Nonlocal cosmological models and tests of modified gravity with gravitational waves

BELGACEM, Enis

Abstract

Dans cette thèse, nous étudions deux sujets en gravité modifiée: les modifications infrarouges non locales de la Relativité Générale et les tests de gravité modifiée avec des sirènes standard à des distances cosmologiques. Dans le Chapitre 1 nous montrons comment construire des termes de masse pour le champ gravitationnel (en particulier pour le mode conforme du tenseur métrique) sans violer l'invariance par difféomorphismes. Ceci est possible grâce aux nonlocalités attendues dans la 'quantum effective action' pour une théorie fondamentale contenant des particules sans masse (comme les gravitons en Relativité Générale). Nous expliquons des aspects conceptuels dans l'étude des théories non locales, en soulignant que la non-localité n'a de sens qu'au niveau de la 'quantum effective action', tandis que l'action fondamentale d'une théorie doit être locale pour éviter des problèmes de causalité. Dans le Chapitre 2, nous montrons les conséquences cosmologiques du modèle non local RT, au niveau du 'background' et des perturbations scalaires. Construire un modèle qui remplit [...]

BELGACEM, Enis. *Nonlocal cosmological models and tests of modified gravity with gravitational waves*. Thèse de doctorat : Univ. Genève, 2020, no. Sc. 5516

DOI : 10.13097/archive-ouverte/unige:149856
URN : urn:nbn:ch:unige-1498568

UNIVERSITÉ DE GENÈVE
Section de Physique
Département de Physique Théorique

FACULTÉ DES SCIENCES
Professeur Michele MAGGIORE

Nonlocal cosmological models and tests of modified gravity with gravitational waves

---

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention physique

par

Enis BELGACEM

de

Pollina (PA) (Italie)

Thèse N° 5516

GENÈVE

Atelier de reproduction de la Section de Physique
2020

UNIVERSITÉ
DE GENÈVE
FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Monsieur Enis BELGACEM

intitulée :

«Nonlocal Cosmological Models and Tests of Modified Gravity with Gravitational Waves»

La Faculté des sciences, sur le préavis de Monsieur M. MAGGIORE, professeur ordinaire et directeur de thèse (Département de physique théorique), Monsieur A. W. RIOTTO, professeur ordinaire (Département de physique thérorique), Monsieur A. KEHAGIAS, professeur (Physics Division, National Technical University of Athens, Athens, Greece), Monsieur G. TASINATO, docteur (Department of Physics, Swansea University, Swansea, United Kingdom), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 26 novembre 2020

Thèse - 5516 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Abstract

This thesis is devoted to the study of a particular class of cosmological modified gravity models known as nonlocal gravity theories and to the possibility of testing, in greater generality, deviations from the standard cosmological model ΛCDM via gravitational-wave detections. In particular, in the first part of this work we investigate the possibility that dark energy could be explained by infrared modifications of general relativity, in the form of nonlocal operators appearing in the quantum effective action for gravity. Despite a complete understanding of the gravity low-energy behavior is still missing, there are indications pointing towards a non-trivial infrared dynamics. Among several phenomenological nonlocal models proposed in recent years, cosmological observations and Solar System tests seem to select one particular model where the nonlocal term describes a mass for the conformal mode. This model fits available cosmological data at the same level of ΛCDM and gives quite similar predictions for cosmological parameters. The study of tensor perturbations reveals a distinctive behavior with considerable departures from ΛCDM and brings us to the second part of the thesis, dealing with modified gravitational-wave propagation. Such an effect, going beyond the specific features of nonlocal gravity, is very general in modified gravity theories and has to be added on top of the usual modifications for the dark energy equation of state. We show how the current and future generations of gravitational-wave detectors can constrain dark energy and test the predictions of modified gravity theories about the propagation of gravitational waves across cosmological distances.

Jury de thèse

- **Alexandros Kehagias**
  Physics Division, National Technical University of Athens, Athens, Greece

- **Michele Maggiore** (supervisor)
  Département de Physique Théorique, Université de Genève, Genève, Switzerland

- **Antonio Riotto**
  Département de Physique Théorique, Université de Genève, Genève, Switzerland

- **Gianmassimo Tasinato**
  Department of Physics, Swansea University, Swansea, UK

Publications

This thesis is based on the following publications about nonlocal cosmological models and tests of modified gravity with gravitational waves:

- *Stability issues of nonlocal gravity during primordial inflation*
  E. Belgacem, G. Cusin, S. Foffa, M. Maggiore and M. Mancarella
  Int.J.Mod.Phys.A 33 (2018) 1, 1850007, Int.J.Mod.Phys.A 33 (2018) 33, 1892005 (erratum), e-Print: 1610.05664 [hep-th]

- *Nonlocal gravity. Conceptual aspects and cosmological predictions*
  E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore
  JCAP 03 (2018) 002, e-Print: 1712.07066 [hep-th]

- *Gravitational-wave luminosity distance in modified gravity theories*
  E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore
  Phys.Rev.D 97 (2018) 10, 104066, e-Print: 1712.08108 [astro-ph.CO]

- *Modified gravitational-wave propagation and standard sirens*
  E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore
  Phys.Rev.D 98 (2018) 2, 023510, e-Print: 1805.08731 [gr-qc]

- *Testing nonlocal gravity with Lunar Laser Ranging*
  E. Belgacem, A. Finke, A. Frassino and M. Maggiore
  JCAP 02 (2019) 035, e-Print: 1812.11181 [gr-qc]

- *Testing modified gravity at cosmological distances with LISA standard sirens*
  E. Belgacem et al. (LISA Cosmology Working Group)
  JCAP 07 (2019) 024, e-Print: 1906.01593 [astro-ph.CO]

- *Cosmology and dark energy from joint gravitational wave-GRB observations*
  E. Belgacem, Y. Dirian, S. Foffa, E. J. Howell, M. Maggiore and T. Regimbau
  JCAP 08 (2019) 015, e-Print: 1907.01487 [astro-ph.CO]

- *Nonlocal gravity and gravitational-wave observations*
  E. Belgacem, Y. Dirian, A. Finke, S. Foffa and M. Maggiore
  JCAP 11 (2019) 022, e-Print: 1907.02047 [astro-ph.CO]

- *Gaussian processes reconstruction of modified gravitational wave propagation*
  E. Belgacem, S. Foffa, M. Maggiore and T. Yang
  Phys.Rev.D 101 (2020) 6, 063505, e-Print: 1911.11497 [astro-ph.CO]

- *Gravity in the infrared and effective nonlocal models*
  E. Belgacem, Y. Dirian, A. Finke, S. Foffa and M. Maggiore
  JCAP 04 (2020) 010, e-Print: 2001.07619 [astro-ph.CO]

I also contributed to the related works:

- *Science Case for the Einstein Telescope*
  M. Maggiore et al.
  JCAP 03 (2020) 050, e-Print: 1912.02622 [astro-ph.CO]
- *Cosmology and the Early Universe*
  B.S. Sathyaprakash et al. (Astro2020 Science White Paper)
  e-Print: 1903.09260 [astro-ph.HE]

More recently, I started working on pulsar-timing arrays as well:

- *Chirality of the gravitational-wave background and pulsar-timing arrays*
  E. Belgacem and M. Kamionkowski
  Phys.Rev.D 102 (2020) 2, 023004, e-Print: 2004.05480 [astro-ph.CO]

Résumé

Dans cette thèse, nous étudions deux sujets en gravité modifiée: les modifications infrarouges non locales de la Relativité Générale et les tests de gravité modifiée avec des sirènes standard à des distances cosmologiques.

Dans le Chapitre 1 nous montrons comment construire des termes de masse pour le champ gravitationnel (en particulier pour le mode conforme du tenseur métrique) sans violer l'invariance par difféomorphismes. Ceci est possible grâce aux non-localités attendues dans la 'quantum effective action' pour une théorie fondamentale contenant des particules sans masse (comme les gravitons en Relativité Générale). Nous expliquons des aspects conceptuels dans l'étude des théories non locales, en soulignant que la non-localité n'a de sens qu'au niveau de la 'quantum effective action' , tandis que l'action fondamentale d'une théorie doit être locale pour éviter des problèmes de causalité.

Dans le Chapitre 2, nous montrons les conséquences cosmologiques du modèle non local RT, au niveau du 'background' et des perturbations scalaires. Construire un modèle qui remplit toutes les contraintes d'observation et donne des prédictions testables dans un proche avenir est en général très difficile. Il est donc remarquable que le modèle RT, qui est actuellement un modèle phénoménologique sans dérivation fondamentale, soit un modèle cosmologique réussi. Ses solutions cosmologiques montrent une expansion accélérée à l'époque actuelle, sans la nécessité d'une constante cosmologique. Donc un terme de masse pour le mode conforme peut fournir une source pour l'expansion accélérée de l'Univers. Les perturbations scalaires sont stables et restent petites pendant toute l'évolution cosmologique. Elles sont très proches de celles en ΛCDM; donc le modèle RT explique bien les données cosmologiques actuelles, tout en restant potentiellement reconnaissable dans le futur.

Une analyse complète avec le Méthode de Monte-Carlo par chaînes de Markov (MCMC) montre que le modèle peut expliquer les données cosmologiques au même niveau que ΛCDM.

Le modèle RT se réduit à la Relativité Générale à de petites distances, sans avoir besoin d'un mécanisme de 'screening' non linéaire et respecte la limite de variation temporelle de la constante de Newton du 'Lunar Laser Ranging'. Il s'agit en général d'une propriété non triviale, même lorsque la solution statique a la limite correcte. Donc le modèle RT passe toutes les contraintes du système solaire et des expériences en laboratoire.

La plus grosse surprise du modèle RT se produit lors de l'étude de la propagation des perturbations tensorielles. En effet, dans le modèle RT, les ondes gravitationnelles se comportent différemment que dans ΛCDM et peuvent conduire à de grandes déviations.

Dans le Chapitre 3, nous expliquons en toute généralité la propagation modifiée des ondes gravitationnelles, et nous étudions en détail comment les générations actuelles et futures des détecteurs d'ondes gravitationnelles peuvent aider à tester les modèles de gravité modifiée. Cette propagation est un observable spécifique aux détecteurs d'ondes gravitationnelles et son effet est présent dans tous les théories de gravité modifiée étudiées les mieux motivées. Dans de nombreux modèles de gravité modifiée, la modification dans la propagation des ondes gravitationnelles est suffisamment important et les détecteurs futurs (LISA, Einstein Telescope, Cosmic Explorer) seront capables de la tester.

Contents

Notation

Units. We use natural units $c = \hbar = 1$ where c is the speed of light and $\hbar = h/(2\pi)$ is the reduced Planck constant.
The Newton gravitational constant G is kept explicitly in the equations or traded for the reduced Planck mass $M_P = 1/\sqrt{8\pi G} \simeq 2.4 \times 10^{18}$ GeV, but we never set $G = 1$.

Metric signature, covariant derivatives and curvature. Unless otherwise specified, we work in $D = 4$ spacetime dimensions. Greek indices $\mu, \nu, \dots$ or $\alpha, \beta, \dots$ refer to spacetime coordinates and take the values 0, 1, 2, 3. Latin indices i, j, … refer to spatial coordinates and take the values 1, 2, 3.
We stick to the General Relativity conventions adopted in the textbook *Gravitation* by Misner, Thorne and Wheeler [1]. The metric signature is $(-,+,+,+)$ and we follow the Einstein summation convention over repeated indices. The metric tensor is denoted by $g_{\mu\nu}$ (with determinant $g < 0$) and the inverse metric tensor is written as $g^{\mu\nu}$.
The covariant derivative of a contravariant vector is

$$\boldsymbol{\nabla}_\nu A^\mu = \partial_\nu A^\mu + \Gamma^\mu{}_{\nu\rho} A^\rho \,,$$

where the Christoffel symbols $\Gamma^\mu{}_{\nu\rho}$ are given by

$$\Gamma^\mu{}_{\nu\rho} = \frac{1}{2} g^{\mu\sigma} \left(\partial_\rho g_{\sigma\nu} + \partial_\nu g_{\sigma\rho} - \partial_\sigma g_{\nu\rho} \right) \,.$$

The Riemann tensor is defined as

$$R^\mu{}_{\nu\rho\sigma} = \partial_\rho \Gamma^\mu{}_{\nu\sigma} - \partial_\sigma \Gamma^\mu{}_{\nu\rho} + \Gamma^\mu{}_{\lambda\rho} \Gamma^\lambda{}_{\nu\sigma} - \Gamma^\mu{}_{\lambda\sigma} \Gamma^\lambda{}_{\nu\rho} \,,$$

and the Ricci tensor is obtained by contracting the first and third indices of the Riemann tensor above, as $R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}$. The Ricci scalar is $R = g^{\mu\nu} R_{\mu\nu}$.
The Einstein equations with a cosmological constant Λ read

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G \, T_{\mu\nu} \,,$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the energy-momentum tensor of matter. They can be derived from the principle of least action applied to the total action made by the Einstein-Hilbert term S_{EH} (with the cosmological constant) plus the matter action S_{m}:

$$S = S_{\mathrm{EH}} + S_{\mathrm{m}} = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \, (R - 2\Lambda) + S_{\mathrm{m}} \,,$$

with the definition

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_{\mathrm{m}}}{\delta g^{\mu\nu}} \,.$$

Chapter 1

Nonlocal gravity: theoretical foundations

1.1 Introduction

The currently accepted cosmological model, known as the ΛCDM model, provides a description of the Universe that is in agreement with a large number of observational data, by only using a quite limited number of free parameters. One of the biggest breakthroughs in cosmology, that gave a solid evidence to the model, was the discovery that the Universe is currently undergoing a phase of accelerated expansion. This milestone in our knowledge was first achieved, at the end of the last century, by the Supernova Search Team [2] and the Supernova Cosmology Project [3], using type Ia Supernovae (SNe) distance measurements. The ΛCDM model has been proven to be consistent with a variety of other independent cosmological observations. These include precision studies of the cosmic microwave backgound (CMB) anisotropies and measurements of the baryon acoustic oscillations (BAO) scale in the correlation function of galaxies using data from galaxy redshift surveys. However, even though it is observationally successful, the model is not fully satisfactory from a theoretical point of view. Indeed, the two ingredients that the ΛCDM model is named after, namely Λ (the cosmological constant) and CDM (cold dark matter), are not fully understood at different levels.

Dark matter is needed as an extra non-relativistic matter component (in addition to ordinary non-relativistic matter, called "baryons" in the cosmological context) and has a crucial role within the model, both at the background and perturbations level. For instance, the presence of dark matter is fundamental for structure formation, as it produces the potential wells for ordinary matter to collapse. Nevertheless the fundamental nature of dark matter and its place in the Standard Model of particle physics, or in some extension of it, are not established yet.

Regarding the cosmological constant Λ, it is responsible for explaining the accelerated expansion within the ΛCDM model, but its value needs a fine tuning to be compatible with observations, because the observed value corresponds to an energy density much smaller than what quantum field theory suggests to be natural for the vacuum energy density [4]. This *cosmological constant problem* can be expressed more precisely by using the language of renormalization in quantum field theory and considering that the bare cosmological constant receives corrections from vacuum loop diagrams. The measurable quantity that is extracted from observations is the total renormalized value of the vacuum energy density. It amounts to $\rho_{\text{vac}} \sim \mathcal{O}(\text{meV})^4$, while the one-loop contribution of each particle species, computed with dimensional regularization[1], is roughly quartic in the mass of the particle (see e.g. Section 2 of [6] for the computation in the case of a scalar field). Since the Standard Model includes particles up to about the TeV scale, the bare cosmological constant must cancel a value of $\mathcal{O}(\text{TeV})^4$ down to the measured $\mathcal{O}(\text{meV})^4$, meaning that a precision of one part over 10^{60} is needed. Such a highly fine tuning and the fact that the observed meV scale associated to vacuum energy does not give any insight for particle physics are probably hints that something is missing in our current understanding. Even if one chooses to accept the fine tuning at one loop, the real obstacle to its viability as a solution to the cosmological constant problem comes when considering higher-order loop corrections, because their contributions are not significantly suppressed with respect to the smaller orders and, as a consequence of this radiative instability, it would be necessary to introduce a fine tuning of similar accuracy at each step in perturbation theory.

The naturalness problem associated to the cosmological constant stimulates the search for alternatives to the ΛCDM model that do not need a cosmological constant for explaining the accelerated expansion. Given that the ΛCDM model is just based on general relativity (GR) in the form of the Einstein-Hilbert action with a cosmological constant term and the matter action, alternatives to ΛCDM have to break this assumption and take some other starting point. A quite logical step along this line of reasoning consists in adding some dark energy component to the energy-momentum tensor sourcing gravity or in attempting a modification of the gravity sector of GR itself. The goal of these modifications is to produce an accelerated expansion at late times without relying on a cosmological constant, with the result that the constant vacuum energy density of ΛCDM is effectively changed into a dark energy contribution which evolves throughout the history of the Universe in a non-trivial way. However, when considering modified gravity theories as a source for dynamical dark energy, one should be careful not to introduce too many free parameters and not to simply change the problem of the unnatural smallness of the cosmological constant into the same problem for the free parameters in the alternative model. Still, even considering these potential difficulties, it is interesting to study modifications of GR by their own right and as a tool for better understanding and appreciating GR itself. Indeed, historically the study of alternatives to GR has already been an active research field independently of the cosmological developments and we have quite old examples for that (see, for instance, the Brans-Dicke theory [7] in 1961). Today we also foresee the possibility to really test some interesting alternative theories of gravity at cosmological scales with current and forthcoming data; this adds another good reason for the relevance of modified gravity studies.

A useful reference point when trying to modify GR is Lovelock's theorem (1971) [8], stating that, in four spacetime dimensions, GR equations (as obtained from the Einstein-Hilbert action with a cosmological constant) are the only local and second-order equations of motion that can be derived from an action containing only the metric tensor. The theorem implies that there are five possible options leading to alternative theories of gravity, namely:

1) questioning the action principle for the metric as the most fundamental level of description for gravity. According to this point of view, the geometrical nature

---

[1] A similar conclusion can be reached, even more simply, by using a cutoff regularization and considering the zero-point energy left by the canonical normalization procedure (see Section 5.7 of [5]). We find it useful to present the discussion in terms of loop diagrams, because this allows us to mention the radiative instability of vacuum energy, which is the central element of the cosmological constant problem.

of gravity is just an emergent phenomenon, that originates statistically from the interaction of many more fundamental degrees of freedom. This approach is known as emergent (or entropic) gravity and illustrations of it are presented in [9] and [10].

2) adding other fields to the gravitational sector. These can be scalars, as in Horndeski theories [11] or in the more general degenerate higher-order scalar-tenosor (DHOST) theories [12, 13, 14, 15], vectors as in Einstein-Aether gravity (reviewed in [16]) that breaks Lorentz invariance, or tensors as in bigravity [17] where an additional metric tensor is present. Local[2] massive gravity theories also belong to this category of modified gravity models as they exhibit extra degrees of freedom with respect to the two gravitational-wave polarizations of the massless GR case, for a total of six in most of non-linear extensions of the Fierz-Pauli action [18] (where a Boulware-Deser ghost appears [19]), reducing to five for a class of ghost-free theories [20] including the de Rham-Gabdadze-Tolley (dRGT) model [21]. A comprehensive review on the theoretical aspects of massive gravity is given in [22].

3) introducing extra dimensions. A theory in this class that will also be useful in the following is the Dvali-Gabadadze-Porrati (DGP) model [23] where the ordinary four-dimensional spacetime metric lives on a brane of the five-dimensional bulk.

4) allowing for higher than second order derivatives in the equations of motion for the metric. Examples are f(R) gravity (see [24] for a review) and Hořava-Lifshitz gravity [25]. The latter also breaks Lorentz invariance. It is important to remember that, because of the Ostrogradsky theorem [26], particular care has to be taken when dealing with higher-order equations of motion. At the quantum level, the appearance of ghost-like degrees of freedom usually imposes to consider the theory to be viable only below some cutoff energy scale, i.e. as an effective field theory.

5) assuming nonlocalities in the equations of motion. This possibility is one the two main subjects of this thesis. We anticipate that nonlocalities are a sensible ingredient when including quantum corrections to general relativity and it is conceivable that they could produce relevant effects in the infrared (IR). As we will discuss in the following, the relevant quantity to consider when taking into account quantum effects is not the classical action, but rather the quantum effective action. For a theory containing massless particles (like gravitons in GR), nonlocalities necessarily appear in its quantum effective action.

When talking about the difficulties in building a quantum theory of gravity, it is usually said that GR is perfectly fine as an effective field theory at energies low enough with respect to the Planck scale. This conclusion can be driven from the simple argument that GR, as a quantum field theory, is power-counting non-renormalizable because the coupling constant $\kappa = (32\pi G)^{1/2} = 2/M_P$ ruling the interaction between gravitons (see Section 2.2.4 of [27]) has dimensions of the inverse of mass. As a consequence of non-renormalizability, the quantum field theory

---

[2] We will see that nonlocal theories require some extra care when counting degrees of freedom. Indeed the RR and RT nonlocal models that will be discussed in this thesis do not introduce extra degreees of freedom with respect to GR, despite they give a mass to the conformal mode of the graviton. This is possible because diffeomorphism invariance is compatible with a mass term, as long as nonlocal structures are used.

of GR certainly needs an ultraviolet (UV) completion at the Planck scale (or below). The statement above seems to imply that in the opposite regime, i.e. at low energies, a perturbative treatment of the quantum field theory for GR is perfectly allowed and gives correct predictions. However, there are several indications of a less trivial IR dynamics due to non-perturbative effects. This is potentially very interesting for cosmology, as it could imply observable effects at very large length scales (the idea that quantum effects in gravity could have cosmological relevance already appeared in old works [28]). We summarize the indications of a non-trivial IR behavior of gravity following [29, 30, 31]. As reviewed in [32], the study of the IR limit of theories containing massless particles shows an intimate connection between soft theorems (describing the limit of zero energy for massless particles in the external lines of Feynman diagrams), memory effects and asymptotic symmetries. In the context of GR, a non-unique vacuum structure is implied by the recovery of the very large BMS group for asymptotically flat spacetimes [33, 34], instead of the simple Poincaré group where the notion of particle is uniquely defined.

Examples of IR divergences have been found in de Sitter space, where some gauge-invariant quantities, relevant for scattering processes and built from the metric fluctuations around the de Sitter background, diverge. The low-energy divergences are produced by the unboundedness of the graviton Feynman propagator at large separations [35, 36], with the strongest effects coming from the propagator of the conformal mode σ of the metric. This result was interpreted as an evidence for quantum instability of de Sitter space, but a general consensus on that has not been reached yet, because it seems that the infrared divergence in the propagator can be eliminated by a nonlocal gauge transformation [37] (see also [38, 39]). Infrared divergences appear as well when performing computations of physical quantities during inflation [40].

de Sitter space has also been used for studying the consequences of quantum fluctuations of massless fields coupled to gravity. In the case of a massless minimally-coupled scalar field with a $\lambda\phi^4$ self-interaction in a fixed de Sitter background, it has been shown in different ways that the scalar field develops dynamically an effective mass $m^2_{\text{eff}} \propto H^2\sqrt{\lambda}$ where the H is Hubble constant of the de Sitter background (see [41, 42, 43, 44, 45, 46, 47]). Besides the effects due to quantum fluctuations of the gravitons itself, when gravity is coupled to matter fields, the quantum fluctuations of the latter also affect the graviton dynamics. In particular, quantum fluctuations of massless matter fields are of great interest for gravity in the IR regime: an example is the non-trivial dynamics of the conformal mode due to the conformal anomaly, which will be reviewed in some detail in Section 1.2.4. In four dimensions and flat spacetime the conformal anomaly generates, in the quantum effective action, a term proportional to $(\Box\sigma)^2$, leading to a propagator for the conformal mode $G_\sigma(x, x') \propto \log(x - x')^2$. Its divergence at large separations suggests a non-trivial large-distance dynamics and the contributions from the conformal anomaly have already been used in reference to the cosmological constant problem [48].

In order to study the effects of quantum fluctuations in gravity and to adequately understand the results cited above, we need to refer to the notion of quantum effective action, which will be discussed in Section 1.2. A crucial difference between the classical action and the quantum effective action is that, while the former is local, the latter has also nonlocal terms whenever the theory contains massless or light particles. In gravity, because of the quantum fluctuations associated to the massless graviton, nonlocal terms are unavoidably present, and could in principle significantly affect the IR behavior of the theory. The techniques for computing these nonlocal terms in the UV regime are well understood [49, 50, 51, 52, 53, 54], while in the IR regime, where they could be cosmologically relevant, the situation is much less clear.

A typical non-perturbative IR effect is dynamical mass generation. We will see through an explicit construction that mass terms for the metric can be written without violating diffeomorphism invariance thanks to nonlocal operators, whose presence is allowed at the level of the quantum effective action. We will also briefly discuss some evidence, both from lattice gravity and from functional renormalization group equations, suggesting the possibility of generating a mass scale. Of course, to confirm or disprove conclusively a mechanism of dynamical mass generation for gravity we would need a non-perturbative understanding of the low-energy regime, that is currently not available. In the absence of a rigorous top-down derivation, an alternative route consists in exploring the structure of possible nonlocal terms associated to a mass scale and study their cosmological predictions; in this way one can assess the viability of nonlocal gravity models and compare them with cosmological datasets.

1.2 The quantum effective action

Every classical theory, once considered at the quantum level, receives corrections from loop diagrams. The result can be summarized by saying that the classical fundamental action of the theory is changed into a quantum effective action (QEA) that, when used at tree level, already takes into account the effect of quantum loops. Following [30], let us review how the QEA construction goes in quantum field theory, by considering a single scalar field $\varphi(x)$, with action $S[\varphi]$, in flat spacetime. We first introduce an auxiliary source $J(x)$ and define the generating functional of Green's functions $Z[J]$ and connected Green's functions $W[J]$, according to

$$Z[J] = e^{iW[J]} \equiv \int D\varphi \; e^{iS[\varphi]+i\int J\varphi} , \tag{1.2.1}$$

with $\int J\varphi \equiv \int d^4x\, J(x)\varphi(x)$. The functional derivatives of $Z[J]$ and $W[J]$ with respect to the external current J can be used to compute the Green's functions and the connected Green's functions, respectively. In particular, the vacuum expectation value of the field $\varphi(x)$ in the presence of the source $J(x)$ is given by

$$\phi[J] \equiv \langle 0|\varphi(x)|0\rangle_J = \frac{\delta W[J]}{\delta J(x)} . \tag{1.2.2}$$

The QEA $\Gamma[\phi]$ is obtained by applying a Legendre transformation to $W[J]$,

$$\Gamma[\phi] \equiv W[J] - \int \phi J . \tag{1.2.3}$$

In the expression above, the current J is to be intended as a function of ϕ, i.e. $J = J[\phi]$, obtained from the formal inversion of $\phi[J]$ given in eq. (1.2.2). The definition in eq. (1.2.3) shows that the QEA is a functional of the vacuum expectation value ϕ, and not of the classical field φ as in the case of the classical action $S[\varphi]$. Using the definitions in eq. (1.2.2) and eq. (1.2.3), the variation of $\Gamma[\phi]$ with respect to ϕ gives

$$\frac{\delta\Gamma[\phi]}{\delta\phi(x)} = \int_y \frac{\delta W[J]}{\delta J(y)}\frac{\delta J(y)}{\delta\phi(x)} - J(x) - \int_y \phi(y)\frac{\delta J(y)}{\delta\phi(x)} = -J(x) , \tag{1.2.4}$$

showing that the variation of the QEA gives the equation of motion for the vacuum expectation value of the field, in the presence of a source. An exact path integral representation for the quantum effective action is obtained from eq. (1.2.1) and eq. (1.2.3) and the property in eq. (1.2.4), reading[3]

$$e^{i\Gamma[\phi]} = e^{iW-i\int\phi J} = \int D\varphi\; e^{iS[\varphi]+i\int(\varphi-\phi)J} = \int D\varphi\; e^{iS[\phi+\varphi]-i\int\frac{\delta\Gamma[\phi]}{\delta\phi}\varphi}\,. \tag{1.2.5}$$

The representation in eq. (1.2.5) indicates that the QEA is a functional of the vacuum expectation value of the field, obtained by integrating out the quantum fluctuations. An iterative procedure for the perturbative evaluation of $\Gamma[\phi]$ consists in starting replacing $\Gamma[\phi]$ with $S[\phi]$ on the right-hand side of eq. (1.2.5) and finding a new refinement for the QEA at each step.

1.2.1 Extension to gravity

In a curved spacetime, a theory is described by the Einstein-Hilbert action $S_{\rm EH}[g_{\mu\nu}]$ and the matter action $S_m[g_{\mu\nu};\varphi]$, where φ denotes the matter fields, so that the total action is $S = S_{\rm EH}[g_{\mu\nu}] + S_m[g_{\mu\nu};\varphi]$. We want to consider the effects of quantum fluctuations of matter fields, but we neglect the quantum fluctuations of gravity itself. Integrating over the quantum fluctuations of matter fields (and not on those of the metric), the procedure described before in flat spacetime (and leading to eq. (1.2.5)) can be easily repeated in curved spacetime to compute the QEA $\Gamma[g_{\mu\nu};\phi]$, which satisfies

$$e^{i\Gamma[g_{\mu\nu};\phi]} = e^{iS_{\rm EH}[g_{\mu\nu}]}\int D\varphi\; e^{iS_m[g_{\mu\nu};\phi+\varphi]-i\int\frac{\delta\Gamma[g_{\mu\nu};\phi]}{\delta\phi}\varphi}\,, \tag{1.2.6}$$

where

$$\int\frac{\delta\Gamma[g_{\mu\nu};\phi]}{\delta\phi}\varphi \equiv \int d^4x\sqrt{-g}\,\frac{\delta\Gamma[g_{\mu\nu};\phi]}{\delta\phi}\varphi\,. \tag{1.2.7}$$

Let us specialize to the case in which the vacuum expectation values of the matter fields vanish i.e. $\phi \equiv \langle 0|\varphi(x)|0\rangle = 0$, and there is no external current that excites them, which means $\delta\Gamma/\delta\phi = J = 0$. The physical reason for this choice is that we are interested to the resulting dynamics of the metric (affected by the quantum behavior of matter fields) and not in the resulting dynamics of matter fields; therefore once we integrate over the quantum fluctuations of the fields, we set them in their vacuum state $\phi = 0$. The corresponding vacuum quantum effective action obtained from eq. (1.2.6) obeys

$$e^{i\Gamma[g_{\mu\nu}]} = e^{iS_{\rm EH}[g_{\mu\nu}]}\int D\varphi\; e^{iS_m[g_{\mu\nu};\varphi]} \equiv e^{iS_{\rm EH}[g_{\mu\nu}]}\,e^{i\Gamma_m[g_{\mu\nu}]}\,. \tag{1.2.8}$$

The definition of the matter energy-momentum tensor $T^{\mu\nu} \equiv \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g_{\mu\nu}}$ combined with the equation for $\Gamma_m[g_{\mu\nu}]$ in eq. (1.2.8) then gives the following result for the vacuum expectation value of $T^{\mu\nu}$,

$$\langle 0|T^{\mu\nu}|0\rangle = \frac{2}{\sqrt{-g}}\frac{\delta\Gamma_m}{\delta g_{\mu\nu}}\,. \tag{1.2.9}$$

As a consequence, the equations of motion derived from the total quantum effective action $\Gamma = S_{\rm EH} + \Gamma_m$ give the Einstein equations $G^{\mu\nu} = 8\pi G\langle 0|T^{\mu\nu}|0\rangle$ where, on

---

[3] In the last step of 1.2.5 we also changed variable setting $\varphi = \varphi' + \phi$ and we renamed the new integration variable φ' as φ.

the right-hand side, all quantum fluctuations due to matter fields are automatically included.

Finally, to get the full QEA one should also integrate over the quantum fluctuations of the metric and include in the path integral formulation the Faddeev-Popov determinant and a gauge fixing term, as usual for quantizing gauge theories.

1.2.2 Two observations: difference with the Wilson effective action and boundary conditions

As stressed in [30], the QEA has not to be confused with the Wilson effective action, which is used in both quantum field theory and statistical physics. The Wilson effective action is obtained by integrating out massive fields from the theory and is a functional of the light fields left, useful for an effective description at energies low compared to the masses of the heavy fields which have been integrated out. This is different from what happens for the QEA, where we integrate out the quantum fluctuations of the fields to obtain a functional of the vacuum expectation values of *all* the fields, without introducing any definition of light or heavy fields. In particular, when considering massless particles, it would not make any sense to integrate them out à la Wilson, but it is perfectly sensible to integrate over their quantum fluctuations, as done in the QEA construction.

Another important point for correctly interpreting the results of the QEA construction, both in the case of flat spacetime that we studied before and in the application to gravity that we have now seen, is to understand what we mean by the different occurrences of the vacuum state $|0\rangle$ and how its different choices affect the vacuum expectation values. The answer to this question lays in the boundary conditions of the path-integral formulation, which were not stated explicitly at the beginning, but determine the meaning of the results. Assigning boundary conditions is a necessary part of the definition of a path integal and different choices correspond to different physical consequences. When using the standard Feynman path integral we get the in-out vacuum expectation values, like $\langle 0_{\rm out}|T^{\mu\nu}|0_{\rm in}\rangle$ for the energy-momentum tensor. In-out matrix elements appear as intermediate steps in QFT computations, e.g. in scattering amplitudes, but by themselves they are not physical quantities. This is clearly proved by the fact that, even if $\hat{\varphi}$ is an hermitian operator (obtaining by promoting the classical field φ to a quantum field), the expectation value $\langle 0_{\rm out}|\hat{\varphi}|0_{\rm in}\rangle$ is not real, and therefore cannot represent any physical quantity. Besides that, it obeys equations of motions in which the Feynman propagator appears and which are therefore acausal (only the retarded propagator corresponds to causal physics). In the corresponding gravitational case, i.e. when we quantize gravity promoting the metric to an operator $\hat{g}_{\mu\nu}$, the vacuum expectation value $\langle 0_{\rm out}|\hat{g}_{\mu\nu}|0_{\rm in}\rangle$ is not real, it does not obey causal equations of motion and therefore cannot be interpreted as a semiclassical metric. To obtain physical vacuum expectation values, the Schwinger-Keldysh path integral has to be used, which corresponds to in-in expectation values, like $\langle 0_{\rm in}|T^{\mu\nu}|0_{\rm in}\rangle$. They represent vacuum expectation values of operators at a given time and do not suffer from the non-reality and acausality problems displayed by the Feynman case. In particular, if $\hat{\varphi}$ is hermitian, then the quantity $\langle 0_{\rm in}|\hat{\varphi}|0_{\rm in}\rangle$ is real, and it obeys causal equations of motions in which the retarded propagator appears [53, 55, 56]. When using the Schwinger-Keldysh path integral the quantity $\langle 0_{\rm in}|\hat{g}_{\mu\nu}|0_{\rm in}\rangle$ plays the role of a semiclassical metric. This observation will be very useful to give the correct interpretation to the equations of motion of the nonlocal models that we will build, which are intended to describe the behavior of the semiclassical metric when including quantum fluctuations.

1.2.3 Nonlocalities in the quantum effective action of QED

As a first example, let us focus on QED and review the results for its quantum effective action following [29, 30]. After integrating out the quantum fluctuations due to the electron, the dynamics of the photon is described by (see [57, 58, 59])

$$\Gamma_{\text{QED}}[A_\mu] = -\frac{1}{4}\int d^4x \left[F_{\mu\nu}\frac{1}{e^2(\Box)}F^{\mu\nu} + \mathcal{O}(F^4)\right] . \tag{1.2.10}$$

In the high-energy limit $|\Box/m_e^2| \gg 1$, the form factor $1/e^2(\Box)$ is given by

$$\frac{1}{e^2(\Box)} \simeq \frac{1}{e^2(\mu)} - \frac{1}{12\pi^2}\log\left(\frac{-\Box}{\mu^2}\right) , \tag{1.2.11}$$

where μ is the renormalization scale and $e(\mu)$ is the renormalized charge at the scale μ. The physical interpretation is that the nonlocality in coordinate space reflects the running of the coupling constant in momentum space. Explicitly, the logarithm of the d'Alembertian is a nonlocal operator defined by

$$\log\left(\frac{-\Box}{\mu^2}\right) \equiv \int_0^\infty dm^2 \left[\frac{1}{m^2+\mu^2} - \frac{1}{m^2-\Box}\right] . \tag{1.2.12}$$

In the low-energy limit, $|\Box/m_e^2| \ll 1$, i.e. when the electron is heavy compared to the relevant energy scales and decouples from the photons, the quantum fluctuations due to the electron only produce local terms, which are suppressed by powers of $|\Box/m_e^2|$,

$$\frac{1}{e^2(\Box)} \simeq \frac{1}{e^2(\mu)} + \frac{4}{15\,(4\pi)^2}\frac{\Box}{m_e^2} . \tag{1.2.13}$$

The QED case shows that nonlocalities appear in the quantum effective action when we integrate over quantum fluctuations of massless or light particles. Indeed the nonlocal structure $\log\left(\frac{-\Box}{\mu^2}\right)$ was originated in the high-energy regime, i.e. when the electrons are light compared to the energy scale in consideration. On the contrary, in the case of heavy electrons the result is the decoupling of the electrons and only local operators appear.

1.2.4 An enlightening case: the anomaly-induced effective action

The construction of the vacuum QEA for gravity illustrated in Section 1.2.1 can be carried out exactly for a theory with massless, conformally-coupled matter fields, in $D = 2$ space-time dimensions, by integrating the conformal anomaly. The same procedure can be used in $D = 4$ to obtain exactly the part of the QEA depending on the conformal mode of the metric. We follow the discussion in Section 3 of [29], see also [36, 48, 49, 52, 53, 54] for reviews.

The anomaly-induced effective action is an explicit example of an exact QEA derivation for gravity starting from the fundamental action of the theory. Knowing both the starting point (a healthy classical theory) and the final point (the anomaly-induced QEA) will allow us to discuss some conceptual points and avoid possible misunderstandings about the degrees of freedom appearing in a QEA. It will be quite simple and very useful to adapt those discussions to a correct understanding of nonlocal gravity models, considering their lack of a fundamental derivation.

Let us consider 2D gravity, including also a cosmological constant λ, coupled to N massless matter fields,

$$S = \int d^2x \sqrt{-g}(\kappa R - \lambda) + S_m \,, \tag{1.2.14}$$

and let us suppose that the $N = N_s + N_f$ matter fields included in S_m are given by N_s conformally-coupled massless scalars[4] and N_f conformally-coupled massless Dirac fermions. In $D = 2$ we use $a = 0, 1$ as Lorentz indices, and the metric signature is $\eta_{ab} = (-,+)$. For conformal matter fields, classically the trace T^a_a of the energy-momentum tensor vanishes. However, at the quantum level the vacuum expectation value of T^a_a is non-zero, and is given by

$$\langle 0|T^a_a|0\rangle = \frac{N}{24\pi} R \,, \tag{1.2.15}$$

where $N = N_s + N_f$. The result in eq. (1.2.15) is the trace anomaly (or conformal anomaly) and its derivation for a conformally-coupled scalar field can be found in Section 14.3 of [53]. Its peculiarity is that, even if it can be obtained with a one-loop computation, it is actually exact at all perturbative orders, see again Section 14.3 of [53] for a proof. In other words, no contribution to the trace anomaly comes from higher loops. Using eq. (1.2.9) we can deduce the QEA corresponding to the vacuum expectation value in eq. (1.2.15). To perform the calculation, let us write

$$g_{ab} = e^{2\sigma} \bar{g}_{ab} \,, \tag{1.2.16}$$

where $\bar{g}_{ab}$ is a fixed reference metric. The function $\sigma(x)$ of spacetime coordinates is known as the conformal mode of the metric. Correspondingly, the Ricci scalar can be written as

$$R = e^{-2\sigma}(\overline{R} - 2\overline{\Box}\sigma) \,, \tag{1.2.17}$$

where the overbars denote the quantities computed with the metric $\bar{g}_{ab}$. In $D = 2$, we can always find a local coordinate transformation such that the reference metric is $\bar{g}_{ab} = \eta_{ab}$ and hence $g_{ab} = e^{2\sigma}\eta_{ab}$. Then the Ricci scalar simplifies to $R = -2e^{-2\sigma}\Box_\eta\sigma$ where $\Box_\eta$ is the flat-space d'Alembertian, while the determinant of the metric gives $\sqrt{-g} = e^{2\sigma}$. From eq. (1.2.9) it follows that

$$\delta\Gamma_m = \frac{1}{2}\int d^2x \sqrt{-g}\ \langle 0|T^{ab}|0\rangle \delta g_{ab} = \int d^2x \sqrt{-g}\ \langle 0|T^{ab}|0\rangle g_{ab}\delta\sigma \,, \tag{1.2.18}$$

and we can express the functional derivative of $\Gamma[\sigma]$ with respect to $\sigma(x)$ as

$$\frac{\delta\Gamma_m}{\delta\sigma} = \sqrt{-g}\,\langle 0|T^a_a|0\rangle = -\frac{N}{12\pi}\Box_\eta\sigma \,. \tag{1.2.19}$$

Finally, the integration of eq. (1.2.19) gives

$$\Gamma_m[\sigma] - \Gamma_m[0] = -\frac{N}{24\pi}\int d^2x\, \sigma\Box_\eta\sigma \,. \tag{1.2.20}$$

---

[4] For reference, we recall that $S_s = -\frac{1}{2}\int d^4x \sqrt{-g}\left(g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + \frac{1}{6}R\varphi^2\right)$ is the action for a conformally-coupled scalar field φ in $D = 4$ spacetime dimensions. The action is invariant under conformal transformations $g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \Omega^2(x) g_{\mu\nu}$, for any arbitrary function of spacetime $\Omega^2(x)$.

Since we have set $D = 2$ we can say that $\Gamma_m[0] = 0$, because when $\sigma = 0$ we have locally $g_{ab} = \eta_{ab}$ and there cannot be any curvature determining a non-zero value of $\Gamma_m[0]$. The exactness of the trace anomaly in eq. (1.2.15) also implies that this derivation of the anomaly-induced QEA in $D = 2$ is exact at all perturbative orders.

The result can also be rewritten in a generally-covariant form by using nonlocal operators. To achieve such a form we observe that $\Box_g = e^{-2\sigma}\Box_\eta$, where $\Box_g$ is the d'Alembertian computed with the full metric $g_{ab} = e^{2\sigma}\eta_{ab}$. Then, the equation $R = -2e^{-2\sigma}\Box_\eta\sigma$ gives $R = -2\Box_g\sigma$, which can be inverted as $\sigma = -(1/2)\Box_g^{-1}R$. The inverse d'Alembertian $\Box_g^{-1}$ is the nonlocal operator that allows us to write the QEA in a manifestly covariant way, known as the Polyakov QEA:

$$\Gamma_m[g_{\mu\nu}] = -\frac{N}{24\pi}\int d^2x\, e^{2\sigma}\sigma\Box_g\sigma = -\frac{N}{96\pi}\int d^2x\,\sqrt{-g}\,R\Box_g^{-1}R\,. \tag{1.2.21}$$

As we said at the end of Section 1.2.1, a study of the full quantum gravitational dynamics also requires to integrate over the quantum fluctuations of the metric. When integrating over those metric flucuations in the path integral, the setting of $g_{ab} = e^{2\sigma}\bar{g}_{ab}$, as in eq. (1.2.16), gives a gauge-fixing condition. Therefore, besides the conformal mode σ, also the Faddeev-Popov ghosts will affect the path-integral integration. It can be shown, see [60, 61, 62], that the Faddeev-Popov ghosts give a contribution -26 to be added to N, while the conformal factor σ gives a contribution $+1$. Therefore, neglecting the topologically-invariant Einstein-Hilbert term and including the cosmological constant λ, the exact quantum effective action of 2D gravity reads

$$\Gamma = -\frac{N-25}{96\pi}\int d^2x\sqrt{-g}\,R\frac{1}{\Box}R - \lambda\int d^2x\,\sqrt{-g}\,. \tag{1.2.22}$$

From eq. (1.2.17) and dropping terms depending only on the reference metric $\bar{g}_{ab}$, and not on σ, we can write a local expression for Γ in terms of the conformal mode:

$$\Gamma = \int d^2x\,\sqrt{-\bar{g}}\left[\frac{N-25}{24\pi}\bar{g}^{ab}\partial_a\sigma\partial_b\sigma + \frac{N-25}{24\pi}\bar{R}\sigma - \lambda e^{2\sigma}\right]. \tag{1.2.23}$$

It is useful to explain a possible confusion that emerges when trying to extract the spectrum of the theory from the QEA in eq. (1.2.23). We are going to see that such an operation is not allowed and only the fundamental action correctly expresses the physical content of the theory. We can also rephrase the situation by saying that the interpretation of a formal expression like that in eq. (1.2.23) changes completely in the two cases in which we take it as a fundamental action or as a quantum effective action. Understanding this point is important for a correct interpretation of the phenomenological nonlocal models that we will study later.

Let us try to read the spectrum of the quantum theory from eq. (1.2.23), treating it as if it were the fundamental action of a QFT; then we would conclude that, for $N \neq 25$, there is one dynamical degree of freedom, σ. Recalling that our signature is $\eta_{ab} = (-,+)$, we would also conclude that for $N > 25$ this degree of freedom is a ghost and for $N < 25$ it has the right sign for the kinetic term. For $N = 25$ there is no dynamics at all.

This conclusion is clearly wrong because we know that eq. (1.2.23) is the QEA of a fundamental theory which is just 2D gravity coupled to N healthy fields, and there cannot be any ghosts in the spectrum of the fundamental theory. If we perform the quantization of the fundamental theory in the gauge of eq. (1.2.16), the fields involved are the matter fields, the Faddeev-Popov ghosts and the conformal mode σ. Each of them has its own creation and annihilation operators, which generate the full Hilbert space of the theory. However, as always in theories with a local invariance (in this case diffeomorphism invariance) the physical Hilbert space is a subset of the full Hilbert space. Similarly to the physical-state condition $\langle s'|\partial_\mu A^\mu|s\rangle = 0$ required in the Gupta-Bleuler quantization of electrodynamics to discard the ghost states associated to A_0, here the condition on physical states $|s\rangle$ and $|s'\rangle$ can be obtained from

$$\langle s'|T^{ab}_{\rm tot}|s\rangle = 0\,, \tag{1.2.24}$$

where $T^{ab}_{\rm tot}$ is the sum of the energy-momentum tensors of matter, ghosts and σ. As explicitly proved in [63], this condition eliminates from the physical spectrum both the states associated with the reparametrization ghosts and the states generated by the creation operators of the conformal mode. The physical spectrum of the fundamental theory is simply given by the quanta of the N healthy matter fields[5], with no ghosts in the theory. Those N matter fields are not visible in the QEA of eq. (1.2.23) because we integrated them out in the path integral. We also observe that the fact that no physical quanta are associated to σ does not mean that the field σ itself has no physical effects. For example in QED there are no physical quanta associated to A_0, but the interaction mediated by A_0 is exactly what is responsible for the Coulomb potential in the static case. In the language of Feynaman diagrams, this is possible because the condition of having no physical states for σ in gravity (or for A_0 in QED) just means that σ cannot appear in external lines, but it can still appear in internal lines and mediate interactions in that way.

In the case $D = 4$ the trace anomaly is

$$\langle 0|T^\mu_\mu|0\rangle = b_1 C^2 + b_2\left(E - \frac{2}{3}\Box R\right) + b_3\Box R\,, \tag{1.2.25}$$

where $C^2 \equiv C^{\mu\nu\rho\sigma}C_{\mu\nu\rho\sigma}$ is the square of the Weyl tensor, $E \equiv R^2 - 4R^{\mu\nu}R_{\mu\nu} + R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ is the Gauss-Bonnet term, and the values of the coefficients b_1, b_2, b_3 depend on the number of massless conformally-coupled scalars, fermions and vector fields. This result for the trace anomaly is exact and receives contribution only at one loop order. When we introduce the conformal mode σ by writing $g_{\mu\nu} = e^{2\sigma}\bar{g}_{\mu\nu}$, we cannot set $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$ as we did in the case $D = 2$. Explicitly, the integration of the conformal anomaly gives

$$\begin{aligned}\Gamma_{\rm anom}[g_{\mu\nu}] \;=\; & \Gamma_{\rm anom}[\bar{g}_{\mu\nu}] - \frac{b_3}{12}\int d^4x\,\sqrt{-g}\,R^2 \\ & + \int d^4x\,\sqrt{-\bar{g}}\left[b_1\sigma\bar{C}^2 + b_2\sigma\left(\bar{E} - \frac{2}{3}\bar{\Box}\bar{R}\right) + 2b_2\sigma\bar{\Delta}_4\sigma\right],\end{aligned} \tag{1.2.26}$$

where Δ_4 is the Paneitz operator

$$\Delta_4 \equiv \Box^2 + 2R^{\mu\nu}\nabla_\mu\nabla_\nu - \frac{2}{3}R\Box + \frac{1}{3}g^{\mu\nu}\nabla_\mu R\nabla_\nu\,. \tag{1.2.27}$$

The quantity $\Gamma_{\rm anom}[\bar{g}_{\mu\nu}]$, corresponding to $\Gamma_{\rm anom}[g_{\mu\nu}]$ evaluated at $\sigma = 0$, cannot be determined from the conformal anomaly alone, because in $D = 4$ we can no longer

---

[5] The conformal mode is never a propagating degree of freedom and, in D=2, there are no graviton polarizations.

set $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$. The maximal information that we can get is the dependence of the QEA on the conformal mode σ.

The covariantization of eq. (1.2.26) is not uniquely determined and a possiblity is given by the Riegert action [64]

$$\begin{aligned}\Gamma_{\rm anom}[g_{\mu\nu}] &= \Gamma_c[g_{\mu\nu}] - \frac{b_3}{12}\int d^4x\,\sqrt{-g}\,R^2 \\ &+\frac{1}{8}\int d^4x\sqrt{-g}\left(E-\frac{2}{3}\Box R\right)\Delta_4^{-1}\left[b_2\left(E-\frac{2}{3}\Box R\right)+2b_1C^2\right].\end{aligned} \tag{1.2.28}$$

It is interesting that, again, a covariant form requires the use of nonlocal operators. In this case the inverse Paneitz operator is needed, while for $D = 2$ the inverse d'Alembertian appeared.

1.3 Nonlocality and mass terms for gauge theories

Now that the notion of quantum effective action has been discussed, we can explain how it may contain nonlocal mass terms for gauge fields (or for the gravitational field), despite preserving gauge (or diffeomorphism) invariance. We follow the presentation in [31].

1.3.1 Massive gauge fields

A simple example is given by massive electrodynamics considered by Dvali in [65]. The Proca action for a massive photon coupled to an external conserved current j^μ is

$$S = \int d^4x\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m_\gamma^2\,A_\mu A^\mu - j_\mu A^\mu\right), \tag{1.3.1}$$

and the corresponding equation of motion reads $\partial_\mu F^{\mu\nu} - m_\gamma^2 A^\nu = j^\nu$. Applying the derivative ∂_ν on both sides and using the current conservation equation $\partial_\nu j^\nu = 0$ we find $m_\gamma^2\,\partial_\nu A^\nu = 0$. Therefore, for $m_\gamma \neq 0$, we obtain $\partial_\nu A^\nu = 0$. Despite this condition is the same as the Lorentz gauge, one should keep in mind that there is no gauge invariance in massive electrodynamics. Thus $\partial_\nu A^\nu = 0$ is just a dynamical consequence of the equations of motion, and does not express any gauge fixing condition, as it would not make any sense. Using $\partial_\nu A^\nu = 0$, the equation of motion becomes $(\Box - m_\gamma^2)A^\mu = j^\mu$, so that we can summarize the equations of motion derived from the Proca action (1.3.1) as

$$(\Box - m_\gamma^2)A^\mu = j^\mu\,, \qquad \partial_\nu A^\nu = 0\,. \tag{1.3.2}$$

As it is well known, the theory exhibits three degrees of freedom and describes a massive spin-1 particle.

Now, let us compare the Proca action with the nonlocal action

$$S = \int d^4x\left[-\frac{1}{4}F_{\mu\nu}\left(1-\frac{m_\gamma^2}{\Box}\right)F^{\mu\nu} - j_\mu A^\mu\right], \tag{1.3.3}$$

whose equation of motion reads

$$\left(1-\frac{m_\gamma^2}{\Box}\right)\partial_\mu F^{\mu\nu} = j^\nu\,. \tag{1.3.4}$$

The action (1.3.3) is nonlocal but gauge invariant, and we can therefore impose $\partial_\mu A^\mu = 0$ as a choice of gauge, so that eq. (1.3.4) reduces to $(\Box - m_\gamma^2)A^\mu = j^\mu$. We ended up with the same two equations in (1.3.2), showing that eqs. (1.3.1) and (1.3.3) are two equivalent formulations of the same classical theory. Another way to show the equivalence of the two theories is to use the Stückelberg trick. We introduce the Stückelberg field φ and replace

$$A_\mu \to A_\mu + (1/m_\gamma)\partial_\mu\varphi\,, \tag{1.3.5}$$

in the Proca action (1.3.1), which becomes

$$S[A_\mu,\varphi] = \int d^4x\left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m_\gamma^2 A_\mu A^\mu - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - m_\gamma A^\mu\partial_\mu\varphi - j_\mu A^\mu\right]. \tag{1.3.6}$$

We have added a new degrees of freedom φ, but we have gained a gauge symmetry, defined by the transformation

$$A_\mu \to A_\mu - \partial_\mu\theta\,, \qquad \varphi \to \varphi + m_\gamma\theta\,, \tag{1.3.7}$$

since the combined transformation leaves invariant the right-hand side of eq. (1.3.5). The equations of motion obtained from the action (1.3.6) are

$$\partial_\mu F^{\mu\nu} = m_\gamma^2 A^\nu + m_\gamma\partial^\nu\varphi + j^\nu\,, \tag{1.3.8}$$

$$\Box\varphi = -m_\gamma\partial_\mu A^\mu\,. \tag{1.3.9}$$

Then eq. (1.3.9) can be formally solved by $\varphi(x) = -m_\gamma\Box^{-1}(\partial_\mu A^\mu)$ and inserting this into eq. (1.3.8) we get eq. (1.3.4) [or, equivalently, inserting it into $S[A_\mu,\varphi]$, eq. (1.3.6), we get eq. (1.3.3)]. It is worth observing that the Ward identities of QED do not forbid a photon mass term [66]. Indeed, they only imply that the photon self-energy $\Sigma_{\mu\nu}$ is transverse, so that, in momentum space, it can be written as

$$\Sigma_{\mu\nu}(p) = \left(g_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right)F(p^2)\,. \tag{1.3.10}$$

If, in the limit $p^2 \to 0$, $F(p^2) \neq 0$, then the photon acquires a nonzero mass, which is precisely the one described by the nonlocal term in eq. (1.3.3).

Following [67] it is also useful to rewrite the nonlocal mass term in a different way that will be enable us to make contact with the gravitational case. We separate the gauge field into its transverse and longitudinal parts,

$$A_\mu = A_\mu^{\rm T} + \partial_\mu\alpha\,, \tag{1.3.11}$$

where $\partial^\mu A_\mu^{\rm T} = 0$. Under a gauge transformations $A_\mu \to A_\mu - \partial_\mu\theta$, we have $\alpha \to \alpha - \theta$ and $A_\mu^{\rm T} \to A_\mu^{\rm T}$, so $A_\mu^{\rm T}$ is gauge invariant. To invert eq. (1.3.11) we take the divergence, which gives $\partial^\mu A_\mu = \Box\alpha$. This can be formally inverted as $\alpha = \Box^{-1}\partial^\mu A_\mu$. Substituting this into $A_\mu^{\rm T} = A_\mu - \partial_\mu\alpha$ we get

$$A_\mu^{\rm T} = A_\mu - \frac{1}{\Box}\partial_\mu\partial^\nu A_\nu \equiv P_\mu^\nu A_\nu\,, \tag{1.3.12}$$

where

$$P_\mu^\nu \equiv \delta_\mu^\nu - \frac{\partial_\mu\partial^\nu}{\Box} \tag{1.3.13}$$

is a nonlocal operator. The transverse part A^{T}_μ is therefore a gauge-invariant and nonlocal functional of the gauge field A_μ. In terms of A^{T}_μ, it is straightforward to check that the action (1.3.3) can be rewritten as[6]

$$S = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} m_\gamma^2 A^{\mathrm{T}}_\mu A^{\mathrm{T}\mu} - j^\mu A_\mu \right) . \tag{1.3.14}$$

In the case of massive electrodynamics the nonlocality is only apparent, since we have seen that the nonlocal term in the equation of motion (1.3.4) can be made local with the gauge choice $\partial_\mu A^\mu = 0$. In this sense, the above manipulations can be performed even at the level of fundamental action, since anyhow the nonlocality can be gauged away. Using non-abelian gauge theories we can build less trivial examples where the nonlocality is genuine and only makes sense at the level of quantum effective action, because, as we will explain in Section 1.3.3, nonlocalities in the fundamental action are not compatible with causality, while they are perfectly fine in the QEA.

The non-abelian generalization of the nonlocal mass term in eq. (1.3.3) is

$$\frac{m_g^2}{2} \mathrm{Tr} \int d^4x \, F_{\mu\nu} \frac{1}{\Box} F^{\mu\nu} , \tag{1.3.15}$$

where $F_{\mu\nu} = F^a_{\mu\nu} T^a$, $\Box^{ab} = D^{ac}_\mu D^{\mu,cb}$ and $D^{ab}_\mu = \delta^{ab} \partial_\mu - g f^{abc} A^c_\mu$ is the covariant derivative. Comparing with eq. (1.3.3) we see that this nonlocal term corresponds to giving a mass m_g to the non-abelian gauge bosons (plus extra nonlocal interaction terms to reconstruct a gauge-invariant quantity). Such a nonlocal mass term cannot be reduced to a local term with a gauge choice, and has been postulated to appear in the quantum effective action of QCD, in order to reproduce lattice QCD non-perturbative results on the running of the strong coupling constant and on the gluon propagator at low energies, see [68, 69, 70].

1.3.2 A mass for the conformal mode of gravity

Nonlocal variables and linearized GR over Minkowski

We now discuss possible generalizations of the above construction to the gravitational field, following [31]. A possible route is to begin with gravity linearized over Minkowski space. First of all, it is useful to see how linearized gravity can be rewritten in terms of nonlocal variables, analogous to A^T_μ of the previous Section (we follow the discussions in [67, 71]). We begin by writing $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$, where $\kappa = (32\pi G)^{1/2}$. To quadratic level, the Einstein-Hilbert action becomes

$$S^{(2)}_{\mathrm{EH}} = \frac{1}{2} \int d^4x \, h_{\mu\nu} \mathcal{E}^{\mu\nu,\rho\sigma} h_{\rho\sigma} , \tag{1.3.16}$$

---

[6] We can further replace A_μ with A^{T}_μ both in the kinetic term and in the interaction with the current of eq. (1.3.14), because α is a pure gauge degree of freedom and it can then be set to zero with a gauge transformation.

where $\mathcal{E}^{\mu\nu,\rho\sigma}$ is the Lichnerowicz operator,[7] while the interaction with matter with energy-momentum tensor $T^{\mu\nu}$, to linear order in $h_{\mu\nu}$, is given by

$$S^{(1)}_{\rm int} = \frac{\kappa}{2}\int d^4x\, h_{\mu\nu}T^{\mu\nu}\,. \tag{1.3.17}$$

The linearized equations of motion derived from $S^{(2)}_{\rm EH}+S^{(1)}_{\rm int}$ are therefore

$$\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma} = -\frac{\kappa}{2}T^{\mu\nu}\,. \tag{1.3.18}$$

We next decompose the metric as

$$h_{\mu\nu} = h^{\rm TT}_{\mu\nu} + \frac{1}{2}(\partial_\mu\epsilon_\nu + \partial_\nu\epsilon_\mu) + \frac{1}{3}\eta_{\mu\nu}s\,, \tag{1.3.19}$$

where $h^{\rm TT}_{\mu\nu}$ is transverse ($\partial^\mu h^{\rm TT}_{\mu\nu}=0$) and traceless ($\eta^{\mu\nu}h^{\rm TT}_{\mu\nu}=0$), and therefore has five independent components. We have therefore decomposed the 10 independent components of the symmetric tensor $h_{\mu\nu}$ into the five components of $h^{\rm TT}_{\mu\nu}$, the four components of ϵ_μ, and the scalar s. Under a linearized diffeomorphism $h_{\mu\nu}\to h_{\mu\nu}-(\partial_\mu\xi_\nu+\partial_\nu\xi_\mu)$ we have $\epsilon_\mu\to\epsilon_\mu-\xi_\mu$ while $h^{\rm TT}_{\mu\nu}$ and s are gauge invariant. Thus ϵ_μ describes the four pure gauge degrees of freedom, while s plus the five components of the TT tensor $h^{\rm TT}_{\mu\nu}$ describe the six gauge-invariant degrees of freedom of the gravitational field. Notice that, at this linearized level, s is equivalent to the conformal mode of the metric. Indeed, restricting to the scalar sector (i.e. setting $\epsilon_\mu=0$ and $h^{\rm TT}_{\mu\nu}=0$) and writing $g_{\mu\nu}=e^{2\sigma}\eta_{\mu\nu}$, comparison with eq. (1.3.19) shows that, at the linear level, $2\sigma=s/3$.

Similarly to the electromagnetic case of section 1.3.1, the quantities that appear in the right-hand side of eq. (1.3.19) are nonlocal functionals of the original metric perturbation $h_{\mu\nu}$. The inversion of eq. (1.3.19) is straightforward [67]. It is convenient to further separate ϵ_μ into its transverse and longitudinal parts, $\epsilon_\mu=\epsilon^{\rm T}_\mu+\partial_\mu\alpha$, where $\partial^\mu\epsilon^{\rm T}_\mu=0$. Then, taking the trace of eq. (1.3.19) we get $h=(4/3)s+\Box\alpha$, while contracting eq. (1.3.19) with $\partial^\mu\partial^\nu$, gives $\partial^\mu\partial^\nu h_{\mu\nu}=\Box[s/3+\Box\alpha]$. Combining these equations we get

$$s=\left(\eta^{\mu\nu}-\frac{1}{\Box}\partial^\mu\partial^\nu\right)h_{\mu\nu}\,,\qquad \alpha=-\frac{1}{3}\frac{1}{\Box}\left(\eta^{\mu\nu}-\frac{4}{\Box}\partial^\mu\partial^\nu\right)h_{\mu\nu}\,. \tag{1.3.20}$$

We can now extract $\epsilon^{\rm T}_\mu$ by applying ∂^μ to eq. (1.3.19) and using the above expressions for α and s. This gives $\epsilon^T_\mu=2\Box^{-1}P^\rho_\mu\partial^\sigma h_{\rho\sigma}$. Finally, substituting these expressions into eq. (1.3.19) we get

$$\begin{aligned} h^{\rm TT}_{\mu\nu} &= h_{\mu\nu}-\frac{1}{3}\left(\eta_{\mu\nu}-\frac{\partial_\mu\partial_\nu}{\Box}\right)h-\frac{1}{\Box}(\partial_\mu\partial^\rho h_{\nu\rho}+\partial_\nu\partial^\rho h_{\mu\rho})+\frac{1}{3}\eta_{\mu\nu}\frac{1}{\Box}\partial^\rho\partial^\sigma h_{\rho\sigma}\\ &\quad+\frac{2}{3}\frac{1}{\Box^2}\partial_\mu\partial_\nu\partial^\rho\partial^\sigma h_{\rho\sigma}\,. \end{aligned} \tag{1.3.21}$$

---

[7] Its precise definition is $\mathcal{E}^{\mu\nu,\rho\sigma}\equiv\frac{1}{2}(\eta^{\mu\rho}\eta^{\nu\sigma}+\eta^{\mu\sigma}\eta^{\nu\rho}-2\eta^{\mu\nu}\eta^{\rho\sigma})\Box+(\eta^{\rho\sigma}\partial^\mu\partial^\nu+\eta^{\mu\nu}\partial^\rho\partial^\sigma)-\frac{1}{2}(\eta^{\mu\rho}\partial^\sigma\partial^\nu+\eta^{\nu\rho}\partial^\sigma\partial^\mu+\eta^{\mu\sigma}\partial^\rho\partial^\nu+\eta^{\nu\sigma}\partial^\rho\partial^\mu)$.

These results can be written more compactly using the projector $P^{\mu\nu}=\eta^{\mu\nu}-(\partial^\mu\partial^\nu/\Box)$. In particular,

$$s = P^{\mu\nu}h_{\mu\nu}\,, \tag{1.3.22}$$

$$h^{\rm TT}_{\mu\nu} = \left(P_\mu^\rho P_\nu^\sigma-\frac{1}{3}P_{\mu\nu}P^{\rho\sigma}\right)h_{\rho\sigma}\,. \tag{1.3.23}$$

The fact that this expression for $h^{\rm TT}_{\mu\nu}$ is indeed transverse and traceless is easily checked by using the properties of $P_{\mu\nu}$, $\partial^\mu P_{\mu\nu}=0$, $\eta^{\mu\nu}P_\mu^\rho P_\nu^\sigma=P^{\rho\sigma}$ and $\eta^{\mu\nu}P_{\mu\nu}=3$.

Plugging the decomposition (1.3.19) into the action (1.3.16) one finds that ϵ_μ cancels (an obvious consequence of the fact that it is a pure gauge mode), and [71]

$$S^{(2)}_{\rm EH}=\frac{1}{2}\int d^4x\left[h^{\rm TT}_{\mu\nu}\Box(h^{\mu\nu})^{\rm TT}-\frac{2}{3}s\Box s\right]\,. \tag{1.3.24}$$

Performing the same decomposition as in (1.3.19) for the energy-momentum tensor, the linearization of the interaction term becomes

$$S^{(1)}_{\rm int}=\frac{\kappa}{2}\int d^4x\left[h^{\rm TT}_{\mu\nu}(T^{\mu\nu})^{\rm TT}+\frac{1}{3}sT\right]\,, \tag{1.3.25}$$

where $T=\eta^{\mu\nu}T_{\mu\nu}$. The equations of motion (1.3.18) derived from $S^{(2)}_{\rm EH}+S^{(1)}_{\rm int}$ can then be rewritten as

$$\Box h^{\rm TT}_{\mu\nu}=-\frac{\kappa}{2}T^{\rm TT}_{\mu\nu}\,,\qquad\qquad \Box s=\frac{\kappa}{4}T\,. \tag{1.3.26}$$

At first, eq. (1.3.26) can be surprising, because it seems to imply that $h^{\rm TT}_{\mu\nu}$ and s describe six radiative gauge-invariant degrees of freedom. Of course, we know that in GR only the two degrees of freedom associated to the helicities ±2 are radiative, while the remaining four gauge-invariant degrees of freedom are non-radiative and satisfy Poisson equations. Furthermore, the sign of the kinetic term of s in eq. (1.3.24) is such that the scalar s seems to be a ghost. As discussed in [71], the resolution of this apparent paradox is related to the nonlocal relation between the original metric perturbation $h_{\mu\nu}$ and the variables $\{h^{\rm TT}_{\mu\nu},s\}$. The fact that this relation is nonlocal in time, and not only in space, implies that the initial data assigned on $h_{\mu\nu}$ on a given time slice are not sufficient to provide initial data on $\{h^{\rm TT}_{\mu\nu},s\}$, so a naive counting of degrees of freedom in terms of $\{h^{\rm TT}_{\mu\nu},s\}$ goes wrong. A simple example to understand what exactly goes wrong, again discussed in [71], is provided by a scalar field ϕ that satisfies a Poisson equation $\nabla^2\phi=\rho$. If one introduces a field $\tilde\phi$ related to ϕ by a nonlocal relation such as $\tilde\phi=\Box^{-1}\phi$, the original Poisson equation can be rewritten as $\Box\tilde\phi=\tilde\rho$, where $\tilde\rho\equiv\nabla^{-2}\rho$, so now $\tilde\phi$ looks like a propagating degree of freedom. However, for $\rho=0$ the original equation $\nabla^2\phi=\rho$ (with vanishing boundary conditions at infinity) only has the solution $\phi=0$. If we want to rewrite this equation in terms of $\tilde\phi$ without introducing spurious degrees of freedom we must therefore supplement the equation $\Box\tilde\phi=\tilde\rho$ with the condition that, when $\tilde\rho=0$, the only acceptable solution is $\tilde\phi=0$, which precisely kills the radiative solution. Notice that the nonlocal relation between the variables $\{h^{\rm TT}_{\mu\nu},s\}$ and the full $h_{\mu\nu}$ is different from what happens in the standard $3+1$ decomposition of the metric perturbations over flat space,

$$\begin{aligned} h_{00} &= 2\psi\,, \qquad\qquad h_{0i} = \beta_i + \partial_i\gamma \\ h_{ij} &= -2\phi\delta_{ij} + \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\boldsymbol{\nabla}^2\right)\lambda + \frac{1}{2}(\partial_i v_j + \partial_j v_i) + H_{ij}^{\rm TT}\,, \end{aligned} \tag{1.3.27}$$

where v^i and β^i are transverse spatial vectors, $\partial_i\beta^i = 0$ and $\partial_i v^i = 0$, and $H_{ij}^{\rm TT}$ is transverse and traceless with respect to the spatial indices, $\partial^j H_{ij}^{\rm TT} = 0$ and $\delta^{ij}H_{ij}^{\rm TT} = 0$. Indeed, the $3+1$ decomposition only involves spatial derivatives and therefore its inversion is nonlocal in space but local in time. From these variables, one can form six variables that are invariant under linearized gauge transformations: the two Bardeen variables, $\Phi = -\phi - (1/6)\boldsymbol{\nabla}^2\lambda$ and $\Psi = \psi - \dot{\gamma} + (1/2)\ddot{\lambda}$, that are scalars under spatial rotations; the spatial vector $\Xi_i = \beta_i - (1/2)\dot{v}_i$, which, being transverse, has only two independent components; and the spatial tensor $H_{ij}^{\rm TT}$, which is already gauge-invariant (again, at the linearized level) and, being transverse and traceless (and carrying only spatial indices, contrary to $h_{\mu\nu}^{\rm TT}$), also has only two independent components. Standard analysis (see e.g. [72, 73] or chapter 18 of [74]) then shows that, after performing the same decomposition for the energy-momentum tensor, namely

$$\begin{aligned} T_{00} &= \rho\,, \qquad\qquad T_{0i} = S_i + \partial_i S\,, \\ T_{ij} &= P\delta_{ij} + \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\boldsymbol{\nabla}^2\right)\Sigma + \frac{1}{2}(\partial_i\Sigma_j + \partial_j\Sigma_i) + \Sigma_{ij}\,, \end{aligned} \tag{1.3.28}$$

where $\partial_i\Sigma^i = 0$, $\partial_i S^i = 0$, $\partial^i\Sigma_{ij} = 0$ and $\delta^{ij}\Sigma_{ij} = 0$, the linearized equations of motion can be rewritten as

$$\boldsymbol{\nabla}^2\Phi = -4\pi G\rho\,, \qquad \boldsymbol{\nabla}^2\Psi = -4\pi G(\rho - 2\boldsymbol{\nabla}^2\Sigma)\,, \tag{1.3.29}$$

$$\boldsymbol{\nabla}^2\Xi_i = -16\pi G S_i\,, \qquad \Box H_{ij}^{\rm TT} = -16\pi G\Sigma_{ij}\,. \tag{1.3.30}$$

We then get the standard result that only the two degrees of freedom of the tensor perturbations obey a wave equation, while the remaining gauge-invariant degrees of freedom described by Φ, Ψ and Ξ_i obey Poisson equations, and therefore are non-radiative.

Comparing the decompositions (1.3.19) and (1.3.27) one finds that the field s can be written explicitly as a nonlocal function of the Bardeen variables as [71]

$$s = 6\Phi - 2\Box^{-1}\boldsymbol{\nabla}^2(\Phi + \Psi)\,. \tag{1.3.31}$$

The apparent radiative nature of s in eq. (1.3.26) is an artifact due to this nonlocal relation, that introduces a spurious degree of freedom associated to the homogeneous equation $\Box s = 0$. Indeed, from eq. (1.3.29), $\boldsymbol{\nabla}^2(\Phi + \Psi)$ is fully determined by the source terms, and vanishes if the latter vanish. Thus, in order to eliminate this spurious degree of freedom we must supplement eq. (1.3.26) with the condition that $s = 0$ when $T = 0$, i.e. we must discard again the homogeneous solution of eq. (1.3.26) (and similarly for the helicities $0, \pm 1$ of $h_{\mu\nu}^{\rm TT}$). At the quantum level, this implies that there are no creation and annihilation operators associated to s, and s cannot appear on the external legs of a Feynman diagram. Therefore, the apparent ghost-like nature of s in eq. (1.3.24) is fictitious and, of course, in General Relativity there is no actual ghost.

Nonlocal mass terms at the linearized level

As we have seen, the use of the variables $\{h_{\mu\nu}^{\rm TT}, s\}$ is not convenient if we want to count the independent degrees of freedom of the theory and determine their radiative/non-radiative nature; for those purposes it is better to work directly with the original metric perturbation $h_{\mu\nu}$, or with the variables of the 3+1 decomposition (1.3.27), or with the ADM decomposition. However, the variables $\{h_{\mu\nu}^{\rm TT}, s\}$ have the advantage that one can very easily see how a diff-invariant nonlocal mass term can be naturally written for different modes of the gravitational field, at the linearized level. Following [31] , we can simply modify eq. (1.3.24) into

$$\Gamma^{(2)} = \frac{1}{2}\int d^4x \left[h_{\mu\nu}^{\rm TT}(\Box - m_1^2)(h^{\mu\nu})^{\rm TT} - \frac{2}{3}s(\Box + m_2^2)s \right] , \tag{1.3.32}$$

for some masses m_1 and m_2, where we have chosen the signs in front of m_1^2 and m_2^2 so that $m_1^2 > 0$ and $m_2^2 > 0$ corresponds to non-tachyonic masses. These mass terms are clearly diff-invariant, since $h_{\mu\nu}^{\rm TT}$ and s are diff-invariant (again, at the linearized level). On the other hand, because of the relations (1.3.22,1.3.23), once rewritten in terms of $h_{\mu\nu}$ they will be nonlocal. We have indeed used the notation Γ, rather than S, to stress that, because of the nonlocality, this modification makes sense at the level of the quantum effective action Γ, rather than for the fundamental action S.

To go beyond the linearized approximation, we can search for covariantizations of these expressions, as we will do in Section 1.3.2. The second term gives a mass to s or, equivalently, to the conformal mode. The models that we will study in the following will be covariantizations of the above expression, with $m_1^2 = 0$ and $m_2^2 \equiv m^2 > 0$. We are therefore assuming that there exists a mechanism that, in the quantum effective action, generates a mass for the conformal mode, while leaving $h_{\mu\nu}^{\rm TT}$ massless. At the phenomenological level this is required by the fact that, among a large class of models explored, only those of this form appear to have a viable cosmological evolution. At the theoretical level, this is also suggested by various arguments, that will be discussed in Section 1.3.4. Thus, we will look for a covariantization of a quantum effective action that, at quadratic level, has the form

$$\Gamma^{(2)} = \frac{1}{2}\int d^4x \left[h_{\mu\nu}^{\rm TT}\Box(h^{\mu\nu})^{\rm TT} - \frac{2}{3}s(\Box + m^2)s \right] , \tag{1.3.33}$$

so that the linearized equations of motion (1.3.26) are modified into

$$\Box h_{\mu\nu}^{\rm TT} = -\frac{\kappa}{2}T_{\mu\nu}^{\rm TT} , \qquad (\Box + m^2)s = \frac{\kappa}{4}T . \tag{1.3.34}$$

To perform the covariantization, it is now convenient to go back to the original metric perturbation $h_{\mu\nu}$. Using eq. (1.3.22), we immediately see that eq. (1.3.33) can be rewritten as

$$\Gamma^{(2)} = \frac{1}{2}\int d^4x \left[h_{\mu\nu}\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma} - \frac{2}{3}m^2(P^{\mu\nu}h_{\mu\nu})^2 \right] , \tag{1.3.35}$$

while eq. (1.3.34) is equivalent to

$$\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma} - \frac{2}{3}m^2P^{\mu\nu}P^{\rho\sigma}h_{\rho\sigma} = -\frac{\kappa}{2}T^{\mu\nu} . \tag{1.3.36}$$

In view of the covariantization, it is also convenient to rescale $h_{\mu\nu} \to h_{\mu\nu}/\kappa$, so that now $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$,

$$\Gamma^{(2)} = \frac{1}{64\pi G} \int d^4x \left[h_{\mu\nu}\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma} - \frac{2}{3}\, m^2 (P^{\mu\nu}h_{\mu\nu})^2 \right] , \tag{1.3.37}$$

and

$$\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma} - \frac{2}{3}\, m^2 P^{\mu\nu}P^{\rho\sigma}h_{\rho\sigma} = -16\pi G\, T^{\mu\nu} . \tag{1.3.38}$$

The RR and RT models

We now look for possible covariantizations of the above expressions, as done in [31]. Covariantizations, when they exists, are in general not unique. However, some choices can be more natural than others. We will see that, starting from the equation of motion (1.3.38) or from the quantum effective action (1.3.37), one ends up quite naturally with two different covariantizations, that define two possible models.

Let us start from the covariantization of eq. (1.3.38). The linearization of the Einstein tensor $G_{\mu\nu}$ is $G^{(1)}_{\mu\nu} = -(1/2)\mathcal{E}_{\mu\nu,\rho\sigma}h^{\rho\sigma}$, so the term $\mathcal{E}_{\mu\nu,\rho\sigma}h^{\rho\sigma} = -2G^{(1)}_{\mu\nu}$ in eq. (1.3.38) is uniquely promoted to $-2G_{\mu\nu}$ in the full covariant theory, by the requirement that we recover GR for $m = 0$. The nontrivial part is the covariantization of the mass term. At linear level the Ricci scalar becomes $R^{(1)} = -(\eta^{\rho\sigma}\Box - \partial^\rho\partial^\sigma)h_{\rho\sigma}$, that can be rewritten as $R^{(1)} = -\Box(P^{\rho\sigma}h_{\rho\sigma})$, so

$$P^{\rho\sigma}h_{\rho\sigma} = -\Box^{-1}R^{(1)} . \tag{1.3.39}$$

Therefore eq. (1.3.36) is equivalent to

$$-2G^{(1)}_{\mu\nu} + \frac{2}{3}\, m^2 P_{\mu\nu}\Box_\eta^{-1}R^{(1)} = -16\pi G\, T_{\mu\nu} , \tag{1.3.40}$$

where the notation $\Box_\eta$ stresses that, until now, the $\Box$ operator was the one with respect to the flat metric $\eta_{\mu\nu}$. After promoting $G^{(1)}_{\mu\nu}$ to $G_{\mu\nu}$, if we want to preserve energy-momentum conservation $\nabla^\mu T_{\mu\nu} = 0$, we must promote $P_{\mu\nu}\Box_\eta^{-1}R^{(1)}$ to a transverse tensor, whose covariant derivative vanishes. To this purpose it is useful to observe that, in a generic Riemannian manifold, any symmetric tensor $S_{\mu\nu}$ can be decomposed as

$$S_{\mu\nu} = S^{\rm T}_{\mu\nu} + \frac{1}{2}(\nabla_\mu S_\nu + \nabla_\nu S_\mu) , \tag{1.3.41}$$

where $\nabla^\mu S^{\rm T}_{\mu\nu} = 0$ [75, 76]. The extraction of the transverse part of a tensor is itself a nonlocal operation. In flat space, where $\nabla_\mu \to \partial_\mu$, proceeding as we have done in the derivation of eqs. (1.3.20)–(1.3.21), one finds that

$$S^{\rm T}_{\mu\nu} = S_{\mu\nu} - \frac{1}{\Box_\eta}(\partial_\mu\partial^\rho S_{\rho\nu} + \partial_\nu\partial^\rho S_{\rho\mu}) + \frac{1}{\Box_\eta^2}\partial_\mu\partial_\nu\partial^\rho\partial^\sigma S_{\rho\sigma} . \tag{1.3.42}$$

Using this expression we can easily check that, in flat space, for a tensor $S_{\mu\nu}$ of the form $S_{\mu\nu}(x) = \eta_{\mu\nu}A(x)$, we have $S^{\rm T}_{\mu\nu} = P_{\mu\nu}A(x)$. Thus, to linear order in an expansion over flat space, the term $P_{\mu\nu}\Box_\eta^{-1}R^{(1)}$ in eq. (1.3.40) is the same as the transverse part of the tensor $(\eta_{\mu\nu}\Box_\eta^{-1}R^{(1)})$, that we denote as $(\eta_{\mu\nu}\Box_\eta^{-1}R^{(1)})^T$, and eq. (1.3.40) is the same as

$$G^{(1)}_{\mu\nu} - \frac{1}{3}\, m^2 (\eta_{\mu\nu}\Box_\eta^{-1}R^{(1)})^{\rm T} = 8\pi G\, T_{\mu\nu} . \tag{1.3.43}$$

In this form, there is a natural covariantization given by

$$G_{\mu\nu} - \frac{1}{3}\, m^2 \left(g_{\mu\nu}\Box^{-1}R\right)^{\rm T} = 8\pi G\, T_{\mu\nu}\,, \tag{1.3.44}$$

where now $\Box^{-1}$ is the inverse of the covariant $\Box$ operator with respect to the generic metric $g_{\mu\nu}$, and the operation of taking the transverse part is the fully covariant operation defined by eq. (1.3.41).

The equation (1.3.44) defines the so-called RT model, where R stands for the occurrence of the Ricci scalar and T for the extraction of the transverse part. This model was first proposed in [77] and, after the study of many alternative possibilities, it turns out to be the only viable one. It is defined at the level of a nonlocal equation of motion rather than by a QEA and there is currently no known nonlocal action from which eq. (1.3.44) can be derived.

A different covariantization emerges naturally if we rather start from the quantum effective action (1.3.37). As usual, $d^4x\,(1/4)h_{\mu\nu}\mathcal{E}^{\mu\nu,\rho\sigma}h_{\rho\sigma}$ becomes $d^4x\sqrt{-g}R$ while, using eq. (1.3.39), $(P^{\mu\nu}h_{\mu\nu})^2$ is the same as $(\Box_\eta^{-1}R^{(1)})^2$, which is naturally covariantized into $(\Box^{-1}R)^2$. Thus, a natural covariantization of eq. (1.3.37) is

$$\begin{aligned}\Gamma_{\rm RR} &= \frac{1}{16\pi G}\int d^4x\sqrt{-g}\left[R - \frac{m^2}{6}(\Box^{-1}R)^2\right] \\ &= \frac{1}{16\pi G}\int d^4x\sqrt{-g}\left[R - \frac{m^2}{6}R\frac{1}{\Box^2}R\right]\,,\end{aligned} \tag{1.3.45}$$

where in the last line we have integrated $1/\Box$ by parts. This gives the model that was first proposed in [78], also known as the RR model, after the two occurrences of the Ricci scalar in the nonlocal term.

Since they can be built as covariantizations of the same nonlocal mass terms over Minkowski, it is clear that the RT and RR model are the sasme at linear order in an expansion over flat space. However, they are in general different, and have different cosmological predictions. The RR model shares most of the phenomenologically attractive properties of the RT model, such as viable cosmological background evolution, stable cosmological perturbations, good fit to Cosmic Microwave Background (CMB), Baryon Acoustic Oscillations (BAO), type Ia Supernovae (SNe) and structure formation data. However, as we will see, it does not pass the constraints from Lunar Laser Ranging; on the contrary, the RT model passes all the observational tests.

1.3.3 Conceptual aspects: localization and causality

A nonlocal QEA can be rewritten in local form by introducing auxiliary fields (see also [79, 80, 81, 82, 83, 84, 85]). This is quite convenient for working out the predictions of the theory (e.g. for studying the equations of motions of the theory, the cosmological perturbations, etc.), but requires some care at the level of interpretation, in order not to confuse the auxiliary fields with actual degrees of freedom of the theory. We follow the discussion in [31] and, as a simple example, we consider the theory of a massive photon discussed in section 1.3.1. We have seen that it can be formulated as a local but non gauge-invariant theory, as in eq. (1.3.1), or as a gauge-invariant theory at the price of nonlocality, as in eq. (1.3.3). One might also get a theory that is at the same time local and gauge-invariant, by introducing an auxiliary anti-symmetric tensor field $U^{\mu\nu}$ defined by $U^{\mu\nu} = \Box^{-1}F^{\mu\nu}$ and implementing this definition into the action (1.3.3) via a Lagrange multiplier. In this way, one gets a local and gauge-invariant action written in terms of the two fields A^μ and $U^{\mu\nu}$. The equations of motion of the theory can then be rewritten as

$$\partial_\mu F^{\mu\nu} = j^\nu + m^2 \partial_\mu U^{\mu\nu} , \qquad \Box U^{\mu\nu} = F^{\mu\nu} . \tag{1.3.46}$$

While the steps leading to eq. (1.3.46) are formally correct, this local and gauge-invariant formulation seems to suggest that the theory has many more degrees of freedom than the Proca theory of a massive photon that was our starting point: we apparently have a massless gauge-invariant vector field A^μ, which carries two degrees of freedom, interacting with an antisymmetric tensor field $U^{\mu\nu}$, which apparently carries six degrees of freedom. This seems very different from the three degrees of freedom of a massive vector field from which we started. Of course, new degrees of freedom cannot pop out from nowhere, and the delicate point here lays in the passage from an equation such as $U^{\mu\nu} = \Box^{-1} F^{\mu\nu}$ to the equation $\Box U^{\mu\nu} = F^{\mu\nu}$, i.e. the inversion of the $\Box$ operator. By itself, the most general solution of an equation such as $\Box U^{\mu\nu} = F^{\mu\nu}$ is given by a solution of the inhomogeneous equation plus the most general solution of the associated homogeneous equation $\Box U^{\mu\nu} = 0$. The latter carries with itself the six degrees of freedom associated to $U^{\mu\nu}$. Clearly, if we want this local and gauge-invariant formulation to be equivalent to the original Proca theory, we cannot accept the most general solution of $\Box U^{\mu\nu} = F^{\mu\nu}$. In other words, the initial condition of the auxiliary field $U^{\mu\nu}$ cannot be taken as independent, but must be fixed in terms of the initial condition of the two transverse and the longitudinal components of A^μ, so that the theory indeed still has three independent degrees of freedom. In this sense, $U^{\mu\nu}$ is just an auxiliary field, and does not carry independent degrees of freedom. In particular, at the quantum level there are no creation/annihilation operators associated to it.

The same situation can also be seen clearly for the Polyakov quantum effective action in $D = 2$, that we discussed in Section 1.2.4. In particular one could localize eq. (1.2.21) by introducing an auxiliary field $U = -\Box^{-1} R$. However, recalling that the conformal mode was $\sigma = -(1/2)\Box^{-1} R$, we recognize that U is not an independent degree of freedom, but rather $U = 2\sigma$. In other words, U is not a generic solution of the equation $\Box U = -R$ with the corresponding infinite number of choices for the homogeneous solution $U_{\rm hom}$ of the equation $\Box U_{\rm hom} = 0$, but rather, only a precise choice of U, i.e. $U = 2\sigma$, gives the correct description. In terms of initial conditions for U, we can then say that they are fixed in terms of the initial conditions of the conformal factor as σ, as $U_{\rm in} = 2\sigma_{\rm in}$ and $\dot{U}_{\rm in} = 2\dot{\sigma}_{\rm in}$. At the quantum level, there are no creation/annihilation operators associated to U because there are no free coefficients coming from $U_{\rm hom}$ that can be promoted to operators.

In the following we will use a similar localization procedure for the RR and RT models. As in the examples above, the auxiliary fields that will be introduced are not new independent degrees of freedom; rather, their initial conditions should be understood as fixed in terms of the initial conditions on the metric, and again there are no creation/annihilation operators associated to them. This excludes issues of ghosts at the quantum level. If one had an explicit derivation of the nonlocal term from a fundamental theory, one would in principle be able to determine explicitly their initial conditions in terms of those on the metric, just like we saw for the Polyakov action. But in practice, lacking such a derivation, the initial conditions on the auxiliary fields of the RR and RT models must be taken as free phenomenological parameters. We will see in Chapter 2 that, in the cosmological context in which we are interested, this ignorance introduces only very limited freedom, both at the level of background evolution and of cosmological perturbations, and the predictive power of the theory is not significantly reduced.

Localization of the RR and RT models

We next show how to write nonlocal gravity in a local form. We write the equations both for the RR model, and for the RT model that will eventually be our main focus, since the comparison between the two models can be instructive, and also the manipulations of the equations of the RR model are somewhat simpler. To write the RR model in a local form we introduce two auxiliary fields U and S, defined by $U = -\Box^{-1}R$ and $S = -\Box^{-1}U$ [78]. This can be implemented at the Lagrangian level by introducing two Lagrange multipliers ξ_1, ξ_2 into eq. (1.3.45),

$$\Gamma_{\rm RR} = \frac{1}{16\pi G}\int d^4x\sqrt{-g}\,\left[R\left(1-\frac{m^2}{6}S\right) - \xi_1(\Box U + R) - \xi_2(\Box S + U)\right] .$$

The variation with respect to $h_{\mu\nu}$ gives $G_{\mu\nu} = (m^2/6)K_{\mu\nu} + 8\pi G T_{\mu\nu}$, where

$$K^{\mu}_{\nu} \equiv 2SG^{\mu}_{\nu} - 2\nabla^{\mu}\partial_{\nu}S + 2\delta^{\mu}_{\nu}\Box_g S + \delta^{\mu}_{\nu}\partial_{\rho}S\partial^{\rho}U - \frac{1}{2}\delta^{\mu}_{\nu}U^2 - \left(\partial^{\mu}S\partial_{\nu}U + \partial_{\nu}S\partial^{\mu}U\right) , \tag{1.3.47}$$

while the variation with respect to the Lagrange multipliers ξ_1, ξ_2 gives $\Box U = -R$ and $\Box S = -U$. Thus, the RR model is formally written as a scalar-tensor theory, with two scalar fields U and S, although, as we have discussed in section 1.3.3, U and S are not independent degrees of freedoms, and their initial conditions are in principle fixed in terms of the initial conditions of the metric. In particular, there are no independent solutions associated to the homogeneous equations $\Box U = 0$ and $\Box S = 0$, and no corresponding quanta at the quantum level.

For the RT model the localization proceeds by defining again $U = -\Box^{-1}R$. We also introduce $S_{\mu\nu} = -Ug_{\mu\nu} = g_{\mu\nu}\Box^{-1}R$ and we extract its transverse part $S^T_{\mu\nu}$ by using eq. (1.3.41). Thus, eq. (1.3.44) is localized in terms of an auxiliary scalar field U and the auxiliary four-vector field S_{μ} that enters through eq. (1.3.41). The equations of motion then read [77, 86]

$$\begin{aligned} G_{\mu\nu} + \frac{m^2}{6}\left(2Ug_{\mu\nu} + \nabla_{\mu}S_{\nu} + \nabla_{\nu}S_{\mu}\right) &= 8\pi G\, T_{\mu\nu} , & (1.3.48)\\ \Box U &= -R , & (1.3.49)\\ (\delta^{\mu}_{\nu}\Box + \nabla^{\mu}\nabla_{\nu})S_{\mu} &= -2\partial_{\nu}U , & (1.3.50) \end{aligned}$$

where eq. (1.3.50) is obtained by taking the divergence of eq. (1.3.41) with $S_{\mu\nu} = -Ug_{\mu\nu}$. The equations of motion of the RT model have a suggestive property in connection with the cosmological constant problem. Let us perform a shift $U(x) \to U(x) + u_0$, with u_0 a constant. Equations (1.3.49) and (1.3.50) are invariant while eq. (1.3.48) becomes

$$G_{\mu\nu} + \frac{m^2}{6}\left(2Ug_{\mu\nu} + \nabla_{\mu}S_{\nu} + \nabla_{\nu}S_{\mu}\right) = 8\pi G\,\left(T_{\mu\nu} - \lambda g_{\mu\nu}\right) . \tag{1.3.51}$$

where $\lambda = m^2 u_0/(24\pi G)$. Thus, u_0 (or, equivalently, the initial condition on U) generates a cosmological constant, and one could choose u_0 to cancel any vacuum energy term in $T_{\mu\nu}$.

It is also instructive to consider the equations of motion of the RR and RT models linearized over flat space, eq. (1.3.38), that were our starting point, and write them in terms of the auxiliary fields and of the metric variables of the $3+1$ decomposition (1.3.27). Since, by construction, the RR and RT model coincide when linearized over flat space, we use the RR model, whose localization is slightly simpler, since it involves two scalar fields U and S, rather than U and S_μ for the RT model. One then finds that eqs. (1.3.29)–(1.3.30) are modified into

$$\begin{aligned}
\nabla^2\left[\Phi-(m^2/6)S\right] &= -4\pi G\rho\,, & \Phi-\Psi-(m^2/3)S &= -8\pi G\Sigma\,, \qquad (1.3.52)\\
\nabla^2\Xi_i &= -16\pi G S_i\,, & \Box H_{ij}^{\rm TT} &= -16\pi G\Sigma_{ij}\,, \qquad (1.3.53)\\
(\Box+m^2)U &= -8\pi G(\rho-3P)\,, & \Box S &= -U\,, \qquad (1.3.54)
\end{aligned}$$

Notice that eq. (1.3.54) is needed to close the system, since S appears in eq. (1.3.52). Equations (1.3.52) and (1.3.53) shows that the original metric perturbations Φ, Ψ and Ξ_i remain non-radiative variables that satisfy Poisson equations, just as in GR. The auxiliary fields U and S satisfy Klein-Gordon equations, but, as we already said, their initial conditions are fixed in terms of the initial conditions on the metric, and therefore are not free radiative degrees of freedom either. From these equations it is also clear that the conformal mode s remains a non-propagating degree of freedom also in the RT or RR models. Indeed, combining the two equations in (1.3.52) we get

$$\begin{aligned}
\nabla^2(\Phi+\Psi) &= 2\nabla^2\left(\Phi-\frac{m^2}{6}S\right)+8\pi G\nabla^2\Sigma\\
&= -8\pi G(\rho-\nabla^2\Sigma)\,. \qquad (1.3.55)
\end{aligned}$$

Then, from eq. (1.3.31) we get (again at the linearized level over flat space)

$$\begin{aligned}
s &= 6\Phi-2\Box^{-1}\nabla^2(\Phi+\Psi)\\
&= 6\Phi+16\pi G\Box^{-1}(\rho-\nabla^2\Sigma)\,. \qquad (1.3.56)
\end{aligned}$$

We see that the nonlocal term in s is fully determined by the energy-momentum tensor, in particular by the density ρ and by the anisotropic stress Σ that enters in T_{ij} through eq. (1.3.28). Thus, s remains a non-radiative degree of freedom, exactly as in GR, and vanishes if $\rho=0$ and $\Sigma=0$.

Causality

We next discuss why nonlocal terms would induce problems with causality if added at the level of a fundamental action, while they do not in a quantum effective action, following [31].

To illustrate the problem with causality of a nonlocal fundamental action, consider for instance an action with a nonlocal term proportional to $(1/2)\int d^4x\,\varphi\Box^{-1}\varphi$ where φ is a scalar field [71]. To complete the definition of this term we must specify the Green's function $G(x,x')$ used to define $\Box^{-1}$, and then

$$\frac{1}{2}\int d^4x\,\varphi(x)(\Box^{-1}\varphi)(x)\equiv\frac{1}{2}\int d^4xd^4x'\,\varphi(x)G(x,x')\varphi(x')\,. \tag{1.3.57}$$

Consider now the contribution of this term to the equation of motion. Taking the variation with respect to φ, we get

$$\frac{1}{2}\frac{\delta}{\delta\varphi(x)}\int dx'dx''\varphi(x')G(x';x'')\varphi(x'')=\frac{1}{2}\int dx'[G(x;x')+G(x';x)]\varphi(x')\equiv\Box^{-1}_{\rm sym}\varphi\,, \tag{1.3.58}$$

where $\Box^{-1}_{\rm sym}$ is the inverse d'Alembertian with respect to the symmetrized Green's function $[G(x;x')+G(x';x)]/2$. Thus, independently of the choice of $G(x,x')$, in the equations of motion we end up with a symmetric Green's function. Since the retarded Green's function is not symmetric, it cannot be obtained from such a variation. The equations of motion obtained from a nonlocal classical action are therefore in general acausal. This is one of the reasons why a fundamental action must be local.

We have already seen in Section 1.2.2 that causality is recovered when using the Schwinger-Keldysh path integral in the definition of the QEA. At the level of the equations of motion for the evolution of vacuum expectation values, the result of the computation with the Schwinger-Keldysh path integral amounts to perform a formal variation of the quantum effective action, without specifying the Green's function used to define $\Box^{-1}$, and then replacing the resulting occurrences of $\Box^{-1}$ in the equations of motion with the $\Box^{-1}$ operator defined using the retarded Green's function (see [51] and Section 12.1.6 of [53]).

1.3.4 Origin of an IR mass scale

It is interesting to discuss possible mechanisms for the generation of nonlocal terms relevant for the IR dynamics and associated to a mass scale. As a starting point, we recall that perturbative loop corrections due to massive matter fields cannot be responsible for producing cosmologically relevant nonlocal structures, as shown in [87]. Let us review how the explanation goes.

Perturbative loop corrections do not affect the IR regime

In gravity the one-loop corrections induced by matter fields can produce nonlocal form factors in the quantum effective action, associated to terms quadratic in the curvature [50, 51, 88, 89, 90, 91] (see also [49, 52, 54] for reviews). The resulting QEA can be expressed using form factors as

$$\Gamma_{\rm one-loop}=\int d^4x\sqrt{-g}\left[\frac{1}{16\pi G}R-R\,k_R(\Box)R-C_{\mu\nu\rho\sigma}k_W(\Box)C^{\mu\nu\rho\sigma}+{\rm GB}\right], \quad (1.3.59)$$

where "GB" denotes a similar nonlocal term that reduces to the topological Gauss-Bonnet term when its form factor is set to one. Consider the contribution to the form factor from a particle of mass M. When the particle is very massive compared to the energies involved (whose scale is set by the Hubble parameter $H(t)$ in a cosmological setting), then the particle decouples and its contribution to the form factor is local and suppressed by a factor $\mathcal{O}(\Box/M^2)\ll 1$. In this case only local terms are produced and the situation, just like we have seen in Section 1.2.3 about QED when $\Box/m_e^2\ll 1$. To get nonlocal operators we should look at the case where the particle of mass M is light with respect to the energy scale involved. In that situation, the form factor k_R has the form [90, 91, 92, 93]

$$k_R\left(\frac{-\Box}{M^2}\right)=\alpha\log\left(\frac{-\Box}{M^2}\right)+\beta\left(\frac{M^2}{-\Box}\right)+\gamma\left(\frac{M^2}{-\Box}\right)\log\left(\frac{-\Box}{M^2}\right)+\delta\left(\frac{M^2}{-\Box}\right)^2+\dots, \quad (1.3.60)$$

and similarly for k_W. In [94] it was observed that the logarithmic terms and the term $(M^2/\Box)$ have little effect on the cosmological evolution in the present epoch, so one might hope that the leading term is actually given by the term proportional to

$M^4/\Box^2$, which is the same operator appearing in eq. (1.3.45) for the RR model. Comparing eq. (1.3.60) with eq. (1.3.45) we see that they match only if $m = \mathcal{O}(M^2/M_P)$, where M_P is the reduced Planck mass. The expansion in eq. (1.3.60) can only be valid today if $M \ll H_0$, and therefore we see that loop corrections from light particles could only generate a nonlocal term $m^2 R \Box^{-2} R$ with $m = \mathcal{O}(M^2/M_P) \ll H_0(H_0/M_P)$. However, we will see in Chapter 2 that m has to be of order H_0 to reproduce the dark energy content today and certainly cannot be suppressed with respect to H_0 by the very small factor H_0/M_P. Hence, loop corrections from light particles cannot provide the required value of m while, on the other hand, we have already seen that heavy particles give only local contributions to the form factors, suppressed by $\mathcal{O}(H_0^2/M^2) \ll 1$.

Dynamical mass generation

In this subsection we will discuss indications, from various non-perturbative techniques, in favor of the possibility of a dynamical mass generation in the IR limit of four-dimensional quantum gravity, in particular in relation to the conformal mode. We follow the presentation in [31].

Functional renormalization group equations. Renormalization group (RG) equations, such as the Polchinski equation [95] and the Wetterich equation [96] are exact equations transform the problem of performing a functional integration for computing the path integral of the QEA into a functional differential equation. In practice, just as the evaluation of the functional integral for an interacting theory requires approximations or numerical methods, one is usually required to truncate the space of action functionals allowed in QEA, to have some hope to solve the functional RG equation. This is the same RG idea used in statistical physics for studying critical phenomena, when one restricts the attention to some couplings among the infinite number of them that is in principle allowed by the symmetries of the microscopic interaction. Therefore the results obtained from functional RG equations can depend on the truncation adopted and require some care in their interpretation. For gravity, functional RG techniques have been developed particularly in connection with the asymptotic safety program, i.e. the search for a non-trivial UV fixed point (see [97] for review). Only more recently, these tools are being applied to the study of the IR behavior of gravity, and a number of functional RG studies have found indications of strong quantum gravity effect in the IR [98, 99, 100, 101]. Dynamical mass generation could be compatible with the findings in [102], where, in a truncation of the theory including only the Einstein-Hilbert term and the cosmological constant, it was found that, evolving the RG flow toward the IR, for some trajectories the running of Newton's constant hits a singularity at a finite momentum scale. But, as we said, the the singularity could also be an artifact of the truncation. Another interesting result in [98] suggests that the cosmological constant could be screened by strong IR effects due to the quantum fluctuations of the trasverse-traceless modes, while the conformal mode fluctuations could generate a new mass scale. This is compatible with the phenomelogical RT and RR models, as they exhibit a mass term for the conformal mode.

Lattice gravity. A possible non-perturbative tool is provided by lattice gravity, based either on a simplicial decomposition of the space-time manifold in Euclidean space (see [103] for review), or on causal dynamical triangulations (CDT) (see [104, 105] for reviews).

Using numerical simulation of CDT it is possible to have some informations on the two-point function of the conformal mode, and ref. [106] showed that the numerical results provide evidence for a massive conformal mode, corresponding to a linearized nonlocal quantum effective action of the form in eq. (1.3.33), whose covariantizations can give the RR or RT models. This numerical result was not studied in the continuum limit, but it is still an indication that quantum effects could generate a mass for the conformal mode of the metric.

Naturalness of the mass scale in nonlocal gravity models.

Let us rewrite the quantum effective action (1.3.45) of the RR model as

$$\begin{aligned}\Gamma_{\rm RR} &= \frac{M_P^2}{2}\int d^4x\sqrt{-g}\left[R-\frac{1}{6}m^2R\frac{1}{\Box^2}R\right]\\ &= \int d^4x\sqrt{-g}\left[\frac{M_P^2}{2}R-R\frac{\Lambda_{\rm RR}^4}{\Box^2}R\right],\end{aligned} \tag{1.3.61}$$

where $\Lambda_{\rm RR}=(1/12)m^2M_P^2$. In this form, it is clear that $\Lambda_{\rm RR}$ should be taken as the fundamental scale generated dynamically, corresponding to a dimensionless form factor $k_R(\Box)=\Lambda_{\rm RR}^4/\Box^2$ in $Rk_R(\Box)R$, while the parameter m is just a derived quantity introduced for convenience. $\Lambda_{\rm RR}$ is a scale generated generated dynamically and its value cannot be predicted, similarly to what happens for QCD where the value of $\Lambda_{\rm QCD}$ can only be obtained by comparison with observations. As we already anticipated, we need $m=\mathcal{O}(H_0)$ in order to have a viable dark energy content in the present epoch. Therefore,

$$\Lambda_{\rm RR}=\mathcal{O}(H_0m_{\rm Pl})^{1/2}=\mathcal{O}(\text{meV})\,. \tag{1.3.62}$$

A similar analysis can be carried out for the RT model to find a scale $\Lambda_{\rm RT}$ of the same order as in the RR case. The appearance of a meV scale is not particularly surprising for QFT and is certainly much more natural than what would be needed to explain dark enrgy by introducing some particle of mass m. Indeed, in the latter case, m is the fundamental scale with a very small value $m\sim H_0\sim 10^{-33}$ eV.

Chapter 2

Nonlocal gravity: cosmological implications

2.1 Cosmology of the RT model

2.1.1 FRW background

Let us study the existence and properties of a Friedman-Robertson-Walker (FRW) background for the RT model, following the presentation given in [31]. We consider a spatially flat background, $ds^2 = -dt^2 + a^2(t)d\mathbf{x}^2$. For symmetry reasons the spatial component S_i of the auxiliary field S_μ must vanish, since there is no preferred spatial direction, and the only variables are $U(t)$ and $S_0(t)$, together with the FRW scale factor $a(t)$. Eqs. (1.3.48)–(1.3.50) then become [77, 31]

$$
\begin{aligned}
H^2 - \frac{m^2}{9}(U - \dot{S}_0) &= \frac{8\pi G}{3}\rho && (2.1.1)\\
\ddot{U} + 3H\dot{U} &= 6\dot{H} + 12H^2\,, && (2.1.2)\\
\ddot{S}_0 + 3H\dot{S}_0 - 3H^2 S_0 &= \dot{U}\,, && (2.1.3)
\end{aligned}
$$

where we have written $T^\mu_\nu = \text{diag}(-\rho, p, p, p)$, and the dot denotes the derivative with respect to cosmic time t. It is convenient to define $Y = U - \dot{S}_0$, $h = H/H_0$, and $\Omega_i(t) = \rho_i(t)/\rho_c(t)$, where $i = \text{M, R, DE}$ labels radiation, matter and dark energy, respectively, and $\rho_c(t) = 3H^2(t)/(8\pi G)$ is the critical density for closing the Universe at time t. We will also use the standard notation $\Omega_M \equiv \Omega_M(t_0)$, $\Omega_R \equiv \Omega_R(t_0)$ and $\Omega_{\text{DE}} \equiv \Omega_{\text{DE}}(t_0)$ (where t_0 is the present value of cosmic time) for the present values of $\Omega_i(t)$. We henceforth use the dimensionless variables

$$x \equiv \ln a(t) \qquad (2.1.4)$$

instead of cosmic time t, and we denote $df/dx = f'$. Then the Friedmann equation (2.1.1) reads

$$h^2(x) = \Omega_M e^{-3x} + \Omega_R e^{-4x} + \gamma Y(x)\,, \qquad (2.1.5)$$

where

$$\gamma \equiv m^2/(9H_0^2)\,. \qquad (2.1.6)$$

This shows that there is an effective DE density

$$\rho_{\text{DE}}(t) = \rho_0 \gamma Y(x)\,, \qquad (2.1.7)$$

where $\rho_0 = 3H_0^2/(8\pi G)$. Using $U(x)$ and $Y(x)$, eqs. (2.1.2) and (2.1.3) take the form

$$Y'' + (3-\zeta)Y' - 3(1+\zeta)Y = 3U' - 3(1+\zeta)U\,, \tag{2.1.8}$$

$$U'' + (3+\zeta)U' = 6(2+\zeta)\,, \tag{2.1.9}$$

where, using eq. (2.1.5),

$$\zeta(x) \equiv \frac{h'}{h} = -\frac{3\Omega_M e^{-3x} + 4\Omega_R e^{-4x} - \gamma Y'}{2(\Omega_M e^{-3x} + \Omega_R e^{-4x} + \gamma Y)}\,. \tag{2.1.10}$$

Initial conditions

As a next step, we discuss the initial conditions on the auxiliary fields, following refs. [29, 107, 31]). To get a first analytic understanding we observe that, in any given cosmological epoch, such as radiation dominance (RD), matter dominance (MD), or an earlier inflationary de Sitter (dS) phase, $\zeta(x)$ has an approximately constant value ζ_0, with $\zeta_0 = 0$ in dS, $\zeta_0 = -2$ in RD and $\zeta_0 = -3/2$ in MD. In the approximation of constant ζ eq. (2.1.9) can be integrated analytically, and has the solution [77]

$$U(x) = \frac{6(2+\zeta_0)}{3+\zeta_0}x + u_0 + u_1 e^{-(3+\zeta_0)x}\,. \tag{2.1.11}$$

The first term on the right-hand side is a particular solution of the inhomogeneous equation, while u_0 and u_1 parametrize the most general solution of the homogeneous equation $\Box U = U'' + (3+\zeta_0)U = 0$. The initial conditions on U, i.e. $U(x_{\rm in})$ and $U'(x_{\rm in})$, are in one-to-one correspondence with the choice of the solutions of homogeneous equation, i.e. with u_0 and u_1. The constant u_0 corresponds to the reintroduction of a cosmological constant, as we have seen in eq. (1.3.51). Our aim is to see if we can obtain a self-accelerated evolution from the nonlocal term, without introducing by hand a cosmological constant, and we will therefore set $u_0 = 0$. A non-vanishing u_0 could always be reintroduced later, and, not surprisingly, produces an evolution that is intermediate between that of the RT model with $u_0 = 0$ and that ΛCDM, see Section 7.4 of [29]. The other solution of the homogeneous equation, proportional to $e^{-(3+\zeta_0)x}$, is instead a decaying mode, in all cosmological phases. Thus, the solution with initial conditions $U(x_{\rm in}) = U'(x_{\rm in}) = 0$ has a marginally stable direction, corresponding to the possibility of reintroducing a cosmological constant, and a stable direction, i.e. is an attractor in the u_1 direction. Consider next eq. (2.1.8). Using eq. (2.1.11) and solving for $Y(x)$ we get

$$\begin{aligned} Y(x) \;=\; & -\frac{2(2+\zeta_0)\zeta_0}{(3+\zeta_0)(1+\zeta_0)} + \frac{6(2+\zeta_0)}{3+\zeta_0}x + u_0 - \frac{6(2+\zeta_0)u_1}{2\zeta_0^2+3\zeta_0-3}e^{-(3+\zeta_0)x} \\ & + a_1 e^{\alpha_+ x} + a_2 e^{\alpha_- x}\,, \end{aligned} \tag{2.1.12}$$

where $\alpha_\pm = (1/2)[-3+\zeta_0 \pm \sqrt{21+6\zeta_0+\zeta_0^2}]$. In both RD and MD we have $\alpha_+ < 0$ and $\alpha_- < 0$, so both modes are decaying. This means that, if we start the evolution deep in the RD phase, with $u_0 = 0$ in order not to have a cosmological constant, and $u_1 \sim a_1 \sim a_2 \sim \mathcal{O}(1)$, the solution will quickly approach the one obtained with initial conditions $U(x_{\rm in}) = U'(x_{\rm in}) = Y(x_{\rm in}) = Y'(x_{\rm in}) = 0$. We will refer to this solution as the 'minimal' RT model.

The situation becomes more interesting if we start the evolution during a primordial phase of de Sitter-like inflation, before RD. In dS there is a growing mode with $\alpha_+ = (-3+\sqrt{21})/2 \simeq 0.79$. Then Y will grow during dS (exponentially in

x, so as a power of the scale factor), and will then decrease again during RD and MD. In general, a growing mode during MD or the late RD phase would be fatal to the viability of the model, because any perturbation of the initial conditions would result in an activation of the unstable mode, and would bring the solution very far from a FRW solution driven by $T_{\mu\nu}$, as in standard cosmology (this is indeed a criterium that ruled out several other nonlocal models). For the evolution during an early dS phase the situation is, however, different [107]. Indeed, let us denote by $x_{\rm in}$ the value of $x = \ln a$ at a time, during inflation, when we set initial conditions $u_1 \sim a_1 \sim a_2 \sim \mathcal{O}(1)$, and by $x_{\rm end}$ the value when inflation ends and RD begins (we neglect for simplicity an intermediate reheating phase). We use the notation

$$x_{\rm end} - x_{\rm in} = \log\left(a_{\rm end}/a_{\rm in}\right) \equiv \Delta N\,, \tag{2.1.13}$$

so ΔN is the number of e-folds from the time where we set initial condition of order one, to the end of a de Sitter phase of inflation. Thus, if $Y(x_{\rm in})$ has a generic value of order one (i.e., is not fine-tuned to zero), by the end of inflation

$$Y(x_{\rm end}) \simeq \exp\{\alpha_+^{\rm dS}\Delta N\} \simeq \exp\{0.79\Delta N\}\,. \tag{2.1.14}$$

The evolution of U can be computed similarly, using eq. (2.1.11). During a quasi-de Sitter phase of inflation, starting from a value of order one, we get

$$U(x_{\rm end}) \simeq 4\Delta N\,. \tag{2.1.15}$$

The important point is that, despite the exponential growth in eq. (2.1.14), even for very large values of ΔN the corresponding DE density $\rho_{\rm DE}(x) = \rho_0\gamma Y(x)$ has no effect on the inflationary dynamics. This is due to the fact that $\rho_0 = 3H_0^2/(8\pi G) \sim (10^{-3}{\rm eV})^4$ is extremely small compared to the energy density during inflation. For instance, if $Y(x_{\rm in}) = \mathcal{O}(1)$ and we take $\Delta N = 60$, at the end of inflation we get $Y(x_{\rm end}) = \mathcal{O}(10^{20})$. Even with such a large value of Y, we have

$$[\rho_0 Y(x_{\rm end})]^{1/4} \sim 10^{-3}{\rm eV} \times Y^{1/4}(x_{\rm end}) \sim 10^2\,{\rm eV}\,. \tag{2.1.16}$$

This is totally negligible compared to the inflationary scale M, that has typical values, say, of order 10^{13} GeV. Thus, during the inflationary phase the evolution of the scale factor is the same as in standard GR without the nonlocal term. So, the important conclusion is that, at the level of background evolution, there is no evident pathology associated with the exponential growth of $Y(x)$. Rather, one will have to study in detail the evolution through dS, RD and MD to see if it gives a viable and interesting background cosmology. As we will recall below, following [29, 107], indeed the corresponding background evolution is viable, and also quite interesting. As discussed in [107], even at the level of cosmological perturbations this growth during de Sitter is innocuous, again because of the smallness of the scale associated to the nonlocal term with respect to the inflationary scale.

Equations (2.1.14) and (2.1.15) give the values of $Y(x)$ and $U(x)$ when they enter the subsequent RD phase (apart from some minor modification due to reheating). As we will see explicitly in Section 2.1.1, even if in the RD and MD phases the solution obtained with vanishing initial conditions is an attractor, the fact that $Y(x)$ enters the RD phase with an exponentially large value gives an evolution that is sensibly different from that of the minimal model, simply because there is not enough time to relax to zero this exponentially large value by the end of the MD phase and the beginning of the current DE-dominated phase, when (having chosen m of order H_0) the energy scale associated to Y eventually becomes comparable to the total energy density. Thus, there is a residual dependence of the dark energy evolution near the present cosmological epoch, from the value ΔN that determines, through eqs. (2.1.14) and (2.1.15), the values of the auxiliary field when the enter the RD phase.

The conclusion is that, at the level of the background cosmological evolution, our ignorance on the initial conditions of the auxiliary fields can be reabsorbed into a single parameter ΔN, that gives the number of e-folds from the moment when these fields have initial conditions $\mathcal{O}(1)$ during inflation, until the end of inflation (plus the parameter u_0, that corresponds to reintroducing a cosmological constant, and that we will set to zero).

As discussed in [30], no further freedom emerges at the level of cosmological perturbations. Indeed, at the perturbation level we must consider all Fourier modes of the perturbations, so in principle we should assign the initial conditions on $\delta U_{\mathbf{k}}(x)$, $\delta Y_{\mathbf{k}}(x)$ and on their first time derivatives, at an initial time x_{in}. The fact that the auxiliary fields do not represent arbitrary degrees of freedom but are fixed in terms of the metric means that the initial conditions for the perturbations of the auxiliary fields will be of order of the metric perturbations. One can therefore ask what happens if we start with initial conditions of this order of magnitude. The explicit numerical study in [30] shows that the effect of such a change in the initial conditions of the perturbations is completely negligible.

We want to study the predictions of the RT model for a few values of ΔN. For this purpose, it is useful to recall that, for inflation taking place at a scale $M_{\text{infl}} = (\rho_{\text{infl}})^{1/4}$, assuming instantaneous reheating, the minimum number of e-folds required to solve the flatness and horizon problems is (see e.g. Section 21.1 of [74])

$$(\Delta N)_{\text{min}} \simeq 64 - \log \frac{10^{16}\,\text{GeV}}{M_{\text{infl}}} . \tag{2.1.17}$$

In the following, beside the 'minimal' model defined by initial conditions of order one during RD, which is equivalent to setting $\Delta N = 0$ (or, equivalently, ΔN of order one), we will study also the cases $\Delta N = 34, 50, 64$ that, according to eq. (2.1.17), approximately correspond to the minimum value of ΔN for $M_{\text{infl}} = \{10^3, 10^{10}, 10^{16}\}$ GeV, respectively. These values are chosen because $M_{\text{infl}} = 10^3$ GeV corresponds to inflation at the electroweak scale, which is on the lower range of possible inflationary scales, while $M_{\text{infl}} = 10^{16}$ GeV is the highest value consistent with the non-detection of tensor perturbations in the CMB anisotropies, and $M_{\text{infl}} = 10^{10}$ GeV is an intermediate value which is quite often considered as a typical inflationary scale.

Of course, the number of e-folds during inflation at a given scale does not need to be the minimum required to solve the flatness and horizon problems and, for a given value of M_{infl}, we could chose a higher value of ΔN. We have therefore studied also how the results change increasing ΔN for a fixed value of M_{infl}. As pointed out in [107], increasing ΔN the results eventually saturate to a limiting curve (as a function of redshift). In particular, setting $M_{\text{infl}} = 10^{16}$ GeV, we find that this limiting curve is reached, within sub-percent level accuracy, already for $\Delta N \simeq 70$. In the following, beside the cases $(M_{\text{infl}} = 10^3\,\text{GeV}, \Delta N = 34)$, $(M_{\text{infl}} = 10^{10}\,\text{GeV}, \Delta N = 50)$ and $(M_{\text{infl}} = 10^{16}\,\text{GeV}, \Delta N = 64)$, we will also show the results for $(M_{\text{infl}} = 10^{16}\,\text{GeV}, \Delta N = 100)$, that represents the limiting curve for the various background quantities as a function of redshift. For brevity, we will refer to these cases as the RT model with $\Delta N = 34, 50, 64$ and 100, respectively. We have checked that the same limiting curve is obtained starting from a different value of M_{infl} and raising again sufficiently ΔN. This behavior is due to a scaling property of the equations when Y

starts from a very large value at the beginning of RD [107].

A quite interesting aspect of the cosmological evolution of the RT model with initial conditions set during inflation, that will emerge clearly from the discussion below, is that the behavior of dark energy at the present epoch depends on the existence and duration (as quantified by ΔN) of a phase of primordial inflation, providing an unexpected connection between early- and late-time cosmology.

Results

Given the initial conditions and a choice of values for the cosmological parameters Ω_M and h_0 (defined as usual from $H_0 = 100 h_0 \,\mathrm{km\,s^{-1} Mpc^{-1}}$), the numerical integration of the equations for the background evolution, eqs. (2.1.5)–(2.1.9), is straightforward. In practice, in the numerical implementation of our integration routine, we consider that the transition between inflation and RD takes place when, extrapolating backward in time the present energy density in radiation $\rho_{R,0}$, the energy density in radiation $\rho_{R,0}/a^4$ becomes equal to $M_{\rm infl}^4$, i.e. when the scale factor has the value $a_\star$ given by $a_\star = \rho_{R,0}^{1/4}/M_{\rm infl}$ (we are working in the approximation of instantaneous reheating). Using $\rho_{R,0}^{1/4} \simeq 2.41 \times 10^{-4}$ eV, the corresponding value of $x = \log a$ is $x_\star \simeq -65.9 + \log(10^{16}\,\mathrm{GeV}/M_{\rm infl})$. Assuming that initial conditions of order one have been set ΔN e-folds earlier, at the inflation-RD transition we take $Y = \exp\{0.79\Delta N\}$ and $U = 4\Delta N$. The numerical integration through the full RD phase would be numerically difficult, and not necessary, since we know that, until we are deep in RD, the solution for Y evolves according to the slowest-decaying mode, which decays as $\exp\{-0.70x\}$ and the solution for U stays constant. Thus, at a value x_0 still deep into RD (we take $x_0 = -15$; RD-MD equilibrium is at $x \simeq -8.1$) we have $U(x_0) = 4\Delta N$, $U'(x_0) = 0$, $Y(x_0) = \exp\{0.79\Delta N - 0.70(x_0 - x_\star)\}$ and $Y'(x_0) = -0.70 Y(x_0)$. At this point we start the numerical evolution with these initial conditions. To produce Fig. 2.1, for the minimal model and for $\Delta N = 34, 50, 64$ we have used the respective mean values for Ω_M and h_0 from Table 2.2, obtained from our MCMC chains. For the limiting curves $\Delta N = 100$ we have not rerun our MCMC and we have used the same values as for $\Delta N = 64$, which is an excellent approximation since we see from Table 2.2 that, for large ΔN, the variation in the parameters are very small (and would give effects totally unappreciable on the scale of the figures). A final detail is that, in ΛCDM, assuming flatness and fixing Ω_M and Ω_R, directly fixes Ω_Λ from $\Omega_M + \Omega_R + \Omega_\Lambda = 1$, and one can immediately integrate the evolution equations. In contrast, in the nonlocal model, once fixed Ω_M and Ω_R (and assuming flatness), the remaining parameter in the equations is γ, which is fixed by trials and errors until the value of the dark energy energy fraction today, $\Omega_{\rm DE}$, obtained from the solution of the equations, satisfies the condition $\Omega_M + \Omega_R + \Omega_{\rm DE} = 1$, i.e. $\Omega_{\rm DE} \simeq 0.7$. The corresponding values of γ turn out to be $\gamma \simeq 5.13555 \times 10^{-2}$ for the minimal model, and $\gamma \simeq \{2.69512 \times 10^{-3}, 1.0321 \times 10^{-3}, 3.73915 \times 10^{-4}, 1.94944 \times 10^{-11}\}$ for $\Delta N = 34, 50, 64, 100$, respectively. For the mass m this means $m/H_0 \simeq 0.68$ for the minimal model, and $m/H_0 \simeq \{0.16, 0.10, 0.06, 4.2 \times 10^{-8}\}$ for $\Delta N = 34, 50, 64, 100$.

The upper left panel of Fig. 2.1 shows the evolution of the dark energy density $\rho_{\rm DE}(x)$, normalized to the total energy density $\rho_{\rm tot}(x) = \rho_M(x) + \rho_R(x) + \rho_{\rm DE}(x)$, as a function of x [recall that here $x = \ln a$, and we normalize the scale factor so that $a(t_0) = 1$]. For orientation, matter-radiation equilibrium is at $x \simeq -8.1$, at the present epoch $x = 0$, and $x > 0$ corresponds to the cosmological future. We see from the plot that the DE density due to the nonlocal term is negligible until the relatively recent cosmological epoch, when eventually dominates.

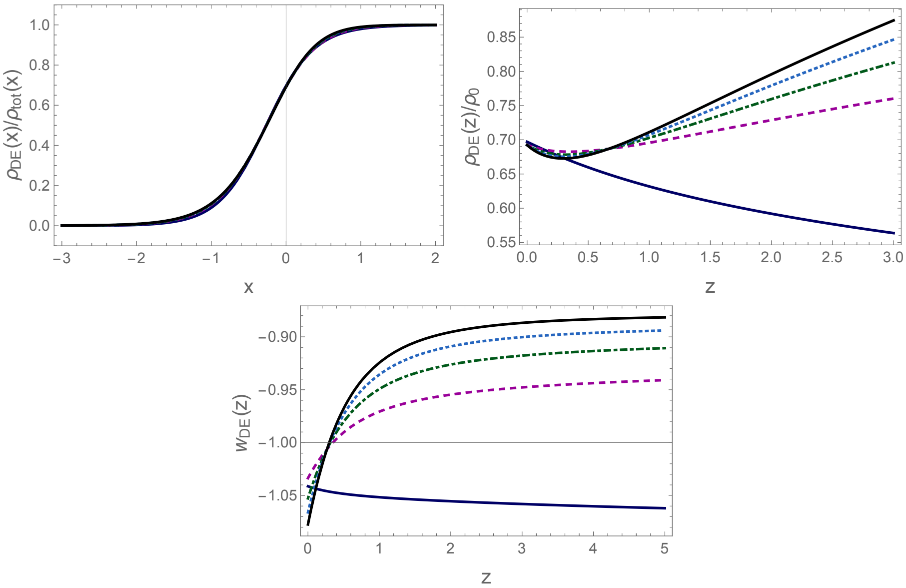

FIGURE 2.1: Upper left: $\rho_{\rm DE}(x)$ normalized to the total energy density $\rho_{\rm tot}(x)$ as a function of x. Upper right: $\rho_{\rm DE}(z)$ normalized to the critical energy density today, ρ_0, as a function of redshift z. Lower panel: the DE equation of state $w_{\rm DE}(z)$ as a function of redshift. The curves correspond to the minimal RT model (blue solid line) and the RT model with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed), $\Delta N = 64$ (cyan, dotted) and $\Delta N = 100$ (black solid line). Figure from [31].

When $\rho_{\rm DE}(x)$ is normalized to $\rho_{\rm tot}(x)$, which includes the contribution of $\rho_{\rm DE}(x)$ itself, the result for the minimal model and for the RT models with large ΔN look all very similar, and the various curves are basically indistinguishable. However, the individual behaviors of $\rho_{\rm DE}(x)$ are quite different. This is shown in the upper right panel of Fig. 2.1, where $\rho_{\rm DE}$ is shown as a function of the redshift z [related to x by $x = -\log(1+z)$], and normalized to the constant critical energy density today ρ_0. We see that, as we approach the present epoch from large z, in the minimal model $\rho_{\rm DE}$ increases, until it reaches the present value $\rho_{\rm DE}/\rho_0 \simeq 0.7$, which is fixed by our choice of $\Omega_M \simeq 0.3$. In contrast, for large ΔN, $\rho_{\rm DE}$ starts from a very large value deep in RD (a consequence of the large value of the auxiliary field Y at the end of inflation), and then decreases for most of its evolution, until the present epoch. This behavior can be understood observing that, for $\Delta N = 0$, the evolution of Y is determined by the particular solution of the inhomogeneous equation (2.1.8), which stays close to zero during RD and then starts to increases with time during MD, until we enter in a regime dominated by DE; in contrast, for large ΔN the solution starts from a very large initial value at the beginning of RD and then decays according to the decaying modes of the associated homogeneous equation, until, close to the recent epoch, the decaying modes have become smaller than the solution of the inhomogeneous equation, that takes over, so the solution for Y starts to rise again.

As mentioned before, for sufficiently large ΔN, the results saturate toward a limiting curve, independent of the chosen value of $M_{\rm infl}$. As explained in [107], this is due to the fact that, for sufficiently large ΔN, an increase in the initial values of Y at the beginning of RD is exactly compensated by a decrease in γ, and we end up on the same solution. This limiting curve is shown as the black solid line in Fig. 2.1, obtained for definiteness setting $(M_{\rm infl} = 10^{16}\,{\rm GeV}, \Delta N = 100)$. For instance, in

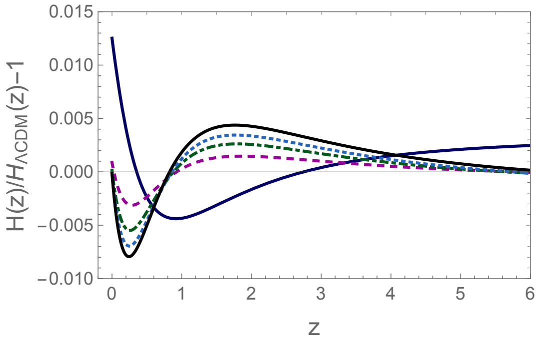

FIGURE 2.2: Relative difference of Hubble rate with respect to ΛCDM for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed), $\Delta N = 64$ (cyan, dotted) and $\Delta N = 100$ (black solid line). Figure from [31].

	RT, minimal	$\Delta N = 34$	$\Delta N = 50$	$\Delta N = 64$	$\Delta N = 100$
w_0	-1.041	-1.034	-1.053	-1.066	-1.077
w_a	-0.023	$+0.127$	$+0.218$	$+0.283$	$+0.335$

TABLE 2.1: Values of w_0 and w_a for the RT model, minimal and with various values of ΔN. Table from [31].

this and in all similar plots below, on the scale of the figure all the curves with $M_{\rm infl} = 10^{16}$ GeV and $\Delta N \gtrsim 70$ are indistinguishable, and fall on this asymptotic curve.

The lower panel in Fig. 2.1 shows the DE equation of state, defined as usual from the conservation equation

$$\dot{\rho}_{\rm DE} + 3H(1 + w_{\rm DE})\rho_{\rm DE} = 0\,. \tag{2.1.18}$$

The different evolutions of $\rho_{\rm DE}$ for the minimal model and for large ΔN result in different, and quite distinctive behaviors of $w_{\rm DE}$ as a function of redshift. For the minimal model $w_{\rm DE}(z)$ is always on the 'phantom' side, $w_{\rm DE}(z) < -1$, while, for large ΔN, the evolution exhibits 'phantom crossing' at $z \simeq 0.30 - 0.35$. In all cases, we see that the DE density starts to dominate near the present cosmological epoch, and its equation of state corresponds to accelerated expansion. Thus, the nonlocal term generates a dynamical DE density that drives an accelerated expansion of the Universe at the current cosmological epoch. This is already a very non-trivial result: it means that giving a mass to the conformal mode, and covariantizing it as discussed in Section 1.3.2, provides an explanation for the observed accelerated expansion of the Universe.

Fig. 2.2 shows the relative difference $[H_{\rm RT}(z) - H_{\Lambda\rm CDM}(z)]/H_{\Lambda\rm CDM}(z)$ between each RT model (minimal and with $\Delta N = 34, 50, 64, 100$) and ΛCDM. Once again, the predictions of each model are computed using the respective mean values of the cosmological parameters in Table 2.2. At $z = 0$ the difference between the various curves is due to the different mean values for H_0, and at large z (but still within MD) it is determined by the different mean values for Ω_M. We see that, at $z = 0$, the minimal RT model differs from ΛCDM by about 1%, while the RT models with large ΔN give a prediction for H_0 basically indistinguishable from that of ΛCDM. Away from $z = 0$, $|\Delta H(z)|/H(z)$ is of order 0.5% or less. The evolution with redshift is, however, quite distinctive, with $\Delta H(z)/H(z)$ oscillating and changing sign

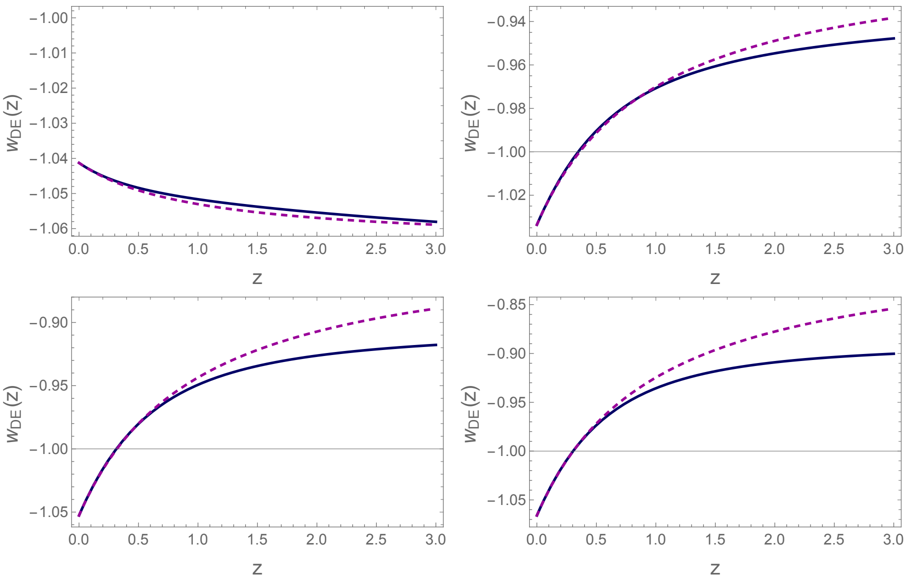

FIGURE 2.3: The DE equation of state $w_{\rm DE}(z)$ from the numerical integration of the equations (blue solid lines), compared with the parametrization (2.1.19) (magenta dashed lines) for RT minimal (upper left panel) and RT with $\Delta N = 34$ (upper right), $\Delta N = 50$ (lower left) and $\Delta N = 64$ (lower right). Figure from [31].

as z increases. These differences with respect to ΛCDM can be compared to a compilation of measurements of $H(z)$ at different redshifts. We will perform this test in Section 2.1.3, after having performed the Bayesian parameter estimation for the models.

It is interesting to compare the actual predictions of the model to the results obtained with the standard (w_0, w_a) parametrization $w_{\rm DE}(a) = w_0 + (1-a)w_a$ [108, 109], or, in terms of redshift,

$$w_{\rm DE}(z) = w_0 + \frac{z}{1+z} w_a \,. \tag{2.1.19}$$

Setting $w_0 \equiv w(a=1)$ and $w_a \equiv -(dw/da)_{|a=1}$ we get the values of w_0 and w_a given in Table 2.1. In Fig. 2.3 we compare the actual numerical result for $w(z)$ to the fit provided by this parametrization. We see that, for large ΔN, the parametrization (2.1.19) is not very accurate beyond some value of z, with the range in z shrinking as ΔN increases.

2.1.2 Scalar perturbations

Perturbation equations

Cosmological scalar perturbations for the RR and RT model (in the minimal case) have been studied in detail in [110, 111] (see also [31, 29] for reviews). Here, after recalling the basic formalism, we will extend the results to the RT model with large ΔN and we will present updated results on various indicators of cosmological perturbations, using the values of the cosmological parameters that will be determined in section 2.1.3 by the comparison with observations. We work in the Newtonian gauge, where, in the scalar perturbation sector, the perturbed FRW metric has the form

$$ds^2 = -(1+2\Psi)dt^2 + a^2(t)(1+2\Phi)\delta_{ij}dx^i dx^j \,, \tag{2.1.20}$$

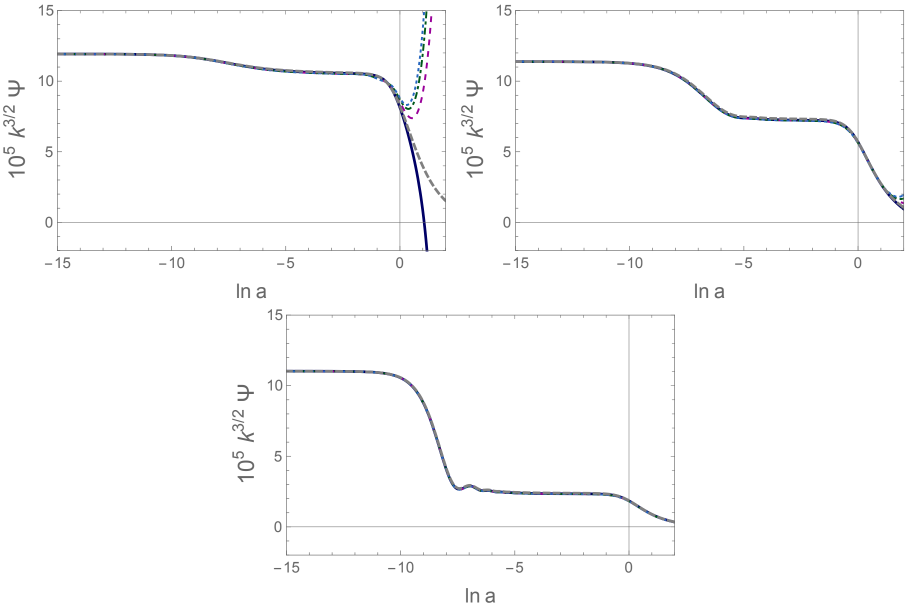

FIGURE 2.4: $k^{3/2}\Psi(a;k)$ in ΛCDM (gray dashed line), the minimal RT model (blue solid line) and RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted) for $\kappa = 0.1$ (upper left panel), $\kappa = 1$ (upper right) and $\kappa = 5$ (lower panel). On the scale of these figures, the results for $\kappa = 5$ are indistinguishable among the models, while for $\kappa = 1$ one can barely distinguish some small differences in the cosmological future, $x > 0$. Observe that the quantity that we plot is $k^{3/2}\Psi(a;k)$ multiplied by a factor 10^5. Matter-radiation equilibrium is at $x \simeq -8.1$, and in this region one sees the usual transition between two different plateaux in Ψ. Figure from [31].

where Φ and Ψ are the Bardeen variables. We similarly perturb the auxiliary fields, writing

$$U(t,\mathbf{x}) = \bar{U}(t) + \delta U(t,\mathbf{x})\,, \qquad S_\mu(t,\mathbf{x}) = \bar{S}_\mu(t) + \delta S_\mu(t,\mathbf{x})\,, \tag{2.1.21}$$

where, in this Section, background quantities are denoted with an overbar. In FRW, $\bar{S}_i$ vanishes because at the background level there is no preferred spatial direction, but its perturbation δS_i is non-vanishing. As with any vector, we can decompose it into a transverse and longitudinal part, $\delta S_i = \delta S_i^{\rm T} + \partial_i(\delta S)$, where $\partial_i(\delta S_i^{\rm T}) = 0$. Since we are considering scalar perturbations, we only retain δS. Thus, in the RT model the metric perturbations in the scalar sector are described by $\Psi, \Phi, \delta U, \delta S_0$ and δS. It is convenient to trade S_0 and S for

$$V = H_0 S_0\,, \qquad Z = H_0^2 S\,, \tag{2.1.22}$$

so we eventually work with the variables $\{\Psi, \Phi, \delta U, \delta V, \delta Z\}$. We similarly perform the usual expansion of the energy-momentum tensor, writing

$$T_0^0 = -(\bar{\rho} + \delta\rho)\,, \qquad T_i^0 = (\bar{\rho} + \bar{p})v_i\,, \qquad T_j^i = (\bar{p} + \delta p)\delta_j^i + \Sigma_j^i\,, \tag{2.1.23}$$

where $\bar{\rho}$ and $\bar{p}$ are the unperturbed density and pressure. The matter perturbation variables are therefore $\delta\rho, \delta p$, v_i, and the anisotropic stress tensor Σ^i_j, which is symmetric and traceless, $\Sigma^i_i = 0$. The pressure perturbations can be written as $\delta p = c_s^2 \delta\rho$, where c_s^2 is the speed of sound of the fluid, and we define as usual $\delta \equiv \delta\rho/\bar{\rho}$ and $\theta \equiv \delta^{ij}\partial_i v_j$, with δ_R, θ_R referring to radiation and δ_M, θ_M to matter. We only consider the contribution to $T_{\mu\nu}$ from radiation and non-relativistic matter, so $\Sigma^i_j = 0$. We transform the perturbation equations to Fourier space and we denote comoving momenta by k. We further define

$$\hat{k} = k/(aH)\,, \qquad \hat{\theta} = \theta/(aH)\,. \tag{2.1.24}$$

We also use

$$\kappa \equiv k/k_{\rm eq}\,, \tag{2.1.25}$$

where $k_{\rm eq} = a_{\rm eq}H_{\rm eq}$ is the wavenumber of the mode that enters the horizon at matter-radiation equilibrium. Numerically, $k_{\rm eq} \simeq 0.014\, h_0\, {\rm Mpc}^{-1} \simeq 0.010\, {\rm Mpc}^{-1}$. To illustrate our numerical results, we use as reference values $\kappa = 0.1, 1$ and 5 (or just $\kappa = 0.1$ and 1, when the results for $\kappa = 5$ turn out to be graphically indistinguishable from $\kappa = 1$). The mode with $\kappa = 5$ entered inside the horizon already during RD, while the mode $\kappa = 1$ reentered at matter-radiation equality. In contrast, the mode with $\kappa = 0.1$ was outside the horizon during RD and most of MD, and re-entered at $z \simeq 1.5$.

In the scalar sector, the perturbed Einstein equations were computed in [111]

$$\hat{k}^2\Phi + 3(\Phi' - \Psi) = \frac{3}{2h^2\rho_0}\left[\delta\rho + \gamma\rho_0\left(\delta U - h\delta V' + 2h\Psi\bar{V}' + h\Psi'\bar{V}\right)\right], \tag{2.1.26}$$

$$\hat{k}^2(\Phi' - \Psi) = -\frac{3}{2h^2\rho_0}\left[\bar{\rho}(1+w)\hat{\theta} + \hat{k}^2\gamma\rho_0\left(h^2\delta Z - \frac{h^2}{2}\delta Z' + h\Psi\bar{V} - \frac{h}{2}\delta V\right)\right], \tag{2.1.27}$$

$$\hat{k}^2(\Psi + \Phi - 3\gamma\delta Z) = \frac{9}{2h^2\rho_0}\bar{\rho}(1+w)e^{2x}\sigma\,, \tag{2.1.28}$$

$$\begin{aligned}&\Phi'' + (3+\zeta)\Phi' - \Psi' - (3+2\zeta)\Psi + \frac{\hat{k}^2}{3}(\Phi + \Psi) \\ &\quad = -\frac{3}{2h^2\rho_0}\left[\delta p - \gamma\rho_0\left(\delta U - h(\Phi' - 2\Psi)\bar{V} - h\delta V - \frac{\hat{k}^2}{3}h^2\delta Z\right)\right],\end{aligned} \tag{2.1.29}$$

where $\hat{k} = k/(aH)$, $\hat{\theta} = \theta/(aH)$, an overbar denotes here a background quantity and, as before, the prime denotes the derivative with respect to $x \equiv \ln a$, and $\zeta(x) \equiv h'/h$. The linearization of the equations for the auxiliary fields gives

$$\begin{aligned}\delta U'' + (3+\zeta)\delta U' + \hat{k}^2\delta U = {}& 2\hat{k}^2(\Psi + 2\Phi) + 6(\Phi'' + (4+\zeta)\Phi') - 6\left[\Psi' + 2(2+\zeta)\Psi\right] \\ & + 2\Psi\bar{U}'' + \left[2\Psi(3+\zeta) + (\Psi' - 3\Phi')\right]\bar{U}'\end{aligned} \tag{2.1.30}$$

$$\begin{aligned}\delta V'' + (3+\zeta)\delta V' + \frac{\hat{k}^2}{2}h(\delta Z' - 4\delta Z) - h^{-1}\delta U' = {}& 2\Psi\bar{V}'' + \left[2(3+\zeta)\Psi + 3(\Psi' - \Phi')\right]\bar{V}' \\ & + \left[\Psi'' + (3+\zeta)\Psi' + 6\Phi'\right]\bar{V} - \left[(1/2)\hat{k}^2 - 3\right](\delta V - 2\Psi\bar{V})\,,\end{aligned} \tag{2.1.31}$$

$$\begin{aligned}\delta Z'' + (1+\zeta)\delta Z' + 2\left(\hat{k}^2 - (3+\zeta)\right)\delta Z = {}& 2h^{-2}\delta U \\ & - h^{-1}\left[\delta V' + 5\delta V - 4\Psi\bar{V}' - 2(\Psi' - \Phi' + 4\Psi)\bar{V}\right].\end{aligned} \tag{2.1.32}$$

In Fig. 2.4 we show the time evolution of the Fourier modes of the Bardeen variable $\Psi_{\mathbf{k}}$ for the RT model (minimal and with $\Delta N = 34, 50, 64$), obtained from the numerical integration of these perturbation equations, and we compare with the result in ΛCDM, for $\kappa = 0.1$, $\kappa = 1$ and $\kappa = 5$. We actually plot $k^{3/2}\Psi_k$, whose square gives the variance of the field per unit logarithmic interval of momentum. We see that, up to the present time $x = 0$, the evolution of the scalar perturbations is well-behaved, and very close to that of ΛCDM, and become closer and closer as k increases. This can be understood from the fact that any instability induced by the nonlocal term on the cosmological evolution can only develop on a timescale t such that mt is (much) larger than one. However, we have seen that m is of order H_0, and in fact numerically smaller, with $m \simeq 0.68 H_0$ for the minimal RT model and even smaller for large ΔN. Thus, any instability induced by the nonlocal term can only develop on a timescale larger or equal than to a few times H_0, and therefore in the cosmological future, where these modes could eventually enter a non-linear regime.

Scalar perturbations during inflation

Here we study the behavior of scalar perurbations during a phase of primordial inflation, following the analysis in [107]. By construction, a successful model of primordial inflation must be such that, during the inflationary phase, all the modes relevant for cosmology eventually exit the horizon. It can be natural to set initial conditions on the auxiliary fields at the time that they exit the horizon (see the discussion below for the proper initial conditions), so we need to study the stability of the perturbation equations for super-horizon modes. In this regime the physical wavelength is much larger than H^{-1} and the physical momentum $k/a \ll H$, so $\hat{k} \ll 1$. Then, in eqs. (2.1.26)–(2.1.32) we keep only the lowest-order terms in $\hat{k}^2$. We also limit ourselves for simplicity to an inflationary phase in the limit of a de Sitter expansion, supported by a constant vacuum energy density. Then the matter perturbations $\delta\rho, \hat{\theta}$ and σ vanish in the exact de Sitter limit.

Furthermore the dimensionless Hubble parameter $h(x) = H(x)/H_0$ becomes constant and is numerically very large for all typical inflationary scales. Indeed,

$$h_{\rm infl} \equiv \frac{H_{\rm infl}}{H_0} = \left(\frac{\rho_{\rm infl}}{\rho_0}\right)^{1/2} \simeq \left(\frac{M_{\rm infl}}{3\times 10^{-3}{\rm eV}}\right)^2 , \tag{2.1.33}$$

where $\rho_{\rm infl}$ is the energy density during a phase of de Sitter inflation, and $M_{\rm infl} \equiv \rho_{\rm infl}^{1/4}$ is the corresponding mass scale. We can then show that all terms in the above equation involving $\bar{V}(x)$ can be set to zero. Indeed, from $Y = U - \dot{S}_0$, it follows that $V' = (U - Y)/h$. Using $\bar{U}(x) = 4x$ and $\bar{Y}(x) \simeq \exp\{0.79x\}$ (apart from an overall constant of order one) it follows that, during inflation, $\bar{V}(x) = \mathcal{O}(h_{\rm infl}^{-1}\exp\{0.79x\})$. Thus $\bar{V}$ reaches a maximum value at the end of inflation of order $h_{\rm infl}^{-1}\exp\{0.79\Delta N\}$. Despite the exponential term, this is still an extremely small number, because $h_{\rm infl}$ is huge. For instance, for $M_{\rm infl} = 10^{13}$ GeV, from eq. (2.1.33) we have $h_{\rm infl} \sim 10^{49}$ and, even setting $\Delta N = 60$, $h_{\rm infl}^{-1}\exp\{0.79\Delta N\} \sim 10^{-49} \times 10^{20}$ is totally negligible. More generally, for inflation at a scale $M_{\rm infl}$, the minimum number of efolds necessary for solving the horizon and flatness problems (neglecting the effects of reheating) is given by

$$\Delta N \simeq 64 - \log\left(\frac{10^{16}\,{\rm GeV}}{M_{\rm infl}}\right) . \tag{2.1.34}$$

Combining this with eq. (2.1.33) we get

$$h_{\rm infl}^{-1}e^{0.79\Delta N}\simeq 10^{-33}\left(\frac{10^{16}\,{\rm GeV}}{M_{\rm infl}}\right)^{1.21}\,, \tag{2.1.35}$$

that even for $M_{\rm infl}$ as low as 1 TeV is at most of order 10^{-18}. Since all occurrences of $\bar{V}$ in the perturbation equations appear multiplied overall by factors $\mathcal{O}(1)$ or even $\mathcal{O}(1/h)$, we can set $\bar{V}(x)=0$ in the above equations. In contrast, from eq. (2.1.11), $\bar{U}'=4$ is a number $O(1)$.

As usual, because of general covariance, in the scalar sector the four linearized Einstein equations (2.1.26)–(2.1.29) are not independent once we take into account the linearized energy-momentum conservation, and we have only two independent equations. We take them to be eqs. (2.1.26) and (2.1.28) which, together with eqs. (2.1.30)–(2.1.32) give a system of five equations for the five functions $\Psi,\Phi,\delta U,\delta V$ and δZ.

To take the limit $\hat{k}^2\ll 1$ in eq. (2.1.28) we observe that, on the left-hand side, we have $\hat{k}^2=k^2/(a^2H^2)$, while on the right-hand side we have $\bar{\rho}e^{2x}/(h^2\rho_0)$. Since $e^{2x}=a^2$ and $\bar{\rho}\sim\rho_0/a^4$, this term is of order $H_0^2/(a^2H^2)$. All modes of cosmological interest are inside the horizon at the present time, so they have $k>H_0$. Thus, even if we take the limit $\hat{k}\to 0$, the factor $\hat{k}^2$ on the left-hand side of eq. (2.1.28) is larger than the factor $\bar{\rho}e^{2x}/h^2\rho_0$ on the right-hand side. On top of this, on the right-hand side we also have the factor $(1+w)e^{2x}\sigma$, which vanishes in an exact de Sitter phase, and is in any case quadratic in the perturbations of an inflaton field (see footnote 8 of [107]). Thus, in this limit eq. (2.1.28) becomes $\Psi+\Phi-3\gamma\delta Z=0$.

It is convenient to introduce rescaled variables u,v,z from $\delta U=h^2u$, $\delta V=hv$ and, for uniformity of notation, $\delta Z=z$. Then the independent equations take the following form, for super-horizon modes during inflation:

$$\Phi'-\Psi = \frac{1}{2}\gamma(u-v')\,, \tag{2.1.36}$$

$$\Psi+\Phi = 3\gamma z\,, \tag{2.1.37}$$

$$u''+3u' = (6\Phi''+12\Phi'-2\Psi')/h^2\,, \tag{2.1.38}$$

$$v''+3v'-3v = u'\,, \tag{2.1.39}$$

$$z''+z'-6z = 2u-v'-5v\,. \tag{2.1.40}$$

We observe that, with this rescaling, h disappears from eq. (2.1.36). Furthermore, in the limit $h\to\infty$, the right-hand side of eq. (2.1.38) vanishes and the equations for the perturbations of the auxiliary fields, u,v and z, decouple from the matter perturbations. Solving the homogeneous equations associated with eqs. (2.1.38)–(2.1.40) we see that z has a mode growing as e^{2x} and v a mode growing as $e^{\alpha_+^{\rm dS}x}\simeq e^{0.79x}$. By the end of inflation, within this linearized approximation, z has grown by a factor $e^{2\Delta N}$ compared to its initial value, while v by a factor $e^{0.79\Delta N}$. Taking for instance $\Delta N=60$, we have $e^{2\Delta N}\sim 10^{52}$. At first sight, this leads to the conclusion that these perturbation variables leave the linear regime. If this were the case then, through their coupling with the metric perturbations in eq. (2.1.36), they would spoil cosmological perturbation theory and the initial conditions $\Phi,\Psi\sim 10^{-5}$ provided by inflation in the standard scenario, leading to an unacceptable cosmology.

However, this conclusion is incorrect because it does not take into account that, because of the rescaling performed, the natural values of the initial conditions on the variables u,v and z are extremely small. In terms of the original auxiliary field $U\equiv -\Box^{-1}R$, a natural initial condition, say set at the time $t_{\rm in}$ that the modes exit the horizon during inflation, could be $U_{\rm in} = \mathcal{O}(1)$. The precise value of $U_{\rm in}$ has no importance and, as we will see from the final result, it is enough that the initial conditions set at this time are not astronomically large numbers, say 10^{55}, see eq. (2.1.44) below. Hence $u(t) \equiv U(t)/h(t)^2$, where $h(t) = H(t)/H_0$. As we have seen in eq. (2.1.33), during a phase of primordial inflation $h^2(t) \equiv h^2_{\rm infl}$ is a huge number. For instance, for $M_{\rm infl} = 10^{13}\,{\rm GeV}$, we have $h^2_{\rm infl} \sim 10^{98}$. Thus, the natural initial condition on u during inflation is

$$u_{\rm in} = h^{-2}_{\rm infl} U_{\rm in}\,, \tag{2.1.41}$$

which, for $M_{\rm infl} = 10^{13}\,{\rm GeV}$, means $u_{\rm in} \sim 10^{-98} U_{\rm in}$. The same happens for the variables v and z (which are those that display instabilities). The natural initial condition on S_0 is such that $(\partial_0 S_0)_{\rm in} \sim U_{\rm in}$ and using $\partial_0 \sim H$ we get $(S_0)_{\rm in} \sim H^{-1}_{\rm infl} U_{\rm in}$. Therefore, for the variable $V = H_0 S_0$, we have $V_{\rm in} \sim U_{\rm in} H_0/H_{\rm infl} = h^{-1}_{\rm infl} U_{\rm in}$. The rescaled variable v is defined by $v = V/h$ and therefore

$$v_{\rm in} \sim h^{-2}_{\rm infl} U_{\rm in}\,. \tag{2.1.42}$$

Similarly, one can see that the natural initial conditions on the momentum modes $S_{\mathbf{k}}$ are given by $Ua^2 \sim k^2 S_{\mathbf{k}}$, where k is the comoving momentum. As we mentioned, a natural time to impose initial conditions is when the mode is crossing the horizon, so that $k^2/a^2 \sim H^2_{\rm infl}$. Then $S_{\rm in} \sim U_{\rm in}/H^2_{\rm infl}$ and for $z \equiv H_0^2 S$ we get initial conditions

$$z_{\rm in} \sim h^{-2}_{\rm infl} U_{\rm in}\,, \tag{2.1.43}$$

just as for u and v. The mode that grows faster is z, which during de Sitter grows as e^{2x}. Thus, starting from an initial value $\sim h^{-2}_{\rm infl}$, it reaches a value at the end of inflation of order $h^{-2}_{\rm infl} e^{2\Delta N}$. Using eq. (2.1.33), we see that this is still an extremely small value,

$$h^{-2}_{\rm infl} e^{2\Delta N} \simeq 10^{-55} \left(\frac{10^{16}\,{\rm GeV}}{M_{\rm infl}} \right)^2 , \tag{2.1.44}$$

that for $M_{\rm infl} = 10^{13}\,{\rm GeV}$ is of order 10^{-49} and even for $M_{\rm infl}$ as low as 1 TeV is still of order 10^{-29}. Thus, the perturbations of the auxiliary fields always remain minuscule during de Sitter inflation, despite their formally growing modes.

A posteriori, the physical interpretation of this result is obvious. The energy scale associated to the nonlocal term is so small compared to the inflationary scale that the nonlocal term has no impact whatsoever on the evolution during inflation. This is precisely what we found in Sect. 2.1.1 for the background evolution during inflation, where again we found an instability in the variable Y, that however has no consequence on the dynamics of the scale factor during inflation.

The RT model therefore has a viable cosmological evolution, both at the background level and at the level of perturbations, during all cosmological phases, despite the presence of an instability during an early inflationary phase. The basic reason is that the energy scale associated to the nonlocal term is totally negligible with respect to the inflationary scale.

Deviations from GR

The full set of perturbation equations computed in [111] is needed for implementing the model into a Boltzmann code and comparing its predictions to CMB, BAO and

SNe observations, as we will do in Section 2.1.3. For a first qualitative understanding, however, it is convenient to introduce some simpler indicators of deviations from ΛCDM. One such quantity is the effective Newton's constant, which is defined so that the modified Poisson equation for the Fourier modes $\Phi_{\mathbf{k}}$ can be rewritten as in GR, with G replaced by $G_{\rm eff}(x,k)$ [recall that here $x \equiv \ln a(t)$ is used to parametrize the time evolution, and should not be confused with a spatial variable],

$$\begin{aligned} k^2\Phi_{\mathbf{k}}(x) &= 4\pi G_{\rm eff}(x;k)a^2\rho_0 \\ &\times\left[\Omega_R e^{-4x}\left(\delta_{R,\mathbf{k}}(x)+\frac{4}{\hat{k}^2}\hat{\theta}_{R,\mathbf{k}}(x)\right)+\Omega_M e^{-3x}\left(\delta_{M,\mathbf{k}}(x)+\frac{3}{\hat{k}^2}\hat{\theta}_{M,\mathbf{k}}(x)\right)\right]. \end{aligned} \tag{2.1.45}$$

Its explicit expression in terms of the perturbed fields is [111]

$$\frac{G_{\rm eff}(x;k)}{G} = 1+\gamma\frac{\delta U_{\mathbf{k}}+h\left(2\Psi_{\mathbf{k}}\bar{V}'+\Psi'_{\mathbf{k}}\bar{V}-\delta V'_{\mathbf{k}}\right)+3h^2\left(\delta Z_{\mathbf{k}}-\frac{1}{2}\delta Z'_{\mathbf{k}}\right)+3h\left(\Psi_{\mathbf{k}}\bar{V}-\frac{1}{2}\delta V_{\mathbf{k}}\right)}{\Omega_R e^{-4x}\left(\delta_{R,\mathbf{k}}(x)+\frac{4}{\hat{k}^2}\hat{\theta}_{R,\mathbf{k}}(x)\right)+\Omega_M e^{-3x}\left(\delta_{M,\mathbf{k}}(x)+\frac{3}{\hat{k}^2}\hat{\theta}_{M,\mathbf{k}}(x)\right)}. \tag{2.1.46}$$

From this expression one finds that, for sub-horizon modes, i.e. in the limit $\hat{k} \gg 1$, we have [111]

$$\frac{G_{\rm eff}(x;k)}{G} = 1+\mathcal{O}\left(\frac{1}{\hat{k}^2}\right). \tag{2.1.47}$$

As we will see in Section 2.1.4, this property, which is not shared by other modified gravity models and in particular by the RR nonlocal model, is crucial, since it allows the RT model to evade limits on the time variation of the (effective) Newton's constant obtained from Lunar Laser Ranging.

Together with $G_{\rm eff}$, a second useful indicator is [112]

$$\eta(x;k) = \frac{\Phi_{\mathbf{k}}(x)+\Psi_{\mathbf{k}}(x)}{\Phi_{\mathbf{k}}(x)}, \tag{2.1.48}$$

which, in GR, vanishes in the absence of anisotropic stress. Alternatively, two useful quantities are the functions $\mu(x;k)$ [113] and $\Sigma(x;k)$ [112] defined by the relations

$$\Psi = [1+\mu(x;k)]\Psi_{\rm GR} \qquad \Psi-\Phi = [1+\Sigma(x;k)](\Psi-\Phi)_{\rm GR}, \tag{2.1.49}$$

where the subscript denotes the same quantities computed in GR, assuming a ΛCDM model with the same value of Ω_M as the modified gravity model. In the literature the quantity that we call $1+\mu$ is sometimes denoted by μ, and similarly our $1+\Sigma$ is sometimes denoted by Σ. Our definitions are such that, in GR, $\mu=\Sigma=0$. The advantage of this parametrization is that it separates the modifications to the motion of non-relativistic particles, which is described by μ, from the modification to light propagation, which is encoded in Σ. Therefore μ is sensitive to structure formation and Σ is sensitive to lensing.

In Fig. 2.5 we show the numerical results for the effective Newton constant as a function of redshift, for the minimal RT model and for the RT model with $\Delta N = 34, 50, 64$, for $\kappa = 0.1$ and 1. We see that, already for $\kappa = 0.1$ (i.e. $k = 0.1k_{\rm eq} \simeq 0.001\,{\rm Mpc}^{-1}$), $G_{\rm eff}$ differs by G by less than 1%, and, for higher values of k, $G_{\rm eff}$ goes quickly to G, in agreements with eq. (2.1.47) (for instance, in the plot for $\kappa = 5$, $|G_{\rm eff}/G|$ would always be below 1.001). For these values of k, there are also some

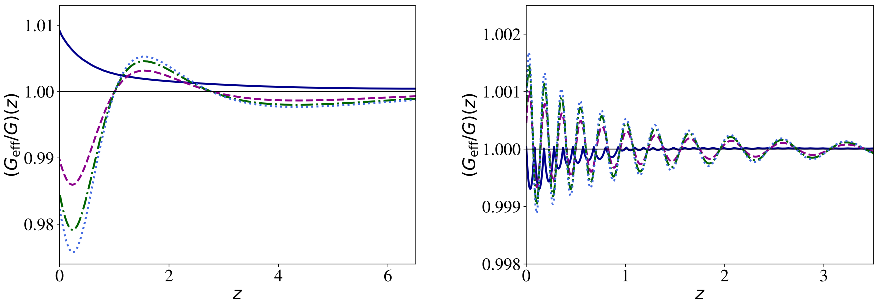

FIGURE 2.5: $G_{\rm eff}/G$ as a function of z for fixed κ, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted), for $\kappa = 0.1$ (left panel) and $\kappa = 1$ (right panel). Figure from [31].

oscillations as a function of z and, for given z, the envelop of the oscillations reproduces the $1/k^2$ behavior found analytically in eq. (2.1.47). Notice that, because of eq. (2.1.47), on small scales $G_{\rm eff}$ reduces to the standard Newton's constant G probed by solar system or by laboratory experiments. However, at typical cosmological scales such as $k \sim k_{\rm eq}$, its value is different, even at $z = 0$. In particular, in the RT models with large ΔN, on these scales $G_{\rm eff} < G$, i.e. gravity is weakened on cosmological scales, while for the minimal RT model it is strengthened. Fig. 2.6 shows, on a logarithmic scale, the dependence of $|(G_{\rm eff}/G) - 1|$ on the wavenumber k, for three different values of the redshift, $z = 0, 0.5$ and 1.

Fig 2.7 shows η as a function of z, again for $\kappa = 0.1$ and 1, while in Figs. 2.8 and 2.9 we show the same results for the indicators Σ and μ. Notice in particular that both Σ and μ have a rather non-trivial dependence on k for cosmological scales $k \sim k_{\rm eq}$. We see from the plots that, at small redshifts, in the RT model with large ΔN, for $k = 0.1 k_{\rm eq}$ both Σ and μ are positive (with $1 + \Sigma$ higher by about 5% than the ΛCDM value of unity, and μ by about 10% in $z = 0$), while for $k = k_{\rm eq}$ or larger the situation is reversed and Σ and μ become negative at small z.

Another useful derived quantity is the growth rate $f(z,k) \equiv d \log \delta_M / d \ln a$. In ΛCDM, for the typical wavenumbers relevant for structure formation, $f(z,k)$ is basically independent of wavenumber k and very well fitted by $f(z) = [\Omega_M(z)]^\gamma$ with γ a constant, numerically close to 0.55. More precisely, writing $f(z) = [\Omega_M(z)]^{\gamma(z)}$, the function $\gamma(z)$ for ΛCDM is shown as the gray solid line in Fig. 2.10, so it is indeed approximately constant and given numerically by $\gamma \simeq 0.55$, within percent level accuracy. We find that the fit $f(z) = [\Omega_M(z)]^{\gamma(z)}$ also holds for the RT model, again with a function $\gamma(z)$ independent of the wavenumber k. The corresponding functions $\gamma(z)$ are shown in Fig. 2.10 for the RT model, minimal and with large ΔN. We see that, for large values of ΔN, $\gamma(z)$ is indeed independent of z within percent level accuracy, just as in ΛCDM, and again is given numerically by $\gamma \simeq 0.55$ (not to be confused with the parameter γ of the background equation for the RT model). For the minimal RT model the variation of $\gamma(z)$ with redshift is somewhat larger, but still it stays between $0.55 - 0.56$ up to $z = 2$. Notice that the growth index γ is a useful quantity only as long as we are in the epoch where DE is still important. When we are deep into MD, $\Omega_M(z) \to 1$, and $[\Omega_M(z)]^\gamma \to 1$ independently of γ.

Two main conclusions emerge from our study of the cosmological perturbations of the RT model in the scalar sector. First, they are well-behaved. This is already a rather non-trivial result. Several modified gravity models have indeed been ruled out by the presence of instabilities in their perturbations. This was for instance the

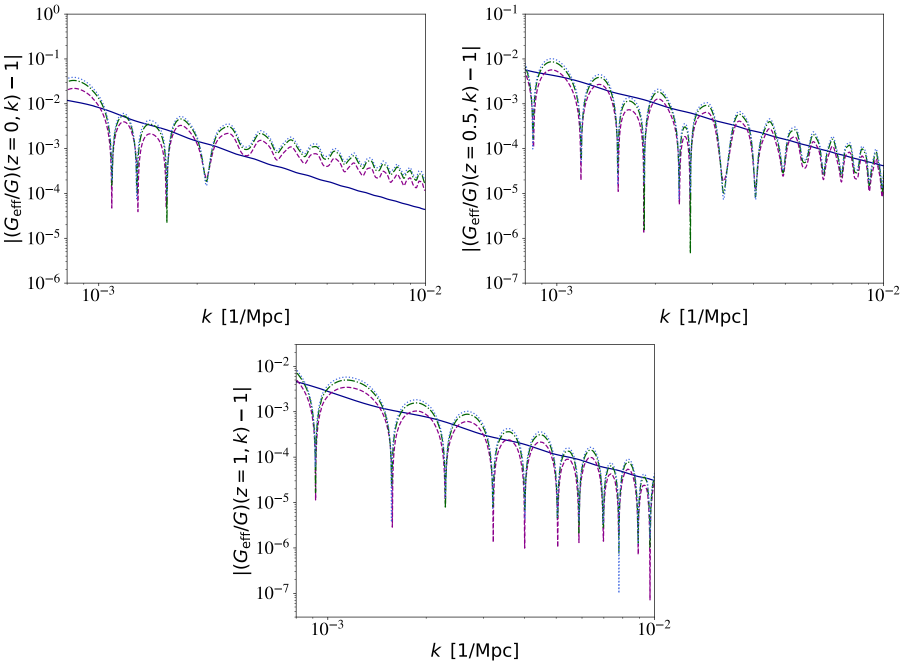

FIGURE 2.6: $|(G_{\rm eff}/G)-1|$ as a function of k for fixed z, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted), on a logarithmic scale. The three panels refers to $z = 0$ (upper left panel), $z = 0.5$ (upper right) and $z = 1$ (lower panel). The sign of $(G_{\rm eff}/G)-1$ is such that, close to the vertical axis, $G_{\rm eff}/G > 1$ for the minimal model and $G_{\rm eff}/G < 1$ for the other cases, and the sign changes each time the logarithmic plot has a downward spike. Figure from [31].

case for the DPG model [23], which opened the way to the study of IR modifications of GR and has a self-accelerated solution [114, 115] but had a ghost-like instability on the self-accelerated branch [116, 117, 118, 119, 120]. Massive gravity [121, 21, 122] has difficulties already in obtaining a viable background FRW evolution [123], while in bigravity [17] a background FRW solutions exist, but, in a branch of solutions that has a dynamical dark energy, the cosmological perturbations have instabilities in both the scalar and tensor sectors [124, 125, 126, 127, 128, 129, 130]. Thus, already the fact of producing quite naturally a viable cosmological background evolution with self-acceleration, and stable scalar perturbations, is a non-trivial results.

The second conclusion that emerges from this study is that, both in the background evolution and in the scalar perturbations, the RT model is very close to ΛCDM, with deviations of at most a few percent, for all ΔN. This already indicates that the model is a good candidate for fitting well the current cosmological observations. In the next Section we will confirm this conclusion by comparing the RT model with ΛCDM from the point of view of the quality of the fit to the cosmological observations, and we will perform Bayesian parameter estimation for the values of the cosmological parameters.

2.1.3 Comparison with cosmological datasets

We now show the details of a comparison with cosmological observations, using the most recent cosmological datasets. We have implemented the perturbations of the

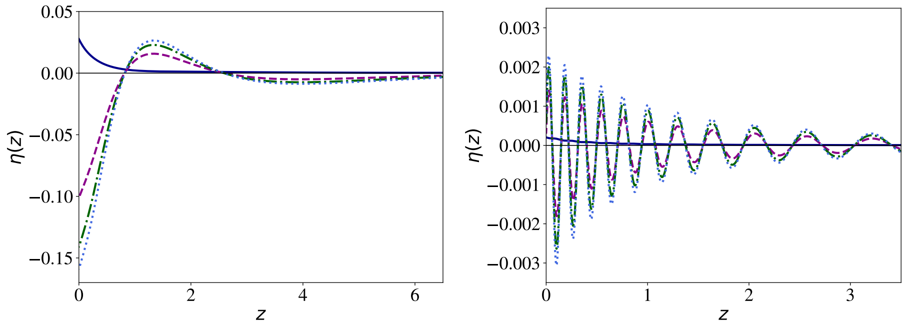

FIGURE 2.7: η as a function of z, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted), for $\kappa = 0.1$ (left panel) and $\kappa = 1$ (right panel). Figure from [31].

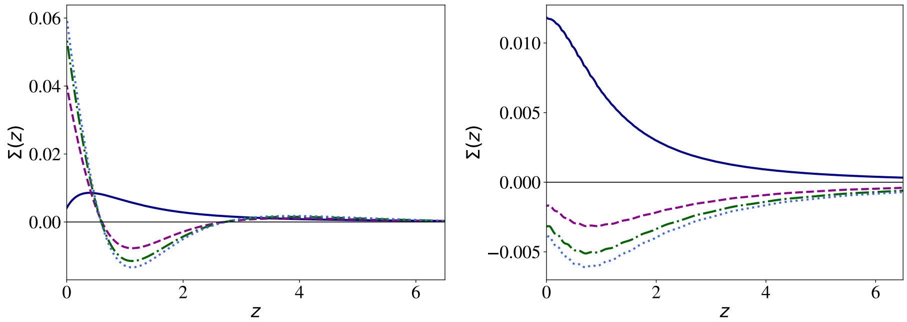

FIGURE 2.8: Σ as a function of z, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted), for $\kappa = 0.1$ (left panel) and $\kappa = 1$ (right panel). Figure from [31].

RT model into the CLASS cosmological Boltzmann code [131], that we have modified so to describe the background evolution and scalar perturbations of the RT model. The modified version of the CLASS code that we used is publicly available on GitHub [132].

Methodology and parameters

For ΛCDM, the *Planck* baseline analysis uses six independent cosmological parameters: the Hubble parameter today $H_0 = 100h_0\,\mathrm{km\,s^{-1}Mpc^{-1}}$, the physical baryon and cold dark matter density fractions today $\omega_b = \Omega_b h_0^2$ and $\omega_c = \Omega_c h_0^2$, respectively, the amplitude A_s and tilt n_s of the primordial scalar perturbations, and the reionization optical depth $\tau_{\rm re}$. Note that, assuming flatness, the energy fraction Ω_Λ associated to a cosmological constant is a derived parameter, fixed by the flatness condition. In the RT model we have a mass scale m [or, equivalently, the dimensionless parameter γ, eq. (2.1.6)] which replaces the cosmological constant, and again can be taken as a derived parameter, fixed by the flatness condition. Thus, for the RT model, we can take the same six independent cosmological parameters, as in ΛCDM.

An important extension, however, is provided by the sum of neutrino masses, $\sum_\nu m_\nu$. As discussed in [133], their inclusion can a priori be important when comparing a modified gravity model to ΛCDM. Oscillation experiments give a lower limit

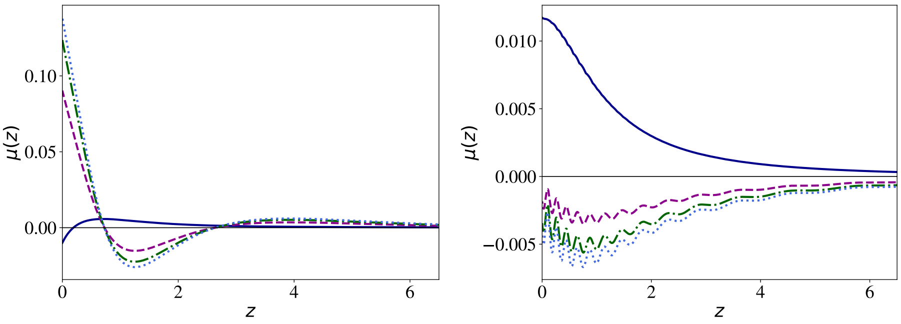

FIGURE 2.9: μ as a function of z, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted), for $\kappa = 0.1$ (left panel) and $\kappa = 1$ (right panel). Figure from [31].

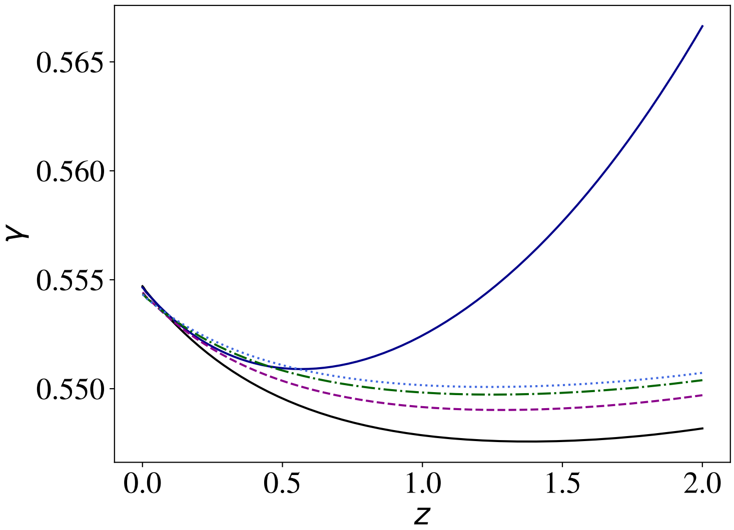

FIGURE 2.10: The function $\gamma(z)$ related to the growth rate $f(z)$ by $f(z) = [\Omega_M(z)]^{\gamma(z)}$, for ΛCDM (gray solid line), for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted). Figure from [31].

$\sum_\nu m_\nu \gtrsim 0.06$ eV [134] (assuming a normal mass hierarchy dominated by the heaviest neutrino mass eigenstate). In the *Planck* baseline analysis the sum of neutrino masses is kept fixed to this minimum allowed value. As discussed in the *Planck* papers [135, 136], there is actually no compelling theoretical reason for this choice, and there are other possibilities, including a degenerate hierarchy with $\sum_\nu m_\nu \gtrsim 0.1$ eV. The choice of fixing the sum of neutrino masses to the minimum allowed values is justified by the fact that, in ΛCDM, letting the sum of neutrino masses as a free parameter, one finds that its marginalized posterior is peaked in zero, and if we let it vary with the prior $\sum_\nu m_\nu \geqslant 0.06$ eV the data drive $\sum_\nu m_\nu$ back to the prior (see Fig. 34 of [136]). In contrast, as was realized in [133], in a modified gravity model the posterior for $\sum_\nu m_\nu$ could be peaked at a value higher than the lower bound 0.06 eV. This happens in the RR nonlocal model [133, 30] and also in the minimal RT model [31]. A uniform comparison of a modified gravity model with ΛCDM therefore requires to let $\sum_\nu m_\nu$ as a free parameter in both models. We denote by $\nu\Lambda$CDM the ΛCDM model in which $\sum_\nu m_\nu$ is added to the list of free parameters.

In summary, we perform Bayesian parameter estimation for both $\nu\Lambda$CDM and the RT model (minimal, and with $\Delta N = 34, 50, 64$), and we will compare the quality of their fits to the datasets discussed below, using, as free parameters,

$$\theta = \{H_0, \omega_b, \omega_c, A_s, n_s, \tau_{\rm re}, \textstyle\sum_\nu m_\nu\} \ . \tag{2.1.50}$$

For CMB, SNe and BAO we use the following likelihoods:

- For CMB we use the *Planck* 2018 data release, using the low-ℓ temperature-only likelihood, the low-ℓ EE likelihood, and the high-ℓ temperature and polarization PLIK likelihood described in ref. [137], as well as the lensing likelihood based on temperature+polarization map-based lensing reconstruction [138].

- For type Ia supernovae we use the likelihood of the Pantheon type Ia supernova sample [139], which includes data from the Pan-STARRS1 (PS1) Medium Deep Survey. Before the Pantheon release, our studies on nonlocal models used the JLA dataset [140].

- For BAO we use the likelihoods of the BAO detection of the 6dF Galaxy Survey [141] and the BAO scale measurement of SDSS DR7 Main Galaxy Sample [142]. We use updated SDSS data for the power spectrum of BAO from Data Release 12 [143].

The MCMC code used for obtaining the results for the RT model showed in the following section is Cobaya 2.0.2, because it already implemented the Planck 2018 likelihood. Before the last Planck release, i.e. when sticking to the Planck 2015 likelihood, we used the MCMC code MontePython 2.2.2.

For the RT model the initial conditions of the perturbations of the auxiliary fields δU and δS_μ are set to zero. As we have explicitly verified for the RR model in [30], taking different initial conditions, of the order of the metric perturbations (which is their natural scale, since, as discussed already in Section 1.3.3, the initial conditions on the auxiliary fields are in principle fixed by the initial conditions on the metric perturbations) has a totally negligible effect.

After having determined in this way the mean values of the parameters of the models (ΛCDM and RT, minimal and with various ΔN), we will use these values to compare the models with further datasets, namely measurements of $H(z)$ ("cosmic chronometers") and $f\sigma_8$ data.

Comparison with CMB, BAO, SNe, $H(z)$ and $f\sigma_8$.

Fit to CMB+BAO+SNe and Bayesian parameter estimation. Table 2.2 shows the results for the Bayesian parameter estimation and the resulting χ^2 for $\nu\Lambda$CDM and the RT model (minimal, and with $\Delta N = 34, 50, 64$), using the combined CMB+SNe+BAO data. Beside the values of the seven fundamental independent parameters given in (2.1.50), we also give some useful derived parameters, namely Ω_M, the reionization redshift $z_{\rm re}$, and the amplitude of matter density fluctuations in spheres of radius $8h^{-1}$ Mpc, σ_8. In the last line we show the differences in χ^2, with respect to the value for $\nu\Lambda$CDM. We recall that, for models with the same number of free parameters, as $\nu\Lambda$CDM and the RT models, the conventional interpretation (based on [144]) is that a difference $|\Delta\chi^2| \leqslant 2$ implies statistical equivalence between the two models, while $2 \lesssim |\Delta\chi^2| \lesssim 6$ suggests weak evidence in favor of the model with lower χ^2, and

Parameter	$\nu\Lambda$CDM	RT, minimal	RT, $\Delta N = 34$	RT, $\Delta N = 50$	RT, $\Delta N = 64$
H_0	67.89 ± 0.47	$68.74^{+0.59}_{-0.51}$	67.95 ± 0.48	67.90 ± 0.47	67.88 ± 0.48
$\sum_\nu m_\nu$ [eV]	< 0.057 (at 1σ)	$0.071^{+0.024}_{-0.066}$	< 0.048 (at 1σ)	< 0.044 (at 1σ)	< 0.041 (at 1σ)
ω_c	0.1193 ± 0.0009	0.1120 ± 0.0009	0.1191 ± 0.0009	0.1190 ± 0.0009	0.1189 ± 0.0009
$100\omega_b$	2.242 ± 0.013	2.237 ± 0.013	2.243 ± 0.013	2.244 ± 0.013	2.244 ± 0.013
$\ln(10^{10} A_s)$	3.045 ± 0.014	3.043 ± 0.014	3.047 ± 0.014	$3.048^{+0.013}_{-0.015}$	3.049 ± 0.014
n_s	0.9665 ± 0.0036	0.9649 ± 0.0036	0.9670 ± 0.0036	0.9673 ± 0.0035	0.9672 ± 0.0035
$\tau_{\rm re}$	0.0555 ± 0.0072	0.0537 ± 0.0072	0.0565 ± 0.0073	$0.0572^{+0.0065}_{-0.0075}$	0.0575 ± 0.0071
Ω_M	0.3085 ± 0.0060	$0.3029^{+0.0061}_{-0.0070}$	0.3075 ± 0.0061	0.3076 ± 0.0060	0.3076 ± 0.0060
$z_{\rm re}$	7.76 ± 0.72	7.60 ± 0.73	7.86 ± 0.72	7.93 ± 0.70	7.96 ± 0.70
σ_8	$0.8164^{+0.0097}_{-0.0068}$	$0.823^{+0.0130}_{-0.0087}$	$0.8141^{+0.0089}_{-0.0067}$	$0.8134^{+0.0088}_{-0.0064}$	$0.8129^{+0.0084}_{-0.0066}$
$\Delta\chi^2$	0	1.30	-0.48	-0.20	-0.00

TABLE 2.2: Mean values and 1σ errors of the parameters for $\nu\Lambda$CDM and the RT model (minimal, and with ΔN= 34, 50, 64), using CMB, BAO and SNe. H_0 is in units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The last line gives the difference in the χ^2 of each given model with respect to $\nu\Lambda$CDM. Table from [31].

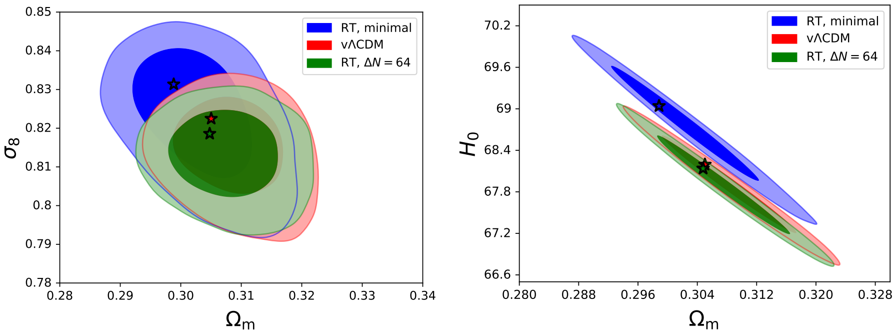

FIGURE 2.11: Left panel: the two-dimensional likelihood in the (Ω_M, σ_8) plane for $\nu\Lambda$CDM (red), the minimal RT model (blue) and the RT model with $\Delta N = 64$ (green). The stars are the best-fit values of the parameters (note that the values reported in Table 2.2 are rather the mean values). Right panel: the same for (Ω_M, H_0). Figure from [31].

$|\Delta\chi^2| \gtrsim 6$ indicates strong evidence in favor of the model with lower χ^2. Thus, all models considered fit the data at a statistically equivalent level.

The result of Bayesian parameter estimation shows that all models with large ΔN give predictions extremely close to those of $\nu\Lambda$CDM, consistently with the analysis of the previous sections, that showed that these models are very close to ΛCDM both in the background evolution and in the cosmological perturbations. The minimal RT model differs a bit more, and in particular predicts a slightly higher value of H_0, which in any case is not enough to significantly relieve the tension with the local H_0 measurement [145, 146]. Indeed, as discussed in [147, 148], it might not be possible to solve the H_0 tension, together with other potential tensions within ΛCDM, with a modification of only the late-Universe dynamics (as in our nonlocal model). The other difference of the minimal RT model is that it predicts a non-zero value for the sum of the neutrino masses, while all other models considered only give an upper bound.

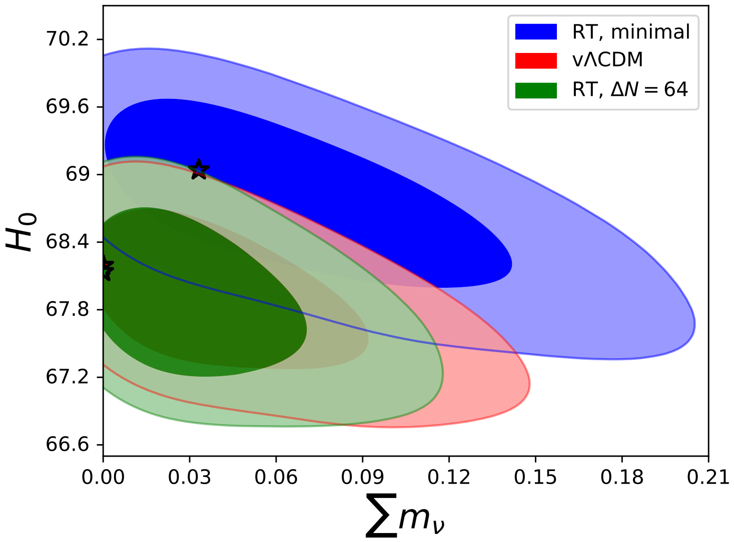

FIGURE 2.12: As in Fig. 2.11, for $(\sum_\nu m_\nu, H_0)$. Figure from [31].

	$\nu\Lambda$CDM	RT, minimal	$\Delta N = 34$	$\Delta N = 50$	$\Delta N = 64$
$\Delta\chi^2$	0	-1.13	0.22	0.42	0.57

TABLE 2.3: Values of $\Delta\chi^2$, with respect to $\nu\Lambda$CDM, for the RT model, minimal and with various values of ΔN, from the fit to a compilation of measurements of $H(z)$. Table from [31].

Figures 2.11 and 2.12 show the two-dimensional likelihoods for (Ω_M, σ_8), (Ω_M, H_0) and $(\sum_\nu m_\nu, H_0)$. The pattern that emerges, from this and similar plots, is that the RT model with large values of ΔN is extremely close to ΛCDM, as we already saw from Table 2.2, while the minimal RT model has some more significant differences, such as a slightly higher value of H_0 (although, as mentioned above, not enough to significantly decrease the tension with local measurements), of σ_8, and of the sum of neutrino masses.

Cosmic chronometers. Another useful observational test is provided by measurements of $H(z)$ at different redshifts ("cosmic chronometers" [149]). We use a compilation of 36 measurements of $H(z)$ between $z = 0.07$ and $z = 2.34$, given in Table I of [150]. Using the respective prediction for $H(z)$ in ΛCDM and in the RT models (with the respective mean values of Ω_M and H_0 from Table 2.2, obtained from the MCMC comparison to CMB+BAO+SNe) to fit these $H(z)$ measurements, we find the difference in χ^2, with respect to $\nu\Lambda$CDM, shown in Table 2.3. The corresponding reduced χ^2, all of order $0.63 - 0.64$, show that, by themselves, all the models fits these data well.

Structure formation and $f\sigma_8$ data. The properties of the models with respect to structure formation are already partly tested by the inclusion of BAO in our MCMC analysis. We further show the comparison of the models with a set of measurements of $f\sigma_8$, using the datapoints that we used in [30, 31]. As said in [31], many more measurement exists, like those collected in [151]. However, many of these datapoints are correlated, due to overlap in the galaxy samples used, and no covariance matrix is available for the full dataset, nor for most of its subsets. Furthermore, different datapoints have been obtained with different fiducial cosmologies, and survey systematics may vary with time of publication and lead to inhomogeneities in the data.

	$\nu\Lambda$CDM	RT, minimal	$\Delta N = 34$	$\Delta N = 50$	$\Delta N = 64$
$\Delta\chi^2$	0	1.41	-0.05	-0.18	-0.28

TABLE 2.4: Values of $\Delta\chi^2$, with respect to $\nu\Lambda$CDM, for the RT model, minimal and with various values of ΔN, from the fit to a compilation of measurements of $f\sigma_8$. Table from [31].

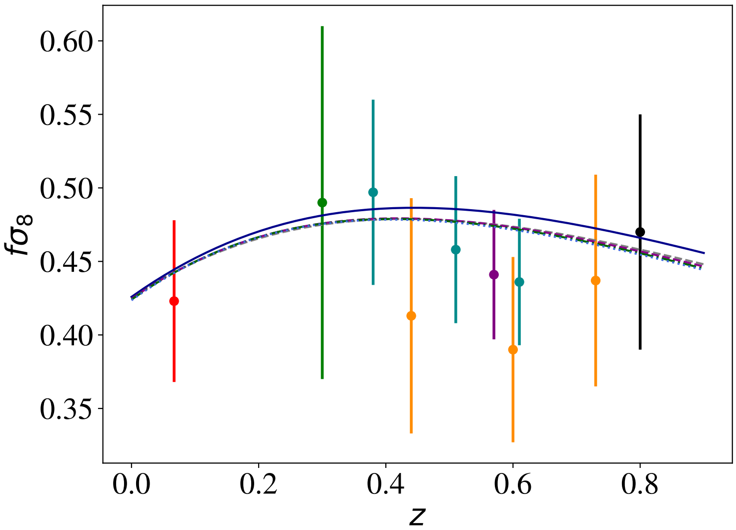

FIGURE 2.13: A collection of measurements of $f\sigma_8$ and the corresponding predictions of $\nu\Lambda$CDM and of the RT model, minimal and with $\Delta N = 34, 50$ and 64. The curve for the minimal RT model is the upper one, while all others are almost indistinguishable on this scale. The data points are from 6dF GRS [152] (red), SDSS LRG [153] (green), BOSS CMASS [154] (purple), WiggleZ [155] (orange), VIPERS [156] (black) and BOSS DR12 [143] (cyan). Figure from [31].

Therefore the use of the full dataset, without the appropriate covariance matrix and corrections, would lead to results of dubious interpretation.

Fig. 2.13 shows the data and the predictions of ΛCDM and of the RT model, minimal and with $\Delta N = 34, \Delta N = 50, \Delta N = 64$, obtained using for each model the respective mean values of Ω_M and H_0 from Table 2.2. The corresponding differences of χ^2, with respect to the value in $\nu\Lambda$CDM, are given in Table 2.4. We see that, once again, the differences between ΛCDM and the RT model with various ΔN are not statistically significant. From the plots of $G_{\rm eff}$ in Fig. 2.5 we see that at low k (upper left panel) the minimal RT model predicts $G_{\rm eff}/G > 1$, while the RT models with large ΔN predict $G_{\rm eff}/G < 1$. The data favor a weakening of gravity at these scales, so the RT models with large ΔN are slightly preferred with respect to ΛCDM, and the minimal RT model is slightly disfavored, but in all cases at a statistically insignificant level.

Finally, Fig. 2.14 shows the relative difference in the linear power spectrum of the RT models with respect to ΛCDM (each one computed using their respective mean values of the cosmological parameters) as a function of k, for $z = 0$ (left panel), and as as a function of z, for the mode with $k = 0.1/\text{Mpc}$ (right panel).

The conclusion of this analysis is that, on the one hand, the RT model, for all values of ΔN, is very close to ΛCDM at the level of background evolution and scalar perturbations, and fits the observations at the same level as ΛCDM. On the other hand, the deviations, which for both the background and scalar perturbations are

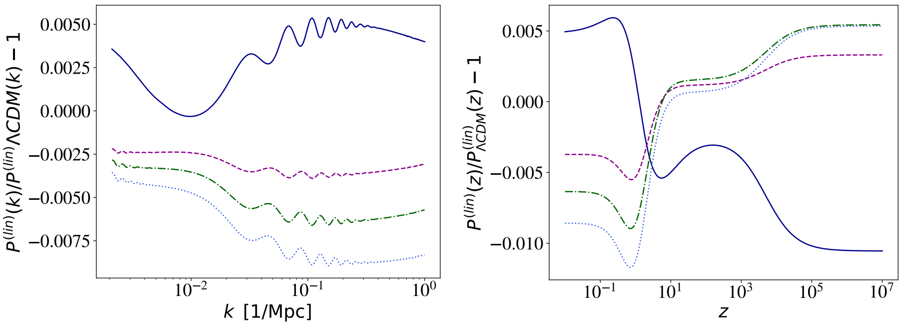

FIGURE 2.14: Relative difference of total linear matter power spectrum, with respect to best-fit ΛCDM, for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted). Left panel: as a function of k, at $z = 0$. Right panel: as a function of redshift, for the mode with $k = 0.1/\text{Mpc}$. Figure from [31].

typically at the percent or sub-percent level, could in principle be within reach for future missions. For instance, assuming that the function $\mu(a, k)$ that characterizes deviations from the Poisson law is scale independent and parametrizing its dependence on the scale factor as $\mu(a) = \mu_s a^s$, a future survey such as EUCLID [157], for fixed cosmological parameters, is expected to measure μ_s with an error $\Delta\mu_s = 0.0046$ for $s = 1$ and $\Delta\mu_s = 0.014$ for $s = 3$ [158]. The RT model has also been selected by the Dark Energy Science Collaboration (DESC) of the Large Synoptic Survey Telescope (LSST), among a few modified gravity models, for further studies and development of dedicated pipelines [159].

2.1.4 Solar System constraints

No vDVZ discontinuity

Any cosmological model that modifies GR on cosmological scales must also be able to reproduce the successes of GR at much smaller scales, such as the solar system and laboratory scales. In theories that introduce extra fields, as in scalar-tensor theories, or extra polarization of the gravitons, as in massive gravity, this is highly non-trivial. The linearized theory does not reduce to GR, and screening mechanisms involving the non-linearities of the theory are needed. As discussed in [86], for the RT model, in contrast, the situation is much simpler: already at the linear level the theory reduces smoothly to GR, and there is no discontinuity such as the vDVZ discontinuity of massive gravity.

Let us consider first the GR limit for the RT model linearized over flat space. In this case eq. (1.3.44) reduces to eq. (1.3.38). In order to compute the matter-matter interaction induced by this coupling of $T_{\mu\nu}$ with $h_{\mu\nu}$ we can proceed as in [77]. We use the gauge invariance of the linearized theory to fix the De Donder gauge $\partial^\mu[h_{\mu\nu} - (1/2)h\eta_{\mu\nu}] = 0$ and then, going in momentum space, the resulting equation can be solved for the Fourier transform $\tilde{h}_{\mu\nu}(k)$, to obtain

$$\tilde{h}_{\mu\nu}(k) = \frac{16\pi G}{k^2}\left[\tilde{T}_{\mu\nu}(k) - \frac{\eta_{\mu\nu}k^2}{2(k^2 - m^2)}\tilde{T}(k) + \frac{m^2}{3(k^2 - m^2)}\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right)\tilde{T}(k)\right], \tag{2.1.51}$$

where $T=\eta^{\mu\nu}T_{\mu\nu}$ is the trace of the energy-momentum tensor. Plugging this result into the linearized interaction term

$$S_{\rm int}=\frac{1}{2}\int d^4x\, h_{\mu\nu}T^{\mu\nu}\,, \tag{2.1.52}$$

and using the conservation equation $k^\mu\tilde{T}_{\mu\nu}(k)=0$, we get

$$S_{\rm int}=8\pi G\int\frac{d^4k}{(2\pi)^4}\,\tilde{T}_{\mu\nu}(-k)\Delta^{\mu\nu\rho\sigma}(k)\tilde{T}_{\rho\sigma}(k)\,, \tag{2.1.53}$$

where

$$\Delta^{\mu\nu\rho\sigma}(k)=\frac{1}{2k^2}\left(\eta^{\mu\rho}\eta^{\nu\sigma}+\eta^{\mu\sigma}\eta^{\nu\rho}-\eta^{\mu\nu}\eta^{\rho\sigma}\right)+\frac{1}{6}\left[\frac{1}{k^2}-\frac{1}{k^2-m^2}\right]\eta^{\mu\nu}\eta^{\rho\sigma}\,. \tag{2.1.54}$$

The first addend is the usual GR result due to the exchange of the helicities ±2 of a massless graviton. The second addend in brackets vanishes for $m\to 0$. Therefore the RT model has no vDVZ discontinuity, and reduces smoothly to GR as $m\to 0$. In the regime where a linearization over flat space is adequate, for modes with $|k^2|\gg m^2$ the predictions of the RT model differ from the predictions of GR by a factor $1+\mathcal{O}(m^2/k^2)$. We have seen that the comparison with cosmological observations fixes $m\sim H_0$ (or even smaller for large ΔN). For $|k|=(1\,\text{a.u.})^{-1}$ (as appropriate to solar system experiments), $m^2/k^2\sim(1\,\text{a.u.}/H_0^{-1})^2\sim 10^{-30}$, and the predictions of the RT model are indistinguishable from that of GR.

The absence of vDVZ discontinuity can also be understood observing that the second addend in brackets in eq. (2.1.54) induces a matter-matter interaction

$$8\pi G\int\frac{d^4k}{(2\pi)^4}\,\frac{1}{6}\tilde{T}(-k)\left[\frac{1}{k^2}-\frac{1}{k^2-m^2}\right]\tilde{T}(k)\,. \tag{2.1.55}$$

Comparing with eq. (1.3.33) one realizes that the two terms in brackets correspond to the exchange of the helicity zero component of $h_{\mu\nu}^{\rm TT}$ and of the trace mode s. In GR, where s is massless and both $h_{\mu\nu}^{\rm TT}$ and s appear with a $\Box$ factor in the quadratic lagrangian [see eq. (1.3.24)] these two terms cancel exactly, while here the cancellation is only partial but is recovered for $m\to 0$. Notice that both the helicity zero component of $h_{\mu\nu}^{\rm TT}$ and s are non-propagating degrees of freedom in GR and remain non-propagating in the RT model. We have indeed seen in eqs. (1.3.52)–(1.3.54) that, in the RR or RT models linearized over flat space, the only radiative degree of freedom of the metric are still given by the helicity ±2 modes described by $H_{ij}^{\rm TT}$ and s remains non-radiative, see eq. (1.3.56). Exactly as in GR, the negative sign in front of the $1/(k^2-m^2)$ term in eq. (2.1.55), which would correspond to a ghost if it were due to a propagating particle, is therefore innocuous from the point of view of quantum vacuum stability. We recall that the helicity zero components of $h_{\mu\nu}^{\rm TT}$ and s are not associated to creation/annihilation operators and cannot appear on the external lines of a Feynman diagram.

Schwarzschild solution

After having checked the recovery of the GR limit in Minkowski space, let us consider the GR limit for the Schwarzschild solution, by studying the static spherically symmetric solution of the RT model. A typical issue of massive gravity theories is that they become non-linear when r is smaller than a distance, the Vainshtein radius, which is parametrically larger than the Schwarzschild radius r_S of the source; e.g. $r_V = (GM/m^4)^{1/5}$ in the theory defined by adding a Fierz-Pauli mass term to the Einstein-Hilbert action [160, 161], and $r_V = (GM/m^2)^{1/3}$ [162] in the dRGT theory [121, 21]. For $m = \mathcal{O}(H_0)$ and $M = M_\odot$, we have $(GM/m^2)^{1/3} \sim 100$ pc. Since linearized theory only holds for $r > r_V$, in massive gravity in the whole range of distances probed by solar system and laboratory experiments the linearized expansion is not valid, and one must show that a Vainshtein mechanism is at work, i.e. that the inclusion of classical non-linearities restore the continuity with GR at $r \ll r_V$. Explicit examples of this type have indeed been found for the dRGT theory [163, 164].

For the RT model, however, the situation is much simpler, and the limit $m \to 0$ of the Schwarzschild solution is smooth. The Schwarzschild solution in the RT model has been worked out in [86]. In the limit $r \gg r_S$, the result for the metric is

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\,, \tag{2.1.56}$$

where

$$A(r) = 1 - \frac{r_S}{r}\left[1 + \frac{1}{3}(1 - \cos mr)\right], \tag{2.1.57}$$

$$B(r) = 1 + \frac{r_S}{r}\left[1 - \frac{1}{3}(1 - \cos mr) + \frac{1}{3}mr\sin mr\right], \tag{2.1.58}$$

In the limit $mr \ll 1$ (but still $r \gg r_S$), eqs. (2.1.57) and (2.1.58) give

$$A(r) \simeq 1 - \frac{r_S}{r}\left(1 + \frac{m^2r^2}{6}\right), \tag{2.1.59}$$

and (to first order in r_S/r) $B(r) = 1/A(r)$. We also mention that the solution for the auxiliary field $U = -\Box^{-1}R$ is simply given by $U(r) = (r_S/r)\cos mr$ (valid for both the RT [86] and RR [165] models). The auxiliary field $S_\mu(x)$ of the RT model, in spherical coordinates, has only one non-vanishing component, which is $S_r(r)$; when solving the equations it is convenient to define $V(r)$ from $S_r(r) = B^{1/2}(r)rV(r)$ and, for $r \gg r_S$, $V(r)$ reduces to $V(r) \simeq -r_S/(2r)$ (see [86]).

For comparison, in massive gravity the analogous computation gives [160, 22]

$$A(r) = 1 - \frac{4}{3}\frac{r_S}{r}\left(1 - \frac{r_S}{12m^4r^5}\right). \tag{2.1.60}$$

The factor $4/3$ in front of r_S/r is due to the extra contribution coming from the exchange of the helicity-0 graviton, and gives rise to the vDVZ discontiuity. In contrast, no vDVZ discontinuity is present in eq. (2.1.59). Furthermore, in eq. (2.1.60) the correction blows up as r decreases, and for $r \sim r_V = (GM/m^4)^{1/5}$ it becomes of the order of the leading term, signaling the breakdown of the linearized approximation. In eq. (2.1.59), in contrast, the correction becomes smaller and smaller as r decreases, and perturbation theory is valid at all scales $r \ll m^{-1}$, until we arrive at $r \simeq r_S$, where eventually also GR becomes non-linear.

In conclusion, in the RT model, in static situations GR is smoothly recovered, with correction $O(m^2r^2)$. The same conclusion holds in the RR model, for which we performed a similar analysis in [165]. Given that m is of order H_0, these corrections are utterly negligible for all r of order of solar system scale or smaller; e.g. $m^2r^2 \sim 10^{-30}$ for r of order of the Earth-Sun distance. Even on galactic scales these corrections to GR are totally irrelevant, with $m^2r^2 \sim 10^{-17}$ for $r = 10$ kpc.

Lunar Laser Ranging test

The above results show that, in a static situation, the RT and RR models recover all successes of GR at short scales. As was pointed out in [166] and then accurately studied in [165], this is not yet sufficient to guarantee that these models are viable at solar system scales. Another crucial test comes from the limit on the time variation of Newton's constant from Lunar Laser Ranging (LLR). The current observational result is $\dot{G}/G = (7.1 \pm 7.6) \times 10^{-14}\,\mathrm{yr}^{-1}$ [167]. This measurement is so accurate that, even if performed at the Earth-Moon scale over the last few decades, it provides significant constraints on cosmological models. Indeed, if we rewrite this limit in terms of the Hubble parameter today, using $H_0 \simeq h_0 \times (9.777752\,\mathrm{Gyr})^{-1}$, we get

$$\frac{\dot{G}}{G} = (0.99 \pm 1.06) \times 10^{-3} \left(\frac{0.7}{h_0}\right) H_0 \,. \tag{2.1.61}$$

Quite generally, in modified gravity models Newton's constant becomes time dependent on cosmological scale. The scale for the time variation today is given by H_0, so on cosmological scales one typically finds $\dot{G}/G \simeq H_0$. If, in a given modified gravity model, this result holds also down to the scale of the solar system and of the Earth-Moon system, then the bound (2.1.61) is violated and the model is ruled out.

In the case of the RT model, however, we have seen in eq. (2.1.47) that G_{eff} reduces to G at small scales. Therefore, it has no time dependence and the RT model satisfies trivially the LLR limit. The situation is different for the RR model and for other modified gravity models like Deser-Woodard model. Indeed, in the RR model, for sub-horizon modes, one finds [165, 111, 166]

$$\frac{G_{\mathrm{eff}}(t)}{G} = \left[1 - \frac{1}{3}m^2\bar{S}(t)\right]^{-1} \left[1 + \mathcal{O}\left(\frac{1}{\hat{k}^2}\right)\right] , \tag{2.1.62}$$

where $\bar{S}(t)$ is the background cosmological solution for the auxiliary field S. This dependence on S can be traced to the term $2SG_{\mu\nu}$ in $K_{\mu\nu}$, see eq. (1.3.47). If one plugs here the solution for $\bar{S}(t)$ corresponding to the FRW background, one finds that $G_{\mathrm{eff}}(t)/G$ is of order H_0, and the bound (2.1.61) is violated. In this case one cannot appeal to non-linear screening mechanisms, since we have seen that the RR model (just as the RT model) has a smooth limit $m \to 0$, so the linearized expansion can be trusted.

Of course, the FRW metric has no direct relevance for the Earth-Moon system. The latter, just as the solar system, does not expand with the Hubble flow. However, the point is that a scalar field, such as S, that evolves on a background that interpolates between the Schwarzschild solution at short scales and the FRW solution at large distances, in general inherits a time dependence on small scales from the matching with the solution at large distances. As an extreme example, in GR one can consider the Einstein-Straus space-time, in which, inside a sphere of radius r_0, the metric is taken to be exactly the static Schwarzschild metric generated by the mass M, while in the exterior it is given by a FRW solution with energy density ρ (see e.g. [168, 169] for review). The two metrics are then matched by requiring that the induced metric on the boundary surface Σ agrees on the two sides. This fixes the matching radius r_0, that, with respect to the Schwarzschild coordinates of the interior, turns out to be given by $M = (4/3)\pi r_0^3\rho$, where ρ is the energy density in FRW. In this case the solution for the metric is exactly static in the interior region, so it describes a limiting case in which the cosmological expansion in the inner region is perfectly screened. Nevertheless, if one studies the propagation of a scalar field obeying the equation $\Box\phi = 0$ in this metric, one finds that the solution for the field in the inner region is time dependent [169]. This is due to the fact that we must impose a matching condition for the scalar field at the surface Σ, and in this way the field inherits a time dependence even in the inner region.

For the RR model, a detailed analysis of the solution for the scalar field S in a background that interpolates between the static solution at short distances and FRW at large distances has been performed in [165]. A useful way of studying the problem is to follow the time evolution of the auxiliary fields U and S of the RR model, starting before the epoch of structure formation. At that time the FRW metric holds everywhere, and U and S evolve with time according to the cosmological background solutions $\bar{U}(t)$ and $\bar{S}(t)$. As structures form and become non-linear, the analysis of [165] shows that the solutions for U and S remain of the form

$$U(t,\mathbf{x}) = \bar{U}(t) + \delta U(t,\mathbf{x})\,, \qquad S(t,\mathbf{x}) = \bar{S}(t) + \delta S(t,\mathbf{x})\,, \tag{2.1.63}$$

where $\delta U(t,\mathbf{x})$ and $\delta S(t,\mathbf{x})$ remain small perturbations of $\bar{U}(t)$ and $\bar{S}(t)$, respectively. In essence, the physical reason behind this result is that, even when structures become non-linear, e.g. in the formation of galaxies, clusters, etc., the metric perturbation Φ never become large. In non-linear structure formation only the second spatial derivatives of Φ become large compared to their values in the linear regime (in particular the Laplacian of Φ, which is related to the density contrast, can become huge); however, the spatial derivatives of Φ never enter in the equations that govern the dynamics of the auxiliary fields U and S. Indeed, in a perturbed FRW metric, to first order in Φ, the explicit expression of the d'Alembertian is

$$\Box U = -(1+2\Phi)(\ddot{U} + 3H\dot{U}) - 4\dot{\Phi}\dot{U} + a^{-2}(1-2\Phi)\nabla^2 U\,, \tag{2.1.64}$$

so spatial derivatives of Φ do not appear. As a result, non-linear structure formation does not stop the time evolution that the auxiliary fields inherited from the earlier epoch described by a spatially homogeneous FRW solutions. Near massive bodies, the perturbations $\delta U(t,\mathbf{x})$, $\delta S(t,\mathbf{x})$ just reduce to the static solutions $U(r)$, $S(r)$ studied in the previous subsection, and remain small as long as r is larger than the Schwarzschild radius of the massive bodies (recall for instance that $U(r) = (r_S/r)\cos mr$, which is much smaller than one for $r \gg r_S$). So, in the end, at the Earth-Moon system scale, the solution for S is, with good approximation, the sum of the cosmological and static solutions, $S(t,\mathbf{x}) = \bar{S}(t) + S_{\text{static}}(r)$. A study of purely static solutions misses the term $\bar{S}(t)$, because assumes from scratch that the solution is time-independent. This time dependence induces a time-dependence of the Newton's constant, such that the RR model violates the limit (2.1.61) and is therefore ruled out by Lunar Laser Ranging.

In contrast, as we have seen in eq. (2.1.47), in the RT model the effective Newton's constant reduces to G at small scales and looses all the dependence on the auxiliary fields. Thus the RT model passes without problems the LLR test.

Chapter 3

Modified gravitational-wave propagation

3.1 Gravitational waves in modified gravity

3.1.1 Modified gravitational-wave propagation

In the previous Chapter we have studied the cosmological consequences of nonlocal gravity (in particular of the successful RT model) at the level of background evolution and scalar perturbations. We now turn to tensor perturbations, i.e. gravitational waves (GWs) propagating in a FRW background. We will see, following [170, 171, 172], that the RT model has striking predictions in the tensor sector, that could be detected in the near future by GW detectors. However, before discussing the specific predictions of the RT model in Section 3.1.2, we want to develop some tools that are completely general for the propagation of GWs across cosmological distances in modified gravity theories and that will lead us to introduce the notion of GW luminosity distance. In this way we will be able to enlarge the scope of our analysis and discuss general constraints (current or future) on modifed gravity theories, obtained from GW detections.

As a first step, it is useful to review GW propagation over a FRW background in the case of GR.

GW propagation in GR

In GR the evolution of tensor perturbations over FRW is governed by the equation

$$\tilde{h}''_A + 2\mathcal{H}\tilde{h}'_A + k^2\tilde{h}_A = 16\pi G a^2 \tilde{\sigma}_A \,, \tag{3.1.1}$$

where $\tilde{h}_A(\eta, \mathbf{k})$ are the Fourier modes of the GW amplitude, and we use the index $A = +, \times$ to label the two polarizations. We are using now conformal time η, related as usual to cosmic time t by $dt = a(\eta)d\eta$, and $a(\eta)$ is the scale factor. In this section the prime denotes the derivative with respect to cosmic time η, and $\mathcal{H} = a'/a$. The source term $\tilde{\sigma}_A(\eta, \mathbf{k})$ is related to the helicity-2 part of the anisotropic stress tensor (see e.g. [74]). In the following we will be interested in the free propagation between source and observer, and we will set it to zero. To get rid of the Hubble friction term, it is convenient to introduce a field $\tilde{\chi}_A(\eta, \mathbf{k})$ from

$$\tilde{h}_A(\eta, \mathbf{k}) = \frac{1}{a(\eta)}\tilde{\chi}_A(\eta, \mathbf{k}) \,. \tag{3.1.2}$$

Then eq. (3.1.2) becomes

$$\tilde{\chi}_A'' + \left(k^2 - \frac{a''}{a}\right)\tilde{\chi}_A = 0\,. \tag{3.1.3}$$

For modes well inside the horizon, such as the GWs targeted by ground-based and space-born detectors, the term $a''/a \sim 1/\eta^2$ is totally negligible with respect to k^2; for instance, for a GW with a frequency $f \sim 10^2$ Hz, as typical of ground-based interferometers, $(k\eta)^{-2} \sim (500\,\mathrm{km}/H_0^{-1})^2 \sim 10^{-41}$. We can then neglect the term a''/a in eq. (3.1.3), which then becomes a standard a wave equation for $\tilde{\chi}_A$, that tells us that GWs propagate at the speed of light (that we have set here equal to unity). The factor $1/a$ in eq. (3.1.2) tells us how the GW amplitude decreases as it propagates across cosmological distances, from the source to the observer.

For a coalescing binary, the factor $1/a$ combines with the standard behavior $1/r$ in the near region, as well as with redshift-dependent factors that arise when transforming frequencies and masses from the source frame to the detector frame. Thus one finds the standard result $h_A \propto 1/d_L(z)$, where d_L is the luminosity distance to the source (see e.g. section 4.1.4 of [27] for a detailed derivation). This is the origin of the fact that coalescing binaries are "standard sirens" [173], i.e. they allow a direct reconstruction of the luminosity distance of the source.

More precisely, in the so-called restricted post-Newtonian (PN) approximation, where one takes into account the PN corrections to the phase but not to the amplitude, one finds

$$h_+(t) = \frac{2(1+\cos^2\iota)}{d_L(z)}(G\mathcal{M}_c)^{5/3}[\pi f(t)]^{2/3}\cos\Phi(t)\,, \tag{3.1.4}$$

$$h_\times(t) = \frac{4\cos\iota}{d_L(z)}(G\mathcal{M}_c)^{5/3}[\pi f(t)]^{2/3}\sin\Phi(t)\,, \tag{3.1.5}$$

where $\Phi(t)$ is the phase, that in general needs to be computed to a high PN order, $\mathcal{M}_c = (1+z)(m_1 m_2)^{3/5}(m_1+m_2)^{-1/5}$ is the redshifted chirp mass (i.e. the quantity actually observed in the detector frame), $f(t)$ is the observed GW frequency, that sweeps upward in time, and ι is the inclination angle of the normal to the orbit with respect to the line of sight. The chirp mass is accurately determined from the time evolution of $f(t)$ [e.g., to lowest order in the PN expansion, $\dot{f} = (96/5)\pi^{8/3}(G\mathcal{M}_c)^{5/3}f^{11/3}$]. Then the amplitude of GWs from a coalescing compact binary provides an absolute measurement of its luminosity distance and in this sense coalescing compact binaries are the GW analogue of standard candles. As we see from eqs. (3.1.4) and (3.1.5), the main uncertainty on the standard siren measurement of $d_L(z)$ comes from the partial degeneracy with $\cos\iota$. This can be broken in particular if both polarizations can be measured, or if we have informations on the inclination angle, e.g. from the observation of an electromagnetic jet. We also stress that, in GR, the luminosity distance $d_L(z)$ measured by standard sirens (with the methodology that we just discussed) is exactly the same as the standard luminosity distance associated to electromagnetic signals.

In GR, for a cosmological model with dark energy density $\rho_{\rm DE}(z)$, the relation between luminosity distance and redshift is

$$d_L(z) = \frac{1+z}{H_0}\int_0^z \frac{d\tilde{z}}{\sqrt{\Omega_M(1+\tilde{z})^3 + \Omega_R(1+\tilde{z})^4 + \rho_{\rm DE}(\tilde{z})/\rho_0}}\,. \tag{3.1.6}$$

Therefore, a simultaneous measurement of d_L and of the redshift z (with an electromagnetic counterpart, or the study of the $d_L - z$ relation with statistical methods) allows us to get cosmological information. In particular, for sources at small redshift, $z \ll 1$, eq. (3.1.6) reduces to the Hubble law $d_L(z) \simeq H_0^{-1} z$, so from a measurement at such redshifts we can get a measurement of H_0. This has indeed been possible with the detection of the binary neutron star (BNS) coalescence GW170817, which is at a redshift $z \simeq 0.01$, and has given the value $H_0 = 70.0^{+12.0}_{-8.0}$ km s^{-1} Mpc^{-1} [174]. The detection of coalescences at higher redshift could in principle allow us to access also the DE equation of state.

Much work as been devoted in the literature to the use of standard sirens for cosmology [175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189].

GW propagation in modified gravity

In many modified gravity theories, including the RT model, the free propagation of tensor perturbations over FRW is governed by an equation of the form

$$\tilde{h}''_A + 2\mathcal{H}[1-\delta(\eta)]\tilde{h}'_A + k^2\tilde{h}_A = 0\,, \tag{3.1.7}$$

with a model-dependent function $\delta(\eta)$. Indeed, in a generic modified gravity model both the “friction term" $2\mathcal{H}\tilde{h}'_A$ and the term $k^2\tilde{h}_A$ in eq. (3.1.7) are modified. As we will recall below, the models that modify the $k^2\tilde{h}_A$ term predict a speed of gravity different from the speed of light. The observation of GW170817 and of the associated GRB has set a limit $|c_{\rm gw} - c|/c < O(10^{-15})$ [190], so such models are ruled out.[1] In particular, a large class of Horndeski theories and other modifications of GR have been ruled out by this limit [192, 193, 194, 195]. It turns out that the models that survive this constraint still modify the friction term. A propagation equation of the form (3.1.7) was indeed first found in some scalar-tensor theories of Horndeski type [196, 197, 198, 199] and in the RR nonlocal model [30, 170]. In [200] it was shown that it also takes place in many other Horndeski-type theories that pass the test on speed of gravity (such as $f(R)$ theories, Jordan-Brans-Dicke, galileon cosmology, etc.), in Degenerate Higher Order Scalar-Tensor (DHOST) theories, and in bigravity. Similar effects take place in theories with extra dimensions, as originally found in [201] (see also [202]), although in this case they are due to the loss of gravitons to the bulk and, in general, are not described by eq. (3.1.7) (see also [203] for a discussion a modified GW propagation within the effective field theory approach to dark energy, [204, 205] for general formalisms for testing gravity with GW propagation, and [206, 207, 208] for further related work in the context of scalar-tensor theories).

Let us then study the general consequences of eq. (3.1.7) (we closely follow the discussion in [170, 171]). We proceed as in the GR case, except that now, to eliminate the friction term, we must introduce $\tilde{\chi}_A(\eta, \mathbf{k})$ from

$$\tilde{h}_A(\eta, \mathbf{k}) = \frac{1}{\tilde{a}(\eta)}\tilde{\chi}_A(\eta, \mathbf{k})\,, \tag{3.1.8}$$

where $\tilde{a}$ now satisfies

$$\frac{\tilde{a}'}{\tilde{a}} = \mathcal{H}[1-\delta(\eta)]\,. \tag{3.1.9}$$

Then we get

$$\tilde{\chi}''_A + \left(k^2 - \frac{\tilde{a}''}{\tilde{a}}\right)\tilde{\chi}_A = 0\,. \tag{3.1.10}$$

---

[1] Although it could still in principle happen that there is dependence on wavenumber that allows for $c_{\rm gw} \neq c$ for modes k well below the frequencies probed by LIGO/Virgo and restore $c_{\rm gw} = c$ to sufficient accuracy at LIGO/Virgo frequencies. This could be motivated in some models [191].

Once again, inside the horizon the term $\tilde{a}''/\tilde{a}$ is totally negligible. The remaining equation,

$$\tilde{\chi}_A'' + k^2 \tilde{\chi}_A = 0\,, \tag{3.1.11}$$

shows that GWs still propagate at the speed of light. This is a consequence of the fact that the term $k^2\tilde{\chi}_A$ in eq. (3.1.7) is the same as in GR. If the coefficient of this term had been different, we would have got a speed of GWs $c_{\rm gw} \neq c$.

As we see from eq. (3.1.9), the effect of the modified friction term is that now the amplitude of $\tilde{h}_A$ is proportional to $1/\tilde{a}$ rather than $1/a$. Then, in the propagation from the source to the observer, the amplitude is multiplied by a factor $\tilde{a}_{\rm emis}/\tilde{a}_{\rm obs} \equiv \tilde{a}(z)/\tilde{a}(0)$, instead of a factor $a_{\rm emis}/a_{\rm obs} = a(z)/a(0)$, where the labels refer to the emission time (at redshift z) and the observation time, at redshift zero, respectively. Therefore

$$\tilde{h}_A \propto \frac{\tilde{a}(z)}{\tilde{a}(0)}\,\frac{a(0)}{a(z)}\,\frac{1}{d_L(z)} = \frac{\tilde{a}(z)}{a(z)}\,\frac{1}{d_L(z)}\,, \tag{3.1.12}$$

where $d_L(z)$ is the usual notion of luminosity distance (note that, since only the ratios $\tilde{a}(z)/\tilde{a}(0)$ and $a(z)/a(0)$ enter, without loss of generality we can choose the normalizations $\tilde{a}(0) = a(0) = 1$). Eq. (3.1.12) motivates the introduction of a 'GW luminosity distance' $d_L^{\rm gw}(z)$ [170], related to the standard luminosity distance appropriate for electromagnetic signals, that we henceforth denote by $d_L^{\rm em}(z)$, by $d_L^{\rm gw}(z) = [a(z)/\tilde{a}(z)]\, d_L^{\rm em}(z)$. Rewriting eq. (3.1.9) as $(\log a/\tilde{a})' = \delta(\eta)\mathcal{H}(\eta)$ and integrating, we get [170]

$$d_L^{\rm gw}(z) = d_L^{\rm em}(z) \exp\left\{-\int_0^z \frac{dz'}{1+z'}\,\delta(z')\right\}\,. \tag{3.1.13}$$

In modified gravity, the quantity extracted from a measurement of the GW amplitude of a coalescing binary is $d_L^{\rm gw}(z)$, rather than $d_L^{\rm em}(z)$. To avoid misunderstandings, notice that the actual distance traveled by GWs from the source to the observer is the same as the distance traveled by electromagnetic signals. Eq. (3.1.13) is simply a convenient way of expressing the fact that, in modified gravity, the amplitude of the GW decreases in a different way during the propagation, so that, for a coalescing binary, the observed amplitude, rather than depending only on $d_L^{\rm em}(z)$ and on the inclination of the orbit, as in GR, it further depends on $\delta(z)$, in such a way that the combined dependence on $d_L^{\rm em}(z)$ and $\delta(z)$ can be reabsorbed into the quantity $d_L^{\rm gw}(z)$.

Conservation of graviton number

The fact that, in modified gravity, the GW amplitude over FRW does not scale as $1/a$ raises a question. As we will recall below, in GR the fact that in FRW $h \propto 1/a$ ensures that the GW energy density $\rho_{\rm GW}$ scales as $1/a^4$; in turn, this is consistent with an interpretation of a GW as a collection of massless graviton, whose comoving number density (i.e. number per unit volume in comoving coordinates) is conserved. Indeed, the fact that the graviton number per comoving volume is conserved means that the graviton number per physical volume scales as $1/a^3$, while the fact that the graviton is massless implies that its energy scales as $1/a$, giving overall the $1/a^4$ behavior of $\rho_{\rm GW}$. One might then wonder whether the scaling $h \propto 1/\tilde{a}$ is an indication that the graviton number is not conserved in modified gravity theories. We will see here that, in fact, in all the theories where GW propagation obeys eq. (3.1.7), the energy density $\rho_{\rm GW}$ scales as $1/a^4$ and therefore the comoving number density of gravitons is still conserved. To this purpose, one must realize that the expression of $\rho_{\rm GW}$ in a generic modified gravity model is different from the GR expression. Let us first recall how things work in GR. We consider tensor perturbations over the FRW metric. Using conformal time, we write

$$ds^2 = a^2 \left[-d\eta^2 + \left(\delta_{ij} + h_{ij}^{\rm TT}\right) dx^i dx^j\right] . \tag{3.1.14}$$

It is convenient to expand the Fourier transform of $h_{ij}^{\rm TT}$ in the basis of the polarization tensors,

$$\tilde{h}_{ij}^{\rm TT}(\eta, \mathbf{k}) = \sum_{A=+,\times} e_{ij}^A(\hat{\mathbf{k}})\tilde{h}_A(\eta, \mathbf{k}) , \tag{3.1.15}$$

where the polarization tensors are normalized as $e_{ij}^A(\hat{\mathbf{k}})e_{ij}^{A'}(\hat{\mathbf{k}}) = 2\delta^{AA'}$. Expanding the Einstein-Hilbert action to second order in $h_{ij}^{\rm TT}$ one then finds (see e.g. sect. 21.3.4 of [74])

$$\begin{aligned} S_2[h] &= \frac{1}{32\pi G} \sum_A \int d^3x d\eta \, a^2 \left[\partial_\eta h_A \partial_\eta h_A - \partial_k h_A \partial_k h_A\right] \\ &= -\frac{1}{2} \sum_A \int d^4x \sqrt{-\bar{g}}\, \bar{g}^{\mu\nu} \partial_\mu \varphi_A \partial_\nu \varphi_A , \end{aligned} \tag{3.1.16}$$

where $\bar{g}_{\mu\nu} = a^2 \eta_{\mu\nu}$ is the background FRW metric in $(\eta, \mathbf{x})$ coordinates, and

$$\varphi_A(\eta, \mathbf{x}) = \frac{1}{\sqrt{16\pi G}} h_A(\eta, \mathbf{x}) . \tag{3.1.17}$$

The action governing the two polarization amplitudes h_A is therefore the same as the curved-space action of two canonically-normalized scalar fields φ_A. The variation of the action (3.1.16) gives eq. (3.1.1) (with the left-hand side equal to zero, unless we add also the matter action). At the same time, from this action we can get the energy-momentum tensor of GWs,

$$\begin{aligned} t_{\mu\nu} &\equiv -\frac{2}{\sqrt{-\bar{g}}} \langle \frac{\delta S_2[h]}{\delta \bar{g}^{\mu\nu}} \rangle = \sum_A \langle \partial_\mu \varphi_A \partial_\nu \varphi_A - g_{\mu\nu} \frac{1}{2} g^{\rho\sigma} \partial_\rho \varphi_A \partial_\sigma \varphi_A \rangle \\ &= \frac{1}{16\pi G} \sum_A \langle \partial_\mu h_A \partial_\nu h_A - g_{\mu\nu} \frac{1}{2} g^{\rho\sigma} \partial_\rho h_A \partial_\sigma h_A \rangle , \end{aligned} \tag{3.1.18}$$

where $\langle \ldots \rangle$ denotes the spatial average over several wavelengths of the GWs, or the temporal average over several periods (see e.g. sect. 1.4 of [27]). We denote by $t_{\eta\eta}$ the $\mu = \nu = 0$ component of $t_{\mu\nu}$ in coordinates $(\eta, \mathbf{x})$ and by $t_{tt} \equiv t_{00}$ the $\mu = \nu = 0$ component of $t_{\mu\nu}$ in coordinates $(t, \mathbf{x})$. From $t_{\eta\eta}(d\eta)^2 = t_{tt}(dt)^2$ and $dt = a d\eta$ if follows that $t_{00} = t_{\eta\eta}/a^2$, so eq. (3.1.18) gives

$$t_{00} = \frac{1}{32\pi G} \frac{1}{a^2} \sum_A \langle (\partial_\eta h_A)^2 + (\partial_i h_A)^2 \rangle . \tag{3.1.19}$$

On a plane wave the terms $\langle (\partial_\eta h_A)^2 \rangle$ and $\langle (\partial_i h_A)^2 \rangle$ are equal. From eq. (3.1.11), for wavelengths well inside the horizon, i.e. for $k\eta \gg 1$, $\tilde{\chi}_A(\eta, \mathbf{k}) \propto \sin(k\eta + \alpha)$, with α a phase. Therefore $\tilde{h}_A(\eta, \mathbf{k}) \propto \sin(k\eta + \alpha)/a(\eta)$ and, again for $k\eta \gg 1$,

$$\partial_\eta h_A(\eta, \mathbf{k}) \propto \frac{k \cos(k\eta + \alpha)}{a(\eta)} \left[1 + O\left(\frac{1}{k\eta}\right)\right] . \tag{3.1.20}$$

In $\langle(\partial_\eta h_A)^2\rangle$ the term $\cos^2(k\eta+\alpha)$, averaged over several periods, simply gives a factor 1/2, so $\langle(\partial_\eta h_A)^2\rangle \propto 1/a^2$ and, from eq. (3.1.19), it then follows that $\rho_{\rm gw}=t_{00}$ is proportional to $1/a^4$, as indeed we expect for any form of radiation.

Let us now see how the situation changes in modified gravity. The propagation equation is now given by eq. (3.1.7). Using eq. (3.1.9) we see that it can be obtained from the GR equation with the replacement $a(\eta)\to\tilde{a}(\eta)$. It can then be formally obtained from the variation of the quadratic action resulting from the replacement $a(\eta)\to\tilde{a}(\eta)$ in eq. (3.1.16), i.e. from

$$S_2^{\rm RT}[h]=\frac{1}{32\pi G}\sum_A\int d^3x d\eta\, \tilde{a}^2\left[\partial_\eta h_A\partial_\eta h_A-\partial_k h_A\partial_k h_A\right]. \tag{3.1.21}$$

Introducing an effective Newton's constant from

$$\frac{1}{\tilde{G}(\eta)}\equiv\frac{1}{G}\frac{\tilde{a}^2(\eta)}{a^2(\eta)} \tag{3.1.22}$$

we can rewrite eq. (3.1.21) as

$$S_2^{\rm RT}[h]=\sum_A\int d^3x d\eta\,\frac{1}{32\pi\tilde{G}(\eta)}\,a^2\left[\partial_\eta h_A\partial_\eta h_A-\partial_k h_A\partial_k h_A\right]. \tag{3.1.23}$$

Thus, as far as tensor perturbations are concerned, at the quadratic level the modified gravity model can be obtained from GR with the replacement $G\to\tilde{G}(\eta)$. Note that $\tilde{G}(\eta)$ plays the role of an effective Newton's constant for tensor perturbations only and is, in general, different from the effective Newton's constant governing scalar perturbations. For example, we have seen in Section 2.1.2 that, in the RT model, scalar perturbations are governed by a different effective Newton's constant, that we denoted as $G_{\rm eff}(\eta,k)$, and which, contrary to $\tilde{G}(\eta)$, depends also on the wavenumber k.

Repeating the above derivation of the energy-momentum tensor of GWs, eq. (3.1.18) becomes

$$t_{\mu\nu}=\frac{1}{16\pi\tilde{G}(\eta)}\sum_A\langle\partial_\mu h_A\partial_\nu h_A-g_{\mu\nu}\frac{1}{2}g^{\rho\sigma}\partial_\rho h_A\partial_\sigma h_A\rangle\,, \tag{3.1.24}$$

simply because the variation $\delta S_2[h]/\delta\bar{g}^{\mu\nu}$ is insensitive to the time dependence of $\tilde{G}(\eta)$. The energy density $\rho_{\rm gw}=t_{00}$ is then given by

$$\rho_{\rm gw}=\frac{1}{16\pi\tilde{G}(\eta)}\frac{1}{a^2}\sum_A\langle(\partial_\eta h_A)^2\rangle\,. \tag{3.1.25}$$

Notice that the $1/a^2$ factor comes from the transformation from $t_{\eta\eta}$ to t_{tt}, i.e. from the relation $dt=ad\eta$. This is determined by the FRW background metric, so it still involves a rather than $\tilde{a}$. In contrast, $h_A\propto\sin(k\eta+\alpha)/\tilde{a}$ and therefore now, for $k\eta\gg1$, $\partial_\eta h_A\propto k\cos(k\eta+\alpha)/\tilde{a}$, which replaces eq. (3.1.20). Again, the term $\cos^2(k\eta+\alpha)$ averages to 1/2, so in the end the time dependence of $\rho_{\rm gw}$ is

$$\rho_{\rm gw}\propto\frac{1}{16\pi\tilde{G}(\eta)}\frac{1}{a^2\tilde{a}^2}=\frac{1}{16\pi G}\frac{1}{a^4}\,. \tag{3.1.26}$$

Thus, once taken into account the fact that the modification of the Einstein equations implies also a modification of the formula for the GW energy-momentum tensor, we find that, in FRW, the GW energy density of the modified gravity model still scales as $1/a^4$, despite modified GW propagation. Therefore, the energy density still corresponds to that of an ensemble of massless gravitons, whose number density in comoving coordinates is constant (so that the number density in physical coordinates scales as $1/a^3$) and whose energy scales as $1/a$. Notice also that the redshift of the graviton frequency $\omega \propto 1/a$, or of the wavelength as $\lambda \propto a$, are kinematical properties that depend only on the background FRW metric, and are the same in GR and in modified gravity. Since the RT model obeys an equation of the class in eq. (3.1.7) (and many other models do, as we already mentioned before), this result about graviton number conservation also gives useful guidance for attempts at deriving the RT model from a fundamental local theory. In particular, it rules out the possibility that the RT model could be derived from a theory with extra dimensions in which gravitons are lost to a higher-dimensional bulk, and rather points towards the dynamical mass generation mechanisms discussed in sect. 1.3.4.

In some models, like the RR model, not only $\rho_{\rm GW} \propto 1/a^4$, but furthermore the effective Newton constant $\tilde{G}$ for the tensor perturbations is the same as the effective Newton's constant $G_{\rm eff}$ in the scalar sector. In particular, it can be shown [171] that this relation holds in any modified gravity theory described by an action of the form

$$S = \frac{1}{8\pi G} \int d^4x \sqrt{-g}\, A(\phi) R + \dots\,, \tag{3.1.27}$$

which is minimally coupled to matter, where A can be a nontrivial functional of extra fields in the gravitational sector, here denoted collectively as ϕ, and the dots denote other possible gravitational interaction terms, that can depend on ϕ but do not contain terms purely quadratic in the gravitational field $h_{\mu\nu}$ nor interactions with ordinary matter.

It is important to observe that the time variation of the effective Newton constant $G_{\rm eff}(\eta, k)$ can only hold on cosmological scales. At such scales, typical cosmological models predict a time dependence such that, today, $|\dot{G}/G| \simeq H_0$, with a coefficient in general of order one. A scenario in which this result holds down to Solar System or Earth-Moon scales would be ruled out by Lunar Laser Ranging experiments, that by now impose a bound $|\dot{G}/G| \lesssim 10^{-3} H_0$ [167]. If a model predicts a non-trivial value for the ratio $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ and its "tensor" Newton constant is equal to the "scalar" Newton constant, then, to be observationally viable, it must have a screening mechanism at short scales, so that the time-dependent effective Newton constant $G_{\rm eff}(\eta, k)$ cannot be extrapolated down to the Earth-Moon scale (this is the case, for instance, for Hordenski, DHOST theories, and bigravity).

Phenomenological parametrization of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$

It is useful to introduce a simple parametrization for the modified GW propagation effect, in the same spirit of what is usually done for the background evolution and for the scalar sector of modified gravity theories. Of course, a parametrization is useful when it encompasses a large class of theories. We know that the deviation of the background evolution from ΛCDM is determined by the DE density $\rho_{\rm DE}(z)$ or, equivalently, by the DE equation of state $w_{\rm DE}(z)$. In principle one could try to reconstruct the whole function $w_{\rm DE}(z)$ from cosmological observations, but current results are unavoidably not very accurate (see e.g. fig. 5 of [209]). The standard approach is rather to use a parametrization for this function, that catches the qualitative features of a large class of models. The most common is the Chevallier–Polarski–Linder parametrization [108, 109], which makes use of two parameters (w_0, w_a),

$$w_{DE}(a) = w_0 + w_a(1-a)\,, \tag{3.1.28}$$

corresponding to the value and the slope of the function at the present time. In terms of redshift,

$$w_{\rm DE}(z) = w_0 + \frac{z}{1+z} w_a\,. \tag{3.1.29}$$

One can then analyze the cosmological data adding (w_0, w_a) to the standard set of cosmological parameters. Similarly, some standard parametrizations are used for describing the modification from GR in the scalar perturbation sector, in order to compare with structure formation and weak lensing, see e.g. [113, 112].

Here we are interested in tensor perturbations, where the effect is encoded in the non-trivial function $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$. Again, rather than trying to reconstruct this whole function from the data, it is more convenient to look for a simple parametrization that catches the main features of a large class of models in terms of a small number of parameters. We now introduce the 2-parameter parameterization that we proposed in [171],

$$\Xi(z) \equiv \frac{d_L^{\rm gw}(z)}{d_L^{\rm em}(z)} = \Xi_0 + \frac{1-\Xi_0}{(1+z)^n}\,, \tag{3.1.30}$$

which depends on the parameters Ξ_0 and n, both taken to be positive. In terms of the scale factor $a = 1/(1+z)$ corresponding to the redshift of the source,

$$\frac{d_L^{\rm gw}(a)}{d_L^{\rm em}(a)} = \Xi_0 + a^n(1-\Xi_0)\,. \tag{3.1.31}$$

The value $\Xi_0 = 1$ corresponds to GR. This parameterization is designed to smoothly interpolate between a unit value

$$\Xi(z \ll 1) = 1\,, \tag{3.1.32}$$

at small redshifts – where cumulative effects of modified gravity wave propagation have not sufficient time to accumulate differences with respect to GR, see eq. (3.1.13) – to a constant value Ξ_0

$$\Xi(z \gg 1) = \Xi_0\,, \tag{3.1.33}$$

at large redshift. The index n determines the rate at which this asymptotic value is reached. Indeed, in the large redshift regime we expect that the effects of modified gravity "turn-off" and $|\delta(z \gg 1)| \ll 1$, since modified gravity should mainly affect late-time evolution (also for ensuring compatibility with CMB observations), in which case the quantity $\Xi(z)$ approaches a constant. This (Ξ_0, n) parametrization was originally proposed in [171], inspired by the fact that it fits extremely well the prediction for $\Xi(z)$ obtained from nonlocal modifications of gravity (see also [30] for a review), but it was then realized that its features are very general, so that it is expected to fit the predictions from a large class of models. Indeed, it has been confirmed in [200] that this parametrization fits well scalar-tensor theories (both models in the Horndeski class, e.g. $f(R)$ and coupled Galileon models, and DHOST theories).

Given this (Ξ_0, n) parametrization, from eq. (3.1.13) we have

$$\begin{aligned}\delta(z) &= -\frac{d\ln\Xi(z)}{d\,\ln(1+z)} \\ &= \frac{n\,(1-\Xi_0)}{1-\Xi_0+\Xi_0(1+z)^n}\,. \end{aligned} \tag{3.1.34}$$

In this parametrization the quantity $\delta(z)$ indeed goes to zero at large redshifts, as desired since at early times gravity propagates as in GR. At late times, instead, $\delta(z \ll 1) = n\,(1-\Xi_0)$.

In general the parametrization (3.1.30) of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ is more robust than the corresponding parametrization (3.1.34) of $\delta(z)$. Indeed, even if $\delta(z)$ should have some non-trivial features as a function of redshift, such as a peak, still these features will be smoothed out by the integral in eq. (3.1.13). Since anyhow $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ must go to one as $z \to 0$ and we expect that in most models it will go asymptotically to a constant at large z, in general the fit (3.1.30) to $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ will work reasonably well even in cases where the corresponding fit (3.1.34) is not too good. This is the case for example in DHOST theories, see Section 3.1.2 of [200].

The main exception to the validity of the (Ξ_0, n) parametrization in eq. (3.1.30) is bigravity, where the effects of graviton oscillations lead to a peculiar behavior for the GW luminosity distance as a function of redshift, see Section 3.3 of [200].

To conclude this Section, let us emphasize the importance of modified GW propagation, as encoded for instance in the parameters (Ξ_0, n) defined by eq. (3.1.30), for studies of dark energy and modified gravity at advanced GW detectors. This stems from two important considerations:

1. We have seen that, in general, a modified gravity model induces deviations from ΛCDM at the background level (through the DE equation of state), in the scalar perturbation sector, and in the tensor perturbation sector. Concerning the background modifications, as encoded for instance in the (w_0, w_a) parameters, several studies [180, 181, 188, 171] have shown that the accuracy that LISA, or third-generation ground-based interferometers such as the Einstein Telescope, could reach on w_0 is not really better than the measurement, at the level of a few percent, that we already have from *Planck* in combination with other cosmological probes such as Baryon Acoustic Oscillations (BAO) and Supernovae (SNe). The results presented in the next Sections will indeed confirm these findings. In contrast, $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ is an observable that is only accessible thanks to GW observations.

2. On top of this, we realized in [170, 171] that, in a generic modified gravity theory, in which the deviation of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ from 1 is of the same order as the deviation of $w_{\rm DE}(z)$ from -1, the effect of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ on standard sirens dominates over the effect of $w_{\rm DE}(z)$. This can be understood from the explicit expression of the standard (“electromagnetic") luminosity distance given in eq. (3.1.6). From this expression one might think, naively, that if one changes the equation of state of DE by, say, 10%, this will induce a corresponding relative variation of d_L again of order 10%. However, this is not true because the cosmological parameters, such as H_0 and Ω_M, that enter in eq. (3.1.6), are not fixed, but must themselves be determined self-consistently within the model,

by comparing with cosmological observations and performing Bayesian parameter estimation, and their best-fit values change if we modify the DE content of the model. The fit to cosmological data basically amounts to requiring that the model reproduces some fixed distance indicators at large redshifts, such as the scale determined by the peaks of the CMB or that from the BAO. Thus, Bayesian parameter estimation has a compensating effect, changing the luminosity distance in a direction opposite to that induced by a change in $\rho_{\rm DE}(z)$, in such a way to keep as small as possible the variation of $d_L^{\rm em}(z)$ at large redshifts. Thus, after performing Bayesian parameter estimation, a relative change in $w_{\rm DE}$ by, say, 10%, would only induce a relative change of order, say, 1% in $d_L(z)$. In contrast, modified GW propagation is an extra effect, that is not compensated by degeneracies with other (fitted) cosmological parameters, and therefore dominates over the effect of $w_{\rm DE}$. We confirm this physical argument in the next Sections, where , where we found that, using standard sirens at LISA [200] or 3G detectors [210] (Einstein Telescope, Cosmic Explorer) in combination with other cosmological probes, Ξ_0 can be measured with a significantly better accuracy than w_0.

Therefore Ξ_0, or more generally modified GW propagation, is a prime observable for dark energy studies with advanced GW detectors.

3.1.2 GW propagation in the RT model

We now discuss GWs in the RT model, focusing on the signal from coalescing binaries at cosmological distances. First of all, notice that this model only changes the gravitational part of the action but not the matter action, so the coupling to matter is unchanged, and at the linearized level, is still given by the usual $h_{\mu\nu}T^{\mu\nu}$ coupling. Thus, the source term in eq. (3.1.1) is not affected. Furthermore we have seen that, at short scales, such as the distance between the two bodies in a coalescing binary, the RT model reduces to GR to huge accuracy, so there is no appreciable modification to the orbital dynamics of a binary system, and the waveform produced by a coalescing binary in the region far from the source (where the $1/r$ GW behavior sets in, but still the expansion of the Universe can be neglected) is the same as in GR. In the signal received by a coalescing binary, the only difference will then come from the free propagation of the GW from the source to the observer, across cosmological distances. The equation governing the free propagation of tensor perturbations in the RT model is [31]

$$\tilde{h}''_A + [2\mathcal{H} - 3\gamma\bar{V}aH_0]\tilde{h}'_A + k^2\tilde{h}_A = 0\,. \tag{3.1.35}$$

So the 'friction term' $2\mathcal{H}\tilde{h}'_A$ is modified with respect to GR, but the term $k^2\tilde{h}_A$ is not. Thus, first of all we see that the RT model passes the constraints from the speed of GWs. As we have mentioned, this is a non-trivial constraint that has ruled out many modified gravity theories. (3.1.35) is of the form (3.1.7), with

$$\delta(\eta) = \frac{3\gamma\bar{V}(\eta)H_0}{2H(\eta)}\,, \tag{3.1.36}$$

where we have used $\mathcal{H} = aH$. Recall that, for the RT model in a FRW background, we have defined the auxiliary field V from $V = H_0 S_0$, where S_0 is the $\mu = 0$ component [in coordinates $(t, \mathbf{x})$] of the auxiliary four-vector field S_μ of the RT model, see

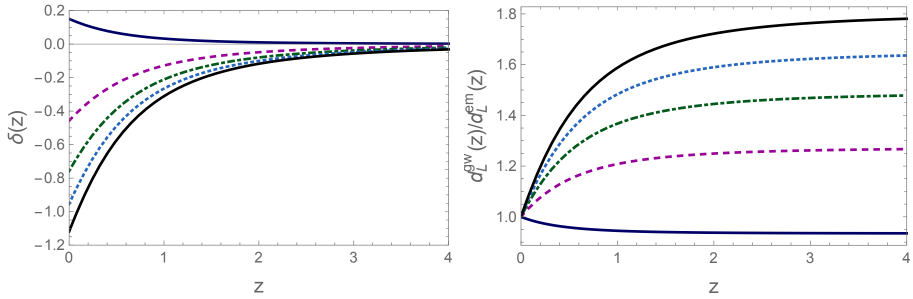

FIGURE 3.1: The functions $\delta(z)$ (left panel) and $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ (right panel), for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted) and $\Delta N = 100$ (black solid line). From [31].

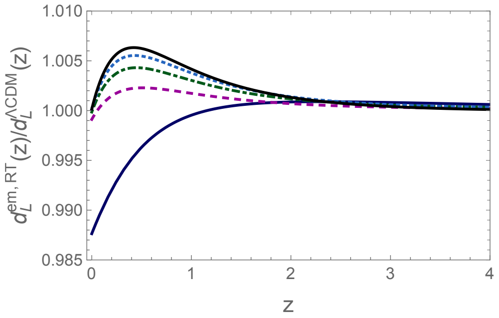

FIGURE 3.2: The ratio of $d_L^{\rm em}(z)$ computed in the RT model to the luminosity distance of ΛCDM for the minimal RT model (blue solid line) and for RT with $\Delta N = 34$ (magenta, dashed), $\Delta N = 50$ (green, dot-dashed) and $\Delta N = 64$ (cyan, dotted) and $\Delta N = 100$ (black solid line), using for each model its own mean values of H_0 and Ω_M. From [31].

eq. (2.1.22). Recalling the definition (2.1.6) of γ, we can also write eq. (3.1.36) as

$$\delta(\eta) = \frac{m^2 \bar{S}_0(\eta)}{6H(\eta)} \,. \tag{3.1.37}$$

Using the numerical solution of the background evolution equation of the RT model studied in Section 2.1.1, we can therefore immediately compute δ and $d_L^{\rm gw}/d_L^{\rm em}$, as functions of the redshift. The results are shown in Fig. 3.1. These results are quite spectacular, in particular at large ΔN. For instance, for $\Delta N = 64$, at large z the ratio $d_L^{\rm gw}/d_L^{\rm em}$ tends asymptotically to a value $\simeq 1.65$, corresponding to a 65% deviation from GR, a truly huge effect. In the limit of large ΔN [exemplified here by the case $(M_{\rm infl} = 10^{16}\,{\rm GeV}, \Delta N = 100)$; as we mentioned, for $M_{\rm infl} = 10^{16}\,{\rm GeV}$ this asymptotic curve is actually reached already at $\Delta N \gtrsim 70$], at large z the ratio $d_L^{\rm gw}/d_L^{\rm em}$ reaches a value $\simeq 1.80$, i.e. a 80% deviation from GR. Similarly, at $z = 0$, $\delta(0)$, in the limit of large ΔN, saturates to a value -1.11, so, in eq. (3.1.7), near $z = 0$ the term $1 - \delta(0) \simeq 2.11$ is more than twice the GR value.

This is very surprising because we have seen that, for all values of ΔN, the RT model differs from ΛCDM by less that 1% at the level of background evolution (see fig. 2.2), and by a few percent to below percent level, depending on wavenumber, for the scalar perturbations, see e.g. Figs. 2.5-2.14. This is indeed what allows the model to fit well the current cosmological observations. One would have then naturally guessed that also in the tensor perturbation sector the differences would be of the same order. Instead, for large ΔN, they are much bigger, a very good news for GW experiments.

For comparing the RT model to ΛCDM the relevant quantity, rather than the ratio of $d_L^{\rm gw}$ to $d_L^{\rm em}$, both computed within the RT model, is actually the ratio of $d_L^{\rm gw}$, computed in the RT model, to the luminosity distance $d_L^{\Lambda{\rm CDM}}$ computed in ΛCDM (for which the notion of electromagnetic and GW luminosity distance coincide), and in which, in each model, the respective mean values of the parameters H_0 and Ω_M are used. However, the results for $d_L^{\rm gw,RT}/d_L^{\Lambda{\rm CDM}}$ turn out to be practically the same as the results shown in the right panel of Fig. 3.1. This can be seen by writing

$$\frac{d_L^{\rm gw,RT}(z)}{d_L^{\Lambda{\rm CDM}}(z)} = \left(\frac{d_L^{\rm gw,RT}(z)}{d_L^{\rm em,RT}(z)}\right) \times \left(\frac{d_L^{\rm em,RT}(z)}{d_L^{\Lambda{\rm CDM}}(z)}\right) , \tag{3.1.38}$$

where, for clarity, we have denoted by $d_L^{\rm gw,RT}(z)$ the GW luminosity distance $d_L^{\rm gw}$ in the RT model. The first factor on the right-hand side is the quantity that we have already shown in the right panel of Fig. 3.1. The second factor is shown in Fig. 3.2, and we see that is very close to one; in particular, for the RT model with large ΔN, it reaches at most a value of order 1.006 for $\Delta N = 100$ near $z \simeq 0.3$, and then quickly goes asymptotically to values of order 1.001. This can be understood observing that the ratio $d_L^{\rm em,RT}(z)/d_L^{\Lambda{\rm CDM}}(z)$ is determined by two factors. First, by the different mean values of H_0 and Ω_M between the RT model with the given ΔN and ΛCDM; second, by the different redshift dependence of the DE density, or, equivalently, the different DE equation of state $w_{\rm DE}(z)$. However, we have seen in Table 2.2 that Bayesian parameter estimation gives for the RT model values of H_0 and Ω_M very close to those of ΛCDM, particularly at large ΔN; furthermore, as discussed in [171], the change in the value of these parameters goes precisely in the direction to cancel the effect in the change of the DE equation of state. This is due to the fact that Bayesian parameter estimation in practice requires the model to fit some fixed distance scales at large redshifts, such as the scales given by the CMB peaks or by the BAO oscillations; thus, if, compared to ΛCDM, one changes $w_{\rm DE}(z)$ in the direction of giving, say, a larger (electromagnetic) luminosity distance at large redshift, H_0 and Ω_M change in the direction such that they partially compensate for this change. As a result the electromagnetic luminosity distance, particularly at moderate to large values z, changes very little. Thus, the difference in the GW luminosity distance of the RT model, compared to ΛCDM, in practice is entirely given by the effect of modified GW propagation, while the DE equation of state and the difference in H_0 and Ω_M among RT and ΛCDM have a negligible effect.

As we already discussed, the z dependence of the ratio $d_L^{\rm gw}/d_L^{\rm em}$ is easily understood observing that, by definition, at $z \to 0$ we must have $d_L^{\rm gw}/d_L^{\rm em} \to 1$ because, if the distance to the source goes to zero, there can be no effect from modified GW propagation. At large z, $d_L^{\rm gw}/d_L^{\rm em}$ goes to a constant because, in the RT model, as in most other modified gravity model, the emergence of dark energy is a relatively recent phenomenon, so the modifications to GR, and hence the function $\delta(z)$ in eq. (3.1.7), go to zero at large redshifts. As a consequence, at large z the integral in eq. (3.1.13) saturates to a constant value. As shown in Fig. 3.3, the numerical results for $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ are extremely well fitted by the simple (Ξ_0, n) parametrization in eq. (3.1.30). The best-fit values of Ξ_0 and n are given in Table 3.1.

Comparing Fig. 3.3 with Fig. 2.3, we see that the (Ξ_0, n) parametrization works much better than the (w_0, w_a) parametrization for the equation of state. This is due

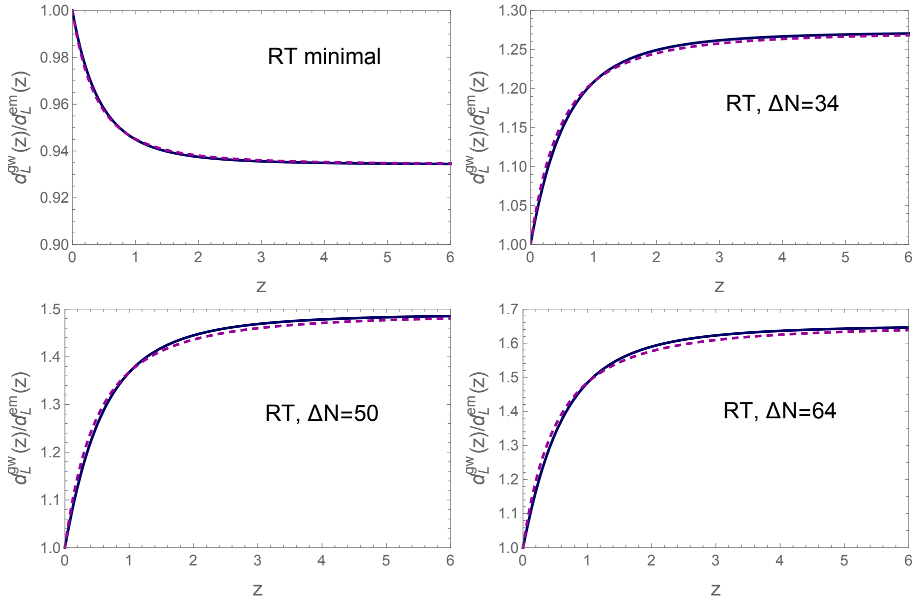

FIGURE 3.3: The function $d_L^{\mathrm{gw}}(z)/d_L^{\mathrm{em}}(z)$ from the numerical integration (blue solid line), compared with the fit (3.1.30) (magenta, dashed). Upper left panel: for the minimal RT model; upper right: RT with $\Delta N = 34$; lower left: RT with $\Delta N = 50$; lower right: RT with $\Delta N = 64$. From [31].

	RT, minimal	$\Delta N = 34$	$\Delta N = 50$	$\Delta N = 64$	$\Delta N = 100$
Ξ_0	0.93	1.27	1.49	1.65	1.80
n	2.59	2.08	2.00	1.95	1.91
$\delta(0)$	0.15	-0.46	-0.76	-0.95	-1.12
$\delta(0)/(1-\Xi_0)$	2.29	1.67	1.54	1.46	1.39

TABLE 3.1: Table from [31].Values of Ξ_0, n, $\delta(0) \equiv \delta(z = 0)$ and $\delta(0)/(1-\Xi_0)$ for the RT model with various values of ΔN. The results have been obtained using for each model its own mean values for Ω_M and h_0 from Table 2.2.

to the fact that eq. (3.1.30) catches correctly both the $z \to 0$ limit and the large z limit.

Fig. 3.4 compares the numerical result for $\delta(z)$ (blue solid line) with the fit of eq. (3.1.34), using the same values of Ξ_0 and n as in Table 3.1 (magenta, dashed lines). We see that the (Ξ_0, n) parametrization provides a fit to $\delta(z)$ less good than to $d_L^{\mathrm{gw}}(z)/d_L^{\mathrm{em}}(z)$, particularly near $z = 0$. This is due to the fact that, for $d_L^{\mathrm{gw}}(z)/d_L^{\mathrm{em}}(z)$, the (Ξ_0, n) parametrization catches correctly both the value in $z = 0$ and the large z limit; thus, as long as $d_L^{\mathrm{gw}}(z)/d_L^{\mathrm{em}}(z)$ is smooth in between, it is natural to find a value of n such that the (Ξ_0, n) parametrization performs well. In contrast, the value of $\delta(z = 0)$ is not automatically reproduced by the parametrization (3.1.34), and indeed we see from the figures that in this region the parametrization is not accurate. For instance, the numerical integration gives the values of $\delta(0) \equiv \delta(z = 0)$ shown in Table 3.1, while the parametrization (3.1.34) would incorrectly predict $\delta(0) \simeq \{0.17, -0.57, -0.98, -1.27, -1.53\}$. Note that, with the parametrization of eq. (3.1.30) and eq. (3.1.34), we have

$$\delta(0) = n(1 - \Xi_0)\,. \tag{3.1.39}$$

This suggests that, after having fixed Ξ_0 so to reproduce exactly the large-z behavior

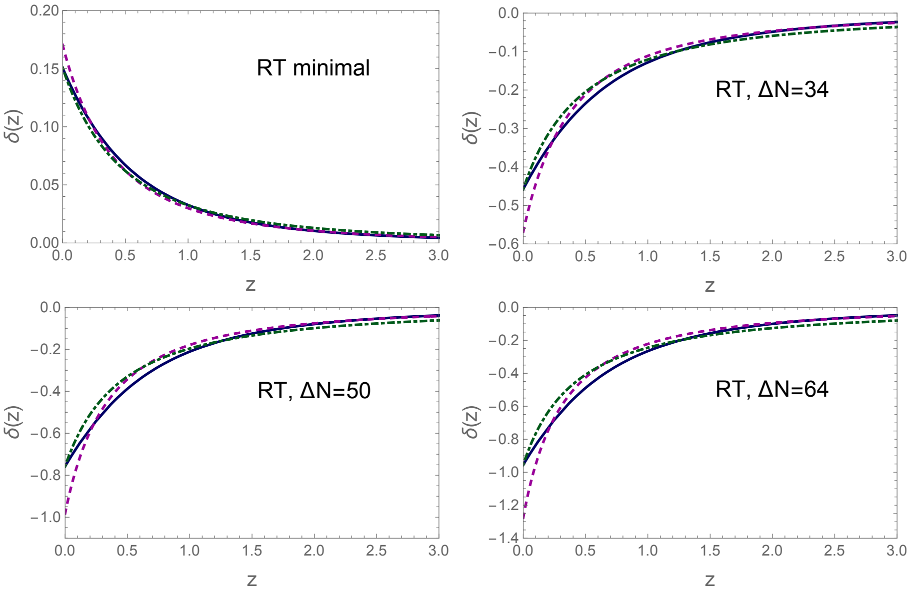

FIGURE 3.4: The function $\delta(z)$ from the numerical integration (blue solid line), compared to the parametrization (3.1.34) with the value of n obtained from the best fit to $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ (magenta, dashed), and with $n = \delta(z=0)/(1-\Xi_0)$ (green, dot-dashed). Upper left panel: for the minimal RT model; upper right: RT with $\Delta N = 34$; lower left: RT with $\Delta N = 50$; lower right: RT with $\Delta N = 64$. From [31].

of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$, rather then choosing n from a best fit to $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$, we could choose $n = \delta(0)/(1-\Xi_0)$, so that the parametrization (3.1.34) reproduces exactly the value of $\delta(0)$. The values obtained in this way are given in the last line of Table 3.1. If one uses these values of n, the fit to $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ significantly degrades, but the fit to $\delta(z)$ becomes more accurate, and is shown as the green dot-dashed lines in Fig. 3.4. In general, since the directly observable quantity is $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$, it is more important to have a simple and accurate analytic representation for it, rather than for $\delta(z)$. Of course, for an accurate comparison with the data, one can also use directly the results of the numerical integration, which are obtained very quickly.

For comparison, the result for the RR model is also of the form (3.1.7), except that the function δ is given by

$$\delta = \frac{3\gamma\, d\bar{V}/d\log a}{2(1-3\gamma\bar{V})}\,, \tag{3.1.40}$$

and now $V = H_0^2 S$, where S is the auxiliary field of the RR model, defined by $U = -\Box^{-1}R$ and $S = -\Box^{-1}U$. The numerical integration then gives again a result very well fitted by eq. (3.1.30), with $\Xi_0 \simeq 0.97$ and $n \simeq 2.5$ [171]. However, contrary to the RT model, here the deviation from GR is only about 3%.

3.2 Cosmology at 2G and 3G detectors

We now analyze the cosmological information that is possible to extract from the current and future generations of gravitational-wave detectors, in particular regarding the dark energy sector of the Universe. We have seen that, already from a theoretical point of view, modified gravitational-wave propagation plays a primary role for testing modified gravity theories that attempt an explanation of dark energy. Here we want to substantiate this statement quantitatively and, furthermore, we want to study the most general cosmological constraints that GW detectors will be able to provide. In this way we can also compare them with the information already available from electromagnetic observations.

The first observations of gravitational waves (GWs) from binary black-holes coalescences [211, 212, 213, 214, 215, 216], as well as the first observation of a neutron star binary coalescence [217], together with the associated γ-ray burst (GRB) [218, 219, 190] and the follow-up studies of the electromagnetic counterpart (see [220] and references therein) have opened the era of GW astronomy. In the near future, with advanced LIGO and advanced Virgo reaching their target sensitivity, and other detectors such as KAGRA and LIGO-India joining the search, it is expected that such detections will take place routinely. On a longer timescale the space interferometer LISA [221], that is expected to fly in 2034, and third-generation (3G) ground-based interferometers currently under study, such as the Einstein Telescope (ET) in Europe [222, 223] and Cosmic Explorer (CE) in the US [224], will have the potential of detecting a large number of coalescing compact binaries at cosmological redshifts.

In this Section we follow [210] and present our contribution to the currently ongoing effort for exploring the scientific potential of 3G interferometers (see the recent ET Science Case paper [225]), by performing an updated study of the cosmological information that can be obtained from the observation of standard sirens. Following [210], we focus on binary neutron stars (BNS) with an electromagnetic counterpart as appropriate for ground-based detectors, since in this case BNS merge within the bandwidth of the detector, and thus can in principle be detected in coincidence with an electromagnetic signal that can allow us to identify the host galaxy and therefore obtain the redshift.

Using the strategy presented in [226, 227, 228, 229, 230, 231], we start by constructing mock catalogs of BNS detections at GW detectors. We consider first a network of 2G detectors composed by advanced LIGO-Hanford, advanced LIGO-Livingston, advanced Virgo, Kagra and LIGO India, assumed to be at their target sensitivity. We will refer to this as the HLVKI network. We will then consider 3G detectors studying two different configurations, namely a single ET detector, and a network made of a single Einstein Telescope plus two Cosmic Explorers.

We will then study the possibility of simultaneous detection of an electromagnetic counterpart, focusing on the case of a joint GW-GRB detection. A single GW detector, even in a triangular configuration as planned for ET, cannot provide the localization of a coalescing binary with a significant accuracy. However, the detection of a temporally coincident GRB can still allow for the measurement of the redshift of the source; indeed, GRB satellites such as Swift regularly obtain redshifts of GRBs, without the need of a GW localization. In particular, for 3G detectors we will estimate the expected number and the redshift distributions of coincidences between GW events and the electromagnetic signal observed at a GRB detector with the characteristics of the proposed THESEUS mission [232, 233, 234], that could be in operation at the same time as 3G detectors.[2]

A network of third-generation GW detectors, such as the ET+CE+CE configuration, would instead localize the source, whose redshift could then be measured also by optical/IR telescopes. For instance, the electromagnetic signal from a kilonova associated to a BNS coalescence could be detected up to $z \simeq 0.55$ by optical imaging at LSST and Subaru, up to $z \simeq 0.76$ with infrared imaging at WFIRST, and up to $z \simeq 0.37$ by optical spectroscopy at ELT. However, it is currently difficult to estimate how much telescope time will be devoted by these facilities to the follow-up of GW events. Here we limit ourselves to the coincidence with GRB detectors, but we should keep in mind that, for a network with significant localization capabilities such as ET+CE+CE, at $z \lesssim 0.5$ a significant number of coincidences with optical/IR telescopes is in principle possible, to the extent that the number of such coincidences could be much larger than those obtained from GRB detections, so for such a network our estimates will be conservative. To construct our mock source catalogs we will examine different possibilities for the local merger rate and for the probability of determining the redshift through the detection of an associated GRB.

3.2.1 Mock source catalogs

GW events

In order to simulate a catalog of binary neutron stars coalescences, we first produce an extra-galactic population of neutron star binaries using the Monte Carlo algorithm developed in [226, 227, 228, 229, 230, 231]. We use the fiducial model of [235] for the distribution of the parameters and we proceed as follows for each source: the location in the sky $\hat{\Omega}$, the cosine of the orientation ι, the polarization ψ and the phase of the signal at coalescence ϕ_0, are drawn from uniform distributions. The redshift is drawn from a (normalized) probability distribution $p(z)$,

$$p(z) = \frac{R_z(z)}{\int_0^{10} R_z(z) dz}\,, \tag{3.2.1}$$

where $R_z(z)$ is the merger rate density per unit redshift, in the observer frame. It can be expressed as

$$R_z(z) = \frac{R_m(z)}{1+z}\frac{dV(z)}{dz}\,, \tag{3.2.2}$$

where dV/dz is the comoving volume element and R_m is the rate per volume in the source frame. The latter is given by

$$R_m(z) = \int_{t_{\rm min}}^{t_{\rm max}} R_f[t(z) - t_d] P(t_d) dt_d\,, \tag{3.2.3}$$

where $R_f(t)$ is the formation rate of massive binaries, $P(t_d)$ is the distribution of the time delay t_d between the formation of the massive progenitors and their merger, and $t(z)$ is the age of the Universe at the time of merger. We assume that $R_f(t)$ in eq. (3.2.3) follows the cosmic star formation rate, for which we use the recent model

---

[2] We stress that, in this thesis, we are referring to the capabilities of the future THESEUS mission as they were estimated when ref. [210] was written. These prospects could be modified with time, according to what sensitivity is judged to be realistic. From internal communications, it seems that the actual capability of THESEUS should be degraded with respect to the one assumed when the study [210] was conducted. As a consequence, the new forecasts for cosmological studies with joint GW/GRB detections at ET/THESEUS could be different from what we discuss here.

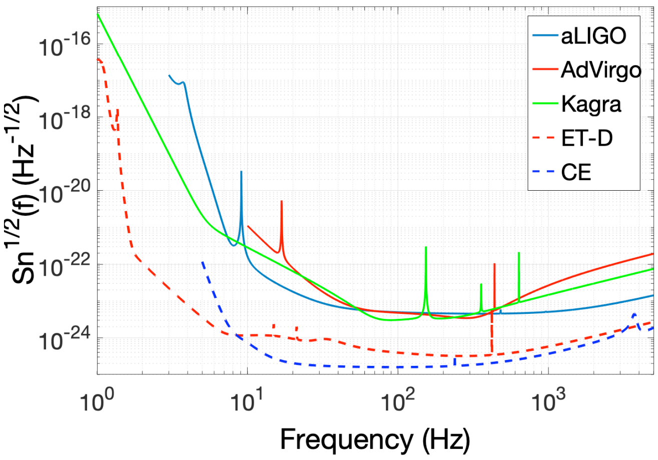

FIGURE 3.5: The strain sensitivities of advanced and 3G GW detectors. For ET we use the ET-D sensitivity curve. From [210].

of [236]. We further assume that the time delay distribution follows $P(t_d) \propto t_d^{\alpha}$, with $\alpha = -1$ for $t_d > t_{\rm min}$, where $t_{\rm min} = 20$ Myr is the minimum delay time for a massive binary to evolve until coalescence, and $t_{\rm max}$ is a maximum time delay, set equal to the Hubble time. The overall normalization is fixed by requiring that the value of R_m at $z = 0$ agrees with the local rate estimated from the O1 LIGO observation run and the O2 LIGO/Virgo observation run [216], using the median rates obtained from the GstLAL pipeline. The result depends on the assumption for the mass distribution of the neutron stars. For a flat mass distribution

$$R_m(z=0) = 662\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}\,, \tag{3.2.4}$$

while for a Gaussian mass distribution

$$R_m(z=0) = 920\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}\,. \tag{3.2.5}$$

In the following we will refer to them just as the “O2 rates”, and we will give our results both for the flat distribution and for the Gaussian distribution. To have a quantitative measure of how the results depend on our astrophysical assumptions, we will also generate alternative catalogs of GW events by assuming a Madau-Dickinson star formation rate [237] and an exponential time delay between formation and merger with an e-fold time of 100 Myr [238]. This will also allow us to compare with the results presented in Table 1 of ref. [239], which also computes the number of BNS detected per year at the HLVKI network and at the ET+CE+CE network under these astrophysical assumptions [and using a local comoving BNS merger rate of 1000 $\rm Gpc^{-3} yr^{-1}$, that is close to our rate for a Gaussian mass distribution, eq. (3.2.5)].

Next, for each BNS generated by this procedure, we determine if its resultant GW emission is detectable with a given GW detector network. We consider three cases: (1) a 2G network composed by advanced LIGO-Hanford+advanced LIGO-Livingston+advanced Virgo+Kagra+LIGO India (HLVKI). (2) A single 3G detector, chosen according to current estimates for the sensitivity of the Einstein Telescope.

(3) A three-detector network made by ET and two CE. The sensitivity curves that we use are shown in Fig. 3.5.[3] The signal-to-noise ratio (SNR), ρ_a, detected by matched filtering with an optimum filter in the ideal case of Gaussian noise, in a detector labeled a, is

$$\rho_a^2 = 4 \int_0^\infty df \, \frac{|F_{+,a}\tilde{h}_+ + F_{\times,a}\tilde{h}_\times|^2}{S_{n,a}} \,, \tag{3.2.6}$$

where f is the GW frequency in the observer frame, $\tilde{h}_+$ and $\tilde{h}_\times$ the Fourier transforms of the GW strain amplitudes of $+$ and $\times$ polarizations, $F_{+,a}$ and $F_{\times,a}$ are the antenna response functions to the GW $+$ and $\times$ polarizations, and $S_{n,a}(f)$ is the one-sided noise power spectral density (PSD) of detector a. The coherent SNR, assuming uncorrelated noises among the detectors, is simply given by the quadrature sum of the individual SNRs, $\rho^2_{\rm tot} = \sum_a \rho_a^2$. The triangular configuration of ET provides three independent differential signals between the arms, equivalent to three detectors, and again the coherent SNR is given by the quadrature sum of the individual SNRs for these three equivalent detectors.

For low-mass systems such as BNS the SNR in one detector is dominated by the inspiral part of the signal and is then given

$$\rho_a^2 = \frac{5}{6} \frac{[G\mathcal{M}(1+z)]^{5/3}\mathcal{F}_a^2}{c^3\pi^{4/3}d_L^2(z)} \int_{f_{\rm min}}^{f_{\rm insp}(z)} df \, \frac{f^{-7/3}}{S_{n,a}(f)} \,. \tag{3.2.7}$$

Here $\mathcal{M}$ is the intrinsic chirp mass, a combination of the two component masses, $d_L(z)$ is the luminosity distance, G is the gravitational constant, c is the speed of light, $f_{\rm min}$ is the low frequency limit of the detector and $f_{\rm insp}(z) = f_{\rm insp}/(1+z)$ is the observed (redshifted) gravitational-wave frequency at the end of the inspiral phase. The factor

$$\mathcal{F}_a^2 = \frac{(1+\cos^2\iota)^2}{4} F^2_{+,a} + \cos^2\iota \, F^2_{\times,a} \,, \tag{3.2.8}$$

characterizes the detector response. In order to decide which detectors contribute to the combined SNR, we assume that each detector has a duty cycle of 80%. We then classify the event as detectable if the combined SNR among the detectors in the network, $\rho_{\rm tot}$, is larger than a SNR threshold level, that we take to be $\rho_{\rm threshold} = 12$. Both for the HLVKI network and for 3G detectors we assume for definiteness 10 years of running (which, given the 80% duty cycle for each detector, corresponds to a shorter stream of actual coincident data). This assumption should be taken as a limiting case (which, for the 2G case, is also necessary to have a sufficient sample of events to obtain the convergence of our MCMC); however, the results for a shorter time span T can be obtained basically by rescaling our results by a factor $\sqrt{T/10\,{\rm yr}}$, corresponding to the fact that, with a large number of events N, the error scales roughly as $1/\sqrt{N}$.

In order to generate our mock catalogs of measured luminosity distances of standard sirens, we assume a fiducial ΛCDM model with $\Omega_M = 0.3087$ and $H_0 = 67.64\,{\rm km\,s^{-1}\,Mpc^{-1}}$, corresponding to the mean values obtained from the CMB + BAO + SNe dataset that we will use, which is presented in detail in Section 3.2.2. Extracting randomly the redshift of the source from the theoretical distribution obtained from eqs. (3.2.1)–(3.2.3) and using our fiducial cosmological model, we obtain a value of $d_L(z)$ for each source. To take into account the observational error in the reconstruction of the luminosity distance from the GW data, we scatter randomly the

---

[3] The ET and CE sensitivity curves, as well as the assumed locations of ET (in Europe) and two CE (in the US) correspond to the choices currently used to develop the Science Case for 3G detectors.

values of $d_L(z)$ according to a Gaussian distribution with a width $\Delta d_L(z)$ equal to the expected error in the reconstruction. For each generated event, this is estimated from $\Delta d_L/d_L = 1/\mathrm{SNR}$, following e.g. ref. [176]. Note that, comparing with the result of an actual mock parameter reconstruction, one finds that, because of the degeneracy with the inclination angle, this can result in an underestimate of the actual value of $\Delta d_L/d_L$, by a factor which has a significant scatter from event to event, but is generically ~ 2 [178]. However, for GW signals detected in coincidence with a GRB (which are the signals that we consider here), assuming that the GRB is beamed within an angle of about 25° one finds that the correlation between distance and inclination is substantially broken, and the above estimate becomes more accurate [178].

Beside the instrumental error, we must consider the error due to lensing. Following [180, 181], we model it as[4]

$$\left(\frac{\Delta d_L(z)}{d_L(z)}\right)_{\rm lensing} \simeq 0.05z\,, \tag{3.2.9}$$

and we add it in quadrature to the instrumental error. However, we will see below that, for the sources at $z < 1.5$, that will largely dominate our results, the lensing error is subdominant with respect to the instrumental error.

If the source is at very low redshift, once determined the measured redshift as discussed in Section 3.2.1, to obtain the cosmological redshift we must correct for the peculiar Hubble flow. This is estimated adding an error on z corresponding to a recessional velocity of the host galaxy of 200 km/s, as in ref. [241].

Electromagnetic counterpart

In general, to identify the counterpart, one can consider two possible strategies. The first, that has been implemented successfully with GW170817, consists in having a network of GW detectors, that allows us to localize the source relatively well. Then, the follow-up with telescopes working, e.g., in the optical or IR can identify the host galaxy and determine its redshift. The second possibility, that can be applied even when no GW localization is available (as, for instance, with a single ET detector) is to use the temporal coincidence of the GW event with a short GRB; for many short GRBs, the redshift has indeed been determined from the X-ray afterglow, that can be accurately localized by *Chandra* or *Swift*/XRT. For instance, in the sample of 67 *Swift* short GRBs discussed in [242], 53 events were rapidly followed up with the on-board X-ray Telescope, leading to 47 detections of the source.

The estimate of the number of coincidences between GW events and electromagnetic observations depends crucially, of course, on the rate of expected GW events, as well as on the network of GRB satellites and telescopes available at the time. We therefore discuss the 2G and 3G cases separately.

GRB coincidences with the HLVKI network. We begin by investigating the possibility of detecting in coincidence a GW signal at the HLVKI network and a GRB with the current generation of GRB satellites. We assume here that the Fermi-GBM can make a coincident detection and that *Swift* can slew to the combined GW/GRB

---

[4] Actually, after publication of [210], we came to know that the linear function in eq. (3.2.9) largely overestimates the lensing error, see Fig. 12 of the ET Science Case paper [225], adapted from Fig. 3 of [240]. As a consequence the orange lines in Figs. 3.8, 3.9, 3.11, 3.12 must be discarded. In any case, all this has no effect on the results that we obtain because the lensing error, even when overestimated from eq. (3.2.9), is negligible with respect to the instrumental error at the redshifts relevant for joint GW/GRB detections (as mentioned right after eq. (3.2.9)).

error box and identify an X-ray counterpart. We note here that for 170817A *Swift* was occulted by Earth at time of Fermi trigger, so imaging by the X-ray telescope (XRT) took place around 1 hr post trigger. At that time it was able to cover 90% of the GW skymap to rule out any bright sources [243].

For a GRB detected in coincidence with a GW signal we require that the peak flux is above the flux limit of the satellite. Based on the modeling of [244] we assume a Gaussian structured jet profile model of GRB170817A given by

$$L(\theta_{\rm V}) = L_{\rm c} \exp\left(-\frac{\theta_{\rm V}^2}{2\theta_{\rm c}^2}\right) , \qquad (3.2.10)$$

with $L(\theta)$ the luminosity per unit solid angle, $\theta_{\rm V}$ the viewing angle and $L_{\rm c}$ and $\theta_{\rm c}$ structure parameters that define the angular profile. The structured jet parameter is given by $\theta_{\rm c} = 4.7^\circ$. The value of $L_{\rm c}$ is given by $L_{\rm c} = L_{\rm p}/4\pi$ erg s^{-1} sr^{-1}, where $L_{\rm p}$ is the peak luminosity of each burst, which is obtained by sampling $\Phi(L_p)dL_p$. We assume the standard broken power-law distribution of the form

$$\Phi(L_{\rm p}) \propto \begin{cases} (L_{\rm p}/L_*)^{\alpha} , & L_{\rm p} < L_* \\ (L_{\rm p}/L_*)^{\beta} , & L_{\rm p} \geqslant L_* \end{cases} \qquad (3.2.11)$$

where $L_{\rm p}$ is the peak luminosity assuming isotropic emission in the rest frame in the 1-10000 keV energy range, L_* is a characteristic value separating the two regimes, and the slopes describing these regimes are given by α and β respectively. Following [245] we use the values $\alpha = -1.95$, $\beta = -3$ and $L_* = 2 \times 10^{52}$ erg sec^{-1}. Given a source at luminosity distance d_L one can convert $4\pi L(\theta_{\rm V})$ to an observed peak flux as a function of viewing angle, $F_P(\theta_{\rm V})$, obtained from the value of the GW inclination angle. A Fermi-GBM detection is recorded if the value of $F_P(\theta_{\rm V})$ is greater than the flux limit of 1.1 ph sec^{-1} cm^{-2} in the 50-300 keV band for Fermi-GBM [244], noting that 95% of the bursts detected in the 64 ms timescale are within this limit. We further assume the total time-averaged observable sky fraction of the Fermi-GBM, which is 0.60 [246]. Using this procedure, among the events in the GW BNS catalog generated as discussed in sect. 3.2.1, we select those that have an observed GRB counterpart.

Table 3.2 shows the number of BNS sources along with the number of coincident GRB detections determined using the procedure above, for the HLVKI network. We see that 10 years of observation would yield of order 14-15 joint detections. In Fig. 3.6 we show the redshift distribution of the GW events (left panel) and of the GW-GRB coincidences (right panel), for a realization of our catalog.

To test the impact of changing our astrophysical assumptions, we have also generated a catalog of GW events assuming a Madau-Dickinson star formation rate and an exponential time delay between formation and merger with an e-fold time of 100 Myr, as in ref. [239]. In this case, assuming again a duty cycle of 80% and a network SNR threshold level $\rho_{\rm threshold} = 12$, we find that the number of BNS detected at the HLVKI network, for the Gaussian mass distribution, is 64/yr, to be compared with the value 48/yr reported in ref. [239].

GRB coincidences with 3G detectors. For 3G detectors the estimates are of course more uncertain. Indeed, the identification of the counterpart depends on the network of GRB satellites and of telescopes at the time when 3G detectors will operate, as well as on issues that are presently difficult to predict, such as the prioritization that will be given by various telescopes to the follow-up of GW signals.

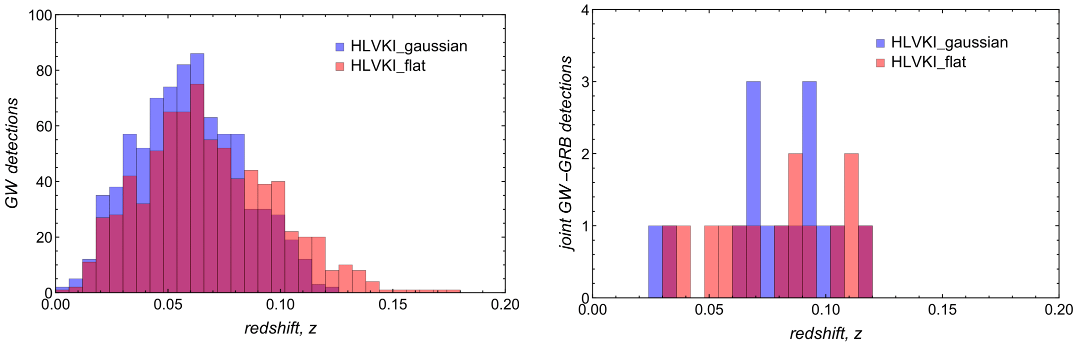

FIGURE 3.6: Left panel: the redshift distribution of 10 yr of GW events from a realization of the mock catalog at the HLVKI network, for the flat and the Gaussian mass distribution. Right panel: the redshift distribution of 10-years of GW-GRB coincidences between the HLVKI network and the current generation of GRB satellites. From [210].

Network	GW events		Joint GW-GRB events	
	Flat	Gaussian	Flat	Gaussian
HLVKI	768	814	14	15

TABLE 3.2: Number of GW events detected by second generation (2G) networks in 10 years, and the expected GW-GRB coincidences obtained by assuming a GRB detector with the characteristics of Fermi-GBM. We show detection rates for BNS populations generated using O2 rates corresponding to both flat and Gaussian mass distributions. From [210].

The proposed THESEUS mission [232, 233, 234] could be particularly useful for performing coincidences between GW events and GRBs, even in the absence of localization from the GW signal. A crucial difference with the 2G case is that, at the sensitivity level of 3G detectors, there will be many more GW events compared to what GRB satellites could detect. The main reason for this is that the GRB instruments are limited by their flux sensitivity for the more distant GRB emissions at wider viewing angles. For instance, it was estimated in [180] that ET will be able to detect $\mathcal{O}(10^5-10^6)$ BNS mergers per year. As we will see below our estimate, given in Table 3.3, is that ET will detect about $(6-7)\times 10^5$ events in 8 yr of actual data taking, corresponding to a rate of order $(0.8-0.9)\times 10^5/\text{yr}$, consistent with previous estimates, although somewhat smaller. In any case, according to the estimate in [233], only about $15-35$ coincident short GRB (sGRB) per year will be detected by THESEUS with its X-Gamma ray Imaging Spectrometer (XGIS); we will see below that our results, using somewhat different assumptions for the GRB luminosity and BNS rate, gives a slightly higher number of coincidences, but still of this order of magnitude. Beside the collimated prompt GRB emission, more isotropic soft X emission is also expected from the afterglow. This could be detected by the Soft X-ray Imager (SXI) on board THESEUS, leading possibly to a few hundred more detections per year [233]. In any case, the number of joint GW-GRB detections will be a very small fraction of the number of GW events.

With a network of at least three GW detectors, accurate localization of the GW signal becomes possible, allowing for electromagnetic follow-up observations, that

Network	GW events		Joint GW-GRB events	
	Flat	Gaussian	Flat	Gaussian
ET	621,700	688,426	389 (128)	511 (169)
ET+CE+CE	5,420,656	7,077,131	644 (213)	907 (299)

TABLE 3.3: Number of GW BNS events detected by third generation (3G) networks in 10 years of data taking (assuming a 80% duty cycle for each detector) and the corresponding GW-GRB coincidences obtained by assuming a GRB detector with the characteristics of THESEUS-XGIS; numbers in parenthesis show the number of sources with arcmin localisation. BNS populations are generated using the O2 rates corresponding to 'flat' and 'Gaussian' mass distributions. From [210].

could determine the redshift of the source. If the source localization is already available through GWs, LSST could detect the counterpart up to $z \simeq 0.55$ and WFIRST up to $z \simeq 0.76$, and many more telescopes in the UV, optical, IR, radio could detect the counterpart at smaller redshifts, say $z \sim 0.1 - 0.3$. However, the follow-up of $O(10^3)$ well localized GW events at $z \sim 0.5$ would require the equivalent of 1 yr of dedicated LSST time, which is not realistic. Currently, a more realistic estimate is that LSST might use of order of 1% of its time for GW follow up, so it will be challenging for LSST to observe more than $\mathcal{O}(10)$ counterparts per year, at $z \sim 0.5$. The localization cost is much smaller at $z \sim 0.1$, where $\mathcal{O}(100)$ events per year could be a more realistic expectation, but this will depend on the science prioritization in the 2030s, when 3G detectors will hopefully operate. Given these large uncertainties, in this analysis we will limit ourselves to the coincidences with GRB detectors.

We repeat our simulations for the coincidence with a single ET detector and with a ET+CE+CE network, assuming that a THESEUS type satellite will be used for coincidence searches. For the GRB detection we assume a duty cycle of 80% due to a reduction of 20% as the satellite passes through the Southern Atlantic Anomaly, a flux limit of $0.2\,\mathrm{ph\,sec^{-1}\,cm^{-2}}$ in the 50–300 keV band and a sky coverage fraction of 0.5 [234]. We note that the XGIS will be able to localise sources to around 5 arcmin only within the central 2 sr of its field of view (FOV); outside this central region localisation will be coarse at best. We therefore consider two scenarios: one, that we will denote as 'optimistic', in which all the events detected by XGIS have a measured redshift, and one, that we will denote as 'realistic', where we assume that only around 1/3 of the sGRBs detected by XGIS could provide redshift estimates.

Events rates, redshift distributions and error $\Delta d_L(z)/d_L(z)$

In our MCMCs we will use a given realization of the catalog of events obtained with the procedure discussed above. It is however useful to describe the qualitative features of these catalogs, such as the redshift distributions of the events and the average value of $\Delta d_L(z)/d_L(z)$ as a function of redshift. This will provide a physical insight into which sources contribute most, to compare with previous works, and to provide ready-to-use formulas that can be applied to future studies.

Events at a single ET detector.

Table 3.3 shows the results of our simulations for the 3G era in terms of the number of GW signals from BNS, along with the number of joint GW/sGRB detections; the number of events with arcmin localisation are shown in parenthesis.

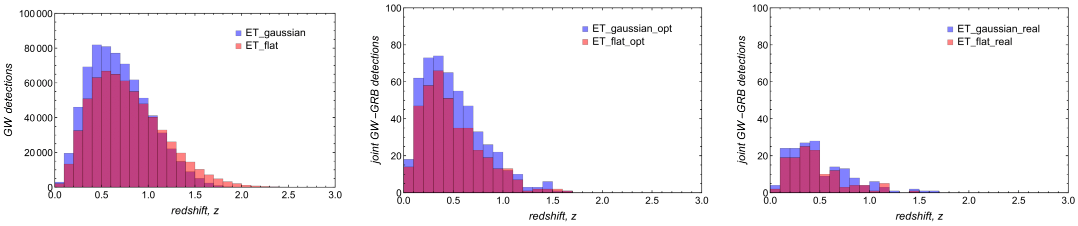

FIGURE 3.7: Left panel: the redshift distributions of 10-years of BNS detections by a ET detector. Middle panel: the coincident detections made by THESEUS in the 'optimistic' scenario for the FOV. Right panel: the coincident detections in the 'realistic' scenario. Notice that the vertical scale for the left panel is very different from that in the middle and right panels. From [210].

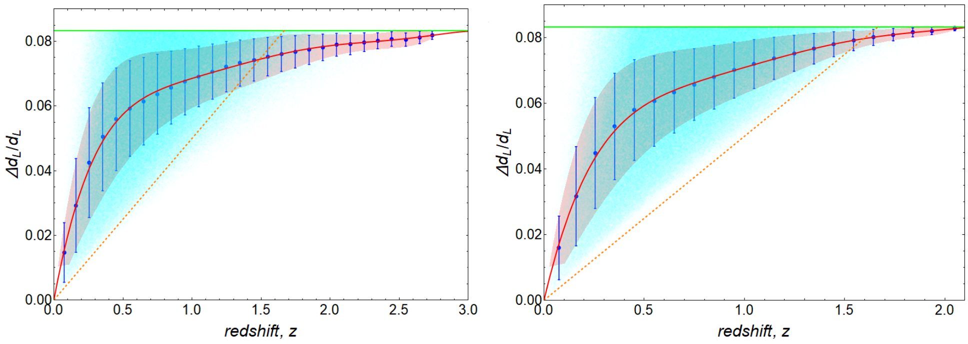

FIGURE 3.8: The ET instrumental contribution to the relative error $\Delta d_L/d_L$ for the specific realization of the catalog of BNS detections shown in Fig. 3.7. All the GW events, with or without a detected EM counterpart, are taken into account. Left panel: for the flat distribution of neutron star masses. Right panel: for the Gaussian distribution of neutron star masses. In each panel, the cyan shaded area corresponds to all the BNS events, while the coordinates of the blue points are given by the mean values of the redshift and of $\Delta d_L/d_L$ in each redshift bin, with the bins chosen as in Fig. 3.7. The blue error bars are the standard deviations of $\Delta d_L/d_L$ in each redshift bin. The horizontal green line at $\Delta d_L/d_L = 1/12$ is fixed by the SNR threshold of 12. The red line and the light red region are the fits to the blue points and error bars, given explicitly in the text. The orange dotted line is the error on $\Delta d_L/d_L$ induced by lensing. The redshift ranges shown in the two panels differ as a result of the different maximum values of redshift reached in the two corresponding catalogs. From [210].

For a single ET detector our estimate of the rate of BNS detection is between 6.2×10^5 and 6.9×10^5 events in 10 yr, having assumed a duty cycle of 80%, which, in the case of a single detector, corresponds to 8 yr of actual data. This corresponds to a rate (normalized to the actual time of data taking)

$$R \simeq (0.8 - 0.9) \times 10^5 \,\mathrm{BNS/yr}\,, \tag{3.2.12}$$

consistent with previous estimates, although somewhat smaller. This can be traced to the fact that we have used a threshold for the network SNR, obtained by combining the three arms of ET, given by $\rho_{\mathrm{threshold}} = 12$, while previous work, e.g. ref. [181], used $\rho_{\mathrm{threshold}} = 8$. We also see from Table 3.3 that, with a single ET detector, we should expect around $39 - 51$ coincident sGRB/GW events in one year of observation using the XGIS and the SXI detectors. These numbers differ from the $15 - 35$ events quoted in [233] for two main reasons. Firstly, the assumed luminosity function and BNS rate differs from that assumed in [233]. Secondly, our calculations assume a structured jet profile based on GRB170817A. From our sample of detections one could expect around $13 - 17$ events yr^{-1} to have arcmin localisations.

Figure 3.7 shows the redshift distributions of 10-years of BNS detections using ET, along with the joint GW-GRB distributions with optimistic and realistic scenarios for the FOV of THESEUS. We see that the vast majority of the joint GW-GRB detections is at $z \lesssim 1$. In this realization of the catalog, the ET event with the highest redshift is at $z \simeq 2.91$ for the flat mass distribution and at $z \simeq 2.10$ for the Gaussian mass distribution, while the joint detection with the highest redshift is at $z \simeq 1.63$ (found both in optimistic and realistic scenario for the THESEUS FOV, and for the Gaussian mass distribution). We find that the higher-z sGRB detections have viewing angles close to the jet axis corresponding with almost face-on BNSs; this selection effect was highlighted in [228].

Figure 3.8 shows the ET instrumental contribution to the relative error on luminosity distance, considering all the events in the specific realization of the catalog of BNS detections presented in the left panel of Fig. 3.7. The events are organized in the same redshift bins as Fig. 3.7 and, for the events belonging to each bin, the mean value of $\Delta d_L/d_L$ and its standard deviation $\sigma(\Delta d_L/d_L)$ are evaluated. Of course, the mean value of $\Delta d_L/d_L$ increases with distance, until we reach the threshold at SNR=12, and therefore the value $\Delta d_L/d_L = 1/12$, beyond which we no longer record the triggers as detections. In contrast, its variance eventually decreases; this is due to the fact that, in a given bin at some intermediate redshift, we have events with different possible orientations with respect to the detector, and therefore different SNR. As the redshift increases toward the horizon of the detector, only the events with optimal orientation can go above the threshold.

It is useful to provide a fit for the mean value and for the standard deviation of $\Delta d_L/d_L$, as a function of the redshift. In order to increase the significance of the fit, in the case of flat mass distribution the two highest redshift bins containing events of the catalog shown in Fig. 3.7 have been excluded (they only contain a very low number of events and we do not show them in the figure).

It is instructive to compare our fits to the instrumental error on $\Delta d_L/d_L$, with the fit to the instrumental error found in [181], which is

$$\overline{\left(\frac{\Delta d_L}{d_L}\right)} = 0.1449z - 0.0118z^2 + 0.0012z^3\,, \tag{3.2.13}$$

as well as with the lensing error (3.2.9). The comparison is shown in Fig. 3.9. We see

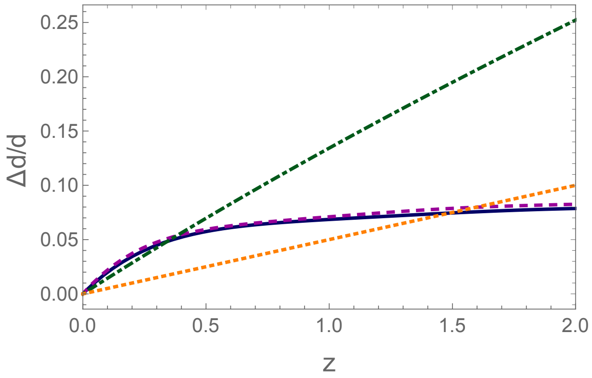

FIGURE 3.9: Our fits to $\Delta d_L/d_L$ for the flat distribution of neutron star masses (blue solid line) and for the Gaussian case (magenta dashed line), compared to the error due to lensing given in eq. (3.2.9) (orange dotted line) and to the fit (3.2.13) to the instrumental error given in ref. [181] (green dot-dashed line). From [210].

first of all that our fits to the instrumental error are consistent with that of ref. [181] at low redshifts, say $z \lesssim 0.3$, but are smaller at large redshift. The sensitivity curve for ET used in ref. [181] [see their eq. (19)] is relatively close to the one that we use, shown in Fig. 3.5. The reason for the difference is rather that we use a SNR threshold $\rho_{\rm threshold} = 12$, while ref. [181] uses $\rho_{\rm threshold} = 8$. On the one hand, this implies that we get less events, and up to smaller redshift, compared to ref. [181]. On the other hand, since $\Delta d_L/d_L$ is estimated as 1/SNR, the events that we retain are those with smaller value of $\Delta d_L/d_L$ and therefore also the average value of $\Delta d_L/d_L$ over the events in a redshift bin is smaller. This effect becomes more important as z increases, as there are many more events that are sub-threshold with respect to $\rho_{\rm threshold} = 12$, but above threshold with respect to $\rho_{\rm threshold} = 8$.

It is also interesting to compare the lensing contribution to $\Delta d_L/d_L$ to the observational error, in the light of this understanding of the dependence of the average observational error on the threshold. Comparing the fit obtained in ref. [181] with $\rho_{\rm threshold} = 8$, given by the green dot-dashed line in Fig. 3.9, to the lensing error, given by the orange dotted line, one would be tempted to conclude that, at ET, the error induced by lensing is negligible at all redshifts. In fact, this statement needs some qualification. The green dot-dashed line represents the average of $\Delta d_L/d_L$ over an ensemble of events, in the same redshift bin, selected by requiring SNR $\geqslant \rho_{\rm threshold} = 8$. However, within each bin, the events that are most useful for cosmological studies are those with the smallest errors on $\Delta d_L/d_L$, i.e. with the highest SNR. Already selecting only the events with SNR $\geqslant \rho_{\rm threshold} = 12$ reduces significantly the instrumental error averaged over such events, and we see that for this ensemble of events lensing becomes larger than the mean value of $\Delta d_L/d_L$ at $z \gtrsim 1.5$.

We have seen above that the vast majority of the GW-GRB coincidences are at $z < 1$. Therefore, in our analysis, lensing will indeed be subdominant. However, it must be kept in mind that, if one restricts the analysis to the 'golden events', i.e. the loudest and best characterized events in each redshift bin (corresponding to the lower edge of the cyan shaded area in Fig. 3.8), the effect of lensing will become more and more important. The contribution to the error on $\Delta d_L/d_L$ from lensing, estimated as in eq. (3.2.9), is shown as the orange dotted line in Fig. 3.8. We see that, for $z < 1$, it is comparable to the lower edge of the distribution of events given by the cyan shaded area, and therefore its inclusion, in quadrature with the observational

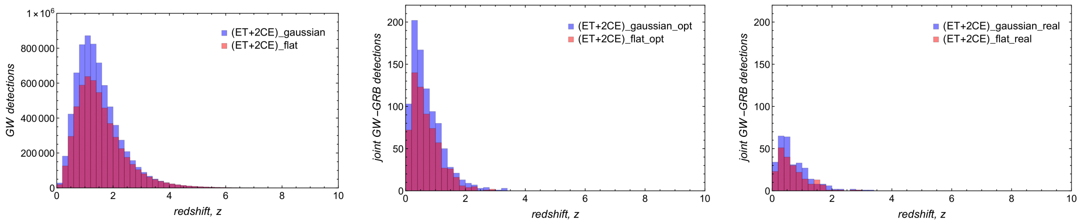

FIGURE 3.10: As in Fig. 3.7 for the ET+CE+CE network and its coincidences with THESEUS. For uniformity, we use the same redshift range in the three panels. From [210].

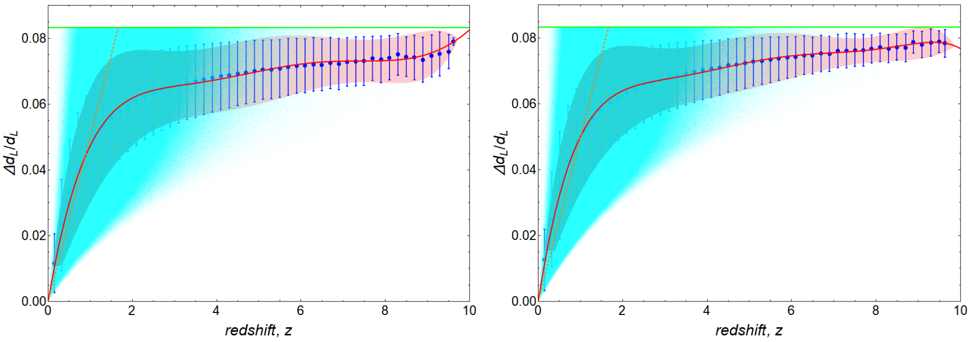

FIGURE 3.11: As in Fig. 3.8, for the ET+CE+CE network. From [210].

error, at most degrades by a factor $\sqrt{2}$ the error on $\Delta d_L/d_L$ on these very few events at the edge of the distribution. These events are interesting because they are those with the smallest error, but for $z < 1$ they are extremely rare. For larger redshifts the situation is different and we see that, say at $z \simeq 1.5$, for a non-negligible fraction of the events the error from lensing can be the limiting factor.

Events at a ET+CE+CE network.

In the second line of Table 3.3 we show the results for the ET+CE+CE network. This configuration, featuring three 3G detectors, can be considered an extreme case. Other cases, such as one ET detector and two advanced 2G detectors, will be intermediate between this case and a single 3G detector. The result for the GW events in Table 3.3 in the Gaussian case corresponds to a detection rate of 710k events/yr. To understand the dependence of these results on the astrophysical assumptions, we have also generated a catalog of GW events, for ET+CE+CE, assuming a Madau-Dickinson star formation rate and an exponential time delay between formation and merger with an e-fold time of 100 Myr. In that case we find 840k events/yr, in good agreement with the value 990k events/yr found, under similar hypothesis, but a slightly larger rate $R_m(z=0) = 1000\,\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$, in ref. [239].

For the coincidences with GRBs we find that, for an ET+CE+CE network, we would get of order $64-90$ coincident events per year, and around $20-30$ events per year will have arcmin localisations. Observe that, even if the ET+CE+CE network has a number of GW detections larger than a single ET by a factor $\mathcal{O}(10)$, the number of coincidences with GRBs is higher only by a factor less than 2. This already tells us that the bottleneck, for joint GW-GRB detections, is on the GRB side, that cannot keep pace with the GW detection rate of 3G detectors.

Figure 3.10 shows the redshift distributions of BNS detections using ET+CE+CE, over 10 yr, along with the joint GW-GRB distributions with optimistic and realistic scenarios for the FOV of THESEUS. In this realization of the catalog, the ET+CE+CE

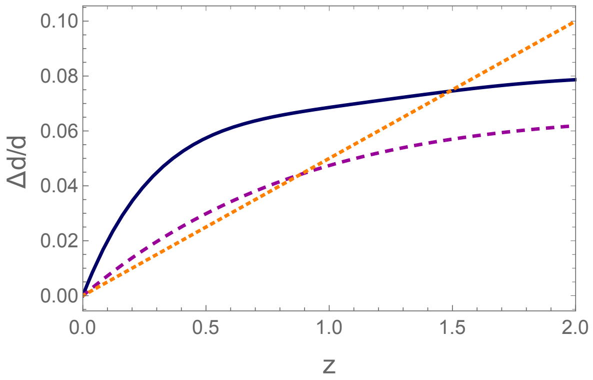

FIGURE 3.12: The fit to the average value of the instrumental error on $\Delta d_L/d_L$ for ET (blue solid line) and for ET+CE+CE (magenta dashed line) (both, for definiteness, for the flat mass distribution), compared to the error due to lensing (orange dotted line). From [210].

event with the highest redshift is at $z \simeq 9.63$ for the flat mass distribution and at $z \simeq 9.66$ for the Gaussian mass distribution; this is due to the fact that a single CE already has a reach to BNS of order $z \simeq 9$ (while we have seen above that, for BNS, ET alone reaches $z \simeq 2-3$). For the joint GW-GRB detections we find that the one with the highest redshift is at $z \simeq 3.38$ (found both in optimistic and realistic scenario for the THESEUS FOV, and for the Gaussian mass distribution). Notice that, in the left panel, events with $z \gtrsim 6$ are not visible on the vertical scale used in the figure, but are indeed present.

Fig. 3.11 shows the instrumental error for the luminosity distance in our catalog of events for ET+CE+CE, with the same meaning of the lines and shaded areas as in Fig. 3.8. The distribution now extends to much higher redshifts, because, as we have seen, CE (assuming the current design configuration, with 40 km arms) has a much larger horizon to BNS.

However, we also see from Fig. 3.11 that the contribution to the error from lensing, given by the dotted orange line, becomes quickly dominant at $z \gtrsim 1.5$ [assuming that the linear extrapolation (3.2.9) is still correct at large redshifts]. For the study of joint GW-GRB detections that we have performed this has a limited impact, as for ET+CE+CE the bulk of the joint GW-GRB detections is at $z \lesssim (1-1.5)$. Given the localization capability of the ET+CE+CE network, one could still hope to extract cosmological information from the very large number of standard sirens at larger redshifts, through statistical methods. However, at these redshifts the dominant contribution to $\Delta d_L/d_L$ will come from lensing (similarly to what happens for LISA). This is clearly seen from Fig. 3.12, where we plot the instrumental error $\Delta d_L/d_L$ at ET and at ET+CE+CE, in the range $z < 2$ where they can be compared, and the error due to lensing. We see that, even if the instrumental error from the ET+CE+CE network is obviously better than for a single ET, above $z \simeq 1$ the error in ET+CE+CE starts to be dominated by lensing, so that at $z \gtrsim 1.5$ both ET and ET+CE+CE are dominated by lensing and therefore eventually this becomes the limiting factor in both configurations.

This limitation could however be turned into a virtue. The situation is indeed similar to the one that was discussed a decade ago in ref. [179], in the context of a study for the Big Bang Observer (BBO). It was found that also in BBO the error is almost entirely dominated by lensing. This means that, once one has determined the dependence of d_L on z from the full ensemble of sources (and possibly by combining standard sirens with CMB, BAO and SNe), the scatter around this mean value for each single BNS event is basically a measurement of the gravitational lensing magnification along that line of sight. Given the very large number of sources, this will produce a map of the lensing magnification across the sky; in particular, the corresponding two-point correlation function gives a measurement of the convergence power spectrum, and provides important information on cosmological structure formation.

3.2.2 Constraints on ΛCDM parameters

In this section we study how the addition of the standard sirens with GRB counterpart that could be observed with the HLVKI network or with 3G detectors, would contribute to the knowledge of cosmological parameters in ΛCDM. The measurement of the luminosity distances from a set of coalescing binaries therefore gives constraints on H_0 and Ω_M.

The most accurate results are obtained by combining the constraints from standard sirens with other cosmological datasets such as CMB, BAO and SNe, to remove the degeneracies between cosmological parameters. A priori it would be interesting to study also the constraints that emerge using only standard sirens, that, even if less constraining, are conceptually interesting because they have systematics completely different from those of electromagnetic observations. However, for 2G detectors with a counterpart identified through a GRB, we find that the number of sources is too small to obtain significant results from standard sirens alone, and we will present only the results obtained by combining standard sirens with CMB+BAO+SNe. For 3G detectors, we will also show the separate results from standard sirens. When combining standard sirens with CMB, BAO and SNe, we use the following datasets:

- *CMB.* We use the 2015 *Planck* [247] measurements of the angular (cross-)power spectra, including full-mission lowTEB data for low multipoles ($\ell \leqslant 29$) and the high-ℓ Plik TT,TE,EE (cross-half-mission) ones for the high multipoles ($\ell > 29$) of the temperature and polarization auto- and cross- power spectra [209]. We also include the temperature+polarization (T+P) lensing data, using only the conservative multipole range $\ell = 40 - 400$ [248, 249].

- *Type Ia supernovae.* We use the JLA data for SN Ia provided by the SDSS-II/SNLS3 Joint Light-curve Analysis [140].

- *Baryon Acoustic Oscillations* (BAO). We use the isotropic constraints provided by 6dFGS at $z_{\rm eff} = 0.106$ [141], SDSS-MGS DR7 at $z_{\rm eff} = 0.15$ [142] and BOSS LOWZ at $z_{\rm eff} = 0.32$ [250], as well as the anisotropic constraints from CMASS at $z_{\rm eff} = 0.57$ [250].

We then run a MCMC, similarly to what we have done for LISA related work, see in particular 3.4.2 for a description of the likelihood used for standard sirens.

Since standard sirens, within ΛCDM, are only sensitive to H_0 and Ω_M, we focus on the two-dimensional likelihoods in the (Ω_M, H_0) plane (although, of course, when we combine standard sirens with CMB+BAO+SNe, the fact that the addition of standard sirens allows a more accurate determination of H_0 and Ω_M also has a beneficial effects on the determination of the other parameters, because it helps to reduce the degeneracies).

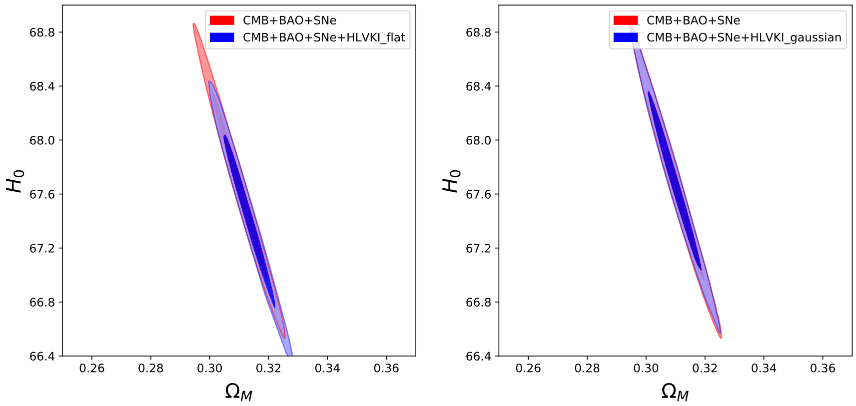

FIGURE 3.13: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, H_0) plane, in ΛCDM, from CMB+BAO+SNe (red), and the result obtained by combining standard sirens at the HLVKI network with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. From [210].

Results for the HLVKI network

The result is shown in Fig. 3.13, where we compare the likelihood in the (Ω_M, H_0) plane obtained from a MCMC using the above CMB+BAO+SNe dataset (red contours) with those obtained from the combined datasets CMB+BAO+SNe+standard sirens (blue contours), for the two distributions of neutron star masses. The contours from standard sirens only are not shown since the MCMC fails to converge due to the small number of sources.

In table 3.4 we show the relative errors $\Delta H_0/H_0$ and $\Delta\Omega_M/\Omega_M$ obtained from the corresponding one-dimensional marginalized likelihood. We see that, for the HLVKI network, the addition of joint GW-GRB detections to the current cosmological dataset does not improve substantially the accuracy on H_0 and Ω_M. This should be traced to the fact that, as we see from Table 3.2, if we perform coincidences of the GW events with GRBs, we only have of order 15 joint detections, even over a 10-yr period. On the other hand, at the typical redshifts of the events seen by 2G detectors, the network of optical and infrared telescopes is expected to provide many more electromagnetic counterparts. In our analysis, even for 2G detectors we have restricted ourselves to a study of coincidences with GRBs only, also as a benchmark for the study of GW-GRB coincidences with 3G detectors. However, it is clear that for 2G detectors the contribution of optical/IR telescopes will be crucial and could raise substantially the number of standard sirens with observed counterpart.

Our results are broadly consistent with the analysis of [241, 251], which show that, to obtain a measurement of H_0 below 1% at 2G detectors with standard sirens only, $\mathcal{O}(50-100)$ standard sirens with counterpart are needed. This would allow to address the discrepancy between the local measurement of H_0 and the value inferred by *Planck* and BAO observations assuming ΛCDM, which has now reached the 4.4σ level [145].

Results for ET

We next consider the case of joint detections between a single ET detector and a GRB detected by THESEUS, again restricting at first to ΛCDM. The result is shown

	CMB+BAO+SNe	combined, flat	combined, gaussian
$\Delta H_0/H_0$	0.72%	0.65%	0.66%
$\Delta\Omega_M/\Omega_M$	2.11%	1.91%	1.96%

TABLE 3.4: Accuracy (1σ level) in the reconstruction of H_0 and Ω_M with CMB+BAO+SNe only, and the combined result CMB+BAO+SNe+standard sirens, using the HLVKI detector network and the flat and gaussian mass distributions. From [210].

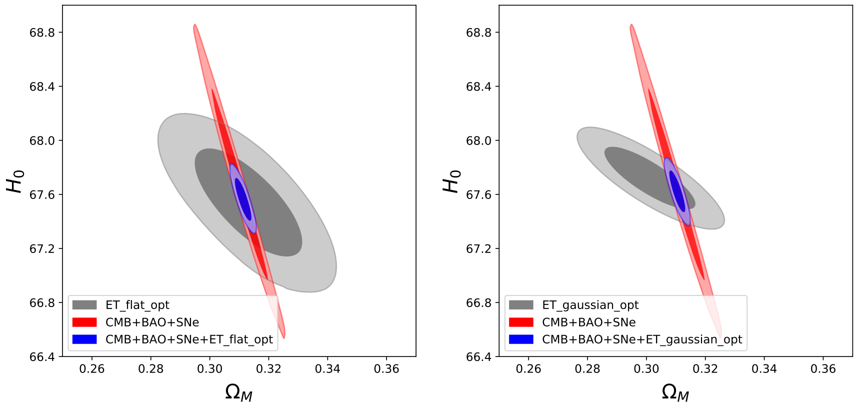

FIGURE 3.14: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, H_0) plane, in ΛCDM, from CMB+BAO+SNe (red), standard sirens at ET with a GRB counterpart determined by THESEUS (gray), and the result obtained by combining standard sirens with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

in Figs. 3.14 and 3.15, where we compare the likelihood in the (Ω_M, H_0) plane obtained from a MCMC using our CMB+BAO+SNe dataset (red contours) with those obtained from standard sirens only (gray contour) and the combined datasets CMB + BAO + SNe + standard sirens (blue contours). In particular, in Fig. 3.14 we show the result for the two distribution of neutron star masses in the optimistic scenario for the FOV of THESEUS, while in Fig. 3.15 we show the result for the two distribution of neutron star masses in the realistic scenario for the FOV.

First of all observe that, despite the fact that the mock catalog of standard sirens has been generated by taking as fiducial cosmological model ΛCDM with the values of H_0 and Ω_M obtained from these CMB+BAO+SNe data, the contour obtained from standard sirens only is not always centered on the mean values given by the CMB+BAO+SNe contour. This is an unavoidable consequence of the fact that, in order to simulate the result from actual observations, we have scattered the values of $d_L(z)$ with a variance $\Delta d_L(z)$ given by the expected observational error, see the discussion in Section 3.2.1. Of course, this is the situation that would take place in an actual observation (although the actual position of the gray contour will depend on the particular realization of the random scattering of the data around their mean value), and all that we should expect is that the contours are consistent at, say, the $(1-2)\sigma$ level, which is indeed the case. However, one should be aware of the fact that, if we combine the standard sirens and the CMB+BAO+SNe dataset in a realization where the two contours do not overlap well, we get a larger error on the

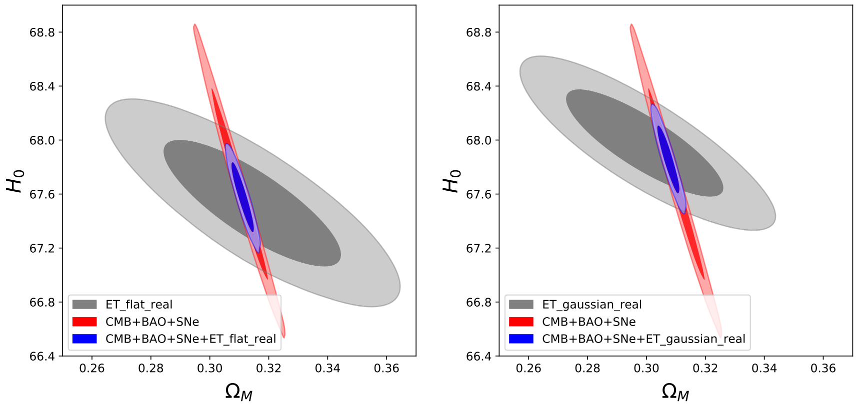

FIGURE 3.15: As in Fig. 3.14, with the realistic estimate for the FOV of THESEUS. From [210].

	CMB+BAO+SNe	ET, flat	ET, gaussian	combined, flat	combined, gaussian
$\Delta H_0/H_0$	0.72%	0.28%	0.23%	0.16%	0.15%
$\Delta\Omega_M/\Omega_M$	2.11%	3.68%	3.38%	0.59%	0.57%

TABLE 3.5: Accuracy (1σ level) in the reconstruction of H_0 and Ω_M with only CMB+BAO+SNe, with only standard sirens (with the flat and gaussian mass distributions, respectively) and the combined results CMB+BAO+SNe+standard sirens, using ET and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

parameters with respect to what is obtained in a realization where the contours happens to overlap well. Once again, this is exactly the situation that will be faced in the actual experiment.

The accuracy on H_0 and Ω_M from the corresponding one-dimensional likelihood are given in Table 3.5 (for the optimistic FOV of THESEUS) and Table 3.6 (for the realistic FOV). From these tables we see that standard sirens at ET, already before combining them with other cosmological datasets, give an accuracy on H_0 between 0.2% and 0.4%, depending on the scenarios considered. This is a very interesting accuracy, that would allow to conclusively arbitrate the tension between the local H_0 measurement and the *Planck*-ΛCDM value, with totally different systematic compared to SNe. In particular, this measurement would have the potential of falsifying ΛCDM.

If instead the value of H_0 from standard sirens should agree with the CMB + BAO + SNe value obtained using ΛCDM, it would then make sense to combine these datasets. As we see from the tables, in that case the overall accuracy on H_0 could reach $(0.15 - 0.25)\%$.

Results for ET+CE+CE

We finally consider the ET+CE+CE network. The two-dimensional likelihoods in the (Ω_M, H_0) plane are shown in Figs. 3.16 and 3.17, and the corresponding 1σ accuracies from the one-dimensional likelihoods are shown in Tables 3.7 (for the optimistic FOV of THESEUS) and Table 3.8 (for the realistic FOV). We see that, in this case, with standard sirens only we get an accuracy of H_0 of about 0.2%, while, combining with the other cosmological datasets we can reach an accuracies of order $(0.07 - 0.12)\%$.

	CMB+BAO+SNe	ET, flat	ET, gaussian	combined, flat	combined, gaussian
$\Delta H_0/H_0$	0.72%	0.42%	0.39%	0.26%	0.25%
$\Delta\Omega_M/\Omega_M$	2.11%	6.17%	5.88%	0.82%	0.82%

TABLE 3.6: As in Table 3.5, assuming the realistic FOV of THESEUS. From [210].

While in itself this would be a remarkable accuracy, still it is not significantly better than that reached with a single ET detector, as we see comparing with Tables 3.5 and 3.6. This result can be understood by looking at the number of events in our catalogs, shown in Table 3.3. Despite the fact that the ET+CE+CE network has a number of GW detections higher by a factor $\mathcal{O}(10)$ compared to a single ET, when we look at joint GW-GRB detections the increase in the number of events is less than a factor of 2. In other worlds, the bottleneck here is on the GRB side. It is crucial to observe, however, that a three detector network such as ET+CE+CE will have excellent localization capabilities. Thus, at least for the events at $z \lesssim 0.5$, the follow-up with optical and IR telescope will be possible, and will probably lead to a significant increase in the number of standard sirens with electromagnetic counterpart. As we discussed in Section 3.2.1, realistic estimates are currently difficult because they also depend on choices such as the amount of telescope time that will be devoted by the various facilities to the follow-up of GW events, and here we have not attempted such an estimate. However, it should be borne in mind that, for a ET+CE+CE network, the joint GW-GRB detections that we are considering in this analysis might provide just a fraction of the whole sample of GW signals with electromagnetic counterpart. Correspondingly, the accuracies that can be obtained on H_0 and Ω_M at ET+CE+CE could be significantly better, compared to the figures that we find.

	CMB+BAO+SNe	ET+CE+CE, flat	ET+CE+CE, gaussian	combined, flat	combined, gaussian
$\Delta H_0/H_0$	0.72%	0.20%	0.22%	0.07%	0.07%
$\Delta\Omega_M/\Omega_M$	2.11%	1.43%	1.31%	0.43%	0.42%

TABLE 3.7: Accuracy (1σ level) in the reconstruction of H_0 and Ω_M with only CMB+BAO+SNe, with only standard sirens (with the flat and gaussian mass distributions, respectively) and the combined results CMB+BAO+SNe+standard sirens, using ET+CE+CE and THESEUS, and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	ET+CE+CE, flat	ET+CE+CE, gaussian	combined, flat	combined, gaussian
$\Delta H_0/H_0$	0.72%	0.24%	0.23%	0.12%	0.11%
$\Delta\Omega_M/\Omega_M$	2.11%	2.12%	2.09%	0.51%	0.52%

TABLE 3.8: As in Table 3.5, assuming the realistic FOV of THESEUS. From [210].

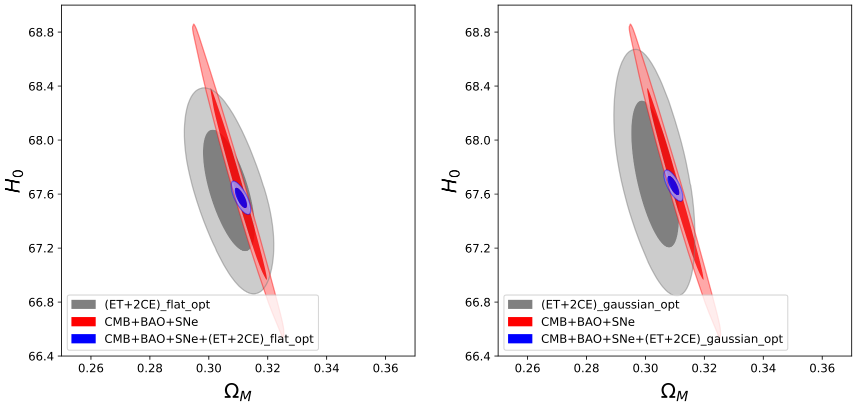

FIGURE 3.16: As in Fig. 3.14, for the ET+CE+CE network and optimistic FOV of THESEUS. From [210].

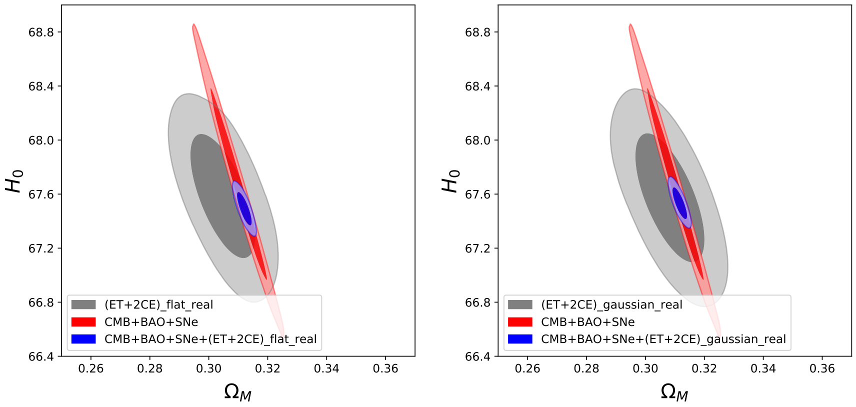

FIGURE 3.17: As in Fig. 3.15, for the ET+CE+CE network and realistic estimate for the FOV of THESEUS. From [210].

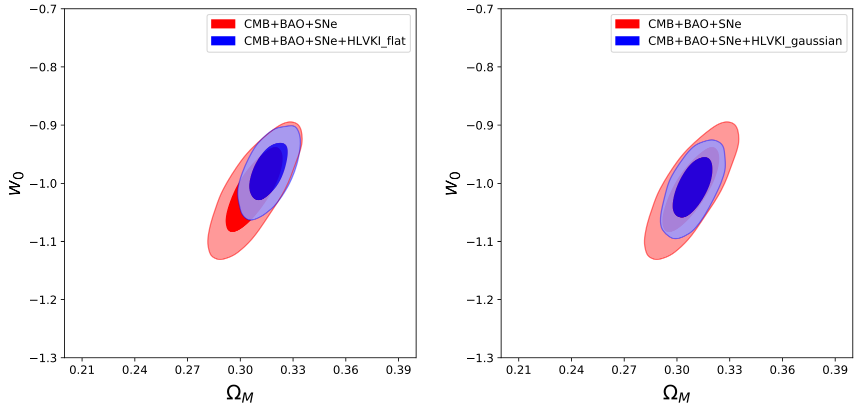

FIGURE 3.18: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, w_0) plane, in wCDM, from CMB+BAO+SNe (red), and the result obtained by combining standard sirens at the HLVKI network with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. From [210].

3.2.3 Constraints on dark energy

Results for the HLVKI network

We begin by presenting the constraints that can be obtained on the DE sector by combining standard sirens at the HLVKI network with the CMB+BAO+SNe dataset described in Section 3.2.2.

Fig. 3.18 shows the likelihood in the (Ω_M, w_0) plane in wCDM, i.e. when we introduce w_0 as the only new parameter that describes the DE sector, while setting $w_a = 0$ and excluding also modified GW propagation, i.e setting $\Xi_0 = 1$, while Table 3.9 shows the error on w_0 (at 1σ, as in all our tables) from the corresponding one-dimensional likelihood. We give the result from CMB+BAO+SNe only, and that obtained by combining CMB+BAO+SNe with standard sirens with flat mass distribution or with the gaussian mass distribution.

The results for the (w_0, w_a) parametrization are shown in Fig. 3.19 and Table 3.10. Of course, enlarging the parameter space with one more parameter w_a results in a larger error on w_0, compared to the results in Table 3.9.

We finally consider the (Ξ_0, w_0) extension of the DE sector. The results are shown in Fig. 3.20 and Table 3.11. The most interesting result is the one for Ξ_0, which (with our rather extreme assumption of 10 yr of data taking) can be measured to an accuracy of order $\Delta\Xi_0 \simeq 0.1$, i.e. (given that our fiducial value has been taken to be the ΛCDM value $\Xi_0 = 1$), a relative accuracy $\Delta\Xi_0/\Xi_0 \simeq 10\%$, which is already in the ballpark of the predictions of interesting modified gravity models [200]. With a shorter but more realistic time of data taking, say 3-4 yr, we still expect to get $\Delta\Xi_0/\Xi_0$ at the level of about 20%.

Of course, CMB, BAO and SNe are blind to modified GW propagation, and the corresponding contour is flat in the Ξ_0 direction. Standard sirens lift this flat direction. In contrast, we see that the improvement on w_0 or w_a from the inclusion of standard sirens is quite modest.

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.033	0.035

TABLE 3.9: Accuracy (1σ level) in the reconstruction of w_0 with CMB+BAO+SNe only, and the combined result CMB+BAO+SNe+standard sirens, using the HLVKI detector network and the flat and gaussian mass distributions. From [210].

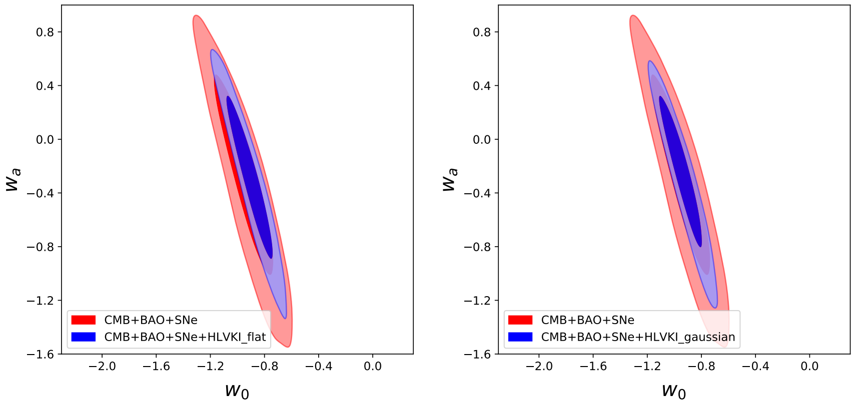

FIGURE 3.19: The 1σ and 2σ contours of the two-dimensional likelihood in the (w_0, w_a) plane, in the (w_0, w_a) extension of the DE sector, from CMB+BAO+SNe (red), and the result obtained by combining standard sirens at the HLVKI network with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. From [210].

Results for ET

We next consider a single ET in coincidence with a GRB detector with the characteristics of THESEUS. Proceeding as before, we show first the results of the w_0 extension, in Tables 3.12 and 3.13 and Figs. 3.21 and 3.22, displaying separately the results for the optimistic and realistic FOV of THESEUS. Observe that, contrary to the HLVKI case, now we can obtain some bounds already using standard sirens alone. However, we see from the figures that the central value of the contour of the standard sirens can happen to be displaced with respect to that from CMB+BAO+SNe. As we already discussed in Section 3.2.2, this is a statistical effect due to the random scattering of the mock GW data according to the error estimate that, depending on the specific realization, can induce a more or less significant displacement between the two contours.

We see that, adding the joint GW-GRB events to the CMB+BAO+SNe dataset, we can improve the accuracy on w_0 by about a factor of 2. This is interesting, although certainly not spectacular. We will see below that the most interesting contribution of a 3G detector such as ET to the exploration of the DE sector rather comes from modified GW propagation.

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.140	0.113	0.106
Δw_a	0.483	0.406	0.380

TABLE 3.10: Accuracy (1σ level) in the reconstruction of w_0 and w_a with CMB+BAO+SNe only, and the combined result CMB+BAO+SNe+standard sirens, using the HLVKI detector network and the flat and gaussian mass distributions. From [210].

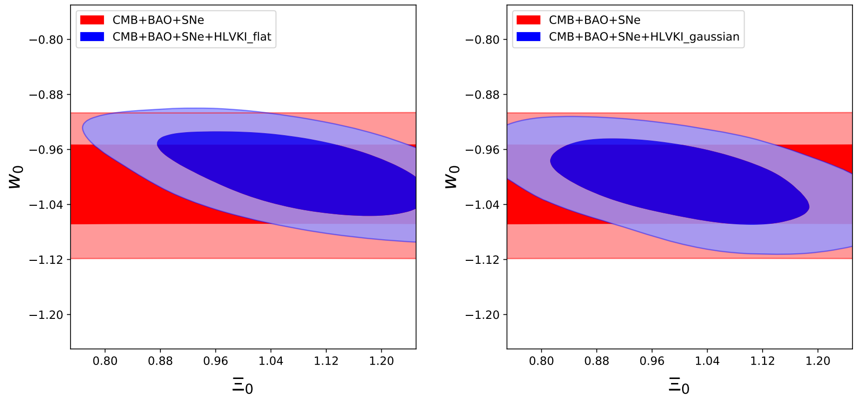

FIGURE 3.20: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ξ_0, w_0) plane, in the (Ξ_0, w_0) extension of the DE sector, from CMB+BAO+SNe (red), and the result obtained by combining standard sirens at the HLVKI network with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. From [210].

We next include (w_0, w_a) as extra parameters. The results are shown in Tables 3.14 and 3.15 and in Figs. 3.23 and 3.24. In this case, with one extra parameter w_a, for some scenarios the MCMC does not converge well with standard sirens only, and we just show the combined result for CMB+BAO+SNe+standard sirens, for all scenarios.

We finally include modified GW propagation, adding to the baseline ΛCDM model the parameters (Ξ_0, w_0), and writing the GW luminosity distance as in eq. (3.1.30). We set for simplicity $n = 5/2$, which is of the order of the value predicted by the RT and RR nonlocal gravity model. However, the precise value of n is of limited relevance for the analysis. Note that, since our catalog of sources has been generated assuming ΛCDM as our fiducial model, our fiducial values for these parameters are $\Xi_0 = 1$ and $w_0 = -1$, and we compute the accuracy $\Delta\Xi_0$ and Δw_0 to which we can find back these values. The results are shown in Tables 3.16 and 3.17 and in Figs. 3.25 and 3.26.

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.042	0.042
$\Delta \Xi_0$	–	0.130	0.125

TABLE 3.11: Accuracy (1σ level) in the reconstruction of w_0 and Ξ_0 with CMB+BAO+SNe only, and the combined result CMB+BAO+SNe+standard sirens, using the HLVKI detector network and the flat and gaussian mass distributions. From [210].

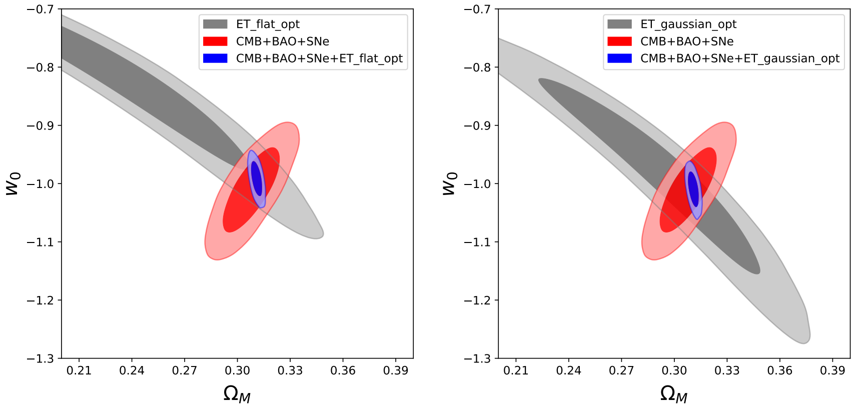

FIGURE 3.21: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, w_0) plane, in wCDM, from CMB+BAO+SNe (red), joint detection of standard sirens at ET and THESEUS (gray) and the result obtained by combining standard sirens with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

It is quite remarkable that, by combining the joint GW-GRB detections from a single ET detector and a GRB detector such as THESEUS, with the current CMB + BAO + SNe dataset, we can reach an accuracy on Ξ_0 that, depending on the scenario, is between 0.7% and 1.1%. By comparison, the 'minimal' RT nonlocal model predicts a deviation from $\Xi_0 = 1$ at the level of 6.6%, almost an order of magnitude larger than this observational sensitivity (and much large values can be obtained with initial conditions set during inflation). This shows that, while the sensitivity of a single ET detector to w_0 will not allow us to obtain a dramatic improvement on the current knowledge of w_0, the sensitivity to Ξ_0 is extremely interesting and well within the prediction of viable modified gravity models. Furthermore, modified GW propagation, as encoded for instance in the Ξ_0 parameter [or, more generally, in the (Ξ_0, n) parameters] is an observable specific to GW detectors, to which electromagnetic observations are blind.

	CMB+BAO+SNe	ET, flat	ET, gaussian	combined, flat	combined, gaussian
Δw_0	0.045	0.109	0.116	0.020	0.021

TABLE 3.12: Accuracy (1σ level) in the reconstruction of w_0 with only CMB+BAO+SNe, with only standard sirens (with the flat and gaussian mass distributions, respectively) and the combined results CMB+BAO+SNe+standard sirens, using ET and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

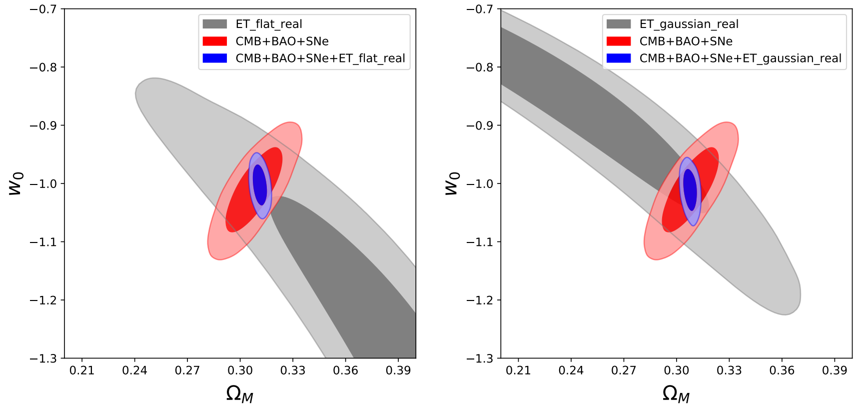

FIGURE 3.22: As in Fig. 3.21, with the realistic estimate for the FOV of THESEUS. From [210].

The results of this section can be compared to those in [171], which, following [180, 181, 187], were obtained under the working hypothesis that ET, over a few years of data taking, will eventually be able to collect $\mathcal{O}(10^3)$ BNS with counterpart (without specifying how the counterpart is actually detected), and assuming a redshift distribution proportional to a simple model for the star formation rate (and neglecting the effect of the delay between binary formation and merger). For the (w_0, Ξ_0) extension of the DE sector, the analysis of ref. [171] then led to the forecast $\Delta\Xi_0 = 0.008$ and $\Delta w_0 = 0.032$, which happens to be very close to the results in Table 3.16.

Comparing with our present results, first of all we see from Table 3.3 that the assumption of 10^3 standard sirens with counterpart, while optimistic, was not unrealistic; in particular, for a gaussian neutron-star mass distribution (and with the optimistic estimate for the FOV of THESEUS), in 10 yr of data we could have $\mathcal{O}(500)$ joint GW-GRB events. Furthermore, as we have already mentioned in Section 3.2.1, in [181] was used a threshold for the network SNR obtained combining the three arms of ET given by $\rho_{\rm threshold} = 8$ while we use $\rho_{\rm threshold} = 12$. Lowering the threshold would lead to an increase in the number of GW events. Still, at first it could be surprising that the result for $\Delta\Xi_0$ that we find in this analysis happens to be practically identical to that of ref. [171], given that the number of sources that we are using here is smaller by a factor $\simeq 2$ compared to the 10^3 sources used in ref. [171]. However, this can be traced to the fact that also the redshift distribution of the sources is different. Indeed, we see from the right-panel in Fig. 3.7 that most of the joint GW-GRB detections are at $z < 0.5$, while in ref. [171] it was assumed that the catalog of sources followed a distribution in redshift determined by the star formation rate; that catalog was peaked at $z \simeq 1$, with long tails at larger z, see Fig. 8 of ref. [171]. On the other hand, it was also found in ref. [171] that the main contribution to the determination of Ξ_0 was given by the sources at $z < 0.7$, that were about one half of the total, so in the end it is not surprising that our catalog, with about a factor of 2

	CMB+BAO+SNe	ET, flat	ET, gaussian	combined, flat	combined, gaussian
Δw_0	0.045	0.301	0.158	0.023	0.024

TABLE 3.13: As in Table 3.12, assuming the realistic FOV of THESEUS. From [210].

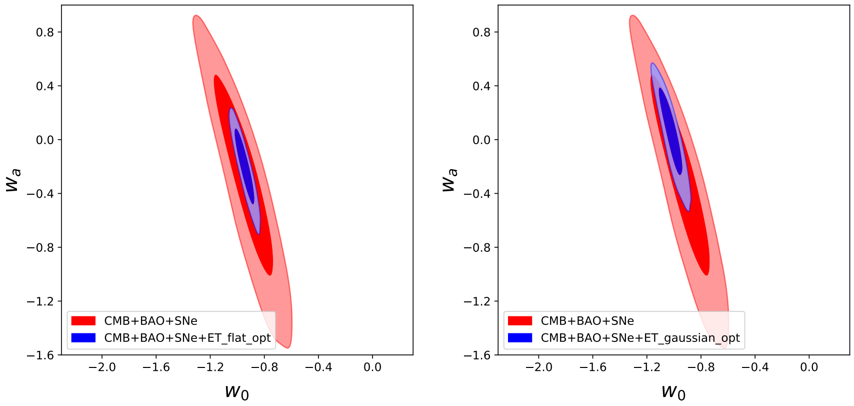

FIGURE 3.23: The 1σ and 2σ contours of the two-dimensional likelihood in the (w_0, w_a) plane from CMB+BAO+SNe (red) and the result obtained by combining standard sirens at ET with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

less sources, but almost all concentrated at $z < 0.7$, gives basically the same results as the catalog used in ref. [171].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.140	0.050	0.058
Δw_a	0.483	0.193	0.224

TABLE 3.14: Accuracy (1σ level) in the reconstruction of (w_0, w_a) with only CMB+BAO+SNe and the combined results CMB+BAO+SNe+standard sirens, using ET and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.140	0.073	0.072
Δw_a	0.483	0.246	0.260

TABLE 3.15: As in Table 3.14, assuming the realistic FOV of THESEUS. From [210].

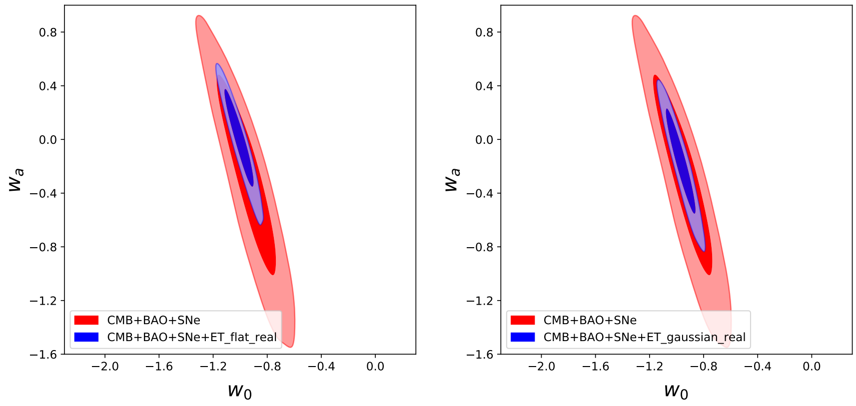

FIGURE 3.24: As in Fig. 3.23, with the realistic estimate for the FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.026	0.024
$\Delta \Xi_0$	–	0.008	0.007

TABLE 3.16: Accuracy (1σ level) in the reconstruction of (w_0, Ξ_0) with only CMB+BAO+SNe and the combined results CMB+BAO+SNe+standard sirens, using ET and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.026	0.026
$\Delta \Xi_0$	–	0.011	0.010

TABLE 3.17: As in Table 3.16, assuming the realistic FOV of THESEUS. From [210].

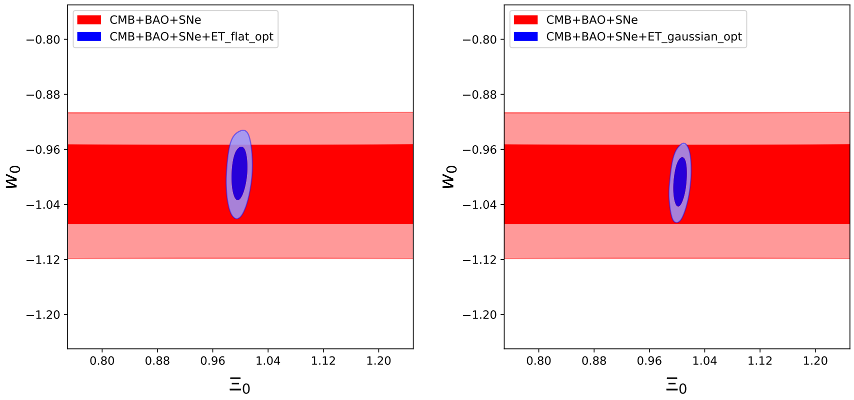

FIGURE 3.25: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ξ_0, w_0) plane from CMB+BAO+SNe (red) and the result obtained by combining joint detections of standard sirens at ET and THESEUS with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

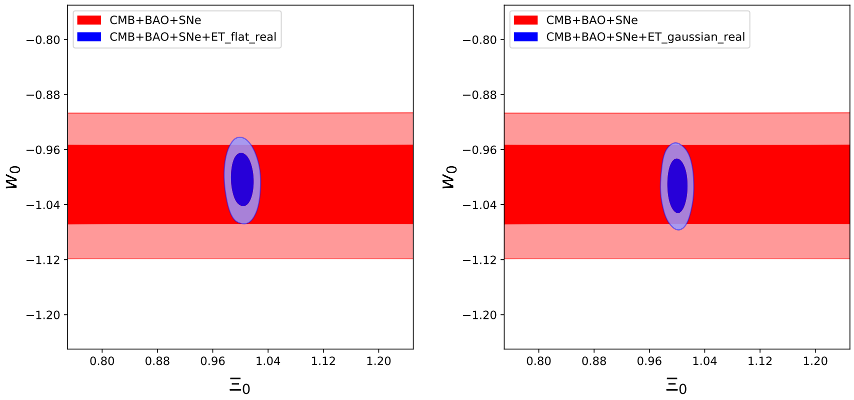

FIGURE 3.26: As in Fig. 3.25, with the realistic estimate for the FOV of THESEUS. From [210].

Results for ET+CE+CE

We finally give the corresponding results for the ET+CE+CE network. For wCDM the results are shown in Figs. 3.27 and 3.28 and Tables 3.18 and 3.19. For the (w_0, w_a) extension the results are shown in Figs. 3.29 and 3.30 and Tables 3.20 and 3.21. For the (Ξ_0, w_0) extension the results are shown in Figs. 3.31 and 3.32 and Tables 3.22 and 3.23.

As we already observed in section 3.2.2 when discussing parameter estimation in ΛCDM, the improvement in the accuracy of the cosmological parameters, compared to the ET-only case, is not very large, because, despite the fact that a network ET+CE+CE detects a number of source larger by an order of magnitude compared to a single ET (and to much larger redshift, see Table 3.3 and Fig. 3.7), the corresponding joint GW-GRB detections do not follow the same increase, because of intrinsic limitations in the GRB detections. In particular, as we discussed in section 3.2.1, despite the fact that the ET+CE+CE network can detect BNS up to $z \simeq 10$ (in our catalog the source with the largest redshift has $z \simeq 9.66$), the joint detections with GRB only reach, in our catalog, $z \simeq 3.38$; all higher redshift sources are lost because their GRB is beyond the flux limit for detection.

To fully exploit the potential of a ET+CE+CE network for standard sirens it is therefore crucial either to have a more powerful network of multi-messenger observations, for instance with IR/optical telescopes (that, guided by the localization capability of the ET+CE+CE network, could provide many more counterparts, at least the region $z < 0.5 - 0.7$ corresponding to their reach), or else one should resort to statistical methods for the determination of the host galaxy of standard sirens without electromagnetic counterpart. This is also an important message of our analysis.

	CMB+BAO+SNe	ET+CE+CE, flat	ET+CE+CE, gaussian	combined, flat	combined, gaussian
Δw_0	0.045	0.041	0.034	0.014	0.013

TABLE 3.18: Accuracy (1σ level) in the reconstruction of w_0 with only CMB+BAO+SNe, with only standard sirens (with the flat and gaussian mass distributions, respectively) and the combined results CMB+BAO+SNe+standard sirens, using ET+CE+CE and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	ET+CE+CE, flat	ET+CE+CE, gaussian	combined, flat	combined, gaussian
Δw_0	0.045	0.074	0.063	0.020	0.018

TABLE 3.19: As in Table 3.18, assuming the realistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.140	0.027	0.025
Δw_a	0.483	0.139	0.137

TABLE 3.20: Accuracy (1σ level) in the reconstruction of (w_0, w_a) with only CMB+BAO+SNe and the combined results CMB+BAO+SNe+standard sirens, using ET+CE+CE and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.140	0.041	0.037
Δw_a	0.483	0.160	0.145

TABLE 3.21: As in Table 3.20, assuming the realistic FOV of THESEUS. From [210].

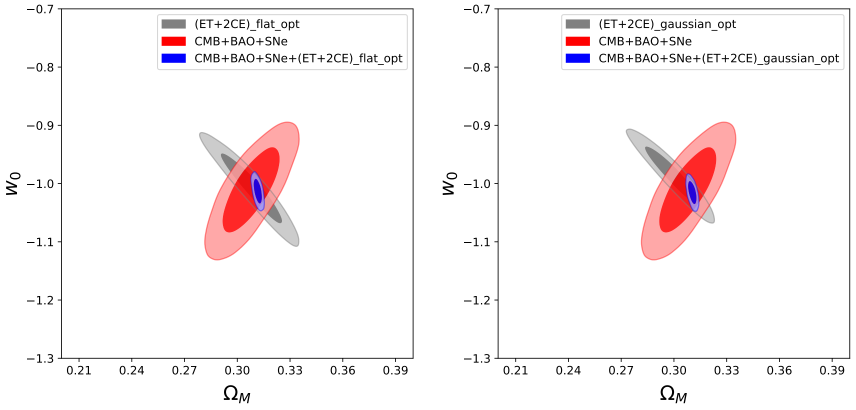

FIGURE 3.27: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, w_0) plane, in wCDM, from CMB+BAO+SNe (red), joint detection of standard sirens at ET+CE+CE and THESEUS (gray) and the result obtained by combining standard sirens with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

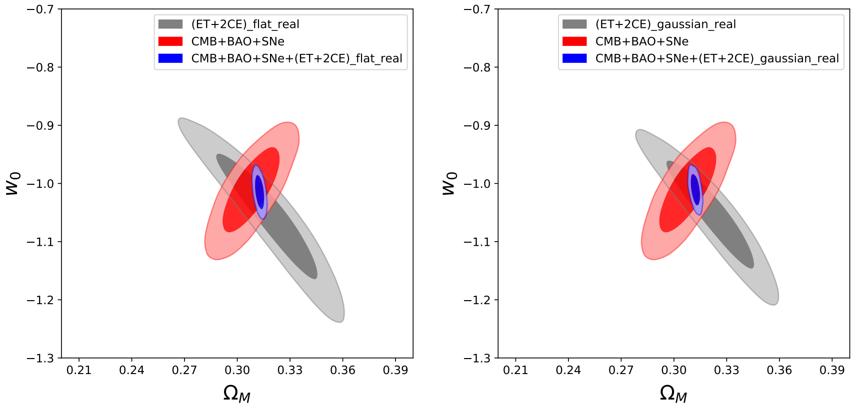

FIGURE 3.28: As in Fig. 3.27, with the realistic estimate for the FOV of THESEUS. From [210].

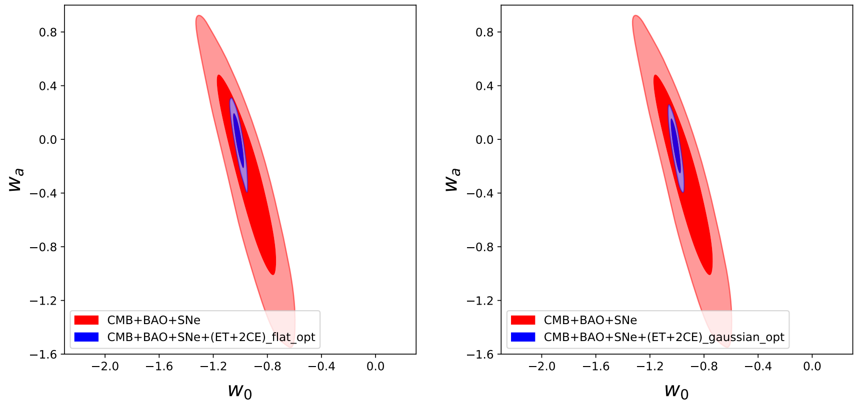

FIGURE 3.29: The 1σ and 2σ contours of the two-dimensional likelihood in the (w_0, w_a) plane from CMB+BAO+SNe (red) and the result obtained by combining standard sirens at ET+CE+CE with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

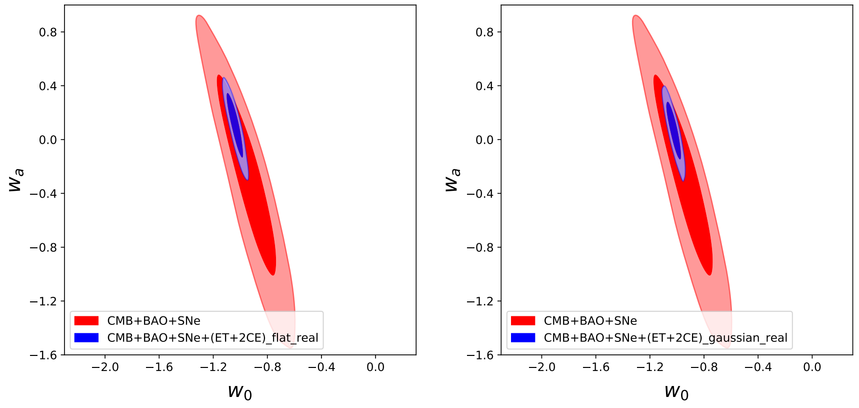

FIGURE 3.30: As in Fig. 3.29, with the realistic estimate for the FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.038	0.042
$\Delta\Xi_0$	–	0.007	0.007

TABLE 3.22: Accuracy (1σ level) in the reconstruction of (w_0, Ξ_0) with only CMB+BAO+SNe and the combined results CMB+BAO+SNe+standard sirens, using ET+CE+CE and THESEUS and assuming the optimistic FOV of THESEUS. From [210].

	CMB+BAO+SNe	combined, flat	combined, gaussian
Δw_0	0.045	0.030	0.033
$\Delta\Xi_0$	–	0.006	0.007

TABLE 3.23: As in Table 3.22, assuming the realistic FOV of THESEUS. From [210].

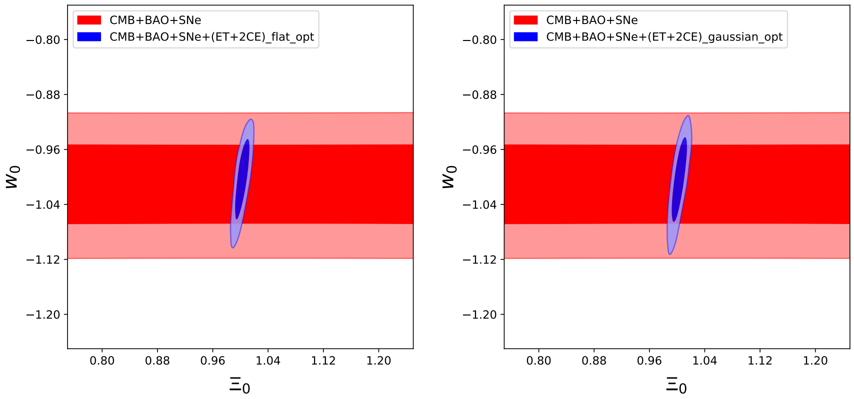

FIGURE 3.31: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ξ_0, w_0) plane from CMB+BAO+SNe (red) and the result obtained by combining standard sirens at ET+CE+CE with CMB+BAO+SNe (blue). Left: in the case of flat neutron star mass distribution. Right: in the case of gaussian neutron star mass distribution. We use the optimistic estimate for the FOV of THESEUS. From [210].

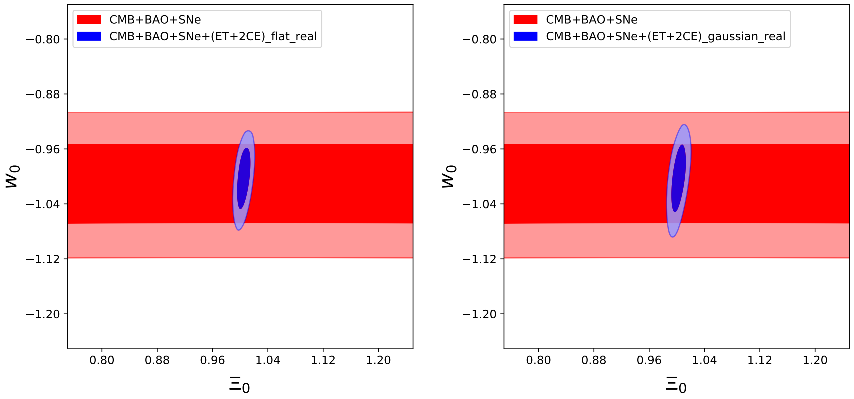

FIGURE 3.32: As in Fig. 3.31, with the realistic estimate for the FOV of THESEUS. From [210].

3.2.4 Summary of the 2G and 3G analysis

One of the main motivations for developing third-generation GW detectors, such as the Einstein Telescope in Europe and Cosmic Explorer in the US, is the possibility of using them for cosmology studies, in a way that will be complementary to what is done with electromagnetic probes such as the CMB, type Ia SNe, BAO, or large-scale structures. In particular, the measurement of the luminosity distance to coalescing binaries through the observation of their GWs, combined with an electromagnetic measurement of the redshift (or with statistical methods), gives access to the Hubble parameter H_0 and, for sources at sufficiently large redshift, also to the dark-energy equation of state. On top of this, a very interesting example of the potential contributions of advanced GW detectors to cosmology is given by the possibility of studying modified GW propagation.

The analysis of this Section improves on the results of ref. [171] for ET by using a more realistic catalog of sources, in the context of a current effort for building the Science Case for ET and, more generally, for 3G detectors. Indeed, in ref. [171] the sensitivity of ET to Ξ_0 was computed assuming that, over a few years of data taking, ET will be able to collect 10^3 standard sirens with observed electromagnetic counterpart [out of the $\mathcal{O}(10^5)$ BNS/yr that ET is expected to detect]. This is the assumption that has been usually made in the literature in this context but, in the absence of a concrete study of how to detect the counterpart, it is nothing more than a reasonable working hypothesis. Furthermore, the actual redshift distribution of the joint GW-electromagnetic detections depends not only on the specifications of ET, but also of the detectors used to observe the electromagnetic counterpart.

To go beyond these simple assumptions one needs a concrete scenario for the detection of the electromagnetic counterpart. Here we have focused on the possibility of a joint GW-GRB detection using, as an electromagnetic partner to 3G detector, the proposed THESEUS mission. As we have repeatedly stressed, this represents only a subset of the possible electromagnetic counterparts that could be observed, particularly for a network with significant localization capabilities such as HLVKI or ET+CE+CE (in which case it will probably not even provide the dominant contribution). However, given the large field of view of GRB satellites, in this case it is possible to give at least some estimates based uniquely on the characteristics of the detector. In contrast, the estimate for the number of counterparts detected at optical/IR telescope also strongly depend on issues, such as the fraction of telescope time that will be devoted to GW follow up, that are currently more difficult to predict.

Our main results are as follows. For the HLVKI network, the number of joint GW-GRB detections (computed assuming that the Fermi-GBM can make a coincident detection and that *Swift* can identify an X-ray counterpart) is quite small, of order 1.5/yr, and is not sufficient to obtains significant cosmological results using standard sirens only. Even when combined with the CMB+BAO+SNe dataset that we have used, it only provides a modest improvement in the accuracy on H_0. Indeed, to get interesting accuracy on H_0 with standard sirens, say of order 1% or better, as needed to solve the discrepancy between *Planck*-ΛCDM value and the local measurement of H_0, a much higher number of events with counterparts are indeed needed, $\mathcal{O}(50-100)$ [241, 251]. However, a 2G network such as HLVKI (or, indeed, already HLV) has significant localization accuracy. This number of sources with counterpart could therefore in principle be obtained through optical/IR follow up. Similarly, joint GW-GRB detections at the HLVKI network cannot be expected to improve significantly the accuracy on the DE equation of state, compared to what already know from electromagnetic observation, see Tables 3.9 and 3.10. The situation is however very different for what concern modified GW propagation, as encoded in the parameter Ξ_0. Since electromagnetic observations are blind to it, any result from 2G detectors will be potentially interesting, and we have found that (after combining with CMB+BAO+SNe to reduce the degeneracy with H_0 and Ω_M), the HLVKI network can measure Ξ_0 to about 13%. This results assumes 10 yr of data, which is probably a very optimistic assumption; however, taking e.g. 5 yr of data the result should be approximately worse by about a factor $\sqrt{2}$, so that a measurement at the 20% level should still be possible. This is already in the range of the predictions from some modified gravity models [200] and is therefore already a very interesting sensitivity.

For a single ET detector, our results for H_0 are given in Tables 3.5 and 3.6 and show that, already using only standard sirens, and limiting ourselves to the coincidences with GRBs (estimated using a detector with the characteristic of THESEUS), ET can reach an extremely interesting accuracy, with $\Delta H_0/H_0$ between 0.2% and 0.4%, depending on the assumptions on the event rate (gaussian vs. flat neutron star mass distribution) and on the FOV of the GRB detector. For the dark energy sector, including both w_0 and Ξ_0, the results are given in Table 3.16. In particular (again, upon combining with CMB+BAO+SNe to reduce the degeneracies), Ξ_0 can be measured to better than 1%, more precisely to $(0.7-0.8)$%. This is an exciting result, since several modified gravity models give predictions significantly higher than this. This result also strengthens the science case for THESEUS. The corresponding result on w_0 is still valuable but less exciting, since it just brings the error down from 4.5% to 2.5%.

We have finally studied the configuration ET+CE+CE. In this case, despite the huge increase in GW detections, by a factor $\mathcal{O}(10)$ compared to a single ET, the final results on the cosmological parameters, limiting ourselves to GW-GRB coincidences, are not significantly better than with a single ET. For instance Ξ_0 can now be measured to $(0.6-0.7)$%. This is due to the fact that the number of joint GW-GRB detections does not increase correspondingly; in fact, from Table 3.3 we see that it does not even increase by a factor of 2. In other words, there is a bottleneck, due to the fact that a single GRB detector cannot keep the pace, and does not have the reach, of GW detections from a ET+CE+CE network. However, this network would have remarkable localization accuracy, and therefore could benefit from the detection of counterparts from optical/IR/radio telescopes. It is clear from our analysis that a strong effort in the follow up of GW signals would be necessary to really exploit the remarkable potential for cosmology of a network of 3G detectors, whether in the ET+CE+CE configuration that we have studied, or in intermediate configurations involving e.g. the planned Voyager upgrade of the advanced LIGO detectors.

3.3 Reconstructing modified gravitational-wave propagation

In Section 3.1.1 we have seen that the parametrizations (3.1.30) and (3.1.34) fit remarkably well the predictions of a large class of models (various Horndeski theories with different choices of the functions that characterize them, or several variant of nonlocal gravity), with two notable exceptions. One is bigravity, where the coupling between the two metrics gives rise to a series of oscillations in $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$. The second are DHOST theories, where the parametrization (3.1.34) misses a bump in $\delta(z)$. Nevertheless, in the latter case eq. (3.1.30) still reproduces reasonably well the behavior of $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$. This is due to the fact that the effect of the bump in $\delta(z)$ is smoothed out by the integration in eq. (3.1.13), so that $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ maintains the monotonic behavior described by eq. (3.1.30). This shows that the parametrization (3.1.30) for $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ is more solid than the corresponding parametrization for $\delta(z)$. From some point of view this is good news, since it shows that eq. (3.1.30) is a good starting point for searching for modified GW propagation in the data, given that $d_L^{\rm gw}(z)$ and $d_L^{\rm em}(z)$ are the directly observable quantities. On the other hand, this might not be sufficient to reconstruct detailed features of $\delta(z)$, which might carry distinct imprints of the cosmological model.

This fact, together with the existence of at least one example, bigravity, where eq. (3.1.30) is not appropriate, stimulates the development of a complementary approach, not based on the use of a specific parametrization. A natural possibility is provided by the method of Gaussian processes (GP). GP is a nonparametric technique which can reconstruct the distribution of a function from the training of the dataset, without assuming any parametrization for it. The GP method is very suitable for reconstructing the functions $d_L^{\rm gw}(z)$ and $d_L^{\rm em}(z)$ and their derivatives directly from the data. Having the distributions and covariances of these distance functions from GP, the reconstruction of $\delta(z)$ can be obtained. Many applications of GP in cosmology can be found in [252, 253, 254, 255, 256, 257, 258, 259, 187, 260]. In particular, in [187] this technique has been applied to the reconstruction of the dark energy equation of state to simulated data from the Einstein Telescope, in combination with other cosmological informations. In this Section we follow [261] to discuss the reconstruction of modified gravitational wave propagation (i.e. the functions $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ and $\delta(z)$) with Gaussian processes.

3.3.1 Mock datasets

To perform our analysis, we considered simulated measurements of the electromagnetic luminosity distance from DES supernovae. The generation of the mock data follows Section III B of [257], where a redshift range from $z_{\rm min}^{\rm (DES)}$=0.05 to $z_{\rm max}^{\rm (DES)}$=1.2 is considered and the errors on luminosity distance are estimated by using Table 14 of [262]. We assume a fiducial ΛCDM cosmology with Hubble parameter $H_0 = 67.64\,{\rm km\,s^{-1}Mpc^{-1}}$ and present fraction of matter energy density $\Omega_M = 0.3087$, corresponding to the mean values obtained from the CMB+SNe+BAO dataset described in Section 3 of [210]. For each SN at given z, the 'measured' value of the luminosity distance is then obtained by randomly scattering the corresponding value of $d_L(z)$ in the fiducial model, according to a Gaussian distribution centered around it and with a standard deviation given by the estimated DES error on luminosity distance. Fig. 3.33 shows the redshift distribution of the resulting catalog, containing 3443 SNe Ia whose light curves are obtained in a time period of five years. Table 3.24 gives a simplified description of the mock data, using the same redshift bins as in

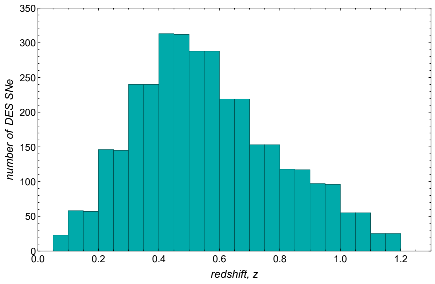

FIGURE 3.33: The redshift distribution of mock DES supernovae. From [261].

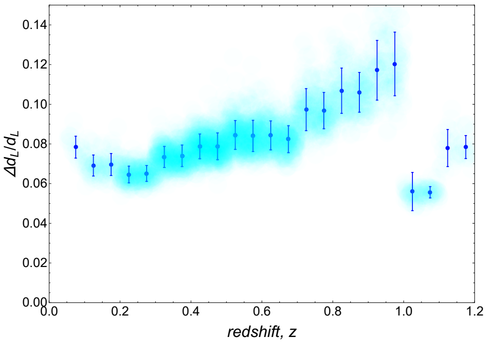

FIGURE 3.34: The relative error $\Delta d_L/d_L$ for the catalog of DES supernovae shown in Fig. 3.33 and Table 3.24. The cyan shaded area corresponds to the single supernovae, while the coordinates of the blue points are given by the mean values of the redshift and of $\Delta d_L/d_L$ in each redshift bin, with the bins chosen as in Fig. 3.33. The blue error bars are the standard deviations of $\Delta d_L/d_L$ in each redshift bin. From [261].

the histogram of Fig. 3.33. The table shows the mean value and the standard deviation of the relative error on luminosity distance $\Delta d_L/d_L$ for each redshift bin and the same quantities are plotted in Fig. 3.34. The drop in the relative error on luminosity distance beyond redshift $z = 1$ is due to selection effects at high redshift and is explained in Section 5.1 of [262].

For the reconstruction of the GW luminosity distance, we consider the two cases already presented in Section 3.2: the 2G network made by advanced LIGO, advanced Virgo, Kagra and LIGO India (HLVKI) at target sensitivity, or a 3G detector such as the Einstein Telescope. We summarize the main points for the resulting GW events. In both the 2G and 3G cases, we consider GW events from binary neutron star mergers with a redshift determined through the joint observation of a gamma-ray burst (GRB), using the mock catalogs of joint GW-GRB events generated in [210]. The generation of those catalogs uses a state-of-the-art treatment of the merger of binary neutron stars, that takes into account both the star formation rate and models of the time delay between the formation of the binary and the merger. The binary formation is assumed to follow the cosmic star formation rate modeled as in [263], while for the time delay distribution we use a power law with a minimum allowed time to coalescence of 20 Myr. Following Section 2.1 of [210], the overall normalization of the merger rate is fixed by matching with the local rate estimated from the O1 LIGO and O2 LIGO/Virgo observation runs [216]. We assume a Gaussian distribution for the neutron star masses and we consider 10 yr of data and an 80% duty cycle for each of the detectors included in the analysis, with a SNR=12 threshold

redshift bin	number of DES SNe	mean redshift	mean $\Delta d_L/d_L$	standard deviation of $\Delta d_L/d_L$
(0.05 , 0.1)	23	0.07500	0.07837	0.00552
(0.1 , 0.15)	58	0.12478	0.06913	0.00530
(0.15 , 0.2)	57	0.17478	0.06963	0.00553
(0.2 , 0.25)	146	0.22491	0.06461	0.00418
(0.25 , 0.3)	145	0.27491	0.06514	0.00398
(0.3 , 0.35)	240	0.32490	0.07345	0.00533
(0.35 , 0.4)	240	0.37490	0.07401	0.00542
(0.4 , 0.45)	313	0.42496	0.07878	0.00623
(0.45 , 0.5)	312	0.47496	0.07882	0.00676
(0.5 , 0.55)	288	0.52491	0.08453	0.00725
(0.55 , 0.6)	288	0.57491	0.08404	0.00788
(0.6 , 0.65)	219	0.62489	0.08434	0.00729
(0.65 , 0.7)	219	0.67489	0.08239	0.00679
(0.7 , 0.75)	153	0.72484	0.09729	0.01058
(0.75 , 0.8)	153	0.77484	0.09683	0.00916
(0.8 , 0.85)	118	0.82490	0.10683	0.01139
(0.85 , 0.9)	117	0.87490	0.10603	0.01003
(0.9 , 0.95)	97	0.92487	0.11714	0.01506
(0.95 , 1)	96	0.97487	0.12033	0.01603
(1 , 1.05)	55	1.02455	0.05602	0.00964
(1.05 , 1.1)	55	1.07455	0.05565	0.00283
(1.1 , 1.15)	25	1.12400	0.07794	0.00936
(1.15 , 1.2)	26	1.17500	0.07842	0.00579

TABLE 3.24: Table from [261]. The mean value and the variance of the relative error on luminosity distance $\Delta d_L/d_L$, averaging over the events in the given redshift bin, for the catalog of DES supernovae shown in Fig. 3.33.

for the total signal-to-noise ratio of a detection (the reader is referred to Section 2.1 of [210] for further details about the SNR calculation). The instrumental contribution to the error on luminosity distance has been estimated as $\Delta d_L/d_L$=1/SNR. In principle, a further contribution to $\Delta d_L/d_L$ comes from weak lensing. In [180, 181] it was modeled as $(\Delta d_L(z)/d_L(z))_{\text{lensing}} = 0.05z$, while the more recent study in [240] gives a significantly smaller effect. In any case, even with the pessimistic estimate $(\Delta d_L(z)/d_L(z))_{\text{lensing}} = 0.05z$, the lensing contribution is negligible for GW detections at HLVKI and it is also subdominant for sources detected at the Einstein Telescope with $z < 1.5$, which constitute 99% of all the events with a detected GRB counterpart in the final ET catalog that we will consider. On the other hand, for low-redshift sources the contribution due to the peculiar Hubble flow is important and, if not corrected for, gives an error on the redshift that can be modeled as due to an unknown peculiar velocity of order $\pm 200\,\text{km/s}$ [241]. Here we assume, conservatively, that this error is not corrected for, and we then propagate it to determine its contribution to the error on luminosity distance. The gravitational wave luminosity distance for the events of the mock catalog is drawn from a Gaussian distribution centered around the value from a fiducial cosmology and with standard deviation given by the sum in quadrature of the errors on luminosity distance described before. In this work we consider two different fiducial cosmologies for GW detections:

- the ΛCDM model with $H_0 = 67.64\,\text{km}\,\text{s}^{-1}\text{Mpc}^{-1}$ and $\Omega_M = 0.3087$ already used for the DES mock catalog;
- the RT nonlocal model with initial conditions set at $\Delta N = 64$ e-folds before the end of inflation, introduced in Section 3.3.

In the following, we will simply refer to those choices as the “ΛCDM" and the “RT" fiducial cosmologies, without repeating again the specifications. In the case of the RT fiducial, the modifications in the cosmological background with respect to ΛCDM are so small that the electromagnetic luminosity distances in the two models differ by just a few parts per thousand. That contribution is utterly negligible when compared to the modified gravitational wave propagation effect in eq. (3.1.13) for the RT model, which amounts to a deviation from GR by more than 60%. As a consequence the fiducial values of the gravitational wave luminosity distance can be simply obtained by multiplying those from ΛCDM by the factor $d_L^{\text{gw}}(z)/d_L^{\text{em}}(z)$ in eq. (3.1.30) with $\Xi_0 = 1.67$ and $n = 1.94$. For the error on luminosity distance, we reasonably assume the relative error to be the same as in the case of the ΛCDM fiducial cosmology, so we also multiply the ΛCDM absolute error by the same factor in eq. (3.1.30). Beside the intrinsic interest of the model, the use of the RT nonlocal model as a fiducial cosmological model can be seen as a case study for exploring the results of Gaussian processes reconstruction when the reference theory has a non-trivial GW propagation equation, despite being fully compatible with electromagnetic observations. Once the mock catalog of GW detections is built, we extract the events whose GRB emission is actually detected. A full explanation for this final stage in the construction of mock data can be found in Section 2.2 of [210], where the criterion for retaining a GW event in the GW-GRB catalog requires the peak flux of the GRB emission to be above the flux limit of the satellite considered for detection, which is Fermi-GBM for GW events at the HLVKI 2G network, and the proposed THESEUS mission [232] for GW events detected at the Einstein Telescope. In the ET/THESEUS catalog we only consider a fraction 1/3 of the events selected by the procedure above, since only the central part of the XGIS spectrometer on board THESEUS will be capable of arcmin localization of sources (this corresponds to the

z	d_L (Mpc)	Δd_L (Mpc)
0.029271	134.815	4.000
0.035195	157.475	5.636
0.060585	283.567	18.706
0.066283	316.373	14.509
0.071053	327.381	20.085
0.071730	342.952	16.957
0.076180	341.595	22.360
0.081819	418.469	30.238
0.088698	396.734	25.757
0.091869	402.590	34.170
0.094237	406.423	31.472
0.095288	432.996	36.423
0.099956	491.071	31.721
0.102531	461.627	36.858
0.114869	626.939	43.010

TABLE 3.25: The events in a realization of joint GW-GRB mock sources, with gravitational wave and gamma-ray burst detections at the HLVKI network and Fermi-GBM respectively. From [210, 261].

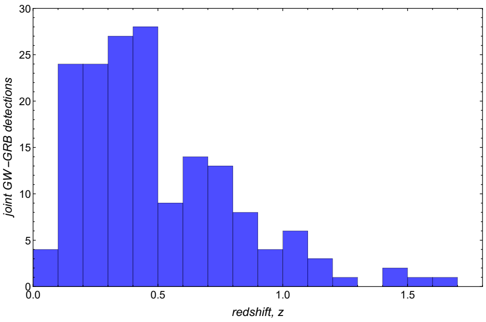

FIGURE 3.35: The redshift distribution of joint GW-GRB mock sources, with GW detections at the Einstein Telescope and gamma-ray burst detections at THESEUS. From [261].

“realistic" assumption for the FOV of THESEUS in Section 2.2.2 of [210]). Table 3.25 shows a realization (containing 15 sources) for the catalog of joint GW-GRB events at the second-generation network HLVKI, assuming ΛCDM as fiducial cosmology. For the case of Einstein Telescope we present the redshift distribution of the sources in Fig. 3.35 and the corresponding description of the instrumental error on luminosity distance in Table 3.26, with the same meaning of columns as in Table 3.24. Fig. 3.36 is a plot for the error on luminosity distance of the 169 events in the catalog realization considered in Table 3.26.

3.3.2 Results

We use the Gaussian processes method to reconstruct the functions $d_L^{\rm em}(z)$ and $d_L^{\rm gw}(z)$, as well as their derivatives with respect to redshift $d_L'^{\rm em}(z)$ and $d_L'^{\rm gw}(z)$, with the mock datasets of electromagnetic and gravitational-wave observations described in Section 3.3.1. Fig. 3.37 shows the result of the reconstruction for the electromagnetic luminosity distance and its derivative with respect to redshift, by using the DES

redshift bin	number of GW-GRB	mean redshift	mean $\Delta d_L/d_L$	standard deviation of $\Delta d_L/d_L$
(0 , 0.1)	4	0.07108	0.00868	0.00244
(0.1 , 0.2)	24	0.15001	0.01784	0.00692
(0.2 , 0.3)	24	0.24043	0.02558	0.00680
(0.3 , 0.4)	27	0.35355	0.03529	0.01004
(0.4 , 0.5)	28	0.44966	0.04843	0.01528
(0.5 , 0.6)	9	0.53785	0.05646	0.01807
(0.6 , 0.7)	14	0.64540	0.05329	0.01318
(0.7 , 0.8)	13	0.73793	0.05493	0.01368
(0.8 , 0.9)	8	0.85497	0.06413	0.00746
(0.9 , 1.0)	4	0.93702	0.06257	0.01228
(1.0 , 1.1)	6	1.05334	0.06494	0.00651
(1.1 , 1.2)	3	1.15162	0.06749	0.00246
(1.2 , 1.3)	1	1.25943	0.07373	0
(1.3 , 1.4)	–	–	–	–
(1.4 , 1.5)	2	1.45375	0.07851	0.00398
(1.5 , 1.6)	1	1.58407	0.07577	0
(1.6 , 1.7)	1	1.62843	0.07947	0

TABLE 3.26: The mean value and the variance of the ET instrumental contribution to $\Delta d_L/d_L$, averaging over the events in the given redshift bin, for the specific realization of the catalog of joint GW-GRB detections shown in Fig. 3.35. From [210, 261].

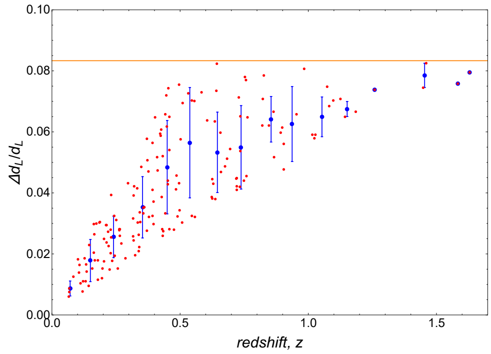

FIGURE 3.36: The ET instrumental contribution to the relative error $\Delta d_L/d_L$ for the catalog of joint GW-GRB detections at the Einstein Telescope and THESEUS shown in Fig. 3.35 and described in Table 3.26. The red points correspond to the individual events, while the coordinates of the blue points are given by the mean values of the redshift and of $\Delta d_L/d_L$ in each redshift bin, with the bins chosen as in Fig. 3.35. The blue error bars are the standard deviations of $\Delta d_L/d_L$ in each redshift bin. The orange horizontal line at $\Delta d_L/d_L$=1/12 corresponds to the SNR=12 threshold for detection. From [261].

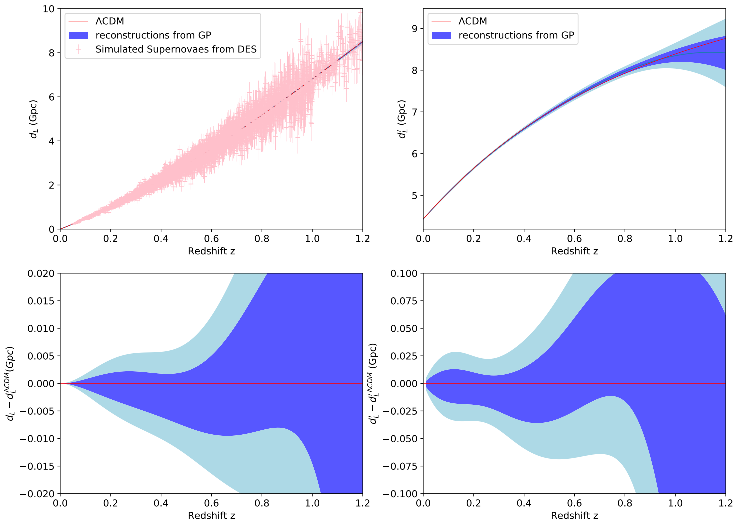

FIGURE 3.37: Reconstruction of the EM luminosity distance from the simulated DES catalog. In all panels, the blue and light blue regions correspond to 68% and 95% confidence levels respectively, while the red curve is the fiducial to be reconstructed. Upper left panel: $d_L^{\rm em}(z)$ and the mock DES supernovae used. Lower left: the difference between the reconstructed $d_L^{\rm em}(z)$ and the fiducial one. Upper right: the derivative $d_L'^{\rm em}(z)$. Lower right: the difference between the reconstructed $d_L'^{\rm em}(z))$ and its fiducial curve. From [261].

mock dataset[5]. For the GW luminosity distance, we show the results in two separate subsections for each mock dataset used (HLVKI or ET). Given a reconstruction of the functions $d_L^{\rm gw}(z)$ and $d_L^{\rm em}(z)$, we can then reconstruct the luminosity distance ratio

$$D(z) \equiv d_L^{\rm gw}(z)/d_L^{\rm em}(z) \tag{3.3.1}$$

as well as the function $\delta(z)$ that is given by

$$\delta(z) = (1+z)\left(\frac{d_L'^{\rm em}(z)}{d_L^{\rm em}(z)} - \frac{d_L'^{\rm gw}(z)}{d_L^{\rm gw}(z)}\right) . \tag{3.3.2}$$

For each dataset used (HLVKI or ET) we will show the result both with ΛCDM as fiducial, and with RT as fiducial. More precisely, when we take the RT model (with $\Delta N = 64$) as fiducial, we will assume that its prediction is exactly given by eq. (3.1.30) with $\Xi_0 = 1.67$ and $n = 1.94$. Actually, for the RT model with large ΔN, eq. (3.1.30) fits extremely well the prediction of the model for $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ (obtained from the numerical integration of the relevant equations, that involve the numerical evolution of the auxiliary fields of the model, see [264]). In contrast, at very small z the numerical result for $\delta(z)$ differs somewhat from that obtained by the parametrization (3.1.34); the correct numerical result for $\delta(z)$ is shown in Fig. 2 of [172]. Again, the difference is due to the fact that small details in $\delta(z)$ gets smoothed

---

[5] We recall from Section 3.3.1 that, contrary to $d_L^{\rm gw}(z)$, the difference in $d_L^{\rm em}(z)$ between ΛCDM and the RT model is very small (a few parts per thousand). Therefore we can safely neglect it and consider the same mock catalog of supernovae for both fiducial cosmological models.

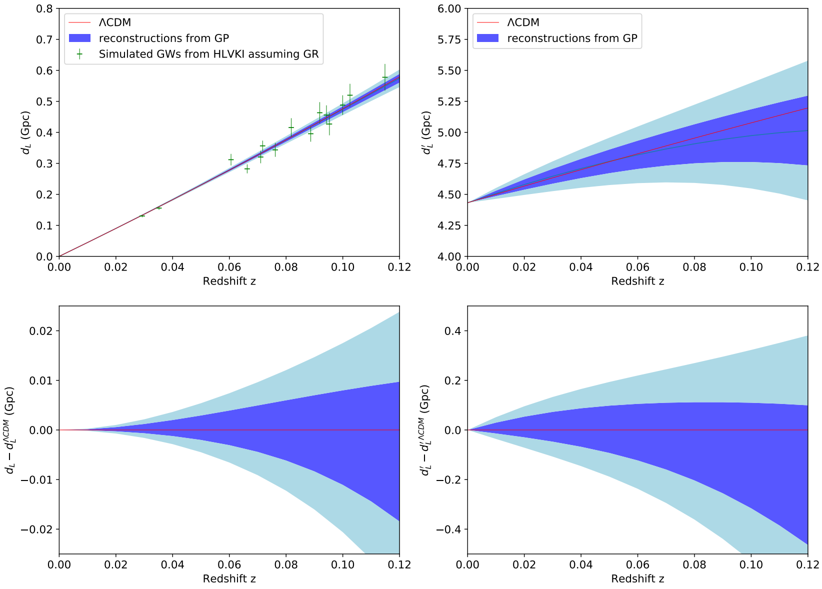

FIGURE 3.38: Reconstruction of the GW luminosity distance from mock detections at HLVKI with electromagnetic counterpart, for the ΛCDM fiducial cosmology. The panels are organized as in Fig. 3.37 and the green points in the upper left panel are the mock data in the HLVKI catalog. Note the difference in vertical scale between the two lower panels. From [261].

out when performing the integration in eq. (3.1.13), as we already discussed above. Here, in order to illustrate the methodology, we simply assume that the result for $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ is exactly given by eq. (3.1.30), so that the result for $\delta(z)$ would also be given by the corresponding equation (3.1.34).

Results for the HLVKI network

In Fig. 3.38 we plot the reconstructed GW luminosity distance obtained from the mock catalog of standard sirens at the second-generation network HLVKI, assuming ΛCDM as fiducial cosmology, while Fig. 3.39 is the analogous plot in the case of the RT fiducial. In the case of ΛCDM fiducial, the reconstructions of the ratio $D(z) = d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ and of $\delta(z)$ [using eq. (3.3.2)] from the HLVKI and DES mock catalogs are plotted in Fig. 3.40. Similarly, Fig. 3.41 shows the results assuming the RT fiducial cosmology. We provide the reconstructions up to the maximum redshift reached in the HLVKI mock catalog $z_{\rm max}^{\rm (HLVKI)} \simeq 0.12$ (see Table 3.25), which is smaller than the maximum redshift $z_{\rm max}^{\rm (DES)}$=1.2 in the simulated DES catalog of supernovae.

It is interesting to compare the results on modified GW propagation from Gaussian processes reconstruction with those obtained by using the (Ξ_0, n) parametrization for the function $D(z) = d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ in eq. (3.1.30). In this case our strategy is to implement the parametrization in the CLASS Boltzmann code and run a MCMC. The parameter n plays a much less important role than Ξ_0 for $D(z)$ because it just regulates the precise interpolation between the value $D(z=0) = 1$ and the asymptotic $D(z \gg 1) = \Xi_0$. Furthermore, chains with both Ξ_0 and n as free parameters

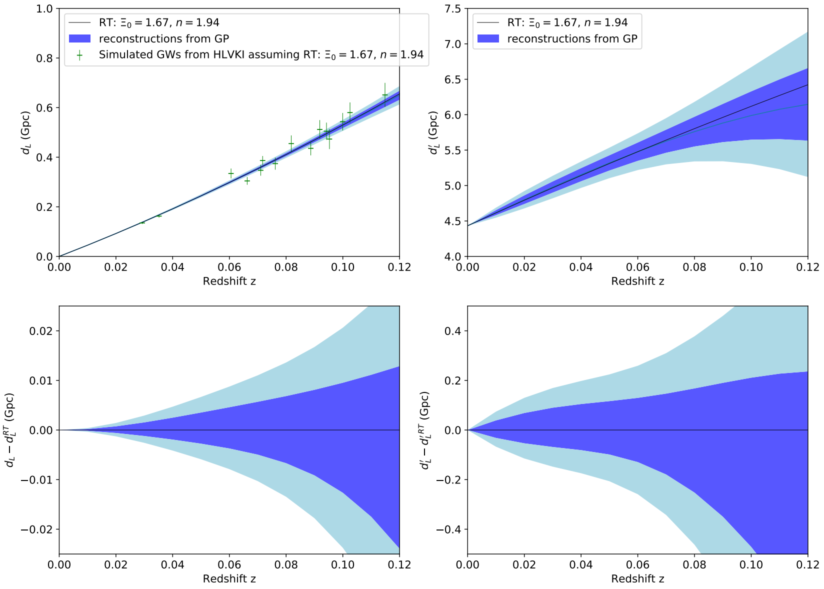

FIGURE 3.39: Reconstruction of the GW luminosity distance from mock detections at HLVKI with electromagnetic counterpart, for the RT fiducial cosmology. Panels as in Fig. 3.38. From [261].

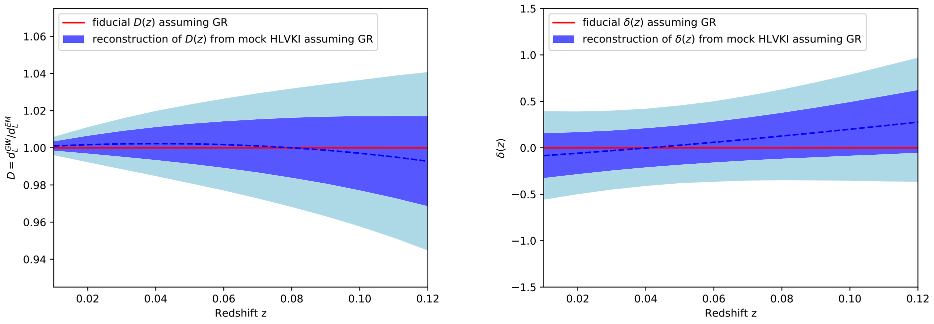

FIGURE 3.40: Results from the mock HLVKI and DES catalogs, assuming the ΛCDM fiducial cosmology. Left panel: reconstruction of the ratio $D(z) = d_L^{\text{gw}}(z)/d_L^{\text{em}}(z)$. Right panel: reconstruction of $\delta(z)$. From [261].

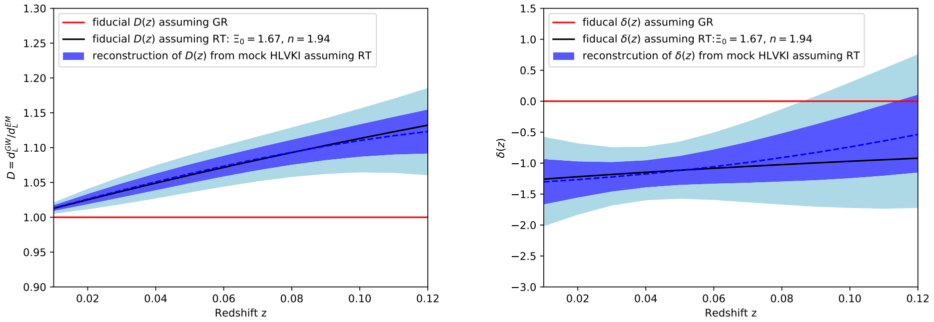

FIGURE 3.41: As in Fig. 3.40, using the mock HLVKI and DES catalogs for the RT fiducial cosmology. From [261].

fail to converge and for these reasons we just keep n fixed and look at the precision reached by the MCMC in obtaining the Ξ_0 value of the fiducial cosmological model. For the RT fiducial cosmology we set n to the actual fiducial value $n^{(\rm RT)} = 1.94$ and we want the MCMC to recover the fiducial $\Xi_0^{(\rm RT)} = 1.67$. In the ΛCDM case, $n^{(\Lambda\rm CDM)}$ is not determined and when running the MCMC we choose to set it to $n = 2.5$, which is in the ballpark of typical values predicted in modified gravity theories (for example in some of the nonlocal gravity models). We then want to recover $\Xi_0^{(\Lambda\rm CDM)} = 1$.

	DES+HLVKI	CMB+BAO+DES+HLVKI
$\Delta\Xi_0$	0.127 (12.7%)	0.127 (12.7%)
$\Delta H_0/H_0$	0.38%	0.21%
$\Delta\Omega_M/\Omega_M$	3.30%	0.71%

TABLE 3.27: Accuracy (1σ level) in the reconstruction of Ξ_0, H_0 and Ω_M with DES+HLVKI and CMB+BAO+DES+HLVKI, assuming ΛCDM as the fiducial cosmology for the HLVKI dataset. The relative error on Ξ_0 is the same as the absolute error, because of the fiducial value $\Xi_0^{(\Lambda\rm CDM)} = 1$. From [261].

	DES+HLVKI	CMB+BAO+DES+HLVKI
$\Delta\Xi_0/\Xi_0$	8.32%	8.32%
$\Delta H_0/H_0$	0.38%	0.21%
$\Delta\Omega_M/\Omega_M$	3.30%	0.74%

TABLE 3.28: As in Table 3.27, assuming the RT model as the fiducial cosmology for the HLVKI dataset. The absolute error on Ξ_0 can be easily found recalling the fiducial value $\Xi_0^{(\rm RT)} = 1.67$. From [261].

For each fiducial cosmology we run MCMCs with two choices of combined datasets:

1. the simulated DES supernovae and the mock catalog of GW detections at HLVKI (with electromagnetic counterpart) described in Section 3.3.1;
2. to reduce the degeneracies between cosmological parameters, in addition to the DES and HLVKI mock catalogs, we also consider the same CMB and BAO data that we used in Section 3.2.

The two combined datasets will be denoted as “DES + HLVKI" and “CMB + BAO + DES + HLVKI", respectively. In the first case the MonteCarlo is sensitive to the set of parameters $\{H_0, \Omega_M, \Xi_0\}$, while in the second case the parameters that come into play are $H_0, \omega_b, \omega_c, A_s, n_s, \tau_{re}, \Xi_0$, while the total matter fraction Ω_M is a derived parameter. The quantities $\omega_b = \Omega_b h^2$ and $\omega_c = \Omega_c h^2$ are the physical baryon and cold dark matter density fractions today, respectively (where h is defined by the relation $H_0 = 100h\,\mathrm{km\,s^{-1}Mpc^{-1}}$). A_s and n_s are the amplitude and tilt of the primordial scalar perturbations, and $\tau_{\rm re}$ is the reionization optical depth. We keep the sum of neutrino masses fixed, at the value $\sum_\nu m_\nu = 0.06$ eV, as in the *Planck* baseline analysis [135].

Fig. 3.42 and Fig. 3.43 show the results for the two-dimensional likelihoods of cosmological parameters, in the case where ΛCDM or RT are assumed as fiducial cosmologies for the mock catalog of GW detections at the HLVKI network. Tables 3.27 and 3.28 contain the errors obtained from the corresponding one-dimensional marginalized likelihoods.

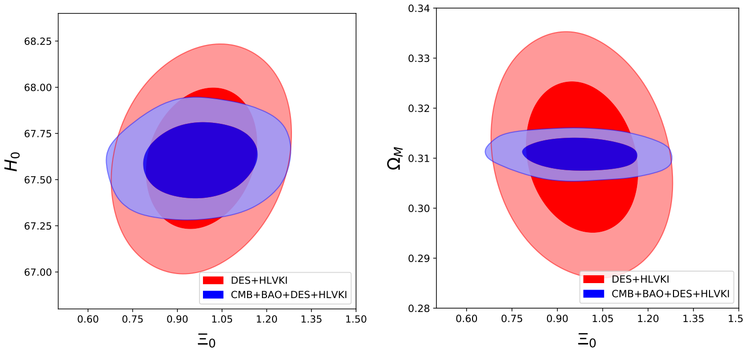

FIGURE 3.42: The 1σ and 2σ contours of two-dimensional likelihoods, from DES+HLVKI (red) and CMB+BAO+DES+HLVKI (blue). The fiducial cosmology for the HLVKI dataset is ΛCDM. Left: in the (Ξ_0, H_0) plane. Right: in the (Ξ_0, Ω_M) plane. From [261].

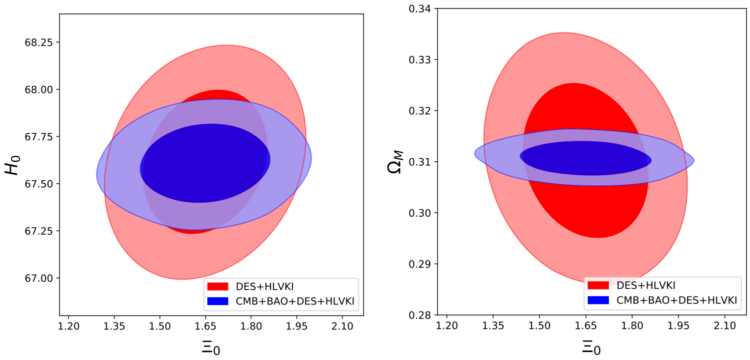

FIGURE 3.43: As in Fig. 3.42, assuming the RT nonlocal model as the fiducial cosmology for the HLVKI dataset. From [261].

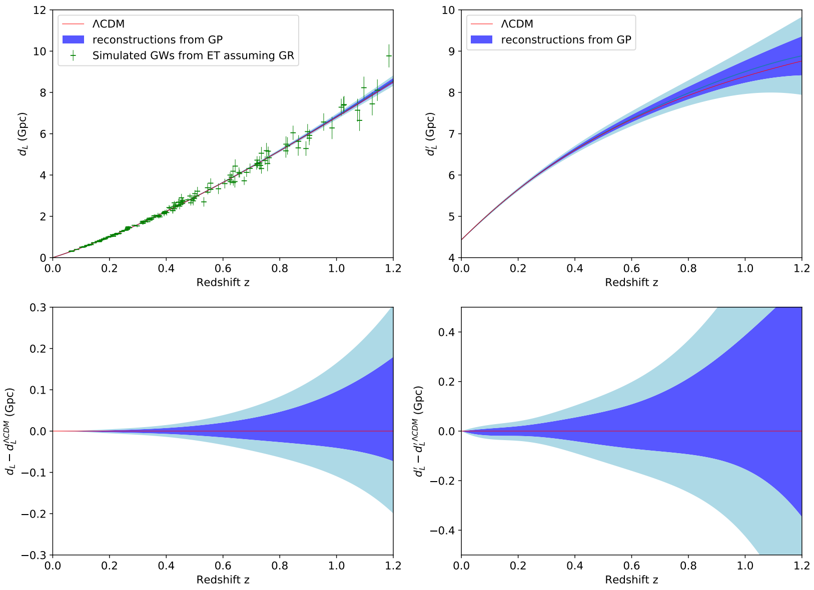

FIGURE 3.44: Reconstruction of the GW luminosity distance from mock detections at the Einstein Telescope with electromagnetic counterpart, assuming ΛCDM as the fiducial cosmological model, again with the same meaning of the panels as in Fig. 3.37. The green points in the upper left panel are the mock data in the ET catalog. From [261].

Results for the Einstein Telescope

Fig. 3.44 and Fig. 3.45 show the reconstruction of the GW luminosity distance from mock detections at the Einstein Telescope, obtained by Gaussian processes, for the ΛCDM and the RT fiducial cosmologies respectively. The final results for the luminosity distance ratio $D(z) = d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ and the funcion $\delta(z)$ are given in Fig. 3.46 assuming the ΛCDM fiducial and in Fig. 3.47 for the RT fiducial. We provide the reconstructions up to the maximum redshift reached in the DES mock catalog of supernovae $z_{\rm max}^{\rm (DES)}$=1.2, which is smaller than the maximum redshift $z_{\rm max}^{\rm (ET)} \simeq 1.63$ in the simulated catalog of GWs from binary neutron stars detected at ET (see Table 3.26).

As we did or the HLVKI network, we also show the results obtained by running MCMCs with the parametrization given by eq. (3.1.30). The two-dimensional likelihoods for the cosmological parameters are shown in Fig. 3.48 and Fig. 3.49, assuming ΛCDM or RT, respectively, as fiducial cosmologies for the mock catalog of GW detections at the ET network. The errors on cosmological parameters are listed in Tables 3.29 and 3.30.

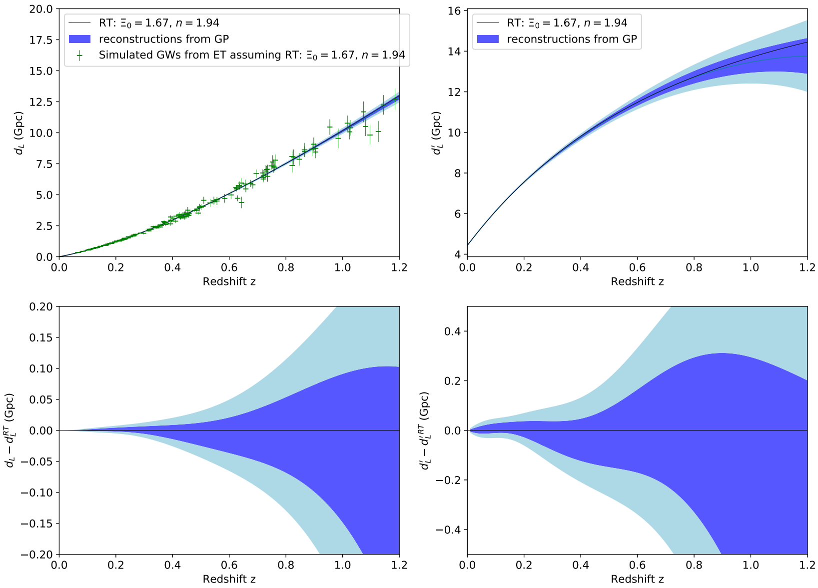

FIGURE 3.45: As in Fig. 3.44, assuming the RT fiducial cosmology. From [261].

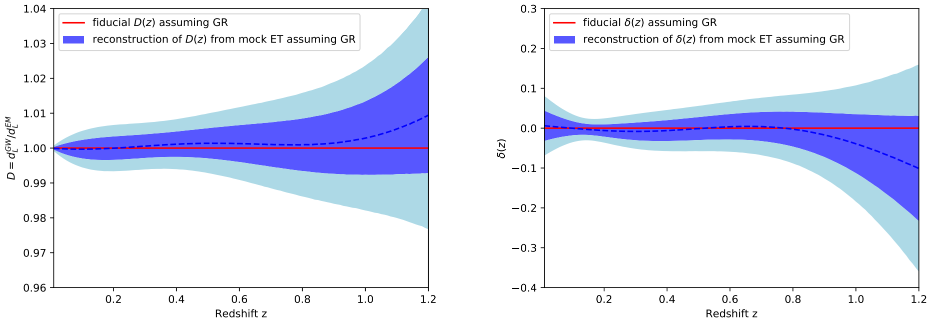

FIGURE 3.46: Results from the mock ET and DES catalogs, assuming the ΛCDM fiducial cosmology. Left panel: reconstruction of the ratio $D(z) = d_L^{\rm gw}(z)/d_L^{\rm em}(z)$. Right panel: reconstruction of $\delta(z)$. From [261].

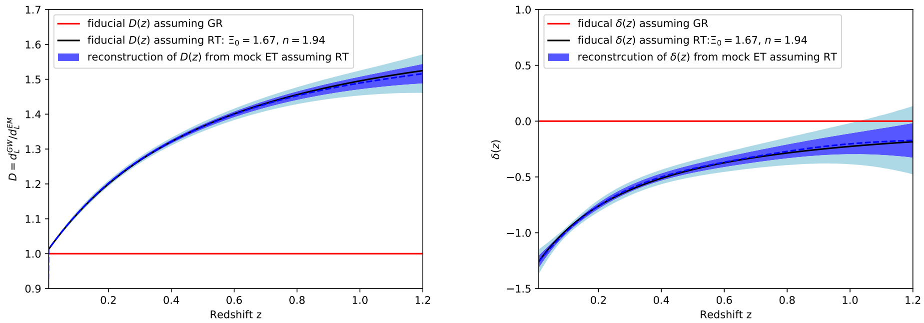

FIGURE 3.47: As in Fig. 3.46, using the mock ET and DES catalogs for the RT fiducial cosmology. From [261].

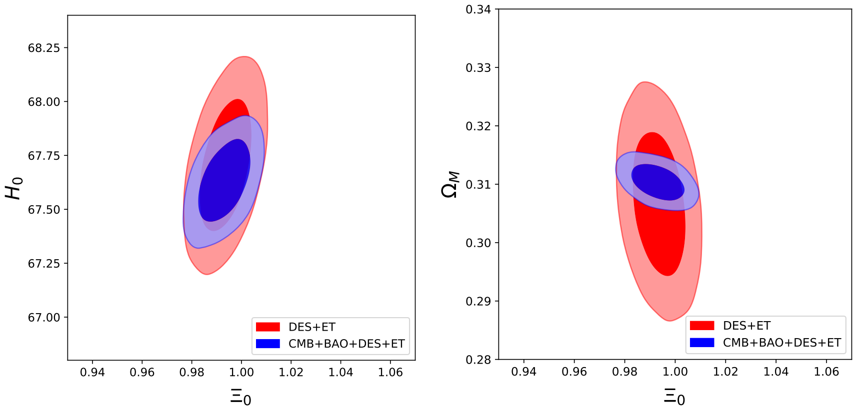

FIGURE 3.48: The 1σ and 2σ contours of two-dimensional likelihoods, from DES+ET (red) and CMB+BAO+DES+ET (blue). The fiducial cosmology for the ET dataset is ΛCDM. Left: in the (Ξ_0, H_0) plane. Right: in the (Ξ_0, Ω_M) plane. From [261].

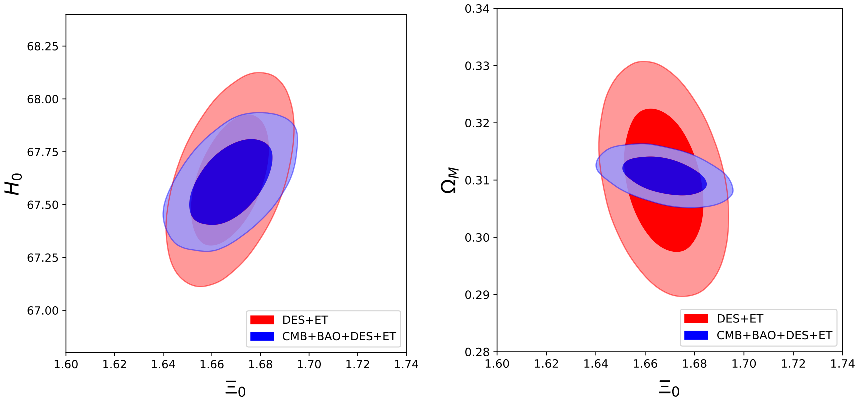

FIGURE 3.49: As in Fig. 3.48, assuming the RT nonlocal model as the fiducial cosmology for the ET dataset. From [261].

	DES+ET	CMB+BAO+DES+ET
$\Delta\Xi_0$	0.008 (0.8%)	0.007 (0.7%)
$\Delta H_0/H_0$	0.31%	0.19%
$\Delta\Omega_M/\Omega_M$	2.78%	0.68%

TABLE 3.29: Accuracy (1σ level) in the reconstruction of Ξ_0, H_0 and Ω_M with DES+ET and CMB+BAO+DES+ET, assuming ΛCDM as the fiducial cosmology for the ET dataset. From [261].

	DES+ET	CMB+BAO+DES+ET
$\Delta\Xi_0/\Xi_0$	0.66%	0.66%
$\Delta H_0/H_0$	0.31%	0.20%
$\Delta\Omega_M/\Omega_M$	2.76%	0.74%

TABLE 3.30: As in Table 3.29, assuming the RT model as the fiducial cosmology for the ET dataset. From [261].

3.3.3 Summary on the reconstruction of modified GW propagation

Modified GW propagation can play a crucial role in telling apart modified gravity models from ΛCDM. We have explored two very convenient ways of extracting the information from the observations, either using the parametrization (3.1.30), or using a parametrization-independent reconstruction method based on Gaussian processes. The first conclusion is that the two methods are quite complementary. The parametrization (3.1.30) of the ratio $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ has been shown to fit extremely well the predictions of most of the best studied modified gravity models; this, together with its simplicity, makes it an extremely convenient tool. On the other hand, the study in [200] has also identified one model, bigravity, where this parametrization is not adequate since it does not catch a series of oscillations that take place in that model; this already makes it worthwhile to test the data also against a parametrization-independent technique. For the function $\delta(z)$ the use of Gaussian processes is even more informative; indeed, the directly observable quantity is $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$, which is related to $\delta(z)$ by eq. (3.1.13). Features of the function $\delta(z)$, such as bumps, can be smoothed out by the integration in eq. (3.1.13), so that eventually the parametrization (3.1.30) could still fit the data relatively well even when some features of the function $\delta(z)$ are not correctly reproduced by the corresponding parametrization (3.1.34). As we already mentioned, this happens indeed in DHOST theories [200]. In this case, a parametrization-independent reconstruction of the function $\delta(z)$ using Gaussian processes might put in evidence structures in $\delta(z)$, and therefore signatures of the underlying model, that would be lost using the parametrization (3.1.30).

The complementarity between the two methods can also be seen in the informations that we get from them. In particular, the parametrization (3.1.30) allows us to get in a very direct manner the asymptotic value of the ratio $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ for large redshift, which is given by the parameter Ξ_0. In contrast, the Gaussian processes reconstruction can identify the correct functional form at moderate redshift without making any assumptions, but becomes less and less accurate with increasing redshift, as we see from all figures showing the result of such reconstructions.

Beside testing the accuracy to which one could confirm the validity of ΛCDM, assuming it as fiducial model, we have also tested the accuracy to which one could validate a modified gravity model. We have focused in particular on a nonlocal modification of gravity, the so-called RT model [77], for which recent work [172] has shown that Ξ_0 can reach values as large as 1.6. The model therefore gives a 60% deviation from ΛCDM in the tensor sector, despite the fact that, both in the background evolution and in scalar perturbations it is very close to ΛCDM, and indeed it fits the existing cosmological datasets at a level statistically indistinguishable from ΛCDM. The result are very interesting, since they show that already with 15 binary neutron stars with counterpart, the second-generation detector network LIGO/Virgo/Kagra (HVLKI) could very clearly detect this effect. Indeed, using the parametrization (3.1.30), the results in Table 3.28 show that Ξ_0 can be determined by HVLKI (combining the results with CMB, BAO and DES SNe) to about 8% level accuracy, well below the 60% deviation predicted by the RT model in the more optimistic case. This is fully consistent with the results obtained in [210] where, using the current JLA SN dataset rather than the mock DES SNe used here, it was found that HVLKI combined with CMB, BAO and JLA can measure Ξ_0 to about 10%. In the present analysis we see that a similar conclusion can be obtained from Gaussian process reconstruction. We see indeed from Fig. 3.41 that, with 15 binary neutron stars with counterpart at HVLKI, it is possible to perform a parametrization-independent reconstruction of the ratio $d_L^{\rm gw}(z)/d_L^{\rm em}(z)$ to a few percent, across the whole range of redshifts considered. For instance, this implies that the prediction of the RT model with $\Delta N = 64$ and the prediction of ΛCDM would be very clearly distinguished.

For third-generation detectors, such as the Einstein Telescope, the situation will be even more exciting: in terms of Ξ_0, ET combined with the CMB+BAO+DES will reach an accuracy on Ξ_0 better than 1%, see Table 3.29, while Fig. 3.47 shows the remarkable accuracy of the Gaussian-processes reconstruction.

3.4 Cosmology with LISA standard sirens

In [200] modified gravitational-wave propagation was extensively discussed for Horndeski theories ($f(R)$ theories, galileon cosmology, etc.), Degenerate Higher Order Scalar-Tensor (DHOST) theories, bigravity and nonlocal cosmological models. Here we discuss the methods and results of the analysis in [200] about the constraints on cosmological parameters, dark energy and modified GW propagation, obtained within the LISA Cosmology Working Group.

3.4.1 Mock catalogs of standard sirens at LISA

The sources that LISA will be able to detect at cosmological distances are massive (and supermassive) binary black holes (MBHB) at high reshift ($z \gtrsim 1$) [265], extreme mass ratio inspirals (EMRIs) at intermediate redshift ($0.1 \lesssim z \lesssim 2-3$) [266] and stellar mass binary black holes at low redshift ($0.01 \lesssim z \lesssim 0.1$) [267]. However, we expect only MBHBs to have powerful EM counterparts, because of the gas rich environment where they merge. In particular EM counterparts are expected in the optical and radio bands as produced by jets or accretion and can enable us to determine precisely the position of the source in the sky. Therefore, as in [200], we focus on mock catalogs of MBHBs. Their construction is explained in detail in Section 4 of [200] elaborating on [188], and here we just summarize the main aspects that finally led to three choices for the scenario of MBHBs formation. The two main ingredients are the initial mass function of the massive black hole seeds, and the inclusion (or not) of delays between galaxy and massive black hole mergers. We distinguish between heavy seeds (bar instabilities of protogalactic disks, with seed masses $\sim 10^5 M_\odot$) and light seeds (population III stars with remnant masses $\sim 100 M_\odot$). The three models that we use are: heavy seeds and no delays ("hnd" for short), heavy seeds with a delay included and the Toomre parameter for the initial bar instability (see [188]) set to $Q = 3$ ("hQ3"), and light seeds due to pop III stars with inclusion of delays ("popIII"). For each LISA detection (using the threshold SNR> 8), the error on the luminosity distance is estimated with a Fisher matrix method and added to the error due to weak lensing and peculiar velocities (see [200] and references therein).

Then, for the EM counterpart detection, first of all only the GW events with a sky localization $\Delta\Omega < 10\,\text{deg}^2$ are selected and then, following [188], one can determine the number of counterparts detected by future EM facilities, like LSST, SKA and ELT. The determination of the EM counterpart will enable us to determine the host galaxy of the MBHB and to obtain the redshift of the GW event using either spectroscopic or photometric techniques. Two different scenarios for the redshift errors were considered in [200] and here we just summarize the errors:

- **Scenario 1 (realistic)**, with relative error on redshift given by $(\Delta z/z)_{\text{photo}} = 0.05(1+z)$ for photometric measurements (as suggested by Euclid related works [268]), while $(\Delta z/z)_{\text{spect}} = 0.01(1+z)^2$ for spectroscopic measurement.
- **Scenario 2 (optimistic)**, with photometric measurements having a relative error on redshift given by $(\Delta z/z)_{\text{photo}} = 0.03(1+z)$, while spectroscopic measurements have $(\Delta z/z)_{\text{spect}} = 0.01$.

Then the redshift errors are simply propagated to the luminosity distance using the fiducial ΛCDM cosmology. As a final result we end up with 3×2 mock catalogs of MBHB events at LISA with EM counterpart (considering 4 years of data taking), with

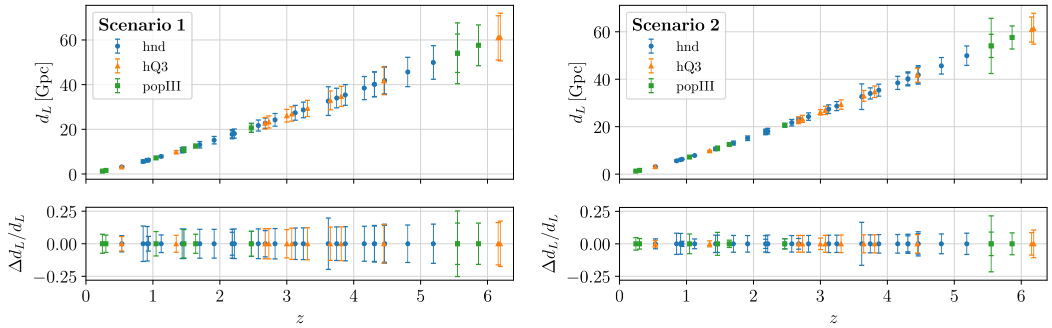

FIGURE 3.50: Luminosity distance as a function of redshift for the sources in the three catalogs "hnd","hQ3" and "pop III" (see the text for their definitions) with errors computed in the two scenarios for the error on redshift: scenario 1 (realistic, left) and scenario 2 (optimistic, right). Bottom: relative error on luminosity distance. Figure from [200].

values of the luminosity distance, redshift and total error on luminosity distance[6]. Fig. 3.50 is a plot of the six final catalogs considered (figure from [200]).

3.4.2 Cosmological parameters and dark energy with LISA

We now use the mock catalogs discussed in the previous section to study the effect of LISA standard sirens on the estimate of the cosmological parameters, with particular emphasis on the DE sector, as described by the parameters (w_0, w_a) for the DE equation of state, eq. (3.1.29), and by the parameters (Ξ_0, n) that enter eq. (3.1.30) for modified GW propagation.

As discussed in section 3.4.1, we provide our results for three models for the formation of MBHBs that we called "hnd", "hQ3" and "popIII". For each formation model, we use the two scenarios (1) and (2) of section 3.4.1 for estimating the error in the redshift measurement, that we propagate to luminosity distance. Thus, overall, we consider six different possibilities.

From the point of view of the cosmological model, we consider different cases. We will start with ΛCDM, to investigate how LISA standard sirens help to constrain its parameters. We will then study the extension of the DE sector obtained introducing a non-trivial DE equation of state, parametrized by w_0 only (without modified GW propagation, i.e. setting $\Xi_0 = 1$); finally, we will introduce modified GW propagation, extending the DE sector through the parameters (w_0, Ξ_0) and fixing $w_a = 0$ and n to a reference value.[7] For each of these three cosmological models we run MCMCs for the six scenarios describing the formation model and the estimate of the redshift error, as discussed above.

Furthermore, for each of these 3×6 cases we run separate MCMCs to compute the constraints that can be obtained with standard sirens only, and those obtained by

---

[6] To obtain representative catalogs for each of the six scenarios, a number of simulated catalogs were ranked according to their errors on ΛCDM cosmological parameters, as obtained from a quick Fisher matrix estimate (more details are available at the end of Section 4 in [200]); then the median catalog for each scenario was selected.

[7] We do not introduce more than two extra parameters at the same time in the DE sector, since the constraining power would decrease and the convergence time of the MCMC chains would become very long. Indeed, already with the two parameters (w_0, w_a), we found that, because of the limited number of sources in the LISA catalogs, the MCMC chains do not resolve the degeneracy between these parameters and do not reach a good convergence.

combining them with CMB, BAO and SNe data. In that case we use the *Planck* 2015 temperature and polarization power spectra [209], the JLA SNe dataset [140] and a set of isotropic and anisotropic BAO data (these are the same datasets that we used, for example, in the analysis for the RR nonlocal model; see Section 3.3.1 of [30] for details on those datasets).

To generate the catalogs discussed in section 3.4.1 we assume a fiducial cosmological model, that we always take to be ΛCDM with $H_0 = 67.64$ and $\Omega_M = 0.3087$, which are the fiducial values obtained from the CMB+BAO+SNe dataset that we use. So, in particular, our fiducial values for the extra parameters in the DE sector are $w_0 = -1$ and $\Xi_0 = 1$, and we use ΛCDM and standard perturbation theory within GR as fiducial theoretical framework for treating and analyzing our dataset. Our aim is to evaluate to what accuracy LISA, alone or in combination with other datasets, can reconstruct these fiducial values. We then generate our simulated catalog of events by assuming that, for a source at redshift z_i, the actual luminosity distance will be $d_L^{\Lambda\mathrm{CDM}}(z_i; H_0, \Omega_M)$, with the above values of H_0 and Ω_M. The "measured" value of the luminosity distance is then randomly extracted from a Gaussian distribution centered on this fiducial value, and with a width obtained from the estimate of the error on the luminosity distance discussed in Section 3.4.1.

The relevant cosmological parameters for evaluating the theoretical values of the GW luminosity distance $d_L^{\mathrm{gw}}(z)$ are Ω_M and H_0 in the ΛCDM case, with the addition of w_0 in wCDM and the further addition of Ξ_0 in the (Ξ_0, w_0) case. Apart from a constant addend coming from normalization, the logarithm of the likelihood assigned to a given set of values $(H_0, \Omega_M, w_0, \Xi_0)$ is given by

$$\ln(L(H_0, \Omega_M, w_0, \Xi_0)) = -\frac{1}{2}\sum_{i=1}^{N_s} \frac{\left[d_L^{\mathrm{gw}}(z_i; H_0, \Omega_M, w_0, \Xi_0) - d_i\right]^2}{\sigma_i^2}, \tag{3.4.1}$$

where N_s is the number of mock sources in the catalog, $d_L^{\mathrm{gw}}(z_i; H_0, \Omega_M, w_0, \Xi_0)$ is the theoretical value of the GW luminosity distance for the i-th source and d_i is the "measured" value of its luminosity distance contained in the catalog. The quantity σ_i is the error on luminosity distance and it also takes into account the error on redshift determination, which is simply propagated to the luminosity distance using the fiducial ΛCDM cosmology. The MCMC code used explores the cosmological parameters space, accepting or rejecting the points according to a Metropolis-Hastings algorithm based on the likelihood specified above. The priors assumed on cosmological parameters are Gaussian (or truncated Gaussian) distributions with mean and standard deviation, in this order, given by: 67.8 and 1.2 for H_0 (in $\mathrm{km\,s^{-1}Mpc^{-1}}$), 0.02225 and 0.00028 for ω_b=$\Omega_b h^2$ (baryon density fraction), 0.1192 and 0.0027 for $\omega_c = \Omega_c h^2$ (cold dark matter density fraction), -1.0 and 1.0 for w_0 (restricted from -3.0 to -0.3), 1.0 and 0.5 for Ξ_0 (restricted to positive values).

The cosmological evolution is studied by using the CLASS Boltzmann code. The modified version of CLASS implementing the GW luminosity distance parametrization in Equation eq. (3.1.30) is available at [132]. The MCMC code obtained by including in MontePython the LISA mock catalogs of sources and implementing the LISA likelihood described in eq. (3.4.1) is available at [269] (it needs the aforementioned modifed version of CLASS).

As we see from Fig. 3.50, the number of total available standard sirens in our sample is quite limited. For instance, in the specific realization of the catalogs that we use there are 32, 12 and 9 sources in the "hnd", "hQ3" and "popIII" scenarios, respectively. It is important to realize that, since these numbers are relatively small, the scatter of the mock data around their fiducial values unavoidably induces fluctuations in the reconstruction of the mean value of the cosmological parameters inferred from standard sirens alone, so that the reconstructed mean values in general will not coincide with those obtained from other cosmological observations such as CMB, BAO and SNe. Thus, depending on the specific realization of the catalog, the contours in parameter space of the likelihood obtained from standard sirens could show mild tensions with those obtained from other datasets. This is unavoidable, and will also happen in the actual experimental situation. Thus, in the plots shown below, the overlap (or lack of it) between different contours has little meaning, as it depends on the specific random realization of the catalog. What carries the important information is the relative size of the contours, that will tell us to what extent the addition of standard sirens to other datasets can improve our knowledge of the cosmological parameters. However, when combining contours from different datasets to obtain a combined estimate on the error on a parameter, one must be careful not to use a specific realization where the separate contours are in mild tension among each other.

ΛCDM

We begin by studying the effect of standard sirens on cosmological parameter estimation in ΛCDM, as a benchmark for our subsequent generalizations, and we will then introduce different extensions of the dark energy sector. As baseline ΛCDM model we use the standard set of six independent cosmological parameters: the Hubble parameter today $H_0 = 100h\,\mathrm{km\,s^{-1}Mpc^{-1}}$, the baryon density fraction today $\omega_b = \Omega_b h^2$, the cold dark matter density fraction today $\omega_c = \Omega_c h^2$, the amplitude A_s and tilt n_s of the spectrum of primordial scalar perturbations, and the reionization optical depth $\tau_{\rm re}$. We keep the sum of neutrino masses fixed, at the value $\sum_\nu m_\nu = 0.06$ eV, as in the *Planck* baseline analysis. We then run a series of MCMC, using the CLASS Boltzmann code [131] (or our modification of it, in the case of non-trivial GW propagation).

In ΛCDM, assuming flatness, $d_L(z)$ depends only on H_0 and Ω_M, so these are the parameters for which the inclusion of standard sirens has the most significant impact. Fig. 3.51 shows the two-dimensional likelihood in the (Ω_M, H_0) plane in ΛCDM, comparing the contribution from CMB+BAO+SNe (red) to the contribution from LISA standard sirens (gray), and the overall combined contours, for the three formation scenarios and the scenario (1) for the determination of the error on the luminosity distance, while Fig. 3.52 shows the results for the scenario (2).

In particular in the most favorable scenario (“hnd" seeds and optimistic errors on d_L), from the corresponding one-dimensional marginalized likelihood we find that, with standard sirens only, the relative error on H_0 is

$$\frac{\Delta H_0}{H_0} = 3.8\%\,, \tag{3.4.2}$$

(which raises to 7.7% in the “realistic" scenario with “hnd" seeds) and the one on Ω_M is $\Delta\Omega_M/\Omega_M = 14.7\%$; using our CMB+BAO+SNe dataset we get instead $\Delta H_0/H_0 = 0.7\%$ and $\Delta\Omega_M/\Omega_M = 2.1\%$; combining CMB+BAO+SNe+standard sirens we get

$$\frac{\Delta H_0}{H_0} = 0.7\%\,, \qquad \frac{\Delta\Omega_M}{\Omega_M} = 2.0\%\,. \tag{3.4.3}$$

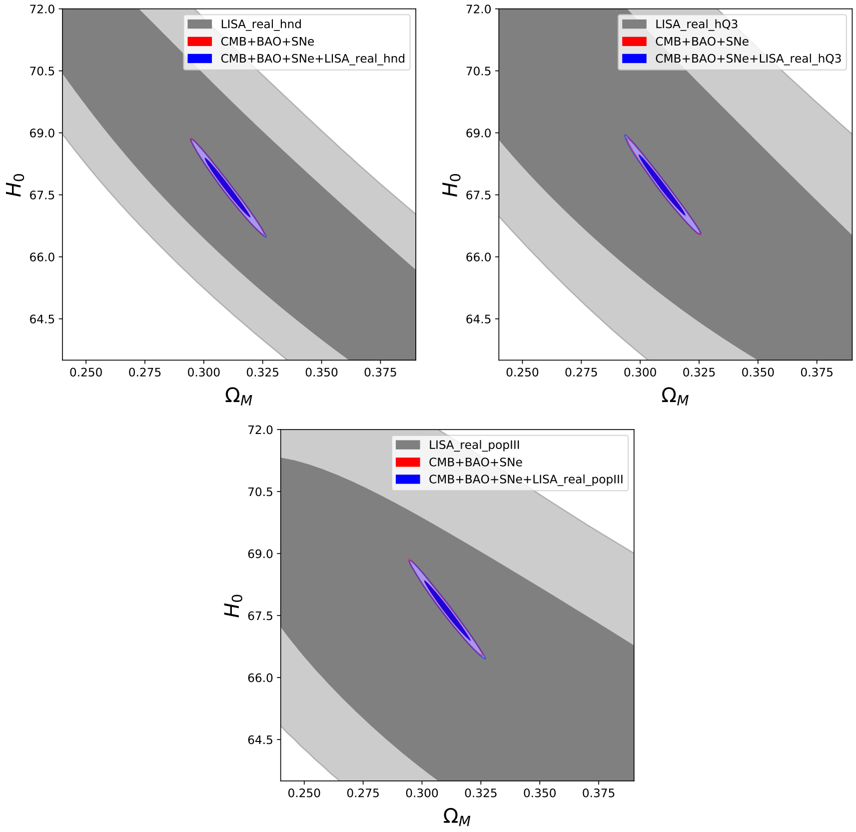

FIGURE 3.51: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, H_0) plane in ΛCDM, with the contribution from CMB+BAO+SNe (red), the contribution from LISA standard sirens (gray) and the overall combined contours (blue), in the scenario (1) (“realistic") for the error on the luminosity distance. Upper left: heavy no-delay (“hnd") scenario; Upper right: “hQ3" scenario; lower panel: “pop III" scenario. H_0 is given in the usual units $\mathrm{km\,s^{-1}Mpc^{-1}}$. Notice that the red contours are almost superimposed to the blue ones. From [200].

Therefore, from MBH binaries at LISA, we do not find a significant improvement on the accuracy on H_0, compared to current results from CMB+BAO+SNe. It should however be observed that the measurement from standard sirens is still useful because it has completely different systematic errors from that obtained with CMB, BAO and SNe.

These results, obtained within the LISA Cosmology Working Group in [200], should be compared with those found a few years ago (2016) in [188]. To perform the comparison, we should however note that the best results quoted in [188] referred to a LISA configuration with 5 Gm arms, and with all other design specifications to their most optimistic possible choices, while our results use the current LISA configuration with 2.5 Gm arms, given in [221]. The corresponding LISA sensitivity curve used in our study is quite different and generally gives worse results. For example the number of standard sirens used in our analysis is roughly half the number used in [188].

Furthermore, at the methodological level there are some differences between our analysis and that in [188]. First, as discussed above, to generate our catalogs we

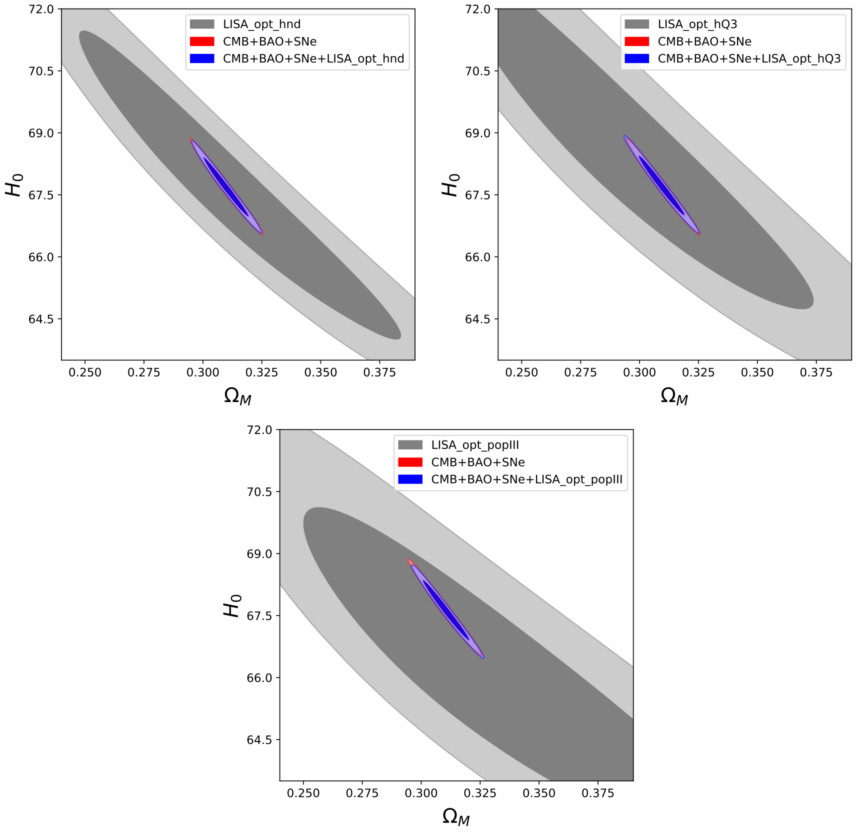

FIGURE 3.52: As in Fig. 3.51, for the scenario (2) ("optimistic") for the error on the luminosity distance. From [200].

have scattered the values $d_L^{\Lambda\mathrm{CDM}}(z_i; H_0, \Omega_M)$, extracting the "measured" values of the luminosity distance from a Gaussian distribution centered on this fiducial value, and with a width Δd_L determined by the projected error estimate. Ref. [188] did not consider scattered data, thus the statistical error due to scatter was not included in the analysis. The Fisher matrix analysis performed in [188] was performed on non-scattered data, while our MCMC analysis takes also into account the statistical uncertainty due to the scattering and thus gives more realistic results. Second, in our MCMC analysis the degeneracies between the cosmological parameters are fully taken into account by freely varying all cosmological parameters of ΛCDM (and, in the following subsections, of its extensions), while in the Fisher matrix analysis of [188] the most stringent results (that eventually, together with all other assumptions mentioned above, led to the estimate that H_0 could be measured to 0.5% with standard sirens only) where obtained by varying only H_0 and assuming a fixed prior on Ω_M. The procedure of the present analysis therefore gives a larger estimate of the error on the cosmological parameters that can be obtained with LISA standard sirens.

On the other hand, one should also be aware of the fact that, before drawing final conclusions on the sensitivity of LISA to H_0, as well as to the DE parameters that will be discussed below, much more work is needed. For instance, the counterpart model used in [188] makes a number of assumptions, which will only be validated when (if) LISA electromagnetic counterparts are actually observed. Another caveat is that the error on the localization used in our analysis is based on the use of the inspiral waveform only, with corrections based on phenomenological waveforms, and

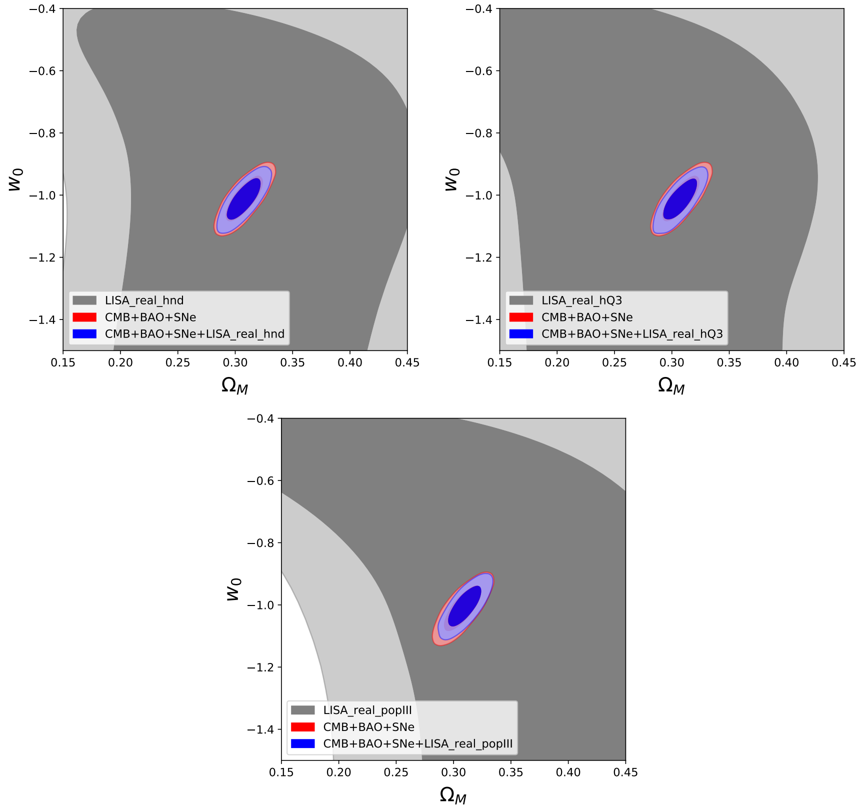

FIGURE 3.53: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ω_M, w_0) plane in wCDM, with the contribution from CMB+BAO+SNe (red), the contribution from LISA standard sirens (gray) and the overall combined contours (blue), in the scenario (1) (“realistic") for the error on the luminosity distance. Upper left: heavy no-delay (“hnd") scenario; Upper right: “hQ3" scenario; lower panel: “pop III" scenario. From [200].

it is still an open issue how this will change when including a full description of the merger and ringdown phases. The estimate of the redshift error also involves other factors, with respect to those that we have discussed, such as the exposure time, that will depend on the availability of telescope time and the pointing accuracy, which are currently difficult to estimate, as well as on the duration of the electromagnetic transient, whose estimation in turn involves the modeling of the counterpart. It should also be observed that the number of sources observed depends strongly on the sensitivity curve at low frequency, where there is potential room for improvement with respect to the sensitivity curve that we have adopted. Finally, we are only using MBHBs with counterpart as standard sirens, and we are not considering stellar mass BHBs and EMRIs, that are not expected to have an EM counterpart, but can still be treated with statistical methods (like correlation with galaxy catalogs) to obtain partial redshift information.

wCDM

We next add the parameter w_0, corresponding to the so-called wCDM model, where the DE equation of state is taken to be constant, i.e. $w_a = 0$ in eq. (3.1.29). A further natural extension would be to the (w_0, w_a) set of parameters. However we have

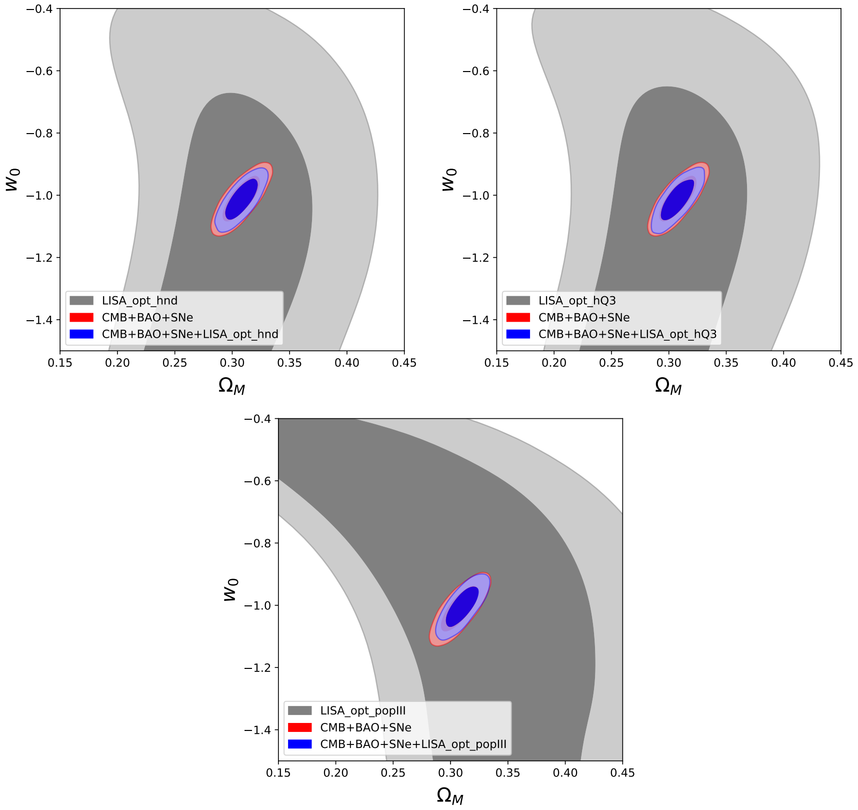

FIGURE 3.54: As in Fig. 3.53, for the scenario (2) ("optimistic") for the error on the luminosity distance. From [200].

found that, with the limited number of sources in the LISA catalogs, it is not possible to disentangle the degeneracy between these two parameters, and the MCMC chains do not reach a good convergence; so we keep $w_a = 0$. Since the model used to generate the catalog of sources is still ΛCDM, we are actually asking to what accuracy we can find back the fiducial value $w_0 = -1$. Figs. 3.53 and 3.54 show the two-dimensional likelihood in the (Ω_M, w_0) plane, displaying the combined contribution from CMB + BAO + SNe (red), and the total combined result (blue), for the six scenarios considered.

Even in the most optimistic scenario that we have considered, we learn from the plots that LISA standard sirens alone do not give any significant constraint on w_0, and, when combined to CMB + BAO + SNe data, they only induce a very marginal improvement. From the corresponding one-dimensional likelihoods, from CMB + BAO + SNe only, we find that w_0 can be reconstructed with the accuracy $\Delta w_0 = 0.045$. Combining CMB+BAO+SNe with standard sirens, in the "optimistic hnd" scenario we get

$$\Delta w_0 = 0.044\,, \tag{3.4.4}$$

so the improvement due to MBHB standard sirens is quite negligible. Basically the same results are obtained in all other scenarios considered.

As we already did in Section 3.4.2, it is interesting to compare the results of our analysis to those in [188]. Indeed, our results significantly degrade the estimates presented in [188]. As discussed in the ΛCDM case, the difference is due to the use of the updated sensitivity curve of LISA, the more realistic assumptions in the construction of the catalogs and estimates of the errors, and the fact that the degeneracies between cosmological parameters are now fully taken into account through a full MCMC. However, as for H_0, one should be aware of all the assumptions and uncertainties that entered in the construction of the source catalogs, and that could significantly alter the results. In particular these estimates can either be improved by combining the cosmological data collected with MBHBs with the ones collected from other LISA sources, such as EMRIs and stellar mass BHBs, or by extending the observational period of LISA with respect to the four years that we assumed in our analysis. In the first case the different redshift ranges where EMRIs and stellar mass BHBs are expected to be observed, will help breaking some degeneracies in the cosmological parameters, such as for example the degeneracy between H_0 and Ω_M in ΛCDM. In the second case instead the estimated errors are expected to improve roughly as $\sim \sqrt{N}$ or better, where N is the number of MBHB merger with EM counterpart observed by LISA, which linearly depends on the observational period. Both these improvements should reduce the error associated with the statistical scatter of the standard sirens data, which was not considered in [188], while in the our analysis presented here, which takes into account more realistic SBHB catalogs and an updated LISA noise sensitivity curve, appears to be the most relevant source of error. Other improvements can be achieved by a more realistic and detailed characterization of the emission and detection of the EM counterparts of MBHB mergers, which both here and in [188] has been modeled only with optical and radio EM emissions, but the consideration of other EM signatures (X-rays, γ-rays, ...) could lead to a higher number of cosmologically useful events.

(Ξ_0, w_0)

We next extend the DE sector by introducing the parameter Ξ_0. In order to keep under control the number of new parameters, which is necessary to ensure the convergence of the MCMC chains, we only take (Ξ_0, w_0) as the parameters that describe the DE sector of the theory, fixing $w_a = 0$ and $n = 2.5$; the latter value is of the order of n suggested by the RR and RT nonlocal models. However, the precise value of n is not very important for the forecasts that we present, since the uncertainty on this quantity is large for $|\Xi_0 - 1| \ll 1$ [171]. Again, we assume ΛCDM as the fiducial model used to generate the catalog of sources, so the fiducial values for the parameters Ξ_0 and w_0 are $\Xi_0 = 1$ and $w_0 = -1$.

Fig. 3.55 shows the two-dimensional likelihood in (Ξ_0, w_0) plane, for the realistic scenario for the error on the luminosity distance, and the three seed scenarios, while Fig. 3.56 shows the results obtained using the optimistic scenario for the error on the luminosity distance. We plot the limit from CMB+BAO+SNe and the combined limit from CMB+BAO+SNe+LISA standard sirens while, as before, using only standard sirens, with the addition of these new parameters with respect to ΛCDM, the MCMC chains fail to converge because of the limited number of sources. Note that CMB, BAO and SNe, as any other electromagnetic probe, are blind to Ξ_0, and therefore the corresponding contour from CMB+BAO+SNe are flat in the Ξ_0 direction. Standard sirens, however, lift this flat direction. The errors on Ξ_0 and w_0 from the corresponding one-dimensional likelihoods are shown in Tables 3.31.

Exactly as in the case of wCDM discussed above, we see that the accuracy that LISA, combined with CMB+BAO+SNe, can reach on w_0 is only of about 4.4% (at least using only MBHBs with counterpart as standard sirens), which is basically entirely determined by the current CMB+BAO+SNe observations.

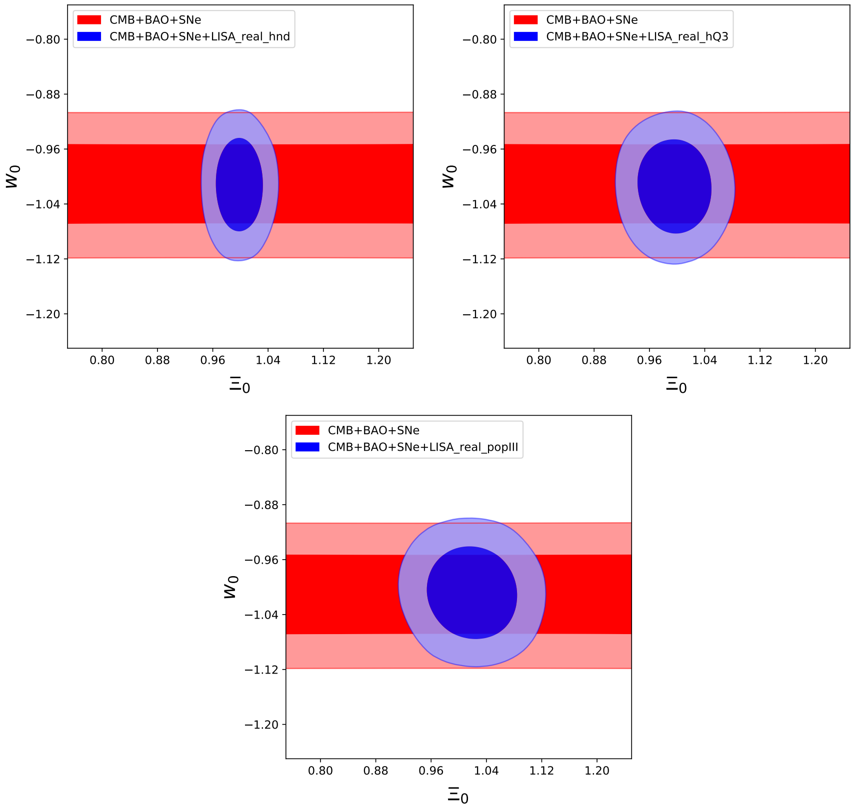

FIGURE 3.55: The 1σ and 2σ contours of the two-dimensional likelihood in the (Ξ_0, w_0) plane, with the combined contribution from CMB+BAO+SNe (red) and the combined contours from CMB+BAO+SNe+LISA standard sirens (blue), in the scenario (1) ("realistic") for the error on the luminosity distance. Upper left: heavy no-delay ("hnd") scenario; Upper right: "hQ3" scenario ; lower panel: "pop III" scenario. From [200].

In contrast, even under the current set of assumptions (see the discussion in sect. 3.4.2), that have led to relatively large errors on H_0 and on w_0 from LISA standard sirens, Ξ_0 still turns out to be an extremely interesting observable for LISA. First of all it can be observed only with GW experiments, and second, Ξ_0 can be measured more accurately than w_0. This is confirmed by the results of our MCMC, which shows that, in the best case, Ξ_0 can be measured to 1% accuracy, and even in the worst case we still have a 4.4% accuracy, see Table 3.31. By comparison, for instance, the RT minimal nonlocal model predicts a deviation from the GR value $\Xi_0 = 1$ at the level of 7% (and much larger, up to 80%, when its initial conditions are set during inflation). Values of $\mathcal{O}(\text{few }\%)$ can be obtained for scalar-tensor theories like $f(R)$, Galileon models and DHOST theories (see the plots in Section 3.1 of [200]). Thus, the accuracy that LISA can reach on modified GW propagation is extremely interesting for testing modified gravity.

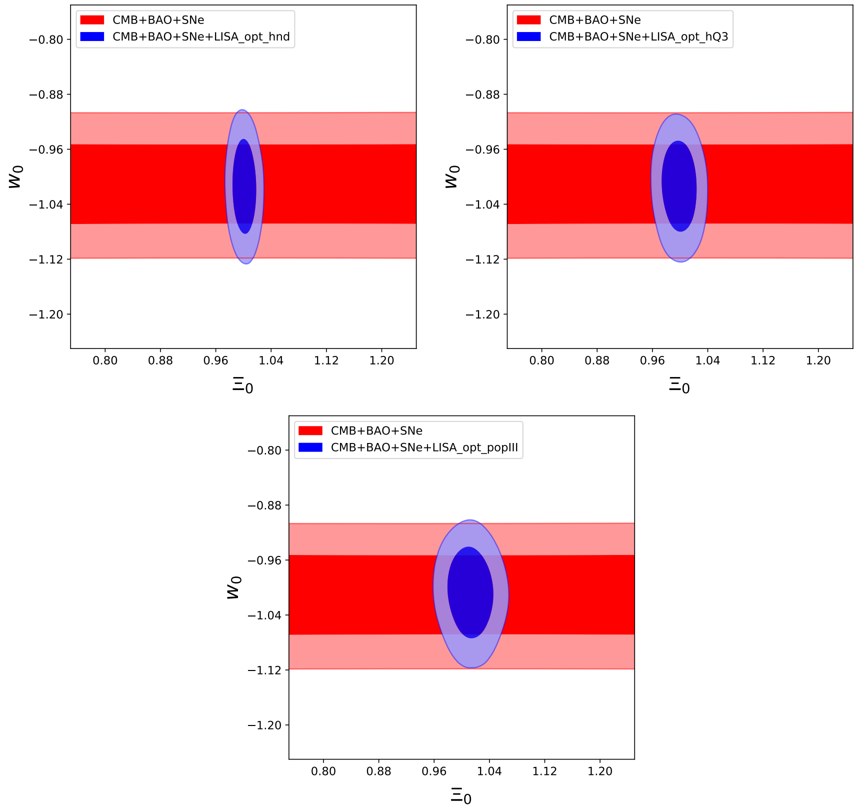

FIGURE 3.56: As in Fig. 3.55, for the scenario (2) ("optimistic") for the error on the luminosity distance. From [200].

seeds	$\Delta\Xi_0$	Δw_0
hnd	0.023	0.045
hQ3	0.036	0.046
popIII	0.044	0.045

seeds	$\Delta\Xi_0$	Δw_0
hnd	0.011	0.045
hQ3	0.017	0.044
popIII	0.022	0.044

TABLE 3.31: Forecasts for the 1σ errors on Ξ_0 and w_0 from CMB+BAO+SNe+LISA standard sirens, for the three seed scenarios. Left: realistic scenario for the error on the luminosity distance. Right: optimistic scenario. From [200].

Chapter 4

Conclusions

Let us summarize the main lessons that can be drawn from the results discussed in this thesis.

1. We have seen that it is possible to build nonlocal cosmological models which are theoretically well motivated. One specific model in this class, the RT model, passes all the currently available tests. A few points can be useful to specify better these findings and to understand their meaning in relation to the difficulties in constructing a good cosmological alternative to ΛCDM:
    - The viability of the RT model at cosmological scales is a highly non-trivial result, considering that other nonlocal theories have instabilities emerging at the background level or at the perturbations one (either in the scalar or tensor sector).
    - The same can be said for its success at Solar System scales, where the RT model smoothly reduces to GR without the need of any screening mechanism. Furthermore, the model complies with the constraints on the time variation of the effective Newton's constant from Lunar Laser Ranging; this is again a test that has excluded other nonlocal models.
    - The RT model does not simply "survive" the tests against cosmological data, but it works at the same level of ΛCDM when tested against CMB, BAO, SNe and structure formation data.
    - The RT model passes the GW170817/GRB 170817A test, because it predicts a speed of propagation of gravitational waves equal to the speed of light.

2. The predictions of the RT model at the background level and in the scalar perturbations sector are close to ΛCDM, but, surprisingly, tensor perturbations show very large deviations, up to the level of 80%, due to modifications in the friction term of the tensor propagation equation.

3. For standard sirens, a modification in the friction term implies that it is necessary to introduce a notion of 'GW luminosity distance', different from the usual electromagnetic one.

4. It should be appreciated how, studying the cosmological consequences of specific models (that could be nonlocal gravity as in this thesis, or scalar-tensor theories as in other works in the literature), one is led to investigate very general features like modified gravitational-wave propagation, which can be used to test the whole idea of corrections to GR occurring at cosmological scales.

5. For dark energy studies, modified GW propagation can be measured better than the dark energy equation of state.

6. We have seen, through a detailed analysis, that the predictions for modified GW propagation of many modified gravity theories are within the reach of 3G ground-based detectors and LISA.
7. We expect that the new observation window given by GW detectors can lead to exciting news for cosmology, and modified GW propagation will be a primary observable in that context.

Acknowledgements

First, I want to thank my supervisor Michele Maggiore for his keen guidance during the last four years and for always having been a reference point, both as a scientist and as a person. I am greatly indebted to him for all the nice ideas he shared with me, for his great enthusiasm and for his friendly advice in difficult moments.

I thank the other members of the thesis committee Alexandros Kehagias, Antonio Riotto and Gianmassimo Tasinato, for having accepted to be part of it.

Special thanks go to Stefano Foffa for his precious mentorship in many occasions and his friendly collaboration.

I am grateful to Yves Dirian for his help during the first years of my PhD, especially with the codes and the numerical work.

I thank Andreas Finke for his friendship, his frankness and for all the physics discussions we had in our office.

I want to express my gratitude to Michele Mancarella for having helped me whenever I needed it and for having shared together the teaching duties in the last semester.

My appreciation goes to Mona Jalilvand and Farbod Hassani for always being nice and helpful to me.

I also thank Davide Lombardo and Valerio De Luca for the time spent playing football on Sundays.

Last, but certainly not least, I want to thank my family for their support, especially my brother Nizar, whose point of view is never trivial and has often changed the way I looked at things, and my sister Lilia for the optimism she always gives me.

Chapter 5

Appendix: The Δ_4 model

5.1 Definition and qualitative properties

Besides the RT and RR model that have beeen the two reference nonlocal models in this thesis, there is another nonlocal gravity model associated to a mass scale that we have studied in detail, namely the Δ_4 model. In this Appendix we follow the analysis given in App. A of [30]. The Δ_4 model is defined by the quantum effective action

$$\Gamma_{\Delta_4} = \frac{M_P^2}{2} \int d^4x \sqrt{-g} \left[R - \frac{m^2}{6} R \frac{1}{\Delta_4} R \right] , \tag{5.1.1}$$

where Δ_4 is the Paneitz operator,

$$\Delta_4 \equiv \Box^2 + 2R^{\mu\nu} \nabla_\mu \nabla_\nu - \frac{2}{3} R \Box + \frac{1}{3} g^{\mu\nu} \nabla_\mu R \nabla_\nu , \tag{5.1.2}$$

As we have seen in Section 1.2.4, the Paneitz operator enters in the anomaly-induced quantum effective action in four dimensions. From the point of view of conformal invariance, is the natural generalization of the d'Alembertian (which appears in the 2D Polyakov quantum effective action) from two to four dimensions.

The linearization over Minkowski space of this model is the same as the RT or RR models (so, in the end, it is another covariantization of the same linearized equations, athough it has a less immediate form).

When studying its cosmology, the Δ_4 model shows a distinctive behavior already at the FRW background level and, as first found in [270], it has a phantom dark energy equation of state with $w_{\rm DE}(0) \simeq -1.34$. Comparison with the *Planck* limits on $w_{\rm DE}(z)$ [209] then already suggested that the model would not fit well the data, although, as was mentioned, a full analysis of the perturbations is necessary to reach a definite conclusion. We have performed such analysis and we summarize here the main methods and results.

Actually, we realized in [30] that the Δ_4 model is ruled out by the behavior of tensor perturbations, that do not propagate with the speed of light, but a sensibly lower space near the present epoch. However, the study of the background and scalar sectors of this model is also methodologically interesting. In fact, in general, in modified gravity models, a DE equation of state on the phantom side has the effect of rising the value of H_0 obtained from parameter estimation, and indeed the value of H_0 obtained from local measurements could be obtained in a wCDM model with $w \simeq -1.3$ [271].

So, the study of the scalar sector of the Δ_4 model is interesting because its DE equation of state is expected to be about the most phantom that one can have in order to fit reasonably the CMB+BAO+SNe data, and then its prediction for H_0 will give an idea of the maximum value of H_0 that could be obtained from models of this type.

It turns out that, despite the H_0 prediction in the Δ_4 model is higher than in the RT model, it is still not large enough to be helpful in the context of the Hubble tension. [Of course, this is just an hypothetical statement, that would make sense if the Δ_4 model had passed all available tests. As we said, the model fails the GW170817/GRB170817 A test for the speed of propagation of gravitational waves.]

It is also conceptually interesting to observe that the Δ_4 model (5.1.1) interpolates between the RR model and a model with a non-trivial form factor for Newton's constant. This can be seen more easily in de Sitter space, where R is constant and $R_{\mu\nu} = (1/4)Rg_{\mu\nu}$, so eq. (5.1.2) simplifies to

$$\Delta_4 = \Box\left(\Box - \frac{1}{6}R\right) . \tag{5.1.3}$$

Then, after integrations by parts, the effective action (5.1.1) becomes

$$\Gamma_{\Delta_4} = \frac{M_P^2}{2}\int d^4x\sqrt{-g}\left[R - \frac{1}{6}m^2\left(\frac{1}{\Box}R\right)\left(\frac{1}{\Box - \frac{1}{6}R}R\right)\right] , \qquad (\mathrm{deSitter}) \tag{5.1.4}$$

For Fourier modes such that $|\Box| \gg R/6$, this reduces to the RR model. In the opposite limit $|\Box| \ll R/6$, in contrast

$$\Gamma_{\Delta_4} \simeq \frac{M_P^2}{2}\int d^4x\sqrt{-g}\left(1 + \frac{m^2}{\Box}\right)R \qquad (\mathrm{deSitter}, |\Box| \ll R/6) , \tag{5.1.5}$$

corresponding to a running Newton's constant. We next recall the main results on the background evolution of the model, and we work out its cosmological perturbations.

5.2 Background evolution

The covariant equation of motion derived from (5.1.1) are

$$\begin{aligned} &G_{\alpha\beta}\left[1 - \frac{m^2}{3}S + \frac{m^2}{9}(\boldsymbol{\nabla} S)^2\right] + \frac{m^2}{6}\Bigg\{\boldsymbol{\nabla}_\alpha\boldsymbol{\nabla}_\beta\left[2S + \frac{1}{3}(\boldsymbol{\nabla} S)^2\right] \\ &\quad -g_{\alpha\beta}\left[2\Box S + \frac{1}{2}(\Box S)^2 - R^{\rho\sigma}\boldsymbol{\nabla}_\sigma S\boldsymbol{\nabla}_\rho S - \frac{1}{6}\Box\left(\boldsymbol{\nabla} S\right)^2\right] \\ &\quad +S\left(\boldsymbol{\nabla}_\rho R_{\alpha\beta} - \boldsymbol{\nabla}_{(\alpha}R_{\beta)\rho}\right)\boldsymbol{\nabla}^\rho S - 4\left(\boldsymbol{\nabla}^\rho S\right)R_{\rho(\alpha}\boldsymbol{\nabla}_{\beta)}S + 2\Box S\boldsymbol{\nabla}_\alpha\boldsymbol{\nabla}_\beta S \\ &\quad -2\left(\boldsymbol{\nabla}^\rho\boldsymbol{\nabla}_{(\alpha}S\right)\boldsymbol{\nabla}_{\beta)}\boldsymbol{\nabla}_\rho S + \frac{2}{3}R\boldsymbol{\nabla}_{(\alpha}S\boldsymbol{\nabla}_{\beta)}S + S\left(\boldsymbol{\nabla}^\lambda R_{\lambda(\alpha\beta)\rho}\right)\boldsymbol{\nabla}^\rho S \\ &\quad +2\left(\boldsymbol{\nabla}^\lambda S\right)R_{\lambda(\alpha\beta)\rho}\boldsymbol{\nabla}^\rho S - 2\left(\boldsymbol{\nabla}_{(\alpha}S\right)\left(\boldsymbol{\nabla}_{\beta)}\Box S\right)\Bigg\} = 8\pi G T_{\alpha\beta} . \end{aligned} \tag{5.2.1}$$

where S is defined by

$$S = \Delta_4^{-1}R . \tag{5.2.2}$$

Specializing to FRW, the background evolution equations are [30, 270]

$$h^2(x) = \Omega_M e^{-3x} + \Omega_R e^{-4x} + \gamma Y(x) , \tag{5.2.3}$$

where again $\gamma = m^2/(9H_0^2)$, while

$$Y = \frac{1}{2}W'(6 - U' - 2U) + W(3 - 6\zeta + \zeta U' + 2\zeta U) + \frac{1}{4}U^2\,, \tag{5.2.4}$$

and we have introduced two auxiliary fields

$$\begin{aligned} W &\equiv H^2 S\,, && (5.2.5)\\ U &\equiv a^{-2}\frac{\partial^2 S}{\partial\eta^2} = H^2\left[S'' + (1+\zeta)S'\right]\,, && (5.2.6)\end{aligned}$$

where η is conformal time (and, as in Section 2.1.1 for the RT model, the prime denotes d/dx, where $x = \log a$). The introduction of the two fields U, W allows us to split the fourth-order equation $\Delta_4 S = R$ into a couple of second-order equations,

$$\begin{aligned} U'' + (5+\zeta)U' + (6+2\zeta)U &= 6(2+\zeta)\,. && (5.2.7)\\ W'' + (1-3\zeta)W' + 2(\zeta^2 - \zeta - \zeta')W &= U\,. && (5.2.8)\end{aligned}$$

Setting $\zeta(x) = \zeta_0$ constant we find that the most general solution of eq. (5.2.7) is

$$U(x) = \frac{3(2+\zeta_0)}{3+\zeta_0} + u_1 e^{-(3+\zeta_0)x} + u_2 e^{-2x}\,. \tag{5.2.9}$$

Therefore both homogeneous solutions are decaying modes, in all cosmological epochs, and even the inhomogeneous solution is constant, rather than linearly growing in x as in eq. (2.1.11) for the RT model. The same holds for W, since the homogeneous equation $W'' + (1 - 3\zeta_0)W' + 2(\zeta_0^2 - \zeta_0)W = 0$ has the solutions $W = e^{\beta_\pm x}$ with $\beta_+ = 2\zeta_0$ and $\beta_- = -1 + \zeta_0$. The value β_- is negative in all three eras, while β_+ is negative in RD and MD and vanishes, corresponding to a constant solution, in dS. Thus, there is no growing mode and the cosmological evolution is stable. Thus, even if we set the initial conditions of order one for U and W in a earlier inflationary epoch, U and W still enter the RD era with a value of order one. In RD $\zeta_0 = -2$, so the inhomogeneous term in eq. (5.2.9) vanishes, and for U the de Sitter solution is matched to the two decaying modes e^{-x} and e^{-2x}. Thus, the solution is quickly attracted toward the one obtained setting $u_0 = 0$ deep in RD. Similarly, for W the solution that emerges from de Sitter is matched to its two decaying modes in RD. Thus, the solution obtained setting the initial conditions $U = W = U' = W' = 0$ at some initial time deep in RD is an attractor and, in the Δ_4 model there is no free parameter associated to the boundary conditions. This makes the model very predictive.

We can now integrate numerically the equations of motion, with initial conditions $U = W = U' = W' = 0$ at some initial time deep in RD. The result of the numerical integration is shown in Fig. 5.1 [30, 270]. We have an effective dark energy density $\rho_{\rm DE} = \rho_0\gamma Y$. We see that, once again, the effective DE density vanishes deep in RD, and begins to grow as we approach radiation-matter equality (around $x_{\rm eq} \simeq -8.1$). In the Δ_4 model, during MD $\rho_{\rm DE}(x)$ eventually stabilizes to a constant, leading to a long phase where the EoS parameter $w(z) \simeq -1$ as in ΛCDM. However, as the DE density starts to dominate over matter, $\rho_{\rm DE}$ finally grows again, leading near $z = 0$ to an EoS with $w_{\rm DE}(z = 0) \simeq -1.36$ (with the choice $\Omega_M \simeq 0.29$ and $h_0 \simeq 0.71$ obtained from parameter estimation in this model, see below). Using again the parametrization (2.1.19) in the region $-1 < x < 0$, for our best-fit values $\Omega_M \simeq 0.29$ and $h_0 \simeq 0.71$ we get $w_0 \simeq -1.33$ and $w_a \simeq 0.53$.

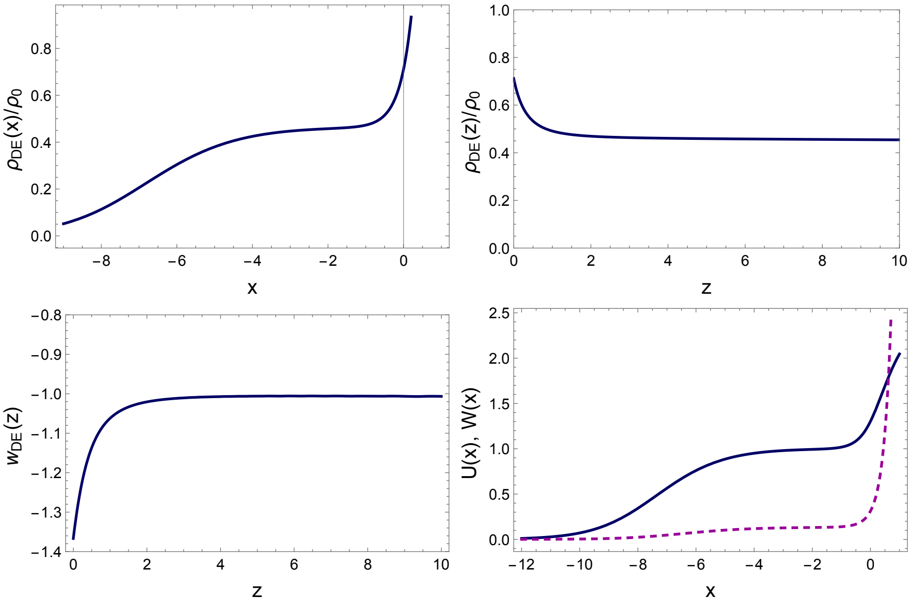

FIGURE 5.1: Figure from [30]. Background evolution of the Δ_4 model, using $\Omega_M \simeq 0.29$ and $h_0 \simeq 0.70$, see Table 5.1.

5.3 Cosmological perturbations in the scalar sector

At the perturbation level we write the metric in the scalar sector as in eq. (2.1.20). The auxiliary field S defined in eq. (5.2.2) satisfies the fourth-order equation $\Delta_4 S = R$. We find convenient to split this fourth-order equation into a pair of second-order equation, introducing a second auxiliary field. In a generic space-time, it is not possible to perform this split covariantly [although this would be possible in de Sitter, see eq. (5.1.3)], so we simply keep the definitions (5.2.5, 5.2.6) also when performing perturbations over an FRW background, i.e. we write $W(\eta, \mathbf{x}) \equiv H^2(\eta) S(\eta, \mathbf{x})$ and $U(\eta, \mathbf{x}) \equiv a^{-2}\partial_\eta^2 S(\eta, \mathbf{x})$. For studying the perturbations it is convenient to use the variable $V = W/h^2 = H_0^2 S$ rather than W. We then expand $V = \bar{V} + \delta V$. As a second perturbation variable for the auxiliary field, instead of δU, it is convenient to chose

$$\delta Z \equiv h^2 \left[\delta V'' + (1+\zeta)\,\delta V' + \bar{V}'\left(\Phi' - \Psi'\right)\right] + 2\left(\Phi - \Psi\right)\bar{U}\,, \qquad (5.3.1)$$

because then higher-derivative terms drop out of the equations. The above equation can be taken as a dynamical equation for δV, while for δZ one finds

$$\delta Z'' + (5+\zeta)\,\delta Z' + \left(6 + 2\zeta + 2\hat{k}^2\right)\delta Z = -h^2\hat{k}^4\delta V + 6\left[\Phi'' + (3+\zeta)\Phi'\right] \qquad (5.3.2)$$
$$+ \left(\bar{U}' + 2\bar{U} - 6 - \frac{\hat{k}^2}{3}h^2\bar{V}'\right)(\Psi' - \Phi') + 2\hat{k}^2(1 - \frac{2}{3}\bar{U})\Psi + 4\left[3(\zeta+2) + \hat{k}^2\left(1 + \frac{\bar{U}}{3}\right)\right]\Phi\,.$$

The analogous of eqs. (2.1.26)–(2.1.29) are

$$\begin{aligned}
&(1-3\gamma\bar{V})\left(\hat{k}^2\Phi+3\Phi'-3\Psi\right)+\frac{3\gamma}{2}\Bigg\{\left(\bar{V}'-\frac{\bar{U}}{2h^2}\right)\delta Z+\frac{1}{2}\bar{V}'\delta Z' \qquad (5.3.3)\\
&\quad-\left[3+\hat{k}^2\left(1+\frac{\bar{U}}{2}\right)\right]\delta V+6\bar{V}'\Psi-3\bar{V}'\Phi'+\frac{\bar{U}^2}{h^2}\Phi+\left(\frac{\bar{U}'}{2}+\bar{U}-3+\frac{5}{6}\hat{k}^2h^2\bar{V}'\right)\delta V'\\
&\quad-\left(\bar{U}'+2\bar{U}\right)\bar{V}'\left(\Phi+\Psi\right)+\frac{2}{3}\hat{k}^2h^2\bar{V}'^2\left(\Phi-\Psi\right)\Bigg\}=\frac{3}{2h^2}\left(\Omega_{\rm R}e^{-4x}\delta_{\rm R}+\Omega_{\rm M}e^{-3x}\delta_{\rm M}\right),\\
&(1-3\gamma\bar{V})\,\hat{k}^2(\Phi'-\Psi)+\frac{\gamma\hat{k}^2}{2}\Bigg[\left(\bar{U}-3\right)\left(\delta V'-\bar{V}'\Psi\right)+\bar{U}\bar{V}'\Phi \qquad (5.3.4)\\
&\quad+\frac{3}{2}\left(2-\bar{U}'-2\bar{U}-\hat{k}^2h^2\bar{V}'\right)\delta V-\frac{1}{2}\bar{V}'\delta Z\Bigg]=-\frac{3}{2h^2}\left(\frac{4}{3}\Omega_{\rm R}e^{-4x}\hat{\theta}_{\rm R}+\Omega_{\rm M}e^{-3x}\hat{\theta}_{\rm M}\right),
\end{aligned}$$

$$\begin{aligned}
&(1-3\gamma\bar{V})\left[\Phi''+(3+\zeta)\Phi'-\Psi'-(3+2\zeta)\Psi+\frac{\hat{k}^2}{3}(\Phi+\Psi)\right]=-\frac{1}{2h^2}\Omega_{\rm R}e^{-4x}\delta_{\rm R}\\
&\quad+\frac{\gamma}{2}\Bigg\{\left[3\Phi'-\left(6+\bar{U}'+2\bar{U}+\frac{2}{3}\hat{k}^2h^2\bar{V}'\right)\Psi+\left(\frac{2}{3}\hat{k}^2h^2\bar{V}'-\bar{U}'-2\bar{U}\right)\Phi\right]\bar{V}'\\
&\quad+\left(\bar{U}-6\right)\frac{\bar{U}}{h^2}\Phi+\left(\bar{V}'-\frac{\bar{U}}{2h^2}+\frac{3}{h^2}\right)\delta Z+\frac{1}{2}\bar{V}'\delta Z'+\left[3(3+2\zeta)+\left(2-\frac{\bar{U}}{2}\right)\hat{k}^2\right]\delta V\\
&\quad+\left(\bar{U}'+2\bar{U}+6+\frac{5}{3}\hat{k}^2h^2\bar{V}'\right)\frac{\delta V'}{2}\Bigg\}, \qquad (5.3.5)\\
&(1-3\gamma\bar{V}+\gamma\bar{V}'^2h^2)\Psi+(1-3\gamma\bar{V}-\gamma\bar{V}'^2h^2)\Phi-3\gamma\delta V(1-\bar{U})+\gamma h^2\bar{V}'\delta V'=0\,. \quad (5.3.6)
\end{aligned}$$

The results of the numerical integration shows that the cosmological perturbations are again stable and relatively close to those of ΛCDM.

5.4 Parameter estimation for the Δ_4 model

We have then implemented the perturbations in our Boltzmann code and performed Bayesian parameter estimation. The results are shown in Table 5.1 (using CMB + BAO + SNe). We also write the results for $\nu\Lambda$CDM and we give the difference in χ^2 compared to it. We see that the Δ_4 model indeed predicts a higher value of H_0, although not sensibly higher than that in the RT model. On the other hand, its χ^2 is significantly worse than than that of $\nu\Lambda$CDM, even including H_0 in the dataset, so the model is already very strongly disfavored by the study of the scalar perturbations.

5.5 Tensor perturbations

On top of this, it was found in [30] that the model is ruled out by the fact that its tensor perturbations do not propagate at the speed of light. This is also interesting from the methodological point of view since it shows that, for these nonlocal models, it is not obvious a priori to satisfy this constraint. The equation for tensor perturbations

	CMB+BAO+SNe	
Parameter	$\nu\Lambda$CDM	Δ_4
H_0	$67.60^{+0.66}_{-0.55}$	$70.27^{+0.95}_{-0.94}$
$\sum_\nu m_\nu$ [eV]	< 0.10 (at 1σ)	$0.185^{+0.087}_{-0.096}$
ω_c	$0.1189^{+0.0011}_{-0.0011}$	$0.1202^{+0.0014}_{-0.0014}$
$100\omega_b$	$2.229^{+0.014}_{-0.015}$	$2.217^{+0.017}_{-0.017}$
$\ln(10^{10}A_s)$	$3.071^{+0.026}_{-0.029}$	$3.080^{+0.034}_{-0.036}$
n_s	$0.9661^{+0.0043}_{-0.0043}$	$0.9637^{+0.0050}_{-0.0050}$
$\tau_{\rm re}$	$0.06965^{+0.01393}_{-0.01549}$	$0.07280^{+0.01769}_{-0.01927}$
Ω_M	$0.3109^{+0.0069}_{-0.0084}$	$0.2925^{+0.0096}_{-0.0101}$
$z_{\rm re}$	$9.150^{+1.396}_{-1.355}$	$9.490^{+1.793}_{-1.656}$
σ_8	$0.8157^{+0.0135}_{-0.0104}$	$0.8240^{+0.0199}_{-0.0177}$
$\chi^2_{\rm min}$	13630.78	13649.98
$\Delta\chi^2_{\rm min}$	0	19.20

TABLE 5.1: Parameter estimation and χ^2 values for $\nu\Lambda$CDM and the Δ_4 model, using the CMB, BAO and SNe datasets. From [30].

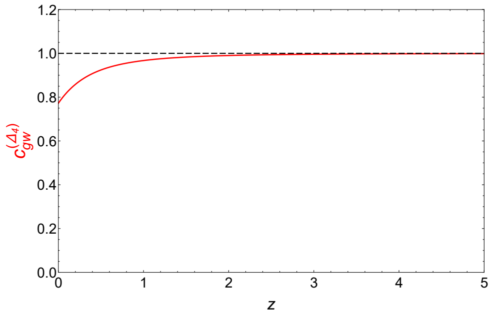

FIGURE 5.2: The speed of gravitational waves in the Δ_4 model as a function of the redshift. From [30].

in the Δ_4 model is

$$\begin{aligned}&(1-3\gamma\bar{V})\left(\partial_\eta^2\tilde{h}_A+2\mathcal{H}\partial_\eta\tilde{h}_A+k^2\tilde{h}_A\right)\\&+\gamma\left[\frac{(\partial_\eta\bar{V})^2}{a^2H_0^2}\left(2\partial_\eta^2\tilde{h}_A-k^2\tilde{h}_A\right)+\left(4\bar{U}-3\right)\partial_\eta\bar{V}\partial_\eta\tilde{h}_A\right]=16\pi Ga^2\tilde{\sigma}_A\,,\end{aligned}\tag{5.5.1}$$

and the corresponding speed of gravitational waves in a FRW background is

$$c_{\rm gw}^{(\Delta_4)}=\sqrt{\frac{1-3\gamma\bar{V}-\frac{\gamma}{a^2H_0^2}(\partial_\eta\bar{V})^2}{1-3\gamma\bar{V}+2\frac{\gamma}{a^2H_0^2}(\partial_\eta\bar{V})^2}}\,.\tag{5.5.2}$$

This quantity is always smaller than one, and in the recent epoch it differs from one significantly, see Fig. 5.2. Thus, the Δ_4 model is ruled out by the GW170817/GRB 170817A test.

Bibliography

[1] C. W. Misner, K. Thorne, and J. Wheeler, *Gravitation*. Freeman, New York, 1973.

[2] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron.J.* **116** (1998) 1009–1038, `arXiv:astro-ph/9805201 [astro-ph]`.

[3] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et al.*, "Measurements of Omega and Lambda from 42 high redshift supernovae," *Astrophys.J.* **517** (1999) 565–586, `arXiv:astro-ph/9812133 [astro-ph]`.

[4] S. Weinberg, "The Cosmological Constant Problem," *Rev. Mod. Phys.* **61** (1989) 1–23.

[5] M. Maggiore, *A Modern Introduction to Quantum Field Theory*. Oxford University Press, 312 p, 2005.

[6] A. Padilla, "Lectures on the Cosmological Constant Problem," `arXiv:1502.05296 [hep-th]`.

[7] C. Brans and R. Dicke, "Mach's principle and a relativistic theory of gravitation," *Phys. Rev.* **124** (1961) 925–935.

[8] D. Lovelock, "The Einstein tensor and its generalizations," *J. Math. Phys.* **12** (1971) 498–501.

[9] T. Padmanabhan, "Emergent Gravity Paradigm: Recent Progress," *Mod. Phys. Lett. A* **30** no. 03n04, (2015) 1540007, `arXiv:1410.6285 [gr-qc]`.

[10] E. P. Verlinde, "On the Origin of Gravity and the Laws of Newton," *JHEP* **04** (2011) 029, `arXiv:1001.0785 [hep-th]`.

[11] G. W. Horndeski, "Second-order scalar-tensor field equations in a four-dimensional space," *Int. J. Theor. Phys.* **10** (1974) 363–384.

[12] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, "Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order," *JHEP* **12** (2016) 100, `arXiv:1608.08135 [hep-th]`.

[13] J. Ben Achour, D. Langlois, and K. Noui, "Degenerate higher order scalar-tensor theories beyond Horndeski and disformal transformations," *Phys. Rev. D* **93** no. 12, (2016) 124005, `arXiv:1602.08398 [gr-qc]`.

[14] M. Crisostomi, K. Koyama, and G. Tasinato, "Extended Scalar-Tensor Theories of Gravity," *JCAP* **04** (2016) 044, `arXiv:1602.03119 [hep-th]`.

[15] D. Langlois, "Dark energy and modified gravity in degenerate higher-order scalar–tensor (DHOST) theories: A review," *Int. J. Mod. Phys. D* **28** no. 05, (2019) 1942006, `arXiv:1811.06271 [gr-qc]`.

[16] C. Eling, T. Jacobson, and D. Mattingly, "Einstein-Aether theory," in *Deserfest: A Celebration of the Life and Works of Stanley Deser*, pp. 163–179. 10, 2004. `arXiv:gr-qc/0410001`.

[17] S. Hassan and R. A. Rosen, "Bimetric Gravity from Ghost-free Massive Gravity," *JHEP* **1202** (2012) 126, `arXiv:1109.3515 [hep-th]`.

[18] M. Fierz and W. Pauli, "On relativistic wave equations for particles of arbitrary spin in an electromagnetic field," *Proc.Roy.Soc.Lond.* **A173** (1939) 211–232.

[19] D. Boulware and S. Deser, "Can gravitation have a finite range?," *Phys.Rev.* **D6** (1972) 3368–3382.

[20] D. Comelli, M. Crisostomi, F. Nesti, and L. Pilo, "Degrees of Freedom in Massive Gravity," *Phys.Rev.* **D86** (2012) 101502, `arXiv:1204.1027 [hep-th]`.

[21] C. de Rham, G. Gabadadze, and A. J. Tolley, "Resummation of Massive Gravity," *Phys.Rev.Lett.* **106** (2011) 231101, `arXiv:1011.1232 [hep-th]`.

[22] K. Hinterbichler, "Theoretical Aspects of Massive Gravity," *Rev.Mod.Phys.* **84** (2012) 671–710, `arXiv:1105.3735 [hep-th]`.

[23] G. Dvali, G. Gabadadze, and M. Porrati, "4-D gravity on a brane in 5-D Minkowski space," *Phys.Lett.* **B485** (2000) 208–214, `arXiv:hep-th/0005016 [hep-th]`.

[24] A. De Felice and S. Tsujikawa, "f(R) theories," *Living Rev. Rel.* **13** (2010) 3, `arXiv:1002.4928 [gr-qc]`.

[25] P. Horava, "Quantum Gravity at a Lifshitz Point," *Phys. Rev. D* **79** (2009) 084008, `arXiv:0901.3775 [hep-th]`.

[26] M. Ostrogradsky, "Mémoires sur les équations différentielles, relatives au problème des isopérimètres," *Mem. Acad. St. Petersbourg* **6** no. 4, (1850) 385–517.

[27] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford Master Series in Physics. Oxford University Press, 2007. `http://www.oup.com/uk/catalogue/?ci=9780198570745`.

[28] T. Taylor and G. Veneziano, "Quantum Gravity at Large Distances and the Cosmological Constant," *Nucl.Phys.* **B345** (1990) 210–230.

[29] M. Maggiore, "Nonlocal Infrared Modifications of Gravity. A Review," *Fundam. Theor. Phys.* **187** (2017) 221–281, `arXiv:1606.08784 [hep-th]`.

[30] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Nonlocal gravity. Conceptual aspects and cosmological predictions," *JCAP* **1803** (2018) 002, `arXiv:1712.07066 [hep-th]`.

[31] E. Belgacem, Y. Dirian, A. Finke, S. Foffa, and M. Maggiore, "Gravity in the infrared and effective nonlocal models," *JCAP* **04** (2020) 010, `arXiv:2001.07619 [astro-ph.CO]`.

[32] A. Strominger, "Lectures on the Infrared Structure of Gravity and Gauge Theory," `arXiv:1703.05448 [hep-th]`.

[33] H. Bondi, M. van der Burg, and A. Metzner, "Gravitational waves in general relativity. 7. Waves from axisymmetric isolated systems," *Proc. Roy. Soc. Lond. A* **269** (1962) 21–52.

[34] R. Sachs, "Gravitational waves in general relativity. 8. Waves in asymptotically flat space-times," *Proc. Roy. Soc. Lond. A* **270** (1962) 103–126.

[35] I. Antoniadis and E. Mottola, "Graviton Fluctuations in De Sitter Space," *J.Math.Phys.* **32** (1991) 1037–1044.

[36] I. Antoniadis and E. Mottola, "4-D quantum gravity in the conformal sector," *Phys.Rev.* **D45** (1992) 2013–2025.

[37] A. Higuchi, D. Marolf, and I. A. Morrison, "de Sitter invariance of the dS graviton vacuum," *Class. Quant. Grav.* **28** (2011) 245012, `arXiv:1107.2712 [hep-th]`.

[38] S. P. Miao, N. C. Tsamis, and R. P. Woodard, "Gauging away Physics," *Class. Quant. Grav.* **28** (2011) 245013, `arXiv:1107.4733 [gr-qc]`.

[39] A. Rajaraman, "de Sitter Space is Unstable in Quantum Gravity," *Phys. Rev.* **D94** no. 12, (2016) 125025, `arXiv:1608.07237 [hep-th]`.

[40] N. C. Tsamis and R. P. Woodard, "Strong infrared effects in quantum gravity," *Annals Phys.* **238** (1995) 1–82.

[41] A. A. Starobinsky and J. Yokoyama, "Equilibrium state of a selfinteracting scalar field in the De Sitter background," *Phys. Rev.* **D50** (1994) 6357–6368, `arXiv:astro-ph/9407016 [astro-ph]`.

[42] A. Riotto and M. S. Sloth, "On Resumming Inflationary Perturbations beyond One-loop," *JCAP* **0804** (2008) 030, `arXiv:0801.1845 [hep-ph]`.

[43] C. Burgess, L. Leblond, R. Holman, and S. Shandera, "Super-Hubble de Sitter Fluctuations and the Dynamical RG," *JCAP* **1003** (2010) 033, `arXiv:0912.1608 [hep-th]`.

[44] A. Rajaraman, "On the proper treatment of massless fields in Euclidean de Sitter space," *Phys. Rev.* **D82** (2010) 123522, `arXiv:1008.1271 [hep-th]`.

[45] M. Beneke and P. Moch, "On dynamical mass generation in Euclidean de Sitter space," *Phys. Rev.* **D87** (2013) 064018, `arXiv:1212.3058 [hep-th]`.

[46] F. Gautier and J. Serreau, "Infrared dynamics in de Sitter space from Schwinger-Dyson equations," *Phys. Lett.* **B727** (2013) 541–547, `arXiv:1305.5705 [hep-th]`.

[47] D. López Nacir, F. D. Mazzitelli, and L. G. Trombetta, "$O(N)$ model in Euclidean de Sitter space: beyond the leading infrared approximation," *JHEP* **09** (2016) 117, `arXiv:1606.03481 [hep-th]`.

[48] I. Antoniadis, P. O. Mazur, and E. Mottola, "Cosmological dark energy: Prospects for a dynamical theory," *New J.Phys.* **9** (2007) 11, `arXiv:gr-qc/0612068 [gr-qc]`.

[49] N. Birrell and P. Davies, *Quantum fields in curved space*. Cambridge University Press, 1982.

[50] A. Barvinsky and G. Vilkovisky, "The Generalized Schwinger-DeWitt Technique in Gauge Theories and Quantum Gravity," *Phys.Rept.* **119** (1985) 1.

[51] A. Barvinsky and G. Vilkovisky, "Beyond the Schwinger-DeWitt Technique: Converting Loops Into Trees and In-In Currents," *Nucl.Phys.* **B282** (1987) 163.

[52] I. L. Buchbinder, S. D. Odintsov, and I. L. Shapiro, *Effective action in quantum gravity*. Institute of Physics, Bristol, UK, 1992.

[53] V. Mukhanov and S. Winitzki, *Introduction to quantum effects in gravity*. Cambridge University Press, Cambridge, 2007.

[54] I. L. Shapiro, "Effective Action of Vacuum: Semiclassical Approach," *Class. Quant. Grav.* **25** (2008) 103001, `arXiv:0801.0216 [gr-qc]`.

[55] R. Jordan, "Effective Field Equations for Expectation Values," *Phys.Rev.* **D33** (1986) 444.

[56] E. Calzetta and B. Hu, "Closed Time Path Functional Formalism in Curved Space-Time: Application to Cosmological Back Reaction Problems," *Phys.Rev.* **D35** (1987) 495.

[57] D. A. Dalvit and F. D. Mazzitelli, "Running coupling constants, Newtonian potential and nonlocalities in the effective action," *Phys.Rev.* **D50** (1994) 1001–1009, `arXiv:gr-qc/9402003 [gr-qc]`.

[58] A. V. Manohar, "Effective field theories," *Lect.Notes Phys.* **479** (1997) 311–362, `arXiv:hep-ph/9606222 [hep-ph]`.

[59] A. Dobado and A. L. Maroto, "Particle production from nonlocal gravitational effective action," *Phys.Rev.* **D60** (1999) 104045, `arXiv:gr-qc/9803076 [gr-qc]`.

[60] V. G. Knizhnik, A. M. Polyakov, and A. B. Zamolodchikov, "Fractal Structure of 2D Quantum Gravity," *Mod. Phys. Lett.* **A3** (1988) 819.

[61] F. David, "Conformal Field Theories Coupled to 2D Gravity in the Conformal Gauge," *Mod. Phys. Lett.* **A3** (1988) 1651.

[62] J. Distler and H. Kawai, "Conformal Field Theory and 2D Quantum Gravity," *Nucl. Phys.* **B321** (1989) 509–527.

[63] J. Polchinski, "A Two-Dimensional Model for Quantum Gravity," *Nucl. Phys.* **B324** (1989) 123–140.

[64] R. J. Riegert, "A Nonlocal Action for the Trace Anomaly," *Phys. Lett.* **B134** (1984) 56–60.

[65] G. Dvali, "Predictive Power of Strong Coupling in Theories with Large Distance Modified Gravity," *New J.Phys.* **8** (2006) 326, `arXiv:hep-th/0610013 [hep-th]`.

[66] M. Porrati, "Higgs phenomenon for 4-D gravity in anti-de Sitter space," *JHEP* **04** (2002) 058, `arXiv:hep-th/0112166 [hep-th]`.

[67] M. Jaccard, M. Maggiore, and E. Mitsou, "A non-local theory of massive gravity," *Phys.Rev.* **D88** (2013) 044033, `arXiv:1305.3034 [hep-th]`.

[68] P. Boucaud, A. Le Yaouanc, J. P. Leroy, J. Micheli, O. Pene, and J. Rodriguez-Quintero, "Testing Landau gauge OPE on the lattice with a $\langle A^2 \rangle$ condensate," *Phys. Rev.* **D63** (2001) 114003, `arXiv:hep-ph/0101302 [hep-ph]`.

[69] M. A. L. Capri, D. Dudal, J. A. Gracey, V. E. R. Lemes, R. F. Sobreiro, S. P. Sorella, and H. Verschelde, "A Study of the gauge invariant, nonlocal mass operator $\mathrm{Tr} \int d^4x F_{\mu\nu} (D^2)^{-1} F_{\mu\nu}$ in Yang-Mills theories," *Phys. Rev.* **D72** (2005) 105016, `arXiv:hep-th/0510240 [hep-th]`.

[70] D. Dudal, J. A. Gracey, S. P. Sorella, N. Vandersickel, and H. Verschelde, "A Refinement of the Gribov-Zwanziger approach in the Landau gauge: Infrared propagators in harmony with the lattice results," *Phys. Rev.* **D78** (2008) 065047, `arXiv:0806.4348 [hep-th]`.

[71] S. Foffa, M. Maggiore, and E. Mitsou, "Apparent ghosts and spurious degrees of freedom in non-local theories," *Phys.Lett.* **B733** (2014) 76–83, `arXiv:1311.3421 [hep-th]`.

[72] E. E. Flanagan and S. A. Hughes, "The Basics of gravitational wave theory," *New J.Phys.* **7** (2005) 204, `arXiv:gr-qc/0501041 [gr-qc]`.

[73] M. Jaccard, M. Maggiore, and E. Mitsou, "Bardeen variables and hidden gauge symmetries in linearized massive gravity," *Phys.Rev.* **D87** (2013) 044017, `arXiv:1211.1562 [hep-th]`.

[74] M. Maggiore, *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*. Oxford University Press, 2018. `https://global.oup.com/academic/product/gravitational-waves-9780198570899?cc=de&lang=en&#`.

[75] S. Deser, "Covariant Decomposition and the Gravitational Cauchy Problem," *Ann.Inst.Henri Poincare* **7** (1967) 149.

[76] J. J. York, "Covariant decompositions of symmetric tensors in the theory of gravitation," *Ann.Inst.Henri Poincare* **21** (1974) 319.

[77] M. Maggiore, "Phantom dark energy from nonlocal infrared modifications of general relativity," *Phys.Rev.* **D89** (2014) 043008, `arXiv:1307.3898 [hep-th]`.

[78] M. Maggiore and M. Mancarella, "Non-local gravity and dark energy," *Phys.Rev.* **D90** (2014) 023005, `arXiv:1402.0448 [hep-th]`.

[79] S. Nojiri and S. D. Odintsov, "Modified non-local-F(R) gravity as the key for the inflation and dark energy," *Phys.Lett.* **B659** (2008) 821–826, `arXiv:0708.0924 [hep-th]`.

[80] S. Jhingan, S. Nojiri, S. Odintsov, M. Sami, I. Thongkool, *et al.*, "Phantom and non-phantom dark energy: The Cosmological relevance of non-locally corrected gravity," *Phys.Lett.* **B663** (2008) 424–428, `arXiv:0803.2613 [hep-th]`.

[81] N. Koshelev, "Comments on scalar-tensor representation of nonlocally corrected gravity," *Grav.Cosmol.* **15** (2009) 220–223, `arXiv:0809.4927 [gr-qc]`.

[82] T. Koivisto, "Newtonian limit of nonlocal cosmology," *Phys.Rev.* **D78** (2008) 123505, `arXiv:0807.3778 [gr-qc]`.

[83] T. S. Koivisto, "Cosmology of modified (but second order) gravity," *AIP Conf.Proc.* **1206** (2010) 79–96, `arXiv:0910.4097 [gr-qc]`.

[84] A. O. Barvinsky, "Serendipitous discoveries in nonlocal gravity theory," *Phys.Rev.* **D85** (2012) 104018, `arXiv:1112.4340 [hep-th]`.

[85] S. Deser and R. Woodard, "Observational Viability and Stability of Nonlocal Cosmology," *JCAP* **1311** (2013) 036, `arXiv:1307.6639 [astro-ph.CO]`.

[86] A. Kehagias and M. Maggiore, "Spherically symmetric static solutions in a non-local infrared modification of General Relativity," *JHEP* **1408** (2014) 029, `arXiv:1401.8289 [hep-th]`.

[87] M. Maggiore, "Perturbative loop corrections and nonlocal gravity," *Phys. Rev.* **D93** (2016) 063008, `arXiv:1603.01515 [hep-th]`.

[88] G. 't Hooft and M. Veltman, "One loop divergencies in the theory of gravitation," *Annales Poincare Phys.Theor.* **A20** (1974) 69–94.

[89] A. O. Barvinsky and G. A. Vilkovisky, "Covariant perturbation theory. 2: Second order in the curvature. General algorithms," *Nucl. Phys.* **B333** (1990) 471–511.

[90] E. V. Gorbar and I. L. Shapiro, "Renormalization group and decoupling in curved space," *JHEP* **0302** (2003) 021, `arXiv:hep-ph/0210388 [hep-ph]`.

[91] E. V. Gorbar and I. L. Shapiro, "Renormalization group and decoupling in curved space. 2. The Standard model and beyond," *JHEP* **0306** (2003) 004, `arXiv:hep-ph/0303124 [hep-ph]`.

[92] J. F. Donoghue, "General relativity as an effective field theory: The leading quantum corrections," *Phys. Rev.* **D50** (1994) 3874–3888, `arXiv:gr-qc/9405057 [gr-qc]`.

[93] A. Codello and R. K. Jain, "On the covariant formalism of the effective field theory of gravity and leading order corrections," *Class. Quant. Grav.* **33** (2016) 225006, `arXiv:1507.06308 [gr-qc]`.

[94] A. Codello and R. K. Jain, "On the covariant formalism of the effective field theory of gravity and its cosmological implications," *Class. Quant. Grav.* **34** (2017) 035015, `arXiv:1507.07829 [astro-ph.CO]`.

[95] J. Polchinski, "Renormalization and Effective Lagrangians," *Nucl. Phys.* **B231** (1984) 269–295.

[96] C. Wetterich, "Exact evolution equation for the effective potential," *Phys. Lett.* **B301** (1993) 90–94, `arXiv:1710.05815 [hep-th]`.

[97] M. Reuter and F. Saueressig, "Quantum Einstein Gravity," *New J. Phys.* **14** (2012) 055022, `arXiv:1202.2274 [hep-th]`.

[98] C. Wetterich, "Graviton fluctuations erase the cosmological constant," *Phys. Lett.* **B773** (2017) 6–19, `arXiv:1704.08040 [gr-qc]`.

[99] B. Knorr, "Infinite order quantum-gravitational correlations," *Class. Quant. Grav.* **35** (2018) 115005, `arXiv:1710.07055 [hep-th]`.

[100] T. R. Morris, "Renormalization group properties in the conformal sector: towards perturbatively renormalizable quantum gravity," *JHEP* **08** (2018) 024, `arXiv:1802.04281 [hep-th]`.

[101] C. Wetterich, "Infrared limit of quantum gravity," *Phys. Rev.* **D98** (2018) 026028, `arXiv:1802.05947 [gr-qc]`.

[102] M. Reuter and H. Weyer, "Quantum gravity at astrophysical distances?," *JCAP* **0412** (2004) 001, `arXiv:hep-th/0410119 [hep-th]`.

[103] H. W. Hamber, *Quantum gravitation: The Feynman path integral approach*. Springer, Berlin, 2009. `http://www.springer.com/978-3-540-85292-6`.

[104] J. Ambjorn, A. Goerlich, J. Jurkiewicz, and R. Loll, "Nonperturbative Quantum Gravity," *Phys. Rept.* **519** (2012) 127–210, `arXiv:1203.3591 [hep-th]`.

[105] R. Loll, "Quantum Gravity from Causal Dynamical Triangulations: A Review," *Class. Quant. Grav.* **37** (2020) 013002, `arXiv:1905.08669 [hep-th]`.

[106] B. Knorr and F. Saueressig, "Towards reconstructing the quantum effective action of gravity," *Phys. Rev. Lett.* **121** (2018) 161304, `arXiv:1804.03846 [hep-th]`.

[107] E. Belgacem, G. Cusin, S. Foffa, M. Maggiore, and M. Mancarella, "Stability issues of nonlocal gravity during primordial inflation," *Int. J. Mod. Phys.* **A33** (2018) 1850007, `arXiv:1610.05664 [hep-th]`. [Erratum: Int. J. Mod. Phys.A33,1892005(2018)].

[108] M. Chevallier and D. Polarski, "Accelerating universes with scaling dark matter," *Int.J.Mod.Phys.* **D10** (2001) 213–224, `arXiv:gr-qc/0009008 [gr-qc]`.

[109] E. V. Linder, "Exploring the expansion history of the universe," *Phys.Rev.Lett.* **90** (2003) 091301, `arXiv:astro-ph/0208512 [astro-ph]`.

[110] S. Nesseris and S. Tsujikawa, "Cosmological perturbations and observational constraints on nonlocal massive gravity," *Phys.Rev.* **D90** (2014) 024070, `arXiv:1402.4613 [astro-ph.CO]`.

[111] Y. Dirian, S. Foffa, N. Khosravi, M. Kunz, and M. Maggiore, "Cosmological perturbations and structure formation in nonlocal infrared modifications of general relativity," *JCAP* **1406** (2014) 033, `arXiv:1403.6068 [astro-ph.CO]`.

[112] L. Amendola, M. Kunz, and D. Sapone, "Measuring the dark side (with weak lensing)," *JCAP* **0804** (2008) 013, `arXiv:0704.2421 [astro-ph]`.

[113] S. F. Daniel *et al.*, "Testing General Relativity with Current Cosmological Data," *Phys.Rev.* **D81** (2010) 123508, `arXiv:1002.1962 [astro-ph.CO]`.

[114] C. Deffayet, "Cosmology on a brane in Minkowski bulk," *Phys.Lett.* **B502** (2001) 199–208, `arXiv:hep-th/0010186 [hep-th]`.

[115] C. Deffayet, G. Dvali, and G. Gabadadze, "Accelerated universe from gravity leaking to extra dimensions," *Phys.Rev.* **D65** (2002) 044023, `arXiv:astro-ph/0105068 [astro-ph]`.

[116] M. A. Luty, M. Porrati, and R. Rattazzi, "Strong interactions and stability in the DGP model," *JHEP* **0309** (2003) 029, `arXiv:hep-th/0303116 [hep-th]`.

[117] A. Nicolis and R. Rattazzi, "Classical and quantum consistency of the DGP model," *JHEP* **0406** (2004) 059, `arXiv:hep-th/0404159 [hep-th]`.

[118] D. Gorbunov, K. Koyama, and S. Sibiryakov, "More on ghosts in DGP model," *Phys.Rev.* **D73** (2006) 044016, `arXiv:hep-th/0512097 [hep-th]`.

[119] C. Charmousis, R. Gregory, N. Kaloper, and A. Padilla, "DGP Specteroscopy," *JHEP* **0610** (2006) 066, `arXiv:hep-th/0604086 [hep-th]`.

[120] K. Izumi, K. Koyama, and T. Tanaka, "Unexorcized ghost in DGP brane world," *JHEP* **0704** (2007) 053, `arXiv:hep-th/0610282 [hep-th]`.

[121] C. de Rham and G. Gabadadze, "Generalization of the Fierz-Pauli Action," *Phys.Rev.* **D82** (2010) 044020, `arXiv:1007.0443 [hep-th]`.

[122] S. Hassan and R. A. Rosen, "Resolving the Ghost Problem in non-Linear Massive Gravity," *Phys.Rev.Lett.* **108** (2012) 041101, `arXiv:1106.3344 [hep-th]`.

[123] G. D'Amico, C. de Rham, S. Dubovsky, G. Gabadadze, D. Pirtskhalava, and A. J. Tolley, "Massive Cosmologies," *Phys. Rev.* **D84** (2011) 124046, `arXiv:1108.5231 [hep-th]`.

[124] F. Koennig, Y. Akrami, L. Amendola, M. Motta, and A. R. Solomon, "Stable and unstable cosmological models in bimetric massive gravity," *Phys. Rev.* **D90** (2014) 124014, `arXiv:1407.4331 [astro-ph.CO]`.

[125] M. Lagos and P. G. Ferreira, "Cosmological perturbations in massive bigravity," *JCAP* **1412** (2014) 026, `arXiv:1410.0207 [gr-qc]`.

[126] G. Cusin, R. Durrer, P. Guarato, and M. Motta, "Gravitational waves in bigravity cosmology," *JCAP* **1505** (2015) 030, `arXiv:1412.5979 [astro-ph.CO]`.

[127] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May, and A. R. Solomon, "Bimetric gravity is cosmologically viable," *Phys. Lett.* **B748** (2015) 37–44, `arXiv:1503.07521 [gr-qc]`.

[128] G. Cusin, R. Durrer, P. Guarato, and M. Motta, "Inflationary perturbations in bimetric gravity," *JCAP* **1509** (2015) 043, `arXiv:1505.01091 [astro-ph.CO]`.

[129] A. Schmidt-May and M. von Strauss, "Recent developments in bimetric theory," *J. Phys.* **A49** (2016) 183001, `arXiv:1512.00021 [hep-th]`.

[130] G. Cusin, R. Durrer, P. Guarato, and M. Motta, "A general mass term for bigravity," *JCAP* **1604** (2016) 051, `arXiv:1512.02131 [astro-ph.CO]`.

[131] D. Blas, J. Lesgourgues, and T. Tram, "The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes," *JCAP* **1107** (2011) 034, `arXiv:1104.2933 [astro-ph.CO]`.

[132] `https://github.com/enisbelgacem/class_public`.

[133] Y. Dirian, "Changing the Bayesian prior: Absolute neutrino mass constraints in nonlocal gravity," *Phys. Rev.* **D96** (2017) 083513, `arXiv:1704.04075 [astro-ph.CO]`.

[134] M. C. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz, "Global fit to three neutrino mixing: critical look at present precision," *JHEP* **12** (2012) 123, `arXiv:1209.3023 [hep-ph]`.

[135] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.* **594** (2016) A13, `arXiv:1502.01589 [astro-ph.CO]`.

[136] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," `arXiv:1807.06209 [astro-ph.CO]`.

[137] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 results. V. CMB power spectra and likelihoods," `arXiv:1907.12875 [astro-ph.CO]`.

[138] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 results. VIII. Gravitational lensing," `arXiv:1807.06210 [astro-ph.CO]`.

[139] D. M. Scolnic *et al.*, "The Complete Light-curve Sample of Spectroscopically Confirmed Type Ia Supernovae from Pan-STARRS1 and Cosmological Constraints from The Combined Pantheon Sample," *The Astrophysical Journal* **859** (June, 2018) 101, `arXiv:1710.00845 [astro-ph.CO]`.

[140] **SDSS** Collaboration, M. Betoule *et al.*, "Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples," *Astron. Astrophys.* **568** (2014) A22, `arXiv:1401.4064 [astro-ph.CO]`.

[141] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, *et al.*, "The 6dF Galaxy Survey: Baryon Acoustic Oscillations and the Local Hubble Constant," *Mon.Not.Roy.Astron.Soc.* **416** (2011) 3017–3032, `arXiv:1106.3366 [astro-ph.CO]`.

[142] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden, and M. Manera, "The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$," *Mon. Not. Roy. Astron. Soc.* **449** (2015) 835–847, `arXiv:1409.3242 [astro-ph.CO]`.

[143] **BOSS** Collaboration, S. Alam *et al.*, "The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample," *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617–2652, `arXiv:1607.03155 [astro-ph.CO]`.

[144] H. Jeffreys, *The theory of probability*. OUP Oxford, 1998.

[145] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, "Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics Beyond LambdaCDM," *Astrophys. J.* **876** (2019) 85, `arXiv:1903.07603 [astro-ph.CO]`.

[146] K. C. Wong *et al.*, "H0LiCOW XIII. A 2.4% measurement of H_0 from lensed quasars: 5.3σ tension between early and late-Universe probes," `arXiv:1907.04869 [astro-ph.CO]`.

[147] V. Poulin, K. K. Boddy, S. Bird, and M. Kamionkowski, "Implications of an extended dark energy cosmology with massive neutrinos for cosmological tensions," *Phys. Rev.* **D97** (2018) 123504, `arXiv:1803.02474 [astro-ph.CO]`.

[148] K. Aylor, M. Joy, L. Knox, M. Millea, S. Raghunathan, and W. L. K. Wu, "Sounds Discordant: Classical Distance Ladder & ΛCDM -based Determinations of the Cosmological Sound Horizon," *Astrophys. J.* **874** (2019) 4, `arXiv:1811.00537 [astro-ph.CO]`.

[149] M. Moresco, R. Jimenez, L. Verde, A. Cimatti, L. Pozzetti, C. Maraston, and D. Thomas, "Constraining the time evolution of dark energy, curvature and neutrino properties with cosmic chronometers," *JCAP* **1612** no. 12, (2016) 039, `arXiv:1604.00183 [astro-ph.CO]`.

[150] S. Nesseris, D. Sapone, and S. Sypsas, "Evaporating primordial black holes as varying dark energy," *Phys. Dark Univ.* **27** (2020) 100413, `arXiv:1907.05608 [astro-ph.CO]`.

[151] L. Kazantzidis and L. Perivolaropoulos, "Evolution of the $f\sigma_8$ tension with the Planck15/ΛCDM determination and implications for modified gravity theories," *Phys. Rev.* **D97** no. 10, (2018) 103503, `arXiv:1803.01337 [astro-ph.CO]`.

[152] F. Beutler *et al.*, "The 6dF Galaxy Survey: $z \approx 0$ measurement of the growth rate and σ_8," *Mon. Not. Roy. Astron. Soc.* **423** (2012) 3430–3444, `arXiv:1204.4725 [astro-ph.CO]`.

[153] A. Oka, S. Saito, T. Nishimichi, A. Taruya, and K. Yamamoto, "Simultaneous constraints on the growth of structure and cosmic expansion from the multipole power spectra of the SDSS DR7 LRG sample," *Mon. Not. Roy. Astron. Soc.* **439** (2014) 2515–2530, `arXiv:1310.2820 [astro-ph.CO]`.

[154] L. Samushia *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: measuring growth rate and geometry with anisotropic clustering," *Mon. Not. Roy. Astron. Soc.* **439** (2014) 3504–3519, `arXiv:1312.4899 [astro-ph.CO]`.

[155] C. Blake *et al.*, "The WiggleZ Dark Energy Survey: Joint measurements of the expansion and growth history at $z < 1$," *Mon. Not. Roy. Astron. Soc.* **425** (2012) 405–414, `arXiv:1204.3674 [astro-ph.CO]`.

[156] S. de la Torre *et al.*, "The VIMOS Public Extragalactic Redshift Survey (VIPERS). Galaxy clustering and redshift-space distortions at z=0.8 in the first data release," *Astron. Astrophys.* **557** (2013) A54, `arXiv:1303.2622 [astro-ph.CO]`.

[157] **Euclid Theory Working Group** Collaboration, L. Amendola *et al.*, "Cosmology and fundamental physics with the Euclid satellite," *Living Rev.Rel.* **16** (2013) 6, `arXiv:1206.1225 [astro-ph.CO]`.

[158] Y. Song *et al.*, "Complementarity of Weak Lensing and Peculiar Velocity Measurements in Testing General Relativity," *Phys.Rev.* **D84** (2011) 083523, `arXiv:1011.2106 [astro-ph.CO]`.

[159] M. Ishak *et al.*, "Modified Gravity and Dark Energy models Beyond $w(z)$CDM Testable by LSST," `arXiv:1905.09687 [astro-ph.CO]`.

[160] A. Vainshtein, "To the problem of nonvanishing gravitation mass," *Phys.Lett.* **B39** (1972) 393–394.

[161] C. Deffayet, G. Dvali, G. Gabadadze, and A. I. Vainshtein, "Nonperturbative continuity in graviton mass versus perturbative discontinuity," *Phys.Rev.* **D65** (2002) 044026, `arXiv:hep-th/0106001 [hep-th]`.

[162] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz, "Effective field theory for massive gravitons and gravity in theory space," *Annals Phys.* **305** (2003) 96–118, `arXiv:hep-th/0210184 [hep-th]`.

[163] K. Koyama, G. Niz, and G. Tasinato, "Analytic solutions in non-linear massive gravity," *Phys.Rev.Lett.* **107** (2011) 131101, `arXiv:1103.4708 [hep-th]`.

[164] K. Koyama, G. Niz, and G. Tasinato, "Strong interactions and exact solutions in non-linear massive gravity," *Phys.Rev.* **D84** (2011) 064033, `arXiv:1104.2143 [hep-th]`.

[165] E. Belgacem, A. Finke, A. Frassino, and M. Maggiore, "Testing nonlocal gravity with Lunar Laser Ranging," *JCAP* **1902** (2019) 035, `arXiv:1812.11181 [gr-qc]`.

[166] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, and S. Pascoli, "Nonlinear structure formation in Nonlocal Gravity," *JCAP* **1409** (2014) 031, `arXiv:1408.1084 [astro-ph.CO]`.

[167] F. Hofmann and J. Müller, "Relativistic tests with lunar laser ranging," *Class. Quant. Grav.* **35** (2018) 035015.

[168] M. Carrera and D. Giulini, "On the influence of global cosmological expansion on the dynamics and kinematics of local systems," `arXiv:0810.2712 [gr-qc]`. [Rev. Mod. Phys.82,169(2010)].

[169] W. G. Laarakkers and E. Poisson, "Radiative falloff in Einstein-Straus space-time," *Phys. Rev.* **D64** (2001) 084008, `arXiv:gr-qc/0105016 [gr-qc]`.

[170] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "The gravitational-wave luminosity distance in modified gravity theories," *Phys. Rev.* **D97** (2018) 104066, `arXiv:1712.08108 [astro-ph.CO]`.

[171] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Modified gravitational-wave propagation and standard sirens," *Phys. Rev.* **D98** (2018) 023510, `arXiv:1805.08731 [gr-qc]`.

[172] E. Belgacem, Y. Dirian, A. Finke, S. Foffa, and M. Maggiore, "Nonlocal gravity and gravitational-wave observations," *JCAP* **1911** (2019) 022, `arXiv:1907.02047 [astro-ph.CO]`.

[173] B. F. Schutz, "Determining the Hubble Constant from Gravitational Wave Observations," *Nature* **323** (1986) 310–311.

[174] B. P. Abbott *et al.*, "A gravitational-wave standard siren measurement of the Hubble constant," *Nature* **551** no. 7678, (2017) 85–88, `arXiv:1710.05835 [astro-ph.CO]`.

[175] D. E. Holz and S. A. Hughes, "Using gravitational-wave standard sirens," *Astrophys. J.* **629** (2005) 15–22, `arXiv:astro-ph/0504616 [astro-ph]`.

[176] N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain, "Short GRB and binary black hole standard sirens as a probe of dark energy," *Phys. Rev.* **D74** (2006) 063006, `arXiv:astro-ph/0601275 [astro-ph]`.

[177] C. L. MacLeod and C. J. Hogan, "Precision of Hubble constant derived using black hole binary absolute distances and statistical redshift information," *Phys. Rev.* **D77** (2008) 043512, `arXiv:0712.0618 [astro-ph]`.

[178] S. Nissanke, D. E. Holz, S. Hughes, N. Dalal, and J. L. Sievers, "Exploring short gamma-ray bursts as gravitational-wave standard sirens," *Astrophys. J.* **725** (2010) 496–514, `arXiv:0904.1017 [astro-ph.CO]`.

[179] C. Cutler and D. E. Holz, "Ultra-high precision cosmology from gravitational waves," *Phys. Rev.* **D80** (2009) 104009, `arXiv:0906.3752 [astro-ph.CO]`.

[180] B. S. Sathyaprakash, B. F. Schutz, and C. Van Den Broeck, "Cosmography with the Einstein Telescope," *Class. Quant. Grav.* **27** (2010) 215006, `arXiv:0906.4151 [astro-ph.CO]`.

[181] W. Zhao, C. Van Den Broeck, D. Baskaran, and T. G. F. Li, "Determination of Dark Energy by the Einstein Telescope: Comparing with CMB, BAO and SNIa Observations," *Phys. Rev.* **D83** (2011) 023005, `arXiv:1009.0206 [astro-ph.CO]`.

[182] W. Del Pozzo, "Inference of the cosmological parameters from gravitational waves: application to second generation interferometers," *Phys. Rev.* **D86** (2012) 043011, `arXiv:1108.1317 [astro-ph.CO]`.

[183] A. Nishizawa, K. Yagi, A. Taruya, and T. Tanaka, "Cosmology with space-based gravitational-wave detectors — dark energy and primordial gravitational waves —," *Phys. Rev.* **D85** (2012) 044047, `arXiv:1110.2865 [astro-ph.CO]`.

[184] S. R. Taylor, J. R. Gair, and I. Mandel, "Hubble without the Hubble: Cosmology using advanced gravitational-wave detectors alone," *Phys. Rev.* **D85** (2012) 023535, `arXiv:1108.5161 [gr-qc]`.

[185] S. R. Taylor and J. R. Gair, "Cosmology with the lights off: standard sirens in the Einstein Telescope era," *Phys. Rev.* **D86** (2012) 023502, `arXiv:1204.6739 [astro-ph.CO]`.

[186] S. Camera and A. Nishizawa, "Beyond Concordance Cosmology with Magnification of Gravitational-Wave Standard Sirens," *Phys. Rev. Lett.* **110** (2013) 151103, `arXiv:1303.5446 [astro-ph.CO]`.

[187] R.-G. Cai and T. Yang, "Estimating cosmological parameters by the simulated data of gravitational waves from the Einstein Telescope," *Phys. Rev.* **D95** (2017) 044024, `arXiv:1608.08008 [astro-ph.CO]`.

[188] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, "Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens," *JCAP* **1604** (2016) 002, `arXiv:1601.07112 [astro-ph.CO]`.

[189] C. Caprini and N. Tamanini, "Constraining early and interacting dark energy with gravitational wave standard sirens: the potential of the eLISA mission," *JCAP* **1610** (2016) 006, `arXiv:1607.08755 [astro-ph.CO]`.

[190] B. P. Abbott *et al.*, "Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J.* **848** (2017) L13, `arXiv:1710.05834 [astro-ph.HE]`.

[191] C. de Rham and S. Melville, "Gravitational Rainbows: LIGO and Dark Energy at its Cutoff," *Phys. Rev. Lett.* **121** (2018) 221101, `arXiv:1806.09417 [hep-th]`.

[192] P. Creminelli and F. Vernizzi, "Dark Energy after GW170817 and GRB170817A," *Phys. Rev. Lett.* **119** (2017) 251302, `arXiv:1710.05877 [astro-ph.CO]`.

[193] J. Sakstein and B. Jain, "Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories," *Phys. Rev. Lett.* **119** (2017) 251303, `arXiv:1710.05893 [astro-ph.CO]`.

[194] J. M. Ezquiaga and M. Zumalacárregui, "Dark Energy After GW170817: Dead Ends and the Road Ahead," *Phys. Rev. Lett.* **119** (2017) 251304, `arXiv:1710.05901 [astro-ph.CO]`.

[195] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, "Strong constraints on cosmological gravity from GW170817 and GRB 170817A," *Phys. Rev. Lett.* **119** (2017) 251301, `arXiv:1710.06394 [astro-ph.CO]`.

[196] I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, "Anisotropic Stress as a Signature of Nonstandard Propagation of Gravitational Waves," *Phys. Rev. Lett.* **113** (2014) 191101, `arXiv:1406.7139 [astro-ph.CO]`.

[197] L. Lombriser and A. Taylor, "Breaking a Dark Degeneracy with Gravitational Waves," *JCAP* **1603** (2016) 031, `arXiv:1509.08458 [astro-ph.CO]`.

[198] S. Arai and A. Nishizawa, "Generalized framework for testing gravity with gravitational-wave propagation. II. Constraints on Horndeski theory," *Phys. Rev.* **D97** (2018) 104038, `arXiv:1711.03776 [gr-qc]`.

[199] L. Amendola, I. Sawicki, M. Kunz, and I. D. Saltas, "Direct detection of gravitational waves can measure the time variation of the Planck mass," *JCAP* **1808** (2018) 030, `arXiv:1712.08623 [astro-ph.CO]`.

[200] **LISA Cosmology Working Group** Collaboration, E. Belgacem *et al.*, "Testing modified gravity at cosmological distances with LISA standard sirens," *JCAP* **1907** (2019) 024, `arXiv:1906.01593 [astro-ph.CO]`.

[201] C. Deffayet and K. Menou, "Probing Gravity with Spacetime Sirens," *Astrophys. J.* **668** (2007) L143–L146, `arXiv:0709.0003 [astro-ph]`.

[202] K. Pardo, M. Fishbach, D. E. Holz, and D. N. Spergel, "Limits on the number of spacetime dimensions from GW170817," *JCAP* **1807** (2018) 048, `arXiv:1801.08160 [gr-qc]`.

[203] J. Gleyzes, D. Langlois, and F. Vernizzi, "A unifying description of dark energy," *Int. J. Mod. Phys.* **D23** (2014) 1443010, `arXiv:1411.3712 [hep-th]`.

[204] A. Nishizawa, "Generalized framework for testing gravity with gravitational-wave propagation. I. Formulation," *Phys. Rev.* **D97** (2018) 104037, `arXiv:1710.04825 [gr-qc]`.

[205] A. Garoffolo, G. Tasinato, C. Carbone, D. Bertacca, and S. Matarrese, "Gravitational waves and geometrical optics in scalar-tensor theories," `arXiv:1912.08093 [gr-qc]`.

[206] E. V. Linder, "No Slip Gravity," *JCAP* **1803** (2018) 005, `arXiv:1801.01503 [astro-ph.CO]`.

[207] C. Dalang and L. Lombriser, "Limitations on Standard Sirens tests of gravity from screening," *JCAP* **1910** (2019) 013, `arXiv:1906.12333 [astro-ph.CO]`.

[208] R. D'Agostino and R. C. Nunes, "Probing observational bounds on scalar-tensor theories from standard sirens," *Phys. Rev.* **D100** (2019) 044041, `arXiv:1907.05516 [gr-qc]`.

[209] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XIV. Dark energy and modified gravity," *Astron. Astrophys.* **594** (2016) A14, `arXiv:1502.01590 [astro-ph.CO]`.

[210] E. Belgacem, Y. Dirian, S. Foffa, E. J. Howell, M. Maggiore, and T. Regimbau, "Cosmology and dark energy from joint gravitational wave-GRB observations," *JCAP* **1908** (2019) 015, `arXiv:1907.01487 [astro-ph.CO]`.

[211] B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** (2016) 061102, `arXiv:1602.03837 [gr-qc]`.

[212] B. P. Abbott *et al.*, "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence," *Phys. Rev. Lett.* **116** no. 24, (2016) 241103, `arXiv:1606.04855 [gr-qc]`.

[213] B. P. Abbott *et al.*, "GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2" *Phys. Rev. Lett.* **118** no. 22, (2017) 221101, `arXiv:1706.01812 [gr-qc]`.

[214] B. P. Abbott *et al.*, "GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence," *Astrophys. J.* **851** no. 2, (2017) L35, `arXiv:1711.05578 [astro-ph.HE]`.

[215] B. P. Abbott *et al.*, "GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence," *Phys. Rev. Lett.* **119** no. 14, (2017) 141101, `arXiv:1709.09660 [gr-qc]`.

[216] **LIGO Scientific, Virgo** Collaboration, B. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X* **9** (2019) 031040, `arXiv:1811.12907 [astro-ph.HE]`.

[217] B. Abbott *et al.*, "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.* **119** (2017) 161101, `arXiv:1710.05832 [gr-qc]`.

[218] A. Goldstein *et al.*, "An Ordinary Short Gamma-Ray Burst with Extraordinary Implications: Fermi-GBM Detection of GRB 170817A," *Astrophys. J.* **848** (2017) L14, `arXiv:1710.05446 [astro-ph.HE]`.

[219] V. Savchenko *et al.*, "INTEGRAL Detection of the First Prompt Gamma-Ray Signal Coincident with the Gravitational-wave Event GW170817," *Astrophys. J.* **848** (2017) L15, `arXiv:1710.05449 [astro-ph.HE]`.

[220] B. P. Abbott *et al.*, "Multi-messenger Observations of a Binary Neutron Star Merger," *Astrophys. J.* **848** no. 2, (2017) L12, `arXiv:1710.05833 [astro-ph.HE]`.

[221] P. Amaro-Seoane *et al.*, "Laser Interferometer Space Antenna," `arXiv:1702.00786 [astro-ph.IM]`.

[222] M. Punturo *et al.*, "The Einstein Telescope: A third-generation gravitational wave observatory," *Class. Quant. Grav.* **27** (2010) 194002.

[223] B. Sathyaprakash *et al.*, "Scientific Objectives of Einstein Telescope," *Class. Quant. Grav.* **29** (2012) 124013, `arXiv:1206.0331 [gr-qc]`. [Erratum: Class. Quant. Grav.30,079501(2013)].

[224] S. Dwyer, D. Sigg, S. W. Ballmer, L. Barsotti, N. Mavalvala, and M. Evans, "Gravitational wave detector with cosmological reach," *Phys. Rev.* **D91** (2015) 082001, `arXiv:1410.0612 [astro-ph.IM]`.

[225] M. Maggiore *et al.*, "Science Case for the Einstein Telescope," *JCAP, to appear* (2020) , `arXiv:1912.02622 [astro-ph.CO]`.

[226] T. Regimbau, T. Dent, W. Del Pozzo, S. Giampanis, T. G. F. Li, C. Robinson, C. Van Den Broeck, D. Meacher, C. Rodriguez, B. S. Sathyaprakash, and K. Wójcik, "Mock data challenge for the Einstein Gravitational-Wave Telescope," *Phys. Rev. D* **86** (2012) 122001, `arXiv:1201.3563 [gr-qc]`.

[227] T. Regimbau, D. Meacher, and M. Coughlin, "Second Einstein Telescope mock science challenge: Detection of the gravitational-wave stochastic background from compact binary coalescences," *Phys. Rev. D* **89** (2014) 084046, `arXiv:1404.1134`.

[228] T. Regimbau, K. Siellez, D. Meacher, B. Gendre, and M. Boër, "Revisiting coincidence rate between Gravitational Wave detection and short Gamma-Ray Burst for the Advanced and third generation," *Astrophys. J.* **799** no. 1, (2015) 69, `arXiv:1410.2739 [astro-ph.HE]`.

[229] D. Meacher, M. Coughlin, S. Morris, T. Regimbau, N. Christensen, S. Kandhasamy, V. Mandic, J. D. Romano, and E. Thrane, "Mock data and

science challenge for detecting an astrophysical stochastic gravitational-wave background with Advanced LIGO and Advanced Virgo," *Phys. Rev. D* **92** (2015) 063002, `arXiv:1506.06744 [astro-ph.HE]`.

[230] D. Meacher, K. Cannon, C. Hanna, T. Regimbau, and B. S. Sathyaprakash, "Second Einstein Telescope mock data and science challenge: Low frequency binary neutron star data analysis," *Phys. Rev. D* **93** no. 2, (Jan, 2016) 024018, `arXiv:1511.01592 [gr-qc]`.

[231] T. Regimbau, M. Evans, N. Christensen, E. Katsavounidis, B. Sathyaprakash, and S. Vitale, "Digging Deeper: Observing Primordial Gravitational Waves below the Binary-Black-Hole-Produced Stochastic Background," *Phys. Rev. Lett.* **118** (2017) 151105, `arXiv:1611.08943`.

[232] L. Amati *et al.*, "The THESEUS space mission concept: science case, design and expected performances," *Adv. Space Res.* **62** (2018) 191–244, `arXiv:1710.04638 [astro-ph.IM]`.

[233] G. Stratta *et al.*, "THESEUS: a key space mission concept for Multi-Messenger Astrophysics," *Adv. Space Res.* **62** (2018) 662–682, `arXiv:1712.08153 [astro-ph.HE]`.

[234] G. Stratta, L. Amati, R. Ciolfi, and S. Vinciguerra, "THESEUS in the era of Multi-Messenger Astronomy," *Mem. Soc. Ast. It.* **89** (2018) 205, `arXiv:1802.01677 [astro-ph.IM]`.

[235] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., "GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences," *Phys. Rev. Lett.* **120** (2018) 091101, `arXiv:1710.05837 [gr-qc]`.

[236] E. Vangioni, K. A. Olive, T. Prestegard, J. Silk, P. Petitjean, and V. Mandic, "The impact of star formation and gamma-ray burst rates at high redshift on cosmic chemical evolution and reionization," *Mon. Not. Roy. Astron. Soc.* **447** (2015) 2575–2587, `arXiv:1409.2462`.

[237] P. Madau and M. Dickinson, "Cosmic Star Formation History," *Ann. Rev. Astron. Astrophys.* **52** (2014) 415–486, `arXiv:1403.0007 [astro-ph.CO]`.

[238] S. Vitale and W. M. Farr, "Measuring the star formation rate with gravitational waves from binary black holes," `arXiv:1808.00901 [astro-ph.HE]`.

[239] B. S. Sathyaprakash *et al.*, "Multimessenger Universe with Gravitational Waves from Binaries," `arXiv:1903.09277 [astro-ph.HE]`.

[240] D. Bertacca, A. Raccanelli, N. Bartolo, and S. Matarrese, "Cosmological perturbation effects on gravitational-wave luminosity distance estimates," *Phys. Dark Univ.* **20** (2018) 32–40, `arXiv:1702.01750 [gr-qc]`.

[241] H.-Y. Chen, M. Fishbach, and D. E. Holz, "A two per cent Hubble constant measurement from standard sirens within five years," *Nature* **562** (2018) 545, `arXiv:1712.06531 [astro-ph.CO]`.

[242] E. Berger, "Short-Duration Gamma-Ray Bursts," *Ann. Rev. Astron. Astrophys.* **52** (2014) 43–105, `arXiv:1311.2603 [astro-ph.HE]`.

[243] P. A. Evans *et al.*, "Swift and NuSTAR observations of GW170817: Detection of a blue kilonova," *Science* **358** no. 6370, (Dec., 2017) 1565–1570. `http://science.sciencemag.org/content/358/6370/1565`.

[244] E. J. Howell, K. Ackley, A. Rowlinson, and D. Coward, "Joint gravitational wave - gamma-ray burst detection rates in the aftermath of GW170817," *Mon. Not. Roy. Astron. Soc.* **485** no. 1, (May, 2019) 1435–1447, `arXiv:1811.09168 [astro-ph.HE]`.

[245] D. Wanderman and T. Piran, "The rate, luminosity function and time delay of non-Collapsar short GRBs," *Mon. Not. Roy. Astron. Soc.* **448** (Apr., 2015) 3026–3037, `arXiv:1405.5878 [astro-ph.HE]`.

[246] E. Burns, V. Connaughton, B.-B. Zhang, A. Lien, M. S. Briggs, A. Goldstein, V. Pelassa, and E. Troja, "Do the Fermi Gamma-Ray Burst Monitor and Swift Burst Alert Telescope see the Same Short Gamma-Ray Bursts?," *Astrophys. J.* **818** (Feb., 2016) 110, `arXiv:1512.00923 [astro-ph.HE]`.

[247] **Planck** Collaboration, R. Adam *et al.*, "Planck 2015 results. I. Overview of products and scientific results," `arXiv:1502.01582 [astro-ph.CO]`.

[248] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters," *Astron. Astrophys.* **594** (2016) A11, `arXiv:1507.02704 [astro-ph.CO]`.

[249] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XV. Gravitational lensing," `arXiv:1502.01591 [astro-ph.CO]`.

[250] **BOSS** Collaboration, L. Anderson *et al.*, "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the Data Release 10 and 11 galaxy samples," `arXiv:1312.4877 [astro-ph.CO]`.

[251] S. M. Feeney, H. V. Peiris, A. R. Williamson, S. M. Nissanke, D. J. Mortlock, J. Alsing, and D. Scolnic, "Prospects for resolving the Hubble constant tension with standard sirens," `arXiv:1802.03404 [astro-ph.CO]`.

[252] M. Seikel, C. Clarkson, and M. Smith, "Reconstruction of dark energy and expansion dynamics using Gaussian processes," *JCAP* **1206** (2012) 036, `arXiv:1204.2832 [astro-ph.CO]`.

[253] M. Seikel, S. Yahya, R. Maartens, and C. Clarkson, "Using H(z) data as a probe of the concordance model," *Phys. Rev.* **D86** (2012) 083001, `arXiv:1205.3431 [astro-ph.CO]`.

[254] S. Yahya, M. Seikel, C. Clarkson, R. Maartens, and M. Smith, "Null tests of the cosmological constant using supernovae," *Phys. Rev.* **D89** no. 2, (2014) 023503, `arXiv:1308.4099 [astro-ph.CO]`.

[255] V. C. Busti, C. Clarkson, and M. Seikel, "Evidence for a Lower Value for H_0 from Cosmic Chronometers Data?," *Mon. Not. Roy. Astron. Soc.* **441** (2014) 11, `arXiv:1402.5429 [astro-ph.CO]`.

[256] V. C. Busti and C. Clarkson, "Dodging the dark matter degeneracy while determining the dynamics of dark energy," *JCAP* **1605** no. 05, (2016) 008, `arXiv:1505.01821 [astro-ph.CO]`.

[257] T. Yang, Z.-K. Guo, and R.-G. Cai, "Reconstructing the interaction between dark energy and dark matter using Gaussian Processes," *Phys. Rev.* **D91** (2015) 123533, `arXiv:1505.04443 [astro-ph.CO]`.

[258] R.-G. Cai, Z.-K. Guo, and T. Yang, "Null test of the cosmic curvature using $H(z)$ and supernovae data," *Phys. Rev.* **D93** no. 4, (2016) 043517, `arXiv:1509.06283 [astro-ph.CO]`.

[259] R.-G. Cai, Z.-K. Guo, and T. Yang, "Dodging the cosmic curvature to probe the constancy of the speed of light," *JCAP* **1608** no. 08, (2016) 016, `arXiv:1601.05497 [astro-ph.CO]`.

[260] R.-G. Cai, N. Tamanini, and T. Yang, "Reconstructing the dark sector interaction with LISA," *JCAP* **1705** no. 05, (2017) 031, `arXiv:1703.07323 [astro-ph.CO]`.

[261] E. Belgacem, S. Foffa, M. Maggiore, and T. Yang, "Gaussian processes reconstruction of modified gravitational wave propagation," `arXiv:1911.11497 [astro-ph.CO]`.

[262] J. P. Bernstein *et al.*, "Supernova Simulations and Strategies For the Dark Energy Survey," *Astrophys. J.* **753** (2012) 152, `arXiv:1111.1969 [astro-ph.CO]`.

[263] E. Vangioni, K. Olive, T. Prestegard, J. Silk, P. Petitjean, and V. Mandic, "The Impact of Star Formation and Gamma-Ray Burst Rates at High Redshift on Cosmic Chemical Evolution and Reionization," *Mon. Not. Roy. Astron. Soc.* **447** (2015) 2575, `arXiv:1409.2462 [astro-ph.GA]`.

[264] Y. Dirian, S. Foffa, M. Kunz, M. Maggiore, and V. Pettorino, "Non-local gravity and comparison with observational datasets. II. Updated results and Bayesian model comparison with ΛCDM," *JCAP* **1605** (2016) 068, `arXiv:1602.03558 [astro-ph.CO]`.

[265] A. Klein *et al.*, "Science with the space-based interferometer eLISA: Supermassive black hole binaries," *Phys. Rev. D* **93** no. 2, (2016) 024003, `arXiv:1511.05581 [gr-qc]`.

[266] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, "Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals," *Phys. Rev. D* **95** no. 10, (2017) 103012, `arXiv:1703.09722 [gr-qc]`.

[267] W. Del Pozzo, A. Sesana, and A. Klein, "Stellar binary black holes in the LISA band: a new class of standard sirens," *Mon. Not. Roy. Astron. Soc.* **475** (2018) 3485–3492, `arXiv:1703.01300 [astro-ph.CO]`.

[268] L. Amendola *et al.*, "Cosmology and fundamental physics with the Euclid satellite," *Living Rev. Rel.* **21** no. 1, (2018) 2, `arXiv:1606.00180 [astro-ph.CO]`.

[269] `https://github.com/enisbelgacem/montepython_public`.

[270] G. Cusin, S. Foffa, M. Maggiore, and M. Mancarella, "Conformal symmetry and nonlinear extensions of nonlocal gravity," *Phys. Rev.* **D93** (2016) 083008, `arXiv:1602.01078 [hep-th]`.

[271] E. Di Valentino, A. Melchiorri, and J. Silk, "Reconciling Planck with the local value of H_0 in extended parameter space," *Phys. Lett.* **B761** (2016) 242–246, `arXiv:1606.00634 [astro-ph.CO]`.