the $k$-bin interval $\Delta k = 0.025 \, \text{Mpc}^{-1}$. The lower panel shows the difference between GR and the MG models normalized by a combined standard deviation $\sigma_{\text{MG+GR}}$, computed as a root-sum-squared error of the two models (reproduced from: Alam et al. [98]).

3.3.2. Convergence-Derived Quantities: Moments of Aperture Mass and Peak Counting

From the convergence (see Equation (32)) one can construct various quantities to be applied in conventional lensing analyses (see Section 3.2.2). Apart from these quantities, one can construct the aperture mass $M_{\text{ap}}(\theta)$ [271,272]

$$M_{\text{ap}}(\theta) = \int d^2\theta' U_\theta(|\theta' - \theta|) \kappa(\theta'),$$

(50)

where $\theta$ is a 2D vector, and $U_\theta(|\theta|)$ is a cylindrical filter function with aperture radius $\theta$. The statistics of aperture mass and its higher-order moments are sensitive to non-Gaussianities in the WL signals in low redshifts.

The variance of the aperture mass $\langle M_{\text{ap}}^2(\theta) \rangle$ is given by the integral over the convergence power spectrum. The third- and fourth-order moment of the aperture mass map, the skewness $\langle M_{\text{ap}}^3(\theta) \rangle$ and kurtosis $\langle M_{\text{ap}}^4(\theta) \rangle$ [273] are complementary to the variance of the aperture mass, which are more sensitive to non-Gaussian information. Skewness quantifies the asymmetry of a random variable’s probability distribution around its mean, and kurtosis describes the “tailedness” of the distribution with respect to a Gaussian one. In theory, they are more sensitive to the deviation from GR than the lower moment counterpart. However, it is also hard for these higher-order moments of aperture mass to achieve a higher discrimination efficiency (the ability to distinguish between different models), in particular in the presence of massive neutrinos.

Peak counting is another powerful statistic in probing non-Gaussianities [274]. Lensing peaks represent local regions of high convergence, a useful proxy for halo abundance (see Section 3.4.2). A peak is defined as a pixel with a larger amplitude than its eight
neighbouring pixels and exceeds the threshold $n_{\text{thres}} \equiv M_{\text{ap}}(\theta)/\sigma(\theta)$, with $\sigma(\theta)$ as the root-mean-square of a given angular scale $\theta$. The peaks trace the largest halos; they probe the tails of the mass function (see Figure 8 and 9 of [275]). Figure 9 shows the impact of the F5 model on the peak count of the aperture mass probability distribution function (PDF) with a threshold $n_{\text{thres}} = 3$, where the peak counts appear to be insensitive to neutrino effects [276]. However, this conclusion depends on the choice of the $|f_{R0}|$ value, and degeneracy with neutrinos and $\sigma_8$ persists for a different MG parameter choice. The authors of [277] also pointed out that peak counting displays degeneracy between $|f_{R0}|$ and $\sigma_8$. Although one can not recover GR by only varying $\sigma_8$, it demonstrates complicated cosmological parameter dependencies with MG effects. Therefore, understanding modelling at non-linear scales is equally important here. Further, when peak counts are applied to the lensing convergence map, the peak positions can be affected by shape noise; source–lens clustering (correlation between source galaxies and lensing potentials; [278]) and intrinsic alignment (see galaxy lensing analysis in Section 3.2.2) are also important systematic effects that need to be studied in the future.

Figure 9. Peak count of aperture mass. The PDF has a clean separation for the GR and MG cases, even in the presence of massive neutrinos (reproduced from [276]).

3.3.3. Wavelet Scattering Transforms

$N$-point functions provide natural extensions to two-point functions to extract information from the non-Gaussian density field. However, modelling $N$-point functions is theoretically challenging (see Section 3.5). Moreover, $N$-point functions involve averages over powers of the density field, which increase the variance of their estimators and turns them into less informative summaries of the density field.

As an alternative to $N$-point functions, wavelet scattering transforms (WSTs) have been introduced in cosmology to analyse both WL data [279,280] and clustering from spectroscopic surveys [281]. The authors of [282] used the Fisher forecast framework and assessed the performance of WSTs. They found that WSTs can improve the cosmological parameter constraints by several factors relative to the power spectrum. The 3D WST stacks a series of operations: convolutions of an input field and its modulus. The convolution operation selects scales, and the modulus operation converts field fluctuations to their local amplitude. The transformation takes an input field, $I_0(x)$, and produces a first-order output field $I_1(x)$. Then it takes the field $I_1(x)$ from the last step as an input to output a second-order field $I_2(x)$. This process can be iterated to generate the $n$th-order field

\begin{align}
I_1(x) &= \left| I_0(x) * \psi_{j_1 l_1} \right|, \\
I_2(x) &= \left| I_1(x) * \psi_{j_2 l_2} \right|, \\
&\vdots
\end{align}

\[(51)\]
where the convolution kernel $\psi_{j,l}$ is defined by a family of wavelets labelled by indices $j$ and $l$, where $j$ defines the scale of the wavelet and $l$ its orientation. A family of wavelet $\psi_{j,l}$ can be found by dilating and rotating the mother wavelet.

The expected values of the fields, $S_n$, are called $n$th-order scattering coefficients, which summarize the 3D information contained in the input field

$$S_n = \langle I_n \rangle,$$

where $\langle \ldots \rangle$ represents the spatial average of the field. Due to its iterative nature, wavelets and $N$-point statistics are highly tied with each other. For example, the first-order scattering coefficient is similar to the power spectrum, while the second-order scattering coefficient can be interpreted as “clustering of clusters” [280], thus encoding part of the four-point statistics information.

In [280], the authors found that wavelet coefficients can obtain constraints very similar to those of the bispectrum on Rubin-like WL data. For 2D data, the authors of [283] introduced wavelet phase harmonics where the harmonic phase operator only acts on the complex phase of a field while keeping the amplitude of the field unchanged. The authors of [284] applied the low-dimensional wavelet phase harmonics to LSS data and numerically forecasted that they provide more stringent constraints on cosmological parameters relative to the joint constraints by power spectrum and bispectrum.

Given that wavelet convolutions are defined locally, this summary statistic has the potential to constrain deviations from gravity that incorporate screening mechanisms on smaller scales. Still, there are currently no applications of the WST to gravity theories other than GR. The major theoretical challenge regarding the WST is that it is hard to analytically predict the WST coefficient due to the modulus operation in the estimator (Equation (51)), and parameter inference relies on simulation-based approaches.

### 3.3.4. Topological Tools

Topological data analysis (TDA; see review, e.g., [285] and [286,287]) refers to a collection of statistical methods that extract information from the structure in data. Often TDA only refers to a particular method, called persistent homology, while a broader definition of TDA includes a large class of data analysis methods that uses notions of shape and connectivity [285]. In this section, we will discuss Minkowski functionals (MFs) and persistent homology as two representative applications of characterising LSS with topological information. They provide complementary information to the $N$-point statistics.

These topological methods can be applied to continuous or discretized cosmological datasets, both in 2D and 3D. When analysing discretized galaxy catalogues, one can construct kernel density or apply tessellation tools to convert the catalogue into a continuous density field. Alternatively, one can also expand the radii of circles (2D) or spheres (3D) centred on each galaxy such that the circles form different topologies as the radii changes.

MFs, sometimes called genus statistics, are a set of statistics that characterize the topological properties of a collection of data points (see early applications to galaxy samples [288,289]). It converts the density contrast map into isodensity contours and includes quantities such as surface area, volume, curvature, and the Euler characteristic.

At linear scales, since the pattern of matter distribution is scale-dependent and does not change in co-moving coordinates, the genus curve of LSS can be used as a standard ruler as it is insensitive to gravitational evolution, galaxy biasing, and RSDs [290]. When including smaller scales, MFs can identify MG models in which structures grow with different rates on different scales [291,292]. The authors of [277] applied MFs to the convergence field. The three MFs in 2D are $V_0$ (area), $V_1$ (total boundary length), and $V_2$ (integral of geodesic curvature along the contour), where $V_2$ is equivalent to the genus statistics and is tightly connected to the peak counts (see Section 3.3.2). Figure 10 shows the response of the three MFs $V_0$, $V_1$, and $V_2$ to the $f(R)$ gravity [277]. In the case of the F5 model, there can be a maximum deviation by $\sim 10\%$ from $\Lambda$CDM.
Figure 10. Impact of $f(R)$ gravity on the Minkowski functionals (MFs) of the convergence field $K$ for $V_0$ (area), $V_1$ (total boundary length), and $V_2$ (integral of geodesic curvature along the contour). The $x$-axis is the signal-to-noise ratio of the smoothed convergence field. The MFs (coloured dots) are the average of 100 simulations with the error bars being the standard error of the mean $\sigma_{\text{std}} / \sqrt{100}$. The analytic prediction for a Gaussian random field (GRF) is shown in the black dashed curve (reproduced from [277]).

While MFs are mathematically well defined, they are inherently non-local due to the integrated quantities, thus posing challenges when handling overlapping or percolating objects [293]. At the same time, MFs are tightly connected to the rich framework studied in TDA, the persistent homology $^{16}$.

Topological analysis encodes complementary information into $N$-point statistics. However, modelling topological characteristics is highly non-trivial. In the case of MFs, while it is possible to obtain analytic predictions beyond GRF approximations, practical challenges require further investigations. For example, the authors of [302] explored the impact of galaxy weights and survey geometries, there they found the analysis is relative stable with respect to these variations, while the impact of RSDs is at percent level on quasi-linear scales. In the case of PDs, there are no analytic predictions for PDs to date (to our knowledge). Moreover, there has been no application of persistence diagrams to MG effects so far. Nevertheless, persistence diagrams can have considerably more information than Euler characteristics, since PDs additionally offer information on how long the specific structures persist.

3.3.5. One-Point Probability Density Distribution

The PDF of the smoothed 3D matter density field [303], also known as the counts-in-cells statistics, describes the likelihood of a given environment density as the outcome of the hierarchical structure formation process. For the density field in the early Universe, this PDF can be described by a Gaussian distribution that is characterized by the variance of the field at a smoothing scale used to define an environment, directly related to the matter power spectrum. However, gravitational evolution introduces non-Gaussianities in the late-time PDF. Contrary to its name, the one-point PDF of the density field encodes information about beyond the two-point statistics.

For the late-time density field, the matter PDF has been shown to be a highly complementary probe of the cosmological parameters to the matter power spectrum within
\( \Lambda \)CDM [303], where the authors also showed that analytical models of the matter PDF obtain sub-percent accurate predictions for smoothing scales of the order of tens of megaparsecs.

Recently, the authors of [304] included the effects of MG in the analytical predictions, achieving a similar accuracy to those obtained in \( \Lambda \)CDM. The authors showed that chameleon theories mainly impact the skewness of the distribution (see Figure 11), given that different density fluctuations evolve under varying gravity conditions. Alternatively, they found that in Vainshtein screening theories the enhanced linear structure growth increases the variance of the distribution while producing more under- and overdense structures (heavier tails), compared to the standard cosmology. In [304], the authors showed through a Fisher forecast the potential of this statistic to increase the significance of a MG signal, with improvements of up to six times compared to the power spectrum alone.

**Figure 11.** Global shape of the matter PDF for F5 (triangle) and \( \Lambda \)CDM (square) at redshift \( z = 0 \) (blue) and \( z = 1 \) (green). The measurement from the simulation is shown as the dashed curve and the model prediction is the solid curve. The \( f(R) \) gravity modifies the skewness of the distribution and leads to an asymmetric enhancement over \( \Lambda \)CDM (reproduced from Cataneo et al. [304]).

However, this statistic has not yet been used to constrain deviations from GR using real data. In practice, the PDF can be estimated from either the 3D galaxy field or 2D WL convergence maps. For modelling gravitational non-linearity, the framework of large-deviation statistics [303,305,306] has been applied to study matter density PDFs and achieved promising results. However, modelling galaxy biasing, galaxy stochasticity, and RSDs still remains a challenge. At the same time, one-point PDFs can be highly complementary to density split statistics (see next paragraph) that match the galaxy density PDF and matter density PDF quantile-by-quantile.

### 3.3.6. Nearest Neighbour Distributions

The \( k \)-nearest neighbour cumulative distribution function (\( k \)NN-CDF) was introduced as an informative summary statistic to constrain cosmology in [307]. It is defined as the cumulative distribution function of distances from a set of volume-filling random points to the \( k \)-nearest tracers, while the two-point correlation function is dominated by overdense regions, the \( k \)NN-CDF is sensitive to all environmental densities since the random points have been chosen to be volume-filling instead of mass-weighted.

In [307], the authors showed that the \( k \)NN-CDF of a set of tracers is related to integrals of its \( N \)-point correlation functions, which implies that the information content of \( k \)NN is related to that of all orders of the correlation function. In [308], the authors found that combining the large-scale power spectrum with \( k \)NN measurements to summarize clustering on smaller scales can improve the constraints by a factor of 3 for \( \sigma_8 \) and 60% for \( \Omega_m \) over the power spectrum only analysis. To assess the information content of \( k \)NNs on smaller scales without an accurate theoretical model, the authors of [307,309] performed a Fisher analysis and found that \( k \)NNs can improve the constraints on the cosmological
parameters compared to the monopole of the two-point correlation function, by breaking degeneracies present in two-point clustering.

In [308], the authors developed a theoretical model to predict \( k \)NN-CDF statistics using hybrid effective field theory (HEFT) to fit two-point correlation functions and \( k \)NN statistics on scales larger than 20 \( h^{-1} \) Mpc, based on averages of the underlying field smoothed on different scales. On larger scales, all CDFs tend to one, and therefore the information from the BAO scale is unfortunately lost. Currently, there are no extensions of this modelling approach to MG theories.

3.4. Environment-Sensitive Estimators

As described in Section 2.3, modifications to GR in the solar neighbourhood are strongly constrained. For an alternative theory of gravity to be viable on larger scales, a so-called screening mechanism needs to be incorporated to recover GR in high-density (high-curvature) regions. To test the existence of screening mechanisms we focus on the effect that the fifth force has in unscreened regions. In this section, we will describe different probes of gravity that leverage the environmental dependence of screening mechanisms to expose unscreened structures and set constraints on the magnitude of the fifth force.

3.4.1. Density Split Statistics

Given that extreme environments have shown promise at constraining MG theories that incorporate environment-based screening mechanisms, recent works have developed summary statistics that can extract information simultaneously from a range of environments from both lensing [310–312] and clustering data [313,314].

In [310–312], the authors used a foreground (low redshift) population of galaxies to divide the sky into patches of equal size but distinct galaxy densities. Then, the background (high redshift) population of galaxies was used to measure the tangential shear around the patches with different densities. The stacked shear signals are tracers of the average profile of density contrast around different environment densities. More recently, this idea was extended to 3D clustering studies [313,314], where environmental densities were estimated using the 3D positions of galaxies. To estimate the density contrast, the galaxy field is smoothed with a tophat or Gaussian filter and then sampled at randomly selected positions. These positions are then ranked according to the filtered density contrast, and split into \( n \) bins. For \( n = 5 \), each bin is also denoted as a “quintile”. The random points on each quintile are then cross-correlated with the redshift space positions of the galaxies. The cross-correlation between the random points belonging to the lowest-density quintile is similar to the void–galaxy cross-correlations, whilst those on the higher end have similar properties to the cluster–galaxy cross-correlations.

By modelling the clustering dependence on the environment and exploiting the effect of RSDs, it was shown that the density split statistics can constrain the cosmological parameters in a \( \Lambda \)CDM Universe with much higher accuracy than the two-point correlation function [314]. In particular, they found that the constraints on the sum of neutrino masses can increase a factor of eight relative to the two-point statistics.

To date, there are no studies of the constraining power of density split RSDs in MG scenarios, but we expect that the environmental dependence would make it a strong test of gravity. One challenge in using the density split technique is in modelling these statistics. First, the definition of quintiles requires smoothing the density field, which can potentially mix the small- and large-scale information. Second, analogous to void identification (see Section 3.4.3), RSDs can produce systematic offsets in identifying different environments, leading to an ambiguity in the quintile definition. Third, as in the case of other two-point statistics, the standard weights are calibrated with respect to the two-point statistics, and these weights can in fact be non-optimal for the density split method. It can generally be hard to model the scale mixing, RSD-induced quintile definition, and residual systematics, while other non-analytic approaches such as simulation-based methods (see Section 4.3) could potentially be interesting to investigate further.
3.4.2. Galaxy Clusters

Galaxy clusters form from the largest peaks in the initial overdensity field (see, e.g., [315]). The fifth force enhances gravity and faster growth and collapse of dark matter halos at non-linear scales. Therefore, MG leaves imprints on galaxy clusters’ abundance, density profile, and internal gas distribution below the Compton wavelength, below which the mediation of the fifth force is allowed. We will discuss the three aspects in the following.

**Abundance:** The co-moving halo number density per virial mass $M$ interval in logarithmic binning $n_{\ln M} \equiv d \ln n / d \ln M$ can be predicted by a spherical collapse model [316]. Structures collapse when the density field exceeds a given threshold $\delta_c$, and a generic prediction is that the massive halo abundance is enhanced in the presence of a fifth force. Figure 12 shows the fractional enhancement of the halo mass function in F5 gravity at three redshift bins [317]. In the case of $f(R)$ chameleon gravity, when the scalar field value is larger compared to the gravitational potential ($|f_R(0)| \gtrsim 10^{-5}$, large field limit), the generic prediction holds and the $\Lambda$CDM-predicted $\delta_c$ is a good description to the ones for the $f(R)$ gravity. In contrast, when the scalar field value is smaller compared to the gravitational potential ($|f_R(0)| \lesssim 10^{-5}$, small field limit), the Compton wavelength shrinks so that the fifth force mediation range becomes shorter, the chameleon mechanism modifies the collapsing threshold $\delta_c$, and reduces the deviation at the high-mass end [318]. Given the non-linear threshold $\delta_c$, cluster counts also contain information about the higher-order moments of the density field (see Section 3.3.2).

![Figure 12](image)

**Density profile:** The impact of the MG effect on the cluster density profile can be split into three regimes: (i) inside the virial radius of halos $\sim O(1)\ h^{-1}\text{Mpc}$; (ii) a few virial radii $\sim O(1 - 10)\ h^{-1}\text{Mpc}$; (iii) splashback radius $\sim O(100)\ h^{-1}\text{Mpc}$. In regime (i) dark matter distribution in the presence of $f(R)$ gravity is described by the Navarro–Frenk–White (NFW [319]) profile. The halo concentration is mass- and environmental-dependent and can be modified relative to the concentration of $\Lambda$CDM halos. The authors of [320] found that there exists a universal fitting formula for the halo concentration of different redshifts and halo masses. In regime (ii) the halo density profile is enhanced compared to GR. The density profile can be probed by measuring, e.g., the projected mass distribution of cluster–galaxy lensing [321]. In regime (iii) for the scales approaching the initial infall stage of the clusters, the splashback radii [322] of the galaxy cluster probes the transition from the screened to the unscreened regime. Accreted material from the unscreened regimes can carry a memory of the fifth force and different dynamics [323].
Internal gas: There are two important observables for clusters, the lensing mass \( M_L \propto r^2 \nabla \Phi(r) \) determined by the gravitational lensing and the dynamical mass \( M_D \propto r^2 \nabla \Psi(r) \) determined by the temperature and pressure of gases measured through X-ray or SZ effects. An inequality between the two metric potentials \( \Phi \) and \( \Psi \) indicates MG. Comparing the lensing and dynamic mass, one can place constraints for chameleon \( f(R) \) gravity [324], Galilean gravity [325], and beyond Horndeski theories [326].

There are a few challenging problems regarding galaxy clusters. The galaxy number counts rely on the mass-observable scaling relation to estimate the clusters’ masses, where the latter is not directly observable. This mass-calibration problem thus involves calibrating the scaling relation and a model that transforms the theoretical mass distribution to the distribution of the clusters [327–330]. These calibration steps can have a strong impact on the cosmological parameters, including the growth rate [331]. Systematics such as mass profile reconstruction can induce spurious MG detection [332]. Generally, disentangling astrophysical processes at cluster scales from gravitational effects is non-trivial.

3.4.3. Properties of 2D and 3D Underdense Regions

In underdense regions, the fifth force may be totally or partially unscreened. Precisely, the fifth force counteracts Newtonian gravity and drives underdensities to expand faster and grow larger [95]. Therefore, basic properties of underdense regions, such as their abundances and density profiles can be powerful in constraining gravity models. In Section 3.2.1 we discussed how the 3D spherical underdense regions can be used to infer the growth rate. In this section, we will further discuss the basic properties of underdense regions identified with more flexible definitions.

Void abundance distribution (number of voids) is usually plotted as a function of void size. The void abundance distribution can be predicted through the excursion set formalism [333], which has been extended to the chameleon and symmetron models of gravity [119,120,334]. Although intuitively there should be more voids in the presence of the fifth force, halos are more likely to form in underdense regions, and the abundance distribution can thus not be a unique indicator of the fifth force. The abundance distribution can be combined with other void properties.

Void density profiles depict the density field as a function of radial distance towards the voids centre [95]. The density contrast field for voids approaches \( \delta \to -1 \) towards the void centre but increases towards the void radii. For smaller voids, there are usually ridges at the void radii which are associated with the surrounding filaments and walls. For larger voids, the ridges are less pronounced. At radii much larger than the size of the voids, the density fluctuations again approach zero. Overall, the density profiles of these underdense regions are modified with respect to GR: the centre will become emptier as more mass outflows towards their high-density surroundings [94,95].

Since the convergence of the WL map is tightly related to the projected matter density (see Equation (32)), the void density profile also manifests itself in the properties related to the void lensing, such as convergence profile, the tangential shear profile [335], as well as the mean excess surface density profile around each void [96,102,336].

Compared to the 3D underdense regions, the 2D regions have a higher signal-to-noise ratio [102,103]. The authors of [98] forecasted the signal-to-noise ratio for the \( f(R) \) and the nDGP model for a DESI-like survey with an LSST-like imaging survey. The forecast used 3D voids as well as 2D tunnels; they found that the tunnels’ tangential shear signal has a higher amplitude and is also more sensitive to the MG models (due to the project effect, the interior area of tunnels has lower density compared to 3D voids), 2D tunnels can thus better distinguish MG from GR compared to the 3D voids.

Although voids are promising probes in constraining gravity, there are a few caveats and challenges. First, the void property can be dependent on the void-finding algorithm, and the optimal void finder is model- and void-statistics-dependent [102,336]. Next, conclusions drawn from simulations can also be dependent on the void-identification: The authors of [96] compared voids identified using both dark matter field and dark matter
halo. The abundance of larger DM-identified voids is boosted due to the fifth force, while
the effect is less pronounced for halo-identified voids. The authors of [334] studied a more
realistic scenario by populating a simulated dark matter halo with galaxies and found the
statistical precision is not enough to distinguish between MG models and GR. The authors
of [337] showed that the void statistics are degenerate with HOD parameters [337], e.g., the
void size function and velocity-related properties. Furthermore, galaxies are often used to
identify voids observationally, where galaxies of different types populate halos differently.
Since halo masses can have a different impact on clustering properties, the resulting void
catalogues, which depend on the minimum halo mass cut, can have different abundances
and void profiles [95,338]. As a complementary approach, peaks of the void lensing map
can also be used to identify voids, also resulting in a higher signal-to-noise ratio relative to
galaxy-identified voids [335,339]. Finally, baryon effects [340,341] and massive neutrinos
also need to be taken into account [342].

3.4.4. Marked Correlation Function

Due to the screening mechanism, MG can induce a non-standard density dependence
in the GR scenario. Therefore, up- or down-weighting the density according to their
environments can enhance the estimators’ environmental sensitivity. A marked correlation
function (mCF) is a generalized 2PCF. It requires weighting each object by an environmental-
dependent “mark” \( m \) [343]:

\[
\mathcal{M}(r) \equiv \frac{1}{n(r)m^2} \sum_{ij} m_i m_j = \frac{1 + W(r)}{1 + \xi(r)},
\]

where \( n(r) \) is the galaxy number density, \( m \) is the mean of the entire galaxy sample
marker, \( W(r) \) is the weighted 2PCF, and \( \xi(r) \) is the standard 2PCF. In the large-scale
limit, the marked correlation function \( M(r) \to 1 \).

The galaxy number density can be used to define a “mark” [344] and a density field is
transformed as

\[
m(\delta) = \left( \frac{\rho + 1}{\rho + \rho_s} \right)^p = \left( \frac{\rho + 1}{\rho + (\delta + 1)} \right)^p,
\]

with \( \rho_s \) and \( p \) being two free parameters, \( \rho \) is a smoothed estimation of the galaxy number
density. Other choices are log transformations [345] and Gaussian transformations of the
density field [346].

Alternatively, gravitational potential \( \Phi_N \) can also be used to define a mark (see,
e.g., [347]):

\[
m(\Phi_N) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(\log_{10}(|\Phi_N|) - \Phi_\star)^2}{2\sigma^2} \right],
\]

where \( \sigma \) and \( \Phi_\star \) are free parameters that characterise the width and mean of the Gaussian,
respectively. Using such a mark, galaxies hosted by a certain halo mass are up-weighted
(see Figure 13).

The theoretical modelling of the marked correlation function in MG theories is under
development (see, e.g., [348]). For instance, the authors of [348] modelled mCF using a PT
approach to up-weight low-density regions in the Universe. The authors combined the
convolution Lagrangian PT with the \( f(R) \) gravity and nDGP models to build a theoretical
framework of the mCF for those models. They tested their analytical model against N-
body simulations (dark matter halos and HOD galaxies) and found good agreement between
the theory and simulations for scales \( > 30 \ h^{-1}\)Mpc.

Marked clustering statistics are a promising tool to test theories of gravity because
the screening mechanisms are environmental-dependent. However, there are challenges
due to the freedom of selecting the marks or weights. Furthermore, the definitions of
the environment are manifold in the literature. Hence, finding an appropriate combination of an environment and mark that increases the MG signal is still a topic that needs further exploration.

![Figure 13](Image)

**Figure 13.** Gravitational potential marked correlation function. (Credit: Alam et al. [98]).

### 3.5. Modelling the Observables

Dynamics of density fluctuation fields $\delta$ become non-linear when approaching smaller scales of order a few tenths of $h^{-1}$ Mpc. For most observables discussed in Section 3, PT (see review [349]) is one of the most common approaches for modelling observables involving large-scale structure (see also developments [350–355]). By perturbatively solving the Vlasov–Poisson system, which governs the dynamics of the density fluctuation field $\delta$, one can analytically achieve the correlator of $N$ density fluctuation fields. In addition to the matter field, galaxies as biased tracers of the matter field are affected by the RSD. The PT approach can also be extended to incorporate galaxy biasing (see review [93]) and RSD [356–360].

The perturbation models mentioned above are all GR-based; there are developments in the perturbation models with MG extensions. The authors of [361] presented a calculation for the matter power spectrum using the standard PT in real space. Given that RSD is a sensitive probe of cosmic structure growth, perturbative MG formalism is extended to redshift space [362] with improved series expansion convergence [363,364]. The authors of [365] presented a Lagrangian PT calculation for the power spectrum in Lagrangian coordinates, helping to improve the integral convergence. Their calculation can also be applied to a wide range of MG models classified as Horndeski theory. The authors of [366] combined convolution Lagrangian PT and a local Lagrangian bias schema to obtain accurate predictions of the redshift space correlation functions in the $f(R)$ and nDGP gravity models.

The accuracy of PT-based methods decreases rapidly from linear/mildly non-linear scales to non-linear scales ($k > 0.2 h^{-1}$ Mpc or $r < 20 h^{-1}$ Mpc). A different approach is constructed to predict the reaction of the halo model-derived [367] power spectrum in the presence of MG effects. Loosely speaking, the non-linear power spectrum can be obtained from the linear power spectrum via a non-linear mapping operator $\mathcal{K}$. MG effects can change the Poisson equation, halo mass function, and halo concentration, which all go into the operator and modify the mapping relation such that $\mathcal{K}_{\text{GR}} \rightarrow \mathcal{K}_{\text{MG}}$. The operator itself can be determined by comparing the power spectrum from simulation with or without MG effects [140]. Using this halo model reaction framework, the authors of [368] provided a model-independent prescription for the non-linear matter power spectrum by parametrizing perturbations using EFTofDE.

For higher-order statistics, the authors of [369] calculated a one-loop matter bispectrum based on PT, while the authors of [370] performed the calculation for the 3PCF for biased...
tracers. The authors of [371] assessed various theoretical modelling approaches, including PT, halo modelling, and fitting approaches for the matter bispectrum in the non-linear regime. However, they found that all the tested models were unlikely to be accurate enough for next-generation lensing surveys, while a corrected formula based on the halo model outperformed the others.

4. Simulations for MG

Cosmological simulations are crucial tools in understanding structure formation at smaller scales [372,373]. They are even more important in understanding the gravitational dynamics in the presence of the fifth force (see, e.g., [374]). Among many gravity models, two models with screening mechanisms $f(R)$ and nDGP (see Section 2) are the most representative, both of which feature modified Poisson equations. Modern cosmological simulations rely on $N$-body codes that solve the modified Poisson equation. Under the weak-field and quasi-static limit, it reads:

$$\nabla^2 \Phi = 4\pi G \delta \rho_m + f(\phi, \nabla \phi, \nabla^2 \phi, \cdots),$$

(56)

in the absence of the second term, Equation (56) is the standard Poisson equation, whereas the modification is given by the second term due to the extra scalar field $\phi$. The dynamics of the scalar field are governed by

$$\mathcal{L}[\phi; \delta \rho_m] = \mathcal{J}(\delta \rho_m),$$

(57)

where $\mathcal{L}$ is a non-linear derivative operator and $\mathcal{J}$ is the source field for the scalar field (cf. Equations (8) and (13)). To solve Equations (56) and (57) a typical method is to represent the extra scalar field on an adaptive grid and solve for their values via relaxation methods [375–377].

Given the complexity of the MG models (i.e., the non-linearity of the equations of motion that governs the evolution of the new scalar degree of freedom, cf. Section 2), just a few numerical codes have been used and adapted to evolve dark matter structures through cosmic time in such cosmologies. For instance, MG $N$-body simulation codes include ECOSMOG [378–381] and ISIS [382], which are extensions of the AMR code RAMSES [383]. MG-GADGET [384] and MG-AREPO [385,386] are based on the TreePM GADGET-2 [387] and the moving-mesh AREPO [388] codes, respectively. MG-GLAM [389,390] is based on the fast full $N$-body code GLAM [391].

Due to the non-linear equation of motion, an MG simulation takes 2 to 10 times the computational time of its GR ($\Lambda$CDM) counterpart. For this reason, the $N$-body gravity solver can be accelerated upon improving the relaxation method (e.g., for ECOSMOG; [392,393]) or FFT-based methods (e.g., MG-GLAM; [389,390]).

Acceleration can also be achieved with approximate methods such as modifications to the co-moving Lagrangian acceleration (COLA; [394]) method to achieve a faster time evolution. These codes, as suggested by their name, compute the dark matter particles’ position in the Lagrangian coordinate and then map to the Eulerian coordinate according to the Lagrangian PT to the $n$th order (e.g., MG-COLA; [395,396]).

4.1. Simulations for Galaxies

The dark matter-only simulations can be post-processed to obtain galaxy catalogues. Typically halo finders are used to define halos from the dark matter simulations (e.g., Subfind and Rockstar; [397,398]). Empirical methods, such as halo occupation distribution (HOD; [399]) or sub-halo abundance matching (SHAM; [400]), can be applied to build the halo–galaxy connection in MG scenarios [98]. For example, the extended lensing physics using analytic ray tracing (ELEPHANT; [102]) simulations were run with the ECOSMOG code [378,379]. The ELEPHANT simulations consist of five independent
realizations of the GR model, three variants of the $f(R)$ gravity ($f_{R0} = -10^{-6} - 10^{-5}$, and $-10^{-4}$) and two nDGP ($H_0 r_c = 5$ and $H_0 r_c = 1$) models. They started at $z = 49$ in a box with size $L = 1024 h^{-1}$Mpc and $1024^3$ particles. Galaxy catalogues were generated with the HOD method where the HOD parameters were tuned to match the number density and projected galaxy clustering of the BOSS-CMASS sample [401]. The HOD catalogues have been used to study the impact of MG in several galaxy statistics (e.g., [98]). The authors of [402] populated the ELEPHANT simulations with galaxies using the SHAM approach to study the environmental dependence of the galaxy luminosity function in MG models. The authors matched the observed SDSS $r$-band luminosity function [403] with the ELEPHANT’s $V_{\text{max}}$ distribution of halos and sub-halos.

More recently, the authors of [385] and [386] implemented the $f(R)$ Hu-Sawicki and nDGP gravity models into the AREPO code [388] to run the SHYBONE (simulating hydrodynamics beyond Einstein) simulations. The authors employed the ILLUSTRIS TNG model of galaxy formation [404,405] to simulate the evolution of gases, stars, black holes and dark matter in MG models, i.e., the interplay between baryonic feedback processes and MG. Due to the high computational cost of these calculations, the SHYBONE simulations were limited to small volumes despite their high resolution. The simulations evolved $2 \times 512^3$ dark matter and gas particles in volumes of $(25 h^{-1}$Mpc)$^3$ and $(62 h^{-1}$Mpc)$^3$ from $z = 127$ to the present epoch. The SHYBONE simulations have been used to explore the impact of MG on the high-redshift distribution of neutral hydrogen [406] and on the SZ effect [211]. The authors of [407] extended the SHYBONE simulations to larger volumes ($L = 302 h^{-1}$Mpc) by retuning the ILLUSTRIS TNG parameters with the aim of studying the gas content in galaxy clusters in MG cosmologies.

On the other hand, current and future galaxy surveys require simulations that cover large cosmological volumes with many realizations to estimate the errors of the different galaxy statistics accurately. More recently, the authors of [389] and [390] adapted the particle–mesh GLAM code [391] to run fast full $N$-body simulations of different MG models, including the $f(R)$ and nDGP models. The authors showed that MG-GLAM is able to run hundreds of large-scale simulations ($L = 512 h^{-1}$Mpc) with $1024^3$ particles using reasonable computational resources. These simulations will assist ongoing and future galaxy surveys (e.g., DESI and Euclid, see §5) by providing large-scale galaxy mocks of MG models to carry out RSD and BAO analyses. For example, the authors of [408] presented a study of the small-scale halo RSDs in MG cosmologies using MG-GLAM simulations as a previous step to modelling redshift-space galaxy clustering.

4.2. Simulations for Radio Sources

Simulating LIM (in particular 21 cm at high redshift) requires Gpc-scale simulation boxes to ensure the statistical modelling of ionized regions and absorption systems. At the same time, it also requires high resolution to guarantee the interplay of complex physical processes, such as ionized regimes and feedback mechanisms. Due to these challenges, LIM simulations so far have been performed by either post-processing dark matter-only simulations or using a semi-numerical approach.

Neutral hydrogen HI assignment schema can be added to the dark matter-only $N$-body simulations. The authors of [409] applied the HI distribution model [410–412] and assigned HI masses to each halo by assuming HI to be confined within the halo with a mass proportional to that of the halo mass. The HI cloud location is converted from real to redshift space (see Section 3.1.3). The HI density is obtained from the HI cloud redshift space position via the cloud-in-cell algorithm. This hybrid $N$-body approach can be extended for different DE models [413]. In the same paper, the authors also applied the hybrid approach to non-standard dark matter simulations of axions and late-forming dark matter models [414]. In addition to $N$-body simulations, particle–mesh-based methods (e.g., FastPM; [415]) can also be post-processed to study LIM [416] via different HI-halo (or HI-stellar mass) connections. The HI-halo connection can be extended to other lines emitted by galaxies such as CO and CII [417].
Alternative to post-processing dark matter-only simulations, a semi-numerical approach was proposed to predict the high-redshift 21 cm signal (21 cm FAST [418]). The semi-numerical code combines PT for evolving the density and velocity gradient field, excursion-set formalism [419] for identifying HII regions, and analytic prescriptions to spin temperature fields. These quantities are used to obtain the brightness temperature for the 21 cm line. The authors of [420] modified 21 cm FAST to incorporate structure growth in the general parametrizations of MG models. The modification results in a change in the 21 cm brightness temperature via non-linear HI density contrast, velocity gradient field, and ionization.

For radio continuum sources, there are two main tracers: star-forming galaxies (SFG) and active galactic nuclei (AGN). The widely used cosmological radio continuum simulations include the observation-based semi-empirical simulation (SKAD; [421]) and the N-body-based semi-analytic simulation (T-RECS; [422]). To date, there are no continuum radio simulations with MG effects as there are a few fundamental challenges in understanding the sources. One difficulty is understanding the supermassive black holes to the level that we can distinguish between radio-active AGN, radio-quiet AGN, and radio-SFG. To date, we can only tell them apart based on morphological structures, e.g., jets and lobes. These morphological properties are, however, not easily accessible even with inputs from other photometric or spectroscopic surveys [423]. In addition, understanding AGN populations and the physical mechanisms behind them require high-resolution simulations. These leave the radio continuum simulations more space to explore in the upcoming radio surveys.

4.3. Simulation-Based Inference

In Section 3.5 we presented different PT-based methods to model two-point statistics. However, these models are limited to producing accurate predictions on larger scales. For instance, the latest applications of PT methods to data restricted their analysis to use scales $k < 0.2$ [424,425] for both the power spectrum and bispectrum analyses. Unlocking the information content of smaller scales in cosmological analysis, where measurements from galaxy surveys are most accurate, requires precise N-body simulations that capture the non-linear growth produced by gravitational evolution. This is especially relevant for MG models, where non-linearities are enhanced by screening mechanisms.

Moreover, modelling summary statistics beyond two-point functions analytically may not be possible. Most of the summary statistics discussed in the previous sections that could potentially detect deviations from GR lack accurate analytical predictions and can only be computed through N-body simulations. Additionally, corrections that account for observational systematics, such as redshift-dependent completeness, survey mask and geometry, and fibre collisions, have only been demonstrated to work for two-point statistics. Their effect on higher-order summaries is unknown. Using N-body simulations, these different systematics can be readily included in the forward model [426,427].

Solving the inverse problem, estimating the posterior over the cosmological and MG parameters given the observed summary statistic, would require the order of $O(10^6)$ N-body simulations to perform Bayesian inference with Markov Chain Monte Carlo. This is currently computationally intractable due to how expensive N-body simulations are, particularly for MG theories. Therefore, most studies rely on modelling the dependence of a summary statistic on cosmology with surrogate models that are trained on a small set of $O(100)$ N-body simulations. The surrogate models are orders of magnitude faster than the original N-body simulations and can then be used to sample the posteriors of the parameters of interest. The most commonly used surrogate models are Gaussian processes [428] and neural networks [429].

For MG models, full N-body simulation suites with varying parameters only exist for $f(R)$ and nDGP gravity models [430], where the authors demonstrate how a Gaussian process emulator can be used to accurately predict the matter power spectrum in these MG models. The same simulations have also been used to obtain MG constraints from lensing observables [431].
Simulation-based models can also directly model the likelihood, or posterior, of a given summary statistic by using flexible density estimators such as Gaussian mixture densities or normalizing flows. An example of their application to cosmology can be found in [427]. These models will be necessary to obtain unbiased constraints on cosmology when the likelihood of a summary statistic is not Gaussian distributed.

Although we have discussed promising summary statistics that could detect deviations from GR, there is no guarantee that all the summaries combined would exhaust the information content of clustering and lensing datasets. Promising alternative venues to maximally constrain gravity models through \( N \)-body simulations are: field-level inference and learning optimal summary statistics with machine learning.

5. Cosmological Surveys of Our Universe

In order to learn about DE and gravity on cosmological scales, we want to map the distribution of matter in our Universe, baryonic and dark, as best as we possibly can. Surveying the Universe can be performed in all wavelengths of the electromagnetic spectrum, from radio to X-rays, using photons from the sky.

In this section, we review past, current, and future relevant photon surveys that map the distribution of matter in our Universe. These surveys are all essential to our understanding of DE, potentially providing evidence that gravity on larger scales deviates from GR.

5.1. Photometric Surveys

Photometric surveys are traditional method for systematically observing the cosmos. A photometric survey comprises a telescope, equipped with a camera and a set of filters, used to observe large portions of the sky with numerous exposures, without pointing at a specific object or location. Photometric surveys provide valuable insights into the intrinsic physical properties of galaxies and their evolutionary processes. Additionally, these surveys contribute to constraining cosmological models via, e.g., weak gravitational lensing measurements. Photometric surveys are also commonly utilized for the purpose of identifying objects of interest that warrant further spectroscopic investigation (see, e.g., review [432]).

There are plenty applications of photometric surveys. Particularly in cosmology, they are necessary for the discovery of type Ia supernovae, and for the study of galaxy clustering and weak gravitational lensing. On the one hand, type-Ia supernovae surveys require a high time cadence of observations, a few days typically, and precise flux measurements. On the other hand, WL surveys require deep imaging to discover large number of galaxies and precise measuring of their shapes. The study of large-scale structures with galaxies can be performed with the angular information provided by the images. However, radial information is limited due to the number of available photometric bands.

Main parameters defining a photometric survey include: area of sky observed, bands, magnitude limit, telescope main mirror size, field-of-view, number of galaxies, and redshift range. Photometric redshift estimates have uncertainties of the order of a few percent on \( \sigma_z \sim 0.05 \) typically.

Early photometric surveys include the automatic plate measuring machine (APM; [433]), the Edinburgh–Durham Southern Galaxy Catalogue (EDSGC; [434]), the Classifying Objects by Medium-Band Observations in 17 Filters project (COMBO-17; [435]), the Canada–Hawaii–France Telescope Legacy Survey (CFHTLS; [436]), the Spitzer Wide-Area Infrared Extragalactic Legacy Survey (SWIRE; [437]), and the Cosmic Evolution Survey (COSMOS; [438]). In addition, there are also the Wide-Field Infrared Survey Explorer [439] and the Two Micron All Sky Survey (2MASS; [440]) operating in the infrared band.

Current and upcoming photometric surveys typically have redshift uncertainty \( \sigma_z \lesssim 0.01 \). There are three ongoing WL surveys: the Kilo-Degree Survey (KiDS; [441]), the Dark Energy Survey (DES; [442]), and the Hyper Suprime-Cam Survey (HSC; [443]) covering an area of 1500 deg\(^2\) with \( \sim 10^8 \) galaxies. The KiDS Data Release 1 to 3 covered an area of
450 deg$^2$ with 14.7 million galaxies [444] and the final Data Release 4 KiDS-1000 covered an area of 1350 deg$^2$ with $\sim3 \times 10^7$ galaxies [445]. The DES Year 1 covered 1321 deg$^2$ with 26 million objects [446] and the DES Year 3 covered 5000 deg$^2$ with $\sim3 \times 10^8$ objects [447].

In the future, stage 4 surveys will probe a larger area and more galaxies. The Vera C. Rubin Observatory (Rubin for short, previously also known as LSST; [448]) will carry out an imaging survey of the Southern Galactic Cap. It will cover a sky area of 20,000 deg$^2$ and reach the $r$-band magnitude limit of $r \sim 27$. The Euclid satellite mission [449] will conduct a photometric survey covering 15,000 deg$^2$ with $10^9$ galaxies for the redshift range $0 < z < 2$. The Nancy Grace Roman Space Telescope (Roman for short; formerly the Wide-Field Infrared Survey Telescope or WFIRST [450]) will provide an imaging survey of 2000 deg$^2$ with $5 \times 10^8$ galaxies within the redshift range of $1 < z < 3$.

5.2. Spectroscopic Surveys

Spectroscopic surveys allow us to map the distribution of galaxies in 3D, where the radial information is recorded from precise redshift measurements. As the name states, a spectroscopic survey is usually equipped with a spectrometer, decomposing the light of objects into thin wavelength bands. A photometric survey is commonly the first step in building a spectroscopic survey, requiring a prior list of pointings, such as fibre-fed surveys. Target lists for such spectroscopic surveys are defined based on the angular positions, fluxes, and colours, which are passed to obtain the spectra. Emission and/or absorption features are easily seen in the observed spectra; they are the main ingredient for redshift measurements, the main goal of a spectroscopic survey.

The main parameters defining the power of a spectroscopic survey are the area of the observed sky, the number density of objects and their redshift range, and the spectral resolution. Spectroscopic surveys typically have redshift uncertainties of $\Delta z/z \sim 10^{-3}$.

A large variety of cosmological spectroscopic surveys have been and will be explored, including deep surveys on small areas, interesting for galaxy population studies, to shallower surveys on large areas, ideal for statistical measurements of the density field of tracers. Early galaxy spectroscopic surveys include the 2dF Galaxy Survey (2dFGS; [451]), 6dF Galaxy Survey (6dFGS; [452]), WiggleZ Dark Energy Survey (WiggleZ for short; [453]), and VIMOS Public Extragalactic Redshift Survey (VIPERS; [454]). Cosmological programs with Sloan Digital Sky Surveys (SDSS; [455,456]), the Baryon Oscillation Spectroscopic Survey (BOSS; [457]) and the extended BOSS (eBOSS; [458]) have provided the largest 3D map to date. BOSS covered 10,000 deg$^2$ with $\sim1.5 \times 10^6$ objects within $0 < z < 0.7$, and eBOSS covered 7500 deg$^2$ with $\sim10^6$ objects for an extended redshift range $0.6 < z < 3.5$ with more tracer types.

The Dark Energy Spectroscopic Survey [459] is a currently ongoing 5-year survey. By the end of the survey it will deliver $\sim30 \times 10^6$ spectra and cover over 14,000 deg$^2$. In the near future, both the Euclid and Roman surveys will deliver spectroscopic data in addition to imaging data. Euclid will provide $\sim5 \times 10^7$ spectra for $0.7 < z < 1.8$, and Roman will provide $\sim2 \times 10^7$ spectra for $1 < z < 3$. The Subaru Prime Focus Spectrograph (PFS; [460]) will be a spectroscopic survey covering 2000 deg$^2$ with $10^7$ emission line OII galaxies for $0.8 < z < 2.4$. The 4-metre Multi-Object Spectroscopic Telescope (4MOST; [461]) will deliver $\sim2.5 \times 10^6$ spectra for over 25,000 deg$^2$ for $0.15 < z < 3.5$.

Figure 14 shows the landscape of the galaxy surveys with photometric redshifts (cross) and spectroscopic surveys (circle) for galaxy number density per area as a function of the survey area. Overall, the future surveys (red) occupy the upper right corner, covering a larger survey area with a higher galaxy number density per area relative to past surveys (black).
5.3. Radio Surveys: Continuum and Intensity Mapping

Radio surveys are a promising tool to test the Universe at unexplored redshifts. Radio continuum at relatively lower wavelengths and HI intensity mapping at high wavelengths are the two main probes in the radio band. A radio continuum survey provides a high-angular resolution of radio galaxies but a low resolution in redshifts. The HI intensity mapping offers a high-radial resolution but a low resolution in the angular direction. The radio continuum and intensity mapping are complementary to each other. In addition, the HI galaxy redshift surveys are similar to what can be obtained from an optical spectroscopic survey where the galaxy 3D coordinates are provided.

The main parameters defining a radio survey are the frequency range and resolution, which translates into the accessible volume in the redshift. We can categorize them mainly by the area covered, angular resolution, and used technique, e.g., interferometry [212] or single-dish approach [462]. Several instruments are either currently recording data, under construction, or being planned. There will be a large amount of radio data available in the coming decades. We summarize the main radio continuum surveys and HI intensity mappings in Figure 15. We gather the results and review the status of the most relevant radio surveys to forecast the constraining power of such observations on DE/MG theories.

Large-scale structure studies with radio continuum sources require surveys with wide areas and high-resolutions in either angular or radial directions (reach sub-arcminute level). In the past decade, surveys such as Faint Images of the Radio Sky at 20 cm (FIRST; [463]), NRAO VLA Sky Survey at 1.4 GHz (NVSS; [464]), and TIFR GMRT Sky Survey (TGSS-ADR; [465]) have improved our understanding of extragalactic radio sources. In addition, there are the LOFAR LBA Sky Survey (LoLSS; [466]), Rapid ASKAP Continuum Survey (RACS; [467]), Karl G. Jansky Very Large Array Sky Survey (VLASS; [468]) and Galactic and Extragalactic All-Sky MWA survey (GLEAM; [469]). Currently, the Low-Frequency Array (LOFAR; [470]) is the only high-angular resolution (6 arcseconds) and high-sensitivity radio telescope with ultra-low frequency. The LOFAR two-metre Sky Survey has published
two data releases (LoTSS DR1, DR2; [471]), and there is the additional LOFAR Low-Band Antenna (LBA) Sky Survey (LoLSS; [466]), which aims to cover the entire northern sky and providing ultra-low-frequency information for $\sim 10^5$ radio sources. In the future, there are plans for an evolutionary map of the Universe (EMU; [472]) that will almost cover the whole southern sky, the precursor of which is the EMU pilot [473] which released data 2022, as well as the Westerbork Observations of the Deep Apertif Northern Sky Survey (WODAN; [474]), which covers the northern sky. Together they will cover the whole sky, thus aiding large-scale structure studies. The Band 1 and 2 of the SKA1-Mid [475] will provide a continuum WL survey and a wide continuum galaxy survey in the redshift range $z = 0.35-3$.

![Figure 15. A summary of radio surveys. The warm red colours denote the radio continuum maps while the cold blue colours represent the 21 cm intensity mapping. The error bars show the frequency ranges for each of the instruments or surveys. While the advantage of radio surveys is their high-angular resolutions, the surveys that measure the 21 cm signal possess good high-radial resolution, i.e., the frequency and its corresponding redshifts (dashed blue vertical lines). We mark the two kinds of resolutions as different colour gradients.](image)

The cosmological principle, which states that the Universe is isotropic and homogeneous, has been challenged by the tension in the cosmic dipole. The dipole measurements from, e.g., the radio continuum surveys NGSS and TGSS, are 2–5 times larger than expected [476]. The results will also potentially favour some of the MG models that lead to inhomogeneity at different scales, with the near-full-sky area of the radio continuum survey in the future, and the homogeneity of the Universe will be tested with much higher confidence. Moreover, with the spectroscopic redshift of the WEAVe-LOFAR survey [477], we will have 3D radio galaxy catalogues for the LOFAR survey, both with high-angular and -radial resolution, allowing us to perform higher-order statistics and more local tests in the future with broader cosmic volumes than the optical surveys.

At present, a plethora of experiments targeting LIM are currently running, under construction, or being proposed. For 21cm IM, the main instruments are CHIME [478,479], HIRAX [480], LOFAR [470], GBT [241,481], FAST [482], BINGO [483,484], CHORD [485], TIANLAI [486,487], SKAO [475] and its precursor MeerKAT [246,488,489]. For a review of the state-of-the-art ongoing and proposed IM surveys, we refer the reader to [212].
Currently, studies are mostly focused on forecasting the constraining power of future observations since there have yet been no LIM detections. For 21cm intensity mapping, a variety of scenarios were taken into account. Some of them focused on the EoR epoch (e.g., [420,490,491]). For the late-time Universe, 21cm IM is expected to improve the constraints on the background evolution [240,492,493], but also the perturbations [494]. Several models have been studied, from phenomenological parametrizations to specific theories. Some studies have tried to forecast the constraints on the $\mu$, $\Sigma$ and $\eta$ MG functions, e.g., [495], while other focused on EFTofDE [496]. Specific studied models are $f(R)$ [219,220], generalised scalar–tensor theories [497], early DE [498,499], interacting DE [500,501], among others [409,502]. Recently, several studies have focused on the cross-correlation between IM and other LSS probes, such as WL, GWs, and galaxy clustering [503–506]. The constraining power on DE of other lines is still mostly uncharted [507].

The main take-home message of the intense work carried out by the community is that LIM observations are expected to improve significantly the constraining power on beyond-GR models, due to their ability to sample wide ranges of redshift. To illustrate this, we collect the available forecasts on $\mu$ and $\Sigma$ in Figure 18.

5.4. Cosmic Microwave Background Surveys

CMB experiments measure the temperature fluctuation and polarization of photons emitted from the last scattering surface. The anisotropies in the CMB temperature can be categorized into effects that occur before or at the surface of the last scattering surface (primary anisotropies) and integrated effects between the last scattering surface and observers (secondary anisotropies). The main parameters defining a CMB survey are sky area, angular resolution (beam), frequency channels, sensitivity to temperature fluctuation, and white noise level at different frequency channels. CMB experiments have different systematics than the LSS surveys due to their distinct observational methods and the nature of the signals they capture. CMB experiments provide insights into the early Universe and complement the LSS probes.

Main CMB experiments in the past include, the Cosmic Background Explorer (COBE; [508]) as a full-sky survey with an angular resolution of 7°. Its two full-sky successors are the Wilkinson Microwave Anisotropy Probe (WMAP; [509]) with a resolution of 0.3° and the Planck satellite [510] with a resolution of 10 arcmin. The Atacama Cosmology Telescope (ACT; [511,512]) covered an area of 1000 deg$^2$ with a resolution of $\sim$1 arcmin, the South Pole Telescope (SPT; [513,514]) covered 500 deg$^2$ with a resolution of $\sim$1 arcmin, BICEP/Keck [515–517], and the Cosmology Large-Angular Scale Surveyor (CLASS; [518]) recorded a 75% sky coverage and a resolution of $<1.5^\circ$. In the future, the Stage 4 CMB observatory (CMB-S4; [519]) will cover 8000 deg$^2$ and achieve a resolution of $<3$ arcmin, while the Simons Observatory [520] will cover 15,000 deg$^2$ with a resolution of $\sim$1.5 arcmin.

5.5. Insights into MG from Current Surveys

Since MG can affect the growth of cosmic structure, the combination of the $f(z)\sigma_8(z)$ (Section 3.1.3), inferred from the anisotropic clustering of the galaxy auto-correlation, can be used test for potential deviation from GR (e.g., see the parametrization of $f\sigma_8$ employed in [521]). Figure 16 shows a compiled measurements from the past galaxy spectroscopic surveys, including 6dFGS [522], WiggleZ [523], MGS sample [524], VIPERS [525], BOSS LRG sample [521], eBOSS LRG [526,527], eBOSS ELG [528,529], and eBOSS QSO samples [181,530]. The $\Lambda$CDM cosmology using the best fitting value from the Planck data [531] with a 5% error bar is shown as the grey shaded area. The precision on the current growth rate constraints in terms of $f\sigma_8(z)$ are up to 8%. Future surveys with increased volume and the number of galaxies will largely improve upon the statistical precision.
Figure 16. A compiled diagram for $f\sigma_8(z)$ as a function of redshift from different galaxy spectroscopic surveys. The solid grey line denotes the $\Lambda$CDM with best fitting parameters from Planck15 [531] with a 5% error bar. The data points are from the following galaxy spectroscopic surveys utilizing the large-scale RSD information: 6dFGS [522], WiggleZ [523], MGS sample [524], VIPERS [525], BOSS LRG sample [521], eBOSS LRG [526,527], eBOSS ELG [528,529], and eBOSS QSO samples [181,530].

While the constraints on the product of $f(z)$ and $\sigma_8(z)$ are consistent with the standard GR scenario, there is a potential tension in $\sigma_8(z)$ [164,532]. Figure 17 shows a compiled measurement of $S_8$ (see Equation (41)). We select a few representative astrophysical probes in later stages, including WL, CMB lensing, galaxy clustering, galaxy clusters, and tSZ. They are compared to the CMB from Planck. The grey band represents Planck18 constraints. The CMB lensing data are in black [151,171,533], the galaxy WL data are in blue [133,165,166], while the galaxy clustering [425,534,535] in red, galaxy clusters [536–540] in yellow, and tSZ [541–543] are in purple. Extensive research has been conducted on systematics related to various probes, aiming to understand the possible systematic-induced tension between the early-time and multiple late-time measurements (see, e.g., review [544]). For example, the CMB estimates are reliant on models via parameters related to amplitude, such as the sum of neutrino mass and the optical depth; they can impact the derived value of $S_8$ [545]. Additionally, there are also anomalies in the Planck lensing amplitude $A_{lens}$ [546] and (see also [514,545,547,548]). However, these considerations do not fully explain the tension observed. In the case of WL, notable attention is given to factors such as intrinsic alignment, non-linear matter power spectrum modelling, photo-z estimation, baryonic effects, and so on, while individually these factors are unlikely to be the sole cause of the $S_8$ tension. The authors of [549] demonstrated that the cumulative impact of these various factors could alleviate the tension. Moreover, further investigation is needed to explore the mass calibration for galaxy cluster counts [550]. In addition to the systematics, parameter interpretation (e.g., using $h^{-1}$Mpc unit [551]) cannot be fully excluded. This mismatch in $S_8$ could also be correlated with the tension in the Hubble parameter $H_0$ (see review [552] and model comparison in [553]).

In the presence of an intriguing $\sim 2\sigma$ tension between CMB measurements and several late-time measurements could potentially indicate deviations from the standard $\Lambda$CDM model, provided that systematic errors can be entirely ruled out. Tentative candidates are: axion monodromy inflation [554–556], sterile neutrinos [557,558], alternative DE models [559–562], and MG models [49,563–565]. However, if this discrepancy is attributed to exotic physics, the proposed model must reconcile the effects in both early and late stages and successfully pass scrutiny from cosmological probes.
Figure 17. A compiled diagram for $S_8 = \sigma_8 \sqrt{\Omega_m/0.3}$ for CMB lensing (black) [151,171,533], galaxy WL data (blue) [133,135,136,549], galaxy clustering (red) [425,534,535], galaxy clusters (yellow) [536–540], and tSZ (purple) [541–543]. The grey band is from Planck using TT+TE+EE+lowE [182]. An intriguing 2σ level tension in several low-redshift experiments exists when compared to the primary CMB data.

Figure 18 shows the constraints from the current survey (black) and the forecast (brown) on the $\mu_0 = \mu|_{z=0}$ and $\Sigma_0 = \Sigma|_{z=0}$ functions (see Equations (21) and (22). In the case of GR, we expect $\mu_0 - 1 = 0$ and $\Sigma_0 - 1 = 0$. The current constraints [164,182,566,567] are all consistent with GR. Future surveys on galaxy clustering, WL, LIM, CMB, and GW will greatly improve upon the constraints [82,495,504,505].

Figure 18. Constraints (black) and forecasts (orange) for $\mu_0 = \mu|_{z=0}$ and $\Sigma_0 = \Sigma|_{z=0}$. From bottom to top we have RSD+WL [566], CMB lensing from Planck [182], or Planck CMB lensing+Supernovae+BAO [182], CMB+ISW+RSD together with a $H_0$ prior with or without WL data from CFHTLenS [567], and DES-Y3+Planck+BAO+RSD+SN [568]. The Euclid and DESI-ELG forecast for galaxy clustering and WL only [82], SKAO radio forecast with Planck and DES data [495], SKAO radio forecast with Planck and galaxy-clustering data [505], joint forecast using WL and clustering of GW+IM+galaxies [504], and the radio continuum surveys WODAN+EMU [192].
Finally, we summarize the current constraints on the $f(R)$ gravity parameter as a function of scale for different probes in Figure 19. While the current surveys can only establish an lower limit for the present-day value of $f_{R0}$ at approximately $10^{-4}$, future galaxy surveys (Sections 5.1–5.3) and CMB experiments (Section 5.4) hold promising potential for enhancing the constraints. Moreover, probes such as 21 cm [222,420] and underdense regions and 2D underdensities (LSST, Euclid; [102]) can be harnessed to offer further promising avenues.

**Figure 19.** Constraints on $f(R)$ from the past and current surveys. CMB lensing (SPT, ACT, Planck; [569–571]), galaxy power spectrum RSD (SDSS; [572]), galaxy–ISW cross-correlation (WMAP, SHOES, SCP, SDSS; [188,189,573]), cluster abundance (Chandra, MaxBCG; [189,574,575]), cluster density profile (MaxBCG; [321]), and WL (CFHTLenS; [576]).

6. Conclusions and Discussion

In this review, we have presented a concise overview of the gravity tests, with a particular emphasis on the interplay between theory and data. We focused on the astrophysical probes sensitive to structure growth at cosmological scales during later stages of our Universe’s evolution. Within this review, we have summarized ideas derived from effective field theories and parametrized frameworks of gravity, which hold promise in providing more robust constraints on gravity in the future, even when extending our investigations into non-linear regimes. The validity of specific models can be tested by comparing their predictions to observations. In particular, we focused on the two most representative MG models: $f(R)$ and nDGP. We discussed how gravity under these models can have a different impact on the potential, density, and velocity fields. As a consequence, derived quantities and their corresponding summary statistics of large-scale structure and CMB photons are altered when compared to the standard ΛCDM model.

While MG models were initially motivated as more natural explanations for the accelerated expansion, it is worth noting that a broader class of MG models, including Horndeski theories, may encounter challenges in accounting for cosmic acceleration [70]. Notably, the $f(R)$ and nDGP models, specifically focused on, are encompassed within this broader class. Considering these circumstances, a fundamental question arises: Why investigate MG from a cosmological standpoint?

Among the various attempts to modify GR, $f(R)$ and nDGP have remained as viable options (see Section 2). Even though these two models might not necessarily describe the true gravity theory of our Universe, their mechanisms are implemented into simulations (Section 4), and most model-specific studies are based on these two models (Section 3). With a model-specific approach, we can look for distinctive features predicted by a given model; it is easier to search for such features when comparing the predictions to observations, particularly when working at a low signal-to-noise regime. From this perspective, these two models provide explanatory examples of how a given observable could be impacted in the presence of a fifth force.
Cosmology has been in a precision era for almost two decades, and future surveys will cover larger areas and produce deeper maps of our Universe. One hopes that, by reducing statistical uncertainties, we can potentially reveal departures from GR-$\Lambda$CDM (Section 5) by nailing down the uncertainties in the parameters of the surviving MG models. Following this route, we list a few common challenges across the probes in the MG studies.

- **Model-dependent vs. model-independent**: above we mentioned the downsides of the model-specific approach. The parametrization framework is more general and can quantify broader classes of gravity models. To date, most MG studies within the parametrization framework are only suitable in the linear regime. Only recently have there been studies that generalized the approach to all scales (Section 2.4). However, extracting physical insights from the parametrized approach can be a difficult task.

- **Analytic templates for observables**: although the null hypothesis can be applied to quantify potential deviations from GR, a first-principle-derived template for a given observable helps to deepen the understanding of structure growth in the presence of a fifth force. However, such templates typically rely on the specifics of MG models, and computing such templates including MG effects is usually challenging. Furthermore, the templates are often motivated by PT. They are more suitable for summary statistics, while templates (even within the standard GR case) that characterize topological properties are yet to be developed.

- **Cosmic degeneracies**: massive neutrino and baryonic processes can both have impacts on the small-scale matter power spectrum (e.g., [577,578]). While various analytic approaches [579–581] have been used to study non-linear effects and baryonic processes, they typically require careful calibration upon hydrodynamical simulations. Meanwhile, the baryonic processes predicted by hydrodynamical simulations [405,582–585] are affected by resolution, calibration strategies, and the choice of sub-grid physics (see review by [142]). These variations can result in uncertainties in predicting the statistics of small-scale phenomena.

- **Simulation degeneracies**: many simulation-based conclusions rely on post-processing of $N$-body simulations. Various assumptions on the construction of observable catalogues can have non-negligible impacts on the conclusions. Typical examples are degeneracies between the HOD parameters or neglect of observational systematics.

This list contains interesting standing-alone problems relevant to MG and generally applied in the context of astrophysics and cosmology. However, there is yet no guarantee that we will find a satisfying answer as to the nature of cosmic acceleration, even if we solve all the problems listed above. Our Universe is a complex system with a huge number of degrees of freedom, and we have to admit that the fact that so many degrees of freedom can be reduced to only a few within the $\Lambda$CDM framework is crucial. The past efforts in studying these representative gravity models taught us how much is left in challenging the phenomenological $\Lambda$.

These lessons encourage those interested in understanding the nature of cosmic acceleration to continue to explore further fundamental theories.


**Funding:** J.H. has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 101025187. C.H.-A. acknowledges support from the Excellence Cluster ORIGINS which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC-2094–390783311. T.T. acknowledges funding from the Swiss National Science Foundation under the Ambizione project PZ00P2_193352. The project leading to this publication has received funding from the Excellence Initiative of Aix-Marseille University—A*MIDEX, a French “Investissements d’Avenir” program (AMX-20-CE-02—DARKUNI). J.Z. is supported by the project “NRW-Cluster for data inten-
sive radio astronomy: Big Bang to Big Data (B3D)” funded through the programme “Profilbildung 2020”, an initiative of the Ministry of Culture and Science of the State of North Rhine-Westphalia. The sole responsibility for the content of this publication lies with the authors. MB acknowledges support from the INFN INDARK PD51 grant.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We thank Dominik Schwarz, Anna Bonaldi, David Bacon, Alex Krolewski, Lucas Lombriser, Enrique Paillas, Hubert Wagner for useful discussions. We especially thank Alex Krolewski for their input on CMB lensing and ISW effects, Enrique Paillas for the density split method, Alessandra Silvestri and Emilio Bellini for the valuable feedback on the theory discussion, Hubert Wagner for dedicated discussions on the topological data analysis, Marta Spinelli for valuable support on technical aspects of IM. We acknowledge Fabian Schmidt for the feedback on the manuscript.

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

Notes

1. Quasi-static limit assumes that the time derivative of the scalar field perturbation is negligible compared to the spatial gradient of the scalar field.
2. Weak field approximation assumes that the amplitudes of the scalar field perturbations and gravitational potentials are much smaller than the speed of light squared.
3. Strictly speaking, $\Psi - \Phi$ can be sourced by an anisotropic pressure and a short-wave correction term $[53,54]$ in the $k$-essence model. However, the absolute difference in the two potentials is shown to be small and can be safely neglected $[54]$.
6. Of course one could try to force $aT = 0$ with very large $\Omega$. However, this would lead to exotic theories incompatible with the Universe we observe.
7. A conformal metric transformation is $g_{\mu\nu} \rightarrow e^{2\omega}g_{\mu\nu}$. At linear order it leads to $\{\Phi, \Psi\} \rightarrow \{\Phi + \omega, \Psi - \omega\}$. Thus, the sum of the potential is conformally invariant.
8. We approximate the optical depth $\tau \ll 1$ and $e^{-\tau(z)} \rightarrow 1$.
9. The authors of $[11]$ worked in the unit $c = \hbar = 1$, and we restored the $c$ in the equation.
10. Here, we only compare to galaxy clustering focusing on BAO and RSD analysis. There is, in principle, no restriction in applying galaxy-clustering analysis to highly non-linear regimes.
11. Recent advances such as the reaction formalism $[140,141]$ have narrowed this gap (see Section 3.5).
12. Mitigation schemes have been shown to pass various null tests. However, one should keep in mind there could still be redshift-sensitive or scale-sensitive residuals.
13. The lensing power spectrum is constructed from the angular power spectrum of the convergence $P_{\kappa\kappa}(k) = \chi^2 C_{\kappa\kappa}(k\chi_1)$, with $\chi_1$ as an effective lens distance in co-moving coordinates.
15. Euler characteristic is also called genus. It is a topological invariant (a property of a topological space when transformed under a bijective and continuous function—homeomorphisms).
16. An intriguing connection between the Euler characteristic and Betti number (represents the “dimension” of a homology group) is discussed in Bobrowski and Skraba $[294]$. The reader can find more information about persistent homology in Wasserman $[285]$, Edelsbrunner et al. $[286]$ and Koplik $[295]$ for a visualized explanation.
17. The radial distance is often normalized by voids’ radii given that the void density profiles are almost universal in terms of the void size.
18. In most of the MG simulations with the $N$-body algorithms, the field is discretized and solved using a finite-difference method; the relaxation method is usually applied to iteratively solve a system of differential equations. The goal of each iteration is to update the discretized field with a new value until the equality of the left- and right-hand sides of Equation (57) converges.
19. The authors of $[543]$ found the $f_{R0}$ is unconstrained in the cosmic shear analysis after marginalizing over nuisance or cosmological parameters.
References


Universe 2023, 9, 302


500. Xu, X.; Ma, Y.Z.; Weltman, A. Constraining the interaction between dark sectors with future HI intensity mapping observations. Phys. Rev. D 2018, 97, 083504.


