



Higher-derivative quantum gravity with purely virtual particles: renormalizability and unitarity

Marco Piva^a

Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland

Received: 2 June 2023 / Accepted: 8 September 2023
© The Author(s) 2023

Abstract We review the formulation of quantum field theories with purely virtual particles, a new type of degrees of freedom that can mediate interactions without ever appear as external on-shell states. This property allows to solve the problem of ghosts in higher-derivative quantum gravity, leading to a renormalizable and unitary theory. The main steps for the BRST quantization of gravity are recalled and renormalizability is discussed. Then, we introduce purely virtual particles in a general quantum field theory and show the derivation of the so-called spectral identities, which are a key ingredient to prove unitarity. Finally, phenomenological consequences and predictions in inflationary cosmology are presented.

1 Introduction

The formulation of a theory of quantum gravity is one of the most important task in theoretical physics. It is expected to shed light on several open problems and, most importantly, it will tell us something new about our understanding of reality, in the same way quantum mechanics made many of our certainties crumble at the beginning of the twentieth century. Many phenomenological questions, such as the understanding of the initial phases of the universe or how to deal with black holes in quantum mechanics, as well as more fundamental ones (do the notions of time, cause, past and future survive at arbitrary scales?) might find answer in the theory of quantum gravity. However, formulating such a theory is as important as it is hard. We will probably need many decades or centuries before to actually answer to all the relevant questions, because it is a very intricate subject and it connects many aspects of physics altogether. Moreover, we do not have enough experimental data to infer much about the quantum nature of gravity. Therefore, in the situation where the complexity of the topic comes together with the lack of data, it is important to stay with our foot on the ground and proceed carefully. The starting point is to look at the past, analyze what has worked so far and try to deviate as least as possible from that path. The risk is to completely loose contact with nature and go astray forever.

Over the last decades there have been several proposals to tackle the problem of quantum gravity. Some of them rely on abandoning the framework that has been very successful so far, i.e. quantum field theory, in favor of something else. In some cases the approach is based on the assumption that gravity is the geometry of spacetime at arbitrary small scales. We have no evidence for this; therefore, those theories (loop quantum gravity, causal dynamical triangulations, spinfoam to name a few) should be viewed as theories of quantum geometry, rather than theories of quantum gravity. Only the contact with nature will tell us whether that assumption is correct or not. We think that a more cautious approach is to study the gravitational field in quantum mechanical terms, as we do for all other fields, in the framework of quantum field theory.

In this review we present an approach to do so in a way that is consistent with the principles that have led to the standard model of particle physics: locality, renormalizability and unitarity. Although conservative, this proposal cannot avoid to introduce something new, albeit minimal. The new idea is the possibility to describe quanta that are purely virtual, i.e. they can be responsible for interactions between other particles, like a virtual photon is responsible for the interaction between two electrons, but cannot appear as on-shell states. These new type of particles are called *purely virtual particles* or *fakeons*, and they are introduced by means of a specific procedure that we call *fakeon prescription* [1, 2]. Their presence is enough to have a local, renormalizable and unitary quantum field theory of gravity.

The approach presented here has its roots in results that were known since many decades. The starting point is that Einstein gravity treated as a quantum field theory is nonrenormalizable [3, 4]. The problem with nonrenormalizability is that, in order to reabsorb the divergences of the theory, we need to introduce infinitely many operators of increasing dimension, each of them with a new independent parameter. Therefore, in principle, if we want to fix all parameters with experiments we would need infinitely many measurements. However, all the known nonrenormalizable theories are effective theories, which means that their

^a e-mail: mpiva@fuw.edu.pl (corresponding author)

nonrenormalizability is due to the fact that some additional (heavy) fields are integrated out. Once those fields are properly included the infinitely many parameters turn out to be functions of a finite subset of them, and the theory is renormalizable. This is not what happens in Einstein gravity, at least not in the same way as in other theories, like the Fermi model [5].

One way of proceed is to just accept that there might exist fundamental theories that are nonrenormalizable and start to work out their properties in order to make them as predictive as they can (see e.g. [6–8]). Along this line is the asymptotic safety program [9, 10], which is a nonperturbative field-theoretic approach that accounts for a generalization of asymptotic freedom, where the coupling constant is assumed to flows to a nonvanishing value in the ultraviolet limit. Then, for the theory to be predictive, such a interacting fixed point needs to have a finite-dimensional critical surface. An asymptotically safe theory is sometimes referred as nonperturbatively renormalizable.

Another possibility is to go beyond Einstein gravity by imposing renormalizability. The natural choice is to add to the classical action the counterterms generated at one loop in Einstein gravity, assume that their couplings are large and redo the analysis. This was done in 1976 by Stelle [11], showing that adding to the action of gravity all the independent terms quadratic in the Riemann tensor, as well as the cosmological term, makes the theory renormalizable. Moreover, some additional degrees of freedom are introduced in this way. Besides the usual graviton modes, one massive scalar and one massive spin-2 particles also propagate. However, the kinetic term of the latter has the wrong sign, causing instabilities that at quantum level can be traded for violations of unitarity. This type of particles are called *ghosts*. At that point the theory, which we call *Stelle gravity*, was (righteously) labeled as physically unacceptable and the research in this direction essentially stopped. Instead of insisting on the principle of renormalizability and try to solve the problem of ghosts, many physicists moved to something else, often different from quantum field theory, to explain quantum gravity. Over the decades only a limited number of people continued to study Stelle gravity and work out its properties, but they often postpone the problem of ghosts for later studies or invoke some nonperturbative effects or theories beyond quantum field theory that will eventually explain the matter [12–16]. These explanations never came. Furthermore, in 1980, Starobinsky shown that the scalar degree of freedom introduced by the R^2 term can explain the accelerated expansion of the early universe and solve a number of phenomenological issues [17]. This is one of the most efficient and simple model of inflation, called *Starobinsky model*, which is included in Stelle gravity. Moreover, since also the cosmological term is generated by renormalization, dark energy might be explained within Stelle gravity.

Despite these results in favor of Stelle gravity, addressing the problem of ghosts has never been the priority of any of the major research lines. On top of this, several potential approaches were already present in the literature (and actually inspired ours), although not all of them were thought for gravity. For example, the Lee-Wick models [18, 19] have been used in the context of quantum gravity only in the recent years [20, 21], or the so-called shadow states [22–24], introduced in 1972, which also share some properties with fakeons, were soon ignored. Finally, another approach to the ghost problem is to formulate quantum gravity as a particular type of nonlocal quantum field theory [25]. There is a decent amount of literature in this topic, but only in the last 10–15 years and only by small groups (see [26] and references therein). Finally, we mention a more recent approach, introduced in [27], which uses the fact that the ghost in Stelle theory is unstable to prove unitarity in situations where it has already decayed, using a generalization of Veltman's argument of [28]. The result is an effective theory, unitary at energy scales where the width of the ghost can be considered large. However, this is far from truly removing the ghost from the spectrum, since it is always possible to find energy ranges where it is long lived and cannot be ignored.

In a nutshell, many interesting ideas to tackle the problem of ghosts were already on the table since a few decades, but each of them has been either abandoned, ignored or poorly studied.

To a relatively young physicist, like the author of this review, the feeling that the historical situation just described gives is that someone did not do their job for fifty years.

Why abandon a successful framework such as quantum field theory? Why (almost) completely ignore the existence of a renormalizable, although not unitary, theory of gravity and move to other approaches? Why pretend not to see that the best inflationary model points in the direction of that renormalizable theory? Probably the answers to these questions lie more on the social aspects of physics rather than the scientific ones. We do not try to give our personal answers in this review, so we do not hurt the feelings of anyone.

Finally, we want to clarify that it might well be that something beyond quantum field theory is necessary to achieve a more fundamental understanding of quantum gravity. However, our opinion is that no stone should be left unturned in quantum field theory before abandoning it and the topic addressed in this review tells us that this was not the case. Furthermore, as stressed above, in the lack of guidance given by nature, we cannot afford to jump into the unknown without any grasp on experimental data. We need to stick to what was successful so far and see where it leads.

The paper is organized as follows. In Sect. 2 we review Stelle gravity, discuss its renormalizability and introduce the ghost problem. In Sect. 3 we recall the basics of unitarity in general quantum field theories, while in Sect. 4 we introduce the spectral identities that are necessary to prove unitarity in theories with purely virtual particles. In Sect. 5 we explain how to apply the idea of purely virtual particles to Stelle gravity and show its predictions in the context of inflationary cosmology. Finally, Sect. 6 contains our conclusions.

Notation and conventions: We use the signature (+, −, −, −) for the metric tensor. The Riemann and Ricci tensors are defined as $R^\mu_{\nu\rho\sigma} = \partial_\rho\Gamma^\mu_{\nu\sigma} - \partial_\sigma\Gamma^\mu_{\nu\rho} + \Gamma^\mu_{\alpha\rho}\Gamma^\alpha_{\nu\sigma} - \Gamma^\mu_{\alpha\sigma}\Gamma^\alpha_{\nu\rho}$ and $R_{\mu\nu} = R^\rho_{\mu\rho\nu}$, respectively. We write the four-dimensional integrals over

spacetime points of a function F of a field $\phi(x)$ as $\int \sqrt{-g}F(\phi) \equiv \int d^4x \sqrt{-g(x)}F(\phi(x))$. We always assume that the integral of the Gauss-Bonnet term vanish, i.e. $\int \sqrt{-g}(R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2) = 0$.

2 Stelle gravity

In this section we introduce Stelle gravity by recalling the general method for the quantization of gravity in quantum field theory. Moreover, we study the propagator of the theory in order to show the presence of the spin-2 ghost. Finally, we give some details about renormalizability.

The Stelle action is

$$S_{HD} = -\frac{1}{2\kappa^2} \int \sqrt{-g} \left[2\Lambda_C + \zeta R + \frac{\alpha}{2} C^2 - \frac{\xi}{6} R^2 \right], \tag{2.1}$$

where $C^2 = C_{\mu\nu\rho\sigma}C^{\mu\nu\rho\sigma}$ is the squared of the Weyl tensor and $\Lambda_C, \zeta, \alpha, \xi$ and κ are real parameters with dimension

$$[\Lambda_C] = 4, \quad [\zeta] = 2, \quad [\alpha] = 0, \quad [\xi] = 0, \quad [\kappa] = 0. \tag{2.2}$$

Note that the parameter κ is redundant, since ζ already accounts for the Planck mass, but it is useful to keep track of loop orders in the perturbative expansion. We can see that no parameters of negative dimension appear, even after the correct normalization of the kinetic terms. In fact, if we expand the metric around Minkowski spacetime as

$$g_{\mu\nu} = \eta_{\mu\nu} + 2\kappa h_{\mu\nu}, \tag{2.3}$$

where $h_{\mu\nu}$ is the graviton field, the quadratic part of the action contains terms with four derivatives that are multiplied by dimensionless parameters. Those terms are dominant in the ultraviolet and after properly normalizing them, the total action still contains parameters with non-negative dimension. In the case of Einstein gravity the four-derivative terms are absent and the kinetic term is multiplied by ζ , which has positive dimension. Once that term is normalized, all the interactions contain powers of $1/\zeta$. Since the counterterms are polynomials in the parameters of the theory (with properly normalized kinetic term), then the counterterms in Einstein gravity are polynomial in $1/\zeta$, which is a parameter with negative dimension. Therefore, it is possible to build counterterms with arbitrary mass dimension, since it can always be compensated by powers of $1/\zeta$. This is why general relativity is said to be nonrenormalizable by power counting. However, to prove the renormalizability of Stelle gravity, we need to show that the counterterms satisfies certain properties, as we explain in the next subsection.

Before to proceed with renormalizability we derive the graviton propagator and study its poles and residues. In this way we establish some notation and review the standard steps for the quantization of gravity. First, we recall that the Becchi–Rouet–Stora–Tyutin (BRST) transformations associated to diffeomorphisms read

$$\begin{aligned} s g_{\mu\nu} &= -\partial_\mu C^\alpha g_{\alpha\nu} - \partial_\nu C^\alpha g_{\mu\alpha} - C^\alpha \partial_\alpha g_{\mu\nu} \\ s C^\rho &= -C^\sigma \partial_\sigma C^\rho \\ s \bar{C}^\sigma &= B^\sigma \\ s B^\tau &= 0. \end{aligned} \tag{2.4}$$

The dimensions of the fields are

$$[g_{\mu\nu}] = 0, \quad [C^\rho] = 0, \quad [\bar{C}^\sigma] = 0, \quad [B^\tau] = 1.$$

It is easy to show that the operator s is nilpotent, i.e. $s^2 = 0$. Therefore, as in the case of gauge theories, we can add to the classical action a gauge-fixing term that is s -exact, obtaining the gauge-fixed action

$$S_{gf} = S_{HD} + s\Psi, \tag{2.5}$$

where Ψ is a fermionic functional with $[\Psi] = -1$. For the purposes of this section, we choose the functional

$$\Psi = \alpha \int \bar{C}^\mu \square (\mathcal{G}_\mu - \lambda \kappa^2 B_\mu), \tag{2.6}$$

where $\square = \eta^{\mu\nu} \partial_\mu \partial_\nu$, λ is a gauge-fixing parameter and \mathcal{G}_μ is the gauge-fixing function. We choose

$$\mathcal{G}_\mu = \eta^{\nu\rho} \partial_\rho g_{\mu\nu}, \tag{2.7}$$

to have a simplified propagator. More general fermionic functionals and gauge-fixing function can help in checking gauge independence of the counterterms (see [29]). With these choices the gauge-fixing term reads

$$s\Psi = \alpha \int B^\mu \square (\mathcal{G}_\mu - \lambda \kappa^2 B_\mu) + S_{gh}, \tag{2.8}$$

where S_{gh} is the action of the Faddeev-Popov ghosts

$$S_{\text{gh}} = \alpha \int \bar{C}^\mu \partial^\nu \square (g_{\mu\rho} \partial_\nu C^\rho + g_{\nu\rho} \partial_\mu C^\rho + C^\rho \partial_\rho g_{\mu\nu}). \tag{2.9}$$

In order to obtain the graviton propagator we substitute B with the solution of its equations of motion, i.e.

$$B_\mu = \frac{1}{2\lambda\kappa^2} \mathcal{G}_\mu, \tag{2.10}$$

and the gauge-fixed action (2.5) becomes

$$S_{\text{gf}} = S_{\text{HD}} + \frac{\alpha}{4\lambda\kappa^2} \int \mathcal{G}^\mu \square \mathcal{G}_\mu + S_{\text{gh}}. \tag{2.11}$$

We write a generic quadratic part of the graviton action as

$$S_{\text{gf}}^{\text{quad}} = \int h_{\mu\nu} T^{\mu\nu\rho\sigma} h_{\rho\sigma}, \tag{2.12}$$

so the propagator $D_{\mu\nu\rho\sigma}$ is the Fourier transform of the Green function of $T^{\mu\nu\rho\sigma}$, i.e.

$$\tilde{T}^{\mu\nu\alpha\beta} D_{\alpha\beta\rho\sigma} = \frac{i}{2} (\delta_\rho^\mu \delta_\sigma^\nu + \delta_\sigma^\mu \delta_\rho^\nu), \tag{2.13}$$

where \tilde{T} is the Fourier transform of T . Considering the symmetries of the indices, a basis for both \tilde{T} and D is given by the spin-2 projectors $\{\Pi^{(2)}, \Pi^{(1)}, \Pi^{(0)}, \bar{\Pi}^{(0)}\}$ plus the tensor $\bar{\bar{\Pi}}^{(0)}$ that takes into account the possible mixing between two different spin-0 component, which are defined as follows

$$\Pi_{\mu\nu\rho\sigma}^{(2)} \equiv \frac{1}{2} (\pi_{\mu\rho} \pi_{\nu\sigma} + \pi_{\mu\sigma} \pi_{\nu\rho}) - \frac{1}{3} \pi_{\mu\nu} \pi_{\rho\sigma}, \tag{2.14}$$

$$\Pi_{\mu\nu\rho\sigma}^{(1)} \equiv \frac{1}{2} (\pi_{\mu\rho} \omega_{\nu\sigma} + \pi_{\mu\sigma} \omega_{\nu\rho} + \pi_{\nu\rho} \omega_{\mu\sigma} + \pi_{\nu\sigma} \omega_{\mu\rho}), \tag{2.15}$$

$$\Pi_{\mu\nu\rho\sigma}^{(0)} \equiv \frac{1}{3} \pi_{\mu\nu} \pi_{\rho\sigma}, \tag{2.16}$$

$$\bar{\Pi}_{\mu\nu\rho\sigma}^{(0)} \equiv \omega_{\mu\nu} \omega_{\rho\sigma}, \tag{2.17}$$

$$\bar{\bar{\Pi}}_{\mu\nu\rho\sigma}^{(0)} \equiv \pi_{\mu\nu} \omega_{\rho\sigma} + \pi_{\rho\sigma} \omega_{\mu\nu}, \tag{2.18}$$

where $\pi_{\mu\nu}$ and $\omega_{\mu\nu}$ are the spin-1 projectors

$$\pi_{\mu\nu} \equiv \eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}, \quad \omega_{\mu\nu} \equiv \frac{p_\mu p_\nu}{p^2}. \tag{2.19}$$

Therefore, we can write (omitting the indices)

$$\tilde{T} = x_2 \Pi^{(2)} + x_1 \Pi^{(1)} + x_0 \Pi^{(0)} + \bar{x}_0 \bar{\Pi}^{(0)} + \bar{\bar{x}}_0 \bar{\bar{\Pi}}^{(0)} \tag{2.20}$$

$$D = y_2 \Pi^{(2)} + y_1 \Pi^{(1)} + y_0 \Pi^{(0)} + \bar{y}_0 \bar{\Pi}^{(0)} + \bar{\bar{y}}_0 \bar{\bar{\Pi}}^{(0)}, \tag{2.21}$$

where x_i are obtained from the quadratic part of the action and y_i are unknowns to be derived by imposing (2.13). For a general theory of gravity, the relations between x_i and y_i are

$$y_1 = \frac{i}{x_1}, \quad y_2 = \frac{i}{x_2}, \quad y_0 = \frac{i\bar{x}_0}{x_0\bar{x}_0 - 3\bar{\bar{x}}_0^2}, \quad \bar{y}_0 = \frac{i x_0}{x_0\bar{x}_0 - 3\bar{\bar{x}}_0^2}, \quad \bar{\bar{y}}_0 = -\frac{i\bar{\bar{x}}_0}{x_0\bar{x}_0 - 3\bar{\bar{x}}_0^2}. \tag{2.22}$$

After applying (2.22) to the action (2.12) and setting Λ_C to zero for simplicity, the graviton propagator reads

$$D_{\mu\nu\rho\sigma} = \frac{i}{p^2 + i\epsilon} \left[\frac{\Pi_{\mu\nu\rho\sigma}^{(2)}}{(\zeta - \alpha p^2)} - \frac{\Pi_{\mu\nu\rho\sigma}^{(0)}}{2(\zeta - \xi p^2)} - \frac{\lambda}{2\alpha p^2} (2\Pi_{\mu\nu\rho\sigma}^{(1)} + \bar{\Pi}_{\mu\nu\rho\sigma}^{(0)}) \right]. \tag{2.23}$$

The expression of the Faddeev-Popov ghost propagator follows straightforwardly from the part of S_{gh} that is quadratic in the fields C, \bar{C} . In momentum space we find

$$\langle C^\mu \bar{C}^\nu \rangle = \frac{i}{p^2 + i\epsilon} \left(\frac{\eta^{\mu\nu} - p^\mu p^\nu / 2p^2}{\alpha p^2} \right). \tag{2.24}$$

We identify the single poles and their residues by splitting the propagator (2.23) into partial fractions. To further simplify, we choose $\lambda = 0$, then we find

$$D_{\mu\nu\rho\sigma}^{\lambda=0} = \frac{i}{2\zeta} \left[\frac{2\Pi_{\mu\nu\rho\sigma}^{(2)} - \Pi_{\mu\nu\rho\sigma}^{(0)}}{p^2 + i\epsilon} - \frac{2\Pi_{\mu\nu\rho\sigma}^{(2)}}{p^2 - \zeta/\alpha + i\epsilon} + \frac{\Pi_{\mu\nu\rho\sigma}^{(0)}}{p^2 - \zeta/\xi + i\epsilon} \right]. \tag{2.25}$$

The first term is the same as in Einstein gravity and describes a massless graviton, while the second and third term are typical of Stelle gravity and describe a massive spin-2 field and a massive scalar field, respectively. The residue at the massive spin-2 pole is negative. This means that the theory propagates a ghost, which violates unitarity. In order to make this clear, we show a simple example by means of a scalar toy model. Consider the action

$$S(\varphi) = \int \left[\frac{1}{2} \partial_\mu \varphi \left(1 + \frac{\square}{M^2} \right) \partial^\mu \varphi - \frac{1}{2} m^2 \varphi \left(1 + \frac{\square}{M^2} \right) \varphi - \frac{\lambda}{3!} \varphi^3 \right], \tag{2.26}$$

where φ is a scalar field, m and M are mass parameters and λ is a coupling constant. Using the Feynman prescription, the propagator reads

$$D(p^2) = \frac{-iM^2}{(p^2 - m^2 + i\epsilon)(p^2 - M^2 + i\epsilon)} = \frac{i}{p^2 - m^2 + i\epsilon} - \frac{i}{p^2 - M^2 + i\epsilon}, \tag{2.27}$$

where in the second step we have appropriately redefined ϵ . A degree of freedom associated to a propagator with negative residue at the pole is called ghost. The propagator (2.27) describes two degrees of freedom, one at $p^2 = m^2$ with a positive residue and one at $p^2 = M^2$ with negative residue. Ghosts should not be confused with tachyons, which are defined as degrees of freedom with negative mass squared¹ and introduce additional pathologies. For the present review it is important just to know that tachyons cannot be cured by the fakeon prescription and the parameters of the theory must be constrained in order to avoid tachyons (see Sect. 5).

The simplest example of a unitarity violation is given by the tree-level 2-to-2 scattering amplitude. Unitarity is encoded in the optical theorem (see next section for the details), which, in the case the process considered, states that

$$2\text{Im}\mathcal{M}_{2\rightarrow 2} = \int d\Pi |\mathcal{M}_{2\rightarrow 1}|^2 \geq 0, \tag{2.28}$$

where $\mathcal{M}_{2\rightarrow 2}$ and $\mathcal{M}_{2\rightarrow 1}$ denote the tree-level amplitudes of the processes $\varphi\varphi \rightarrow \varphi\varphi$ and $\varphi\varphi \rightarrow \varphi$, respectively, while $d\Pi$ is the integration measure over the phase space of final states. The amplitude $\mathcal{M}_{2\rightarrow 2}$ is

$$\mathcal{M}_{2\rightarrow 2} = -\lambda^2 \left(\frac{1}{p^2 - m^2 + i\epsilon} - \frac{1}{p^2 - M^2 + i\epsilon} \right) \tag{2.29}$$

and its imaginary part reads

$$\text{Im}\mathcal{M}_{2\rightarrow 2} = \pi\lambda^2 [\delta(p^2 - m^2) - \delta(p^2 - M^2)], \tag{2.30}$$

which is not nonnegative and therefore violates (2.28). Similar violations appear every time a odd number of ghosts is on shell in the right-hand side of (2.28).

This simple example shows what happens also in Stelle theory. Moreover, note that the pole of the additional scalar in (2.25) has positive residue. Therefore, if we send the mass of the ghost to infinity ($\alpha \rightarrow 0$) we find a unitary theory describing a massless graviton and a massive scalar, the Starobinsky model. However, in that case renormalizability is lost. In fact, the presence of the ghost is crucial to obtain certain cancellations between Feynman diagrams that avoid higher-dimensional counterterms to be generated by renormalization. In the end renormalizability can be obtained at the price of unitarity. These two properties seem to be mutually exclusive in quantum gravity, since they are both entangled to the ghost. If we have the ghost we loose unitarity, while if we get rid of the ghost we loose renormalizability. To define a quantum field theory of gravity that makes sense both physically and mathematically we need to find a way to consistently remove the ghost from the set of physical states and keep its contributions inside Feynman diagrams at the same time. This cannot be achieved without relaxing some assumptions of standard quantum field theory. In our approach we disentangle the on-shell contributions of the ghost from the virtual ones, that would typically be related in the standard case of the Feynman prescription. Once this modification is introduced we can obtain the property we want and remove the ghost from the spectrum at any energy scale without spoiling the consistency of the theory.

Before to proceed with the details of the fakeon prescription we give a sketch of the proof of renormalizability in Stelle gravity for the readers that are not familiar with it. We anticipate that the fakeon prescription do not modify the divergences of the theory, therefore from this point of view there is no difference with the Stelle gravity.

2.1 Renormalization

We give the main ingredients to prove the renormalizability of the theory with the help of the Batalin-Vilkovisky formalism, which we briefly introduce below. In [11] it was assumed that the divergent part of the effective action satisfies the so-called Kluberg-Stern-Zuber conjecture [30, 31]. More recently it has been proved by a theorem in [32] for a large class of models, which includes higher-derivative quantum gravity. We do not review the proof here and the reader is referred to [32] for details.

The Batalin-Vislkovisky formalism is a useful tool to prove renormalizability in gauge theories and gravity. It makes use of an extended action Σ that accounts for the sources of the BRST transformations and allows to write the Ward-Takahashi-Slavnov-Taylor

¹ Note that a ghost can also be a tachyon.

identities in a compact form. The counterterms then satisfy certain properties and the divergences are removed by means of an iterative procedure.

We collect all the fields into the row $\Phi^\alpha = (g_{\mu\nu}, C^\rho, \bar{C}^\sigma, B^\tau)$ and introduce a set of sources $K_\alpha = (K_g^{\mu\nu}, K_C^\rho, K_{\bar{C}}^\sigma, K_B^\tau)$ for the composite BRST operators $s\Phi^\alpha$, conjugate to the fields, and define the *antiparentheses* of two functionals X and Y of Φ and K as

$$(X, Y) \equiv \int \left(\frac{\delta_r X}{\delta \Phi^\alpha} \frac{\delta_l Y}{\delta K_\alpha} - \frac{\delta_r X}{\delta K_\alpha} \frac{\delta_l Y}{\delta \Phi^\alpha} \right), \tag{2.31}$$

where the integral is over the spacetime points associated with repeated indices and the subscripts l, r in δ_l, δ_r denote the left and right functional derivatives, respectively. We introduce the *ghost number*

$$\text{gh}\#(g_{\mu\nu}) = \text{gh}\#(B^\tau) = 0, \quad \text{gh}\#(C^\rho) = 1, \quad \text{gh}\#(\bar{C}^\sigma) = -1, \tag{2.32}$$

so that the gauge-fixed action is invariant under the global symmetry

$$\Phi^\alpha \rightarrow \Phi^\alpha e^{i\beta \text{gh}\#(\Phi^\alpha)}, \tag{2.33}$$

where β is a constant. Finally, the extended action is defined by adding the source term $S_K = -\int s\Phi^\alpha K_\alpha$ to the gauge-fixed action. The dimensions of the sources are

$$[K_g] = 3, \quad [K_C] = 3, \quad [K_{\bar{C}}] = 3, \quad [K_B] = 2, \tag{2.34}$$

while the ghost numbers are fixed by requiring that S_K be invariant under the symmetry (2.33) combined with²

$$K_\alpha \rightarrow K_\alpha e^{i\beta \text{gh}\#(K_\alpha)} \tag{2.35}$$

and read

$$\text{gh}\#(K_g) = -1, \quad \text{gh}\#(K_C) = -2, \quad \text{gh}\#(K_{\bar{C}}) = 0, \quad \text{gh}\#(K_B) = -1. \tag{2.36}$$

Finally, the statistics of the fields Φ^α and sources K_α are opposite to each other.

In the Batalin-Vilkovisky formalism, the gauge-fixing term can be written as

$$s\Psi = (S_K, \Psi) \tag{2.37}$$

and the extended action reads

$$\Sigma(\Phi, K) = S_{\text{HD}} + (S_K, \Psi) + S_K. \tag{2.38}$$

This action is used to define the generating functional Z of the correlation functions and the generating functional W of the connected correlation functions by means of the formula

$$Z(J, K) = \int [d\Phi] \exp\left(i\Sigma(\Phi, K) + i \int \Phi^\alpha J_\alpha \right) = \exp iW(J, K). \tag{2.39}$$

As usual, the generating functional of one-particle irreducible diagrams is defined as the Legendre transform of W with respect to J

$$\Gamma(\Phi, K) = W(J, K) - \int \Phi^\alpha J_\alpha, \tag{2.40}$$

where $\Phi^\alpha = \delta_r W / \delta J_\alpha$. Renormalization is achieved by means of parameter and field redefinitions, which are canonical transformations with respect to the antiparentheses. It is easy to show that Σ and Γ satisfy the *master equations*

$$(\Sigma, \Sigma) = 0, \quad \text{and} \quad (\Gamma, \Gamma) = 0. \tag{2.41}$$

The first one is a generalization of the BRST invariance of the gauge-fixed action. In fact, the operator

$$\sigma \equiv (\Sigma, \cdot) \tag{2.42}$$

reduces to the BRST transformations when acting on Φ , while it further reduces to a diffeomorphism when acting on $g_{\mu\nu}$. The second equation collects the Ward–Takahashi–Slavnov–Taylor identities. Moreover, it contains information about the counterterms. In order to keep track of the loop expansion, we temporarily reintroduce the factors \hbar , the n -th order in \hbar being the $n + 1$ -th order in loops. Then, by expanding Γ in powers of \hbar and using the master equation, it is possible to show that the divergent part of the effective action satisfies the equation

$$\sigma \Gamma_{\text{div}}^{(n)} = 0, \tag{2.43}$$

² This symmetry is not spoiled by radiative corrections since both Feynman rules and the diagrammatics are compatible with it.

where $\Gamma_{\text{div}}^{(n)}$ is the divergent part of the functional $\Gamma^{(n)}$, which we assume to be convergent up to the order \hbar^{n-1} . If a functional X satisfies $\sigma X = 0$, we say that it is σ -closed, while if it is such that $X = \sigma Y$, where Y is another functional, we say that it is σ -exact. The general solution of (2.43) can be written as

$$\Gamma_{\text{div}}^{(n)}(\Phi, K) = \tilde{G}_n(\Phi, K) + \sigma \tilde{X}_n(\Phi, K), \tag{2.44}$$

where \tilde{G}_n is a σ -closed local functional. In order to prove renormalizability we need to ensure that it is possible to move all the dependence on the sources and unphysical fields into the σ -exact term. This means that the solution (2.44) is reorganized in the form

$$\Gamma_{\text{div}}^{(n)}(\Phi, K) = G_n(h) + \sigma X_n(\Phi, K), \tag{2.45}$$

where the functional G_n now depends only on the metric fluctuation and is gauge invariant, being σ -closed. This is the Kluberg-Stern–Zuber conjecture mentioned above, which is now proved by a theorem [32]. We give some additional properties of Σ , $\Gamma_{\text{div}}^{(n)}$ and X_n , which allow us to write $\Gamma_{\text{div}}^{(n)}$ in a more explicit form

- (1) The divergent part cannot depend on B , $K_{\bar{C}}$ and K_B because no vertices of the action (2.38) contain them, so no one-particle irreducible diagrams can be built with B , $K_{\bar{C}}$ and K_B .
- (2) Σ depends on K_g and \bar{C} only through the combination

$$\tilde{K}_g^{\mu\nu} \equiv K_g^{\mu\nu} + \alpha \square \int \frac{\delta \mathcal{G}_\rho}{\delta g_{\mu\nu}} \bar{C}^\rho. \tag{2.46}$$

This property ensures that also $\Gamma_{\text{div}}^{(n)}$ depends on K_g and \bar{C} only through \tilde{K}_g . In fact, given a diagram with an external $K_g^{\mu\nu}$ leg, there exists an almost identical diagram where the $K_g^{\mu\nu}$ leg is replaced by a $\alpha \square \int \frac{\delta \mathcal{G}_\rho}{\delta g_{\mu\nu}} \bar{C}^\rho$ leg and viceversa.

- (3) X_n also depends on $K_g^{\mu\nu}$ and \bar{C}^ρ via the combination $\tilde{K}_g^{\mu\nu}$. The proof of this property is as follows. The dimension of X_n and its ghost number imply that we can parametrize it as

$$X_n(\Phi, K) = \int \Delta g_{\mu\nu} K_g^{\mu\nu} + \int \Delta C^\rho K_\rho^C + \int \bar{C}^\rho L_\rho, \tag{2.47}$$

where L_ρ is a function of dimension 3 and ghost number zero, while $\Delta g_{\mu\nu}$ and ΔC^ρ are the renormalizations of the metric tensor and the Faddeev-Popov ghosts, respectively. Then, from the expression (2.45) and property 1) it follows that also σX_n does not depend on B . In σX_n the terms linear in B are

$$\sigma X_n|_B = \int \left(\frac{\delta_r S}{\delta g_{\mu\nu}} \frac{\delta_l X_n}{\delta K_g^{\mu\nu}} - \frac{\delta_r X_n}{\delta \bar{C}^\rho} \frac{\delta_l S}{\delta K_\rho^C} \right) \Big|_B = \int B^\rho \left[\alpha \square \int \frac{\delta_r \mathcal{G}_\rho}{\delta g_{\mu\nu}} \Delta g_{\mu\nu} + L_\rho \right]. \tag{2.48}$$

Therefore, the expression in the squared bracket in (2.48) must vanish, which implies that the terms proportional to K_g plus those proportional to \bar{C} in X_n are

$$\begin{aligned} X_n(\Phi, K)|_{K_g, \bar{C}} &= \int \left[\Delta g_{\mu\nu} K_g^{\mu\nu} - \bar{C}^\rho \alpha \square \int \frac{\delta_r \mathcal{G}_\rho}{\delta g_{\mu\nu}} \Delta g_{\mu\nu} \right] \\ &= \int \Delta g_{\mu\nu} \tilde{K}_g^{\mu\nu}. \end{aligned} \tag{2.49}$$

From the properties listed above we deduced that

$$X_n(\Phi, K) = \int \Delta g_{\mu\nu} \tilde{K}_g^{\mu\nu} + \int \Delta C^\rho K_\rho^C. \tag{2.50}$$

Although $\Gamma_{\text{div}}^{(n)}$ can be written in the form (2.45), it is still possible that the σ -exact term contains terms that depend only on the physical fields. More explicitly, a straightforward calculation gives

$$\sigma X_n = \int \frac{\delta S_{\text{HD}}}{\delta g_{\mu\nu}} \Delta g_{\mu\nu} - \int \Delta \mathcal{R}_{\mu\nu} \tilde{K}^{\mu\nu} - \int \Delta \mathcal{R}^\rho K_\rho^C, \tag{2.51}$$

where

$$\Delta \mathcal{R}_{\mu\nu} = -\sigma \Delta g_{\mu\nu} + \int \Delta g_{\alpha\beta} \frac{\delta_l (s g_{\mu\nu})}{\delta g_{\alpha\beta}} + \int \Delta C^\tau \frac{\delta_l (s g_{\mu\nu})}{\delta C^\tau}, \tag{2.52}$$

$$\Delta \mathcal{R}^\rho = -\sigma \Delta C^\rho + \int \Delta C^\tau \frac{\delta_l (s C^\rho)}{\delta C^\tau} \tag{2.53}$$

We can see that there is a term proportional to the equations of motion, which depends only on the metric. Then, if we split the metric as in (2.3) we have

$$\Delta g_{\mu\nu} = t_0 g_{\mu\nu} + t_1 h_{\mu\nu} + t_2 \eta_{\mu\nu} h + \mathcal{O}(h^2), \tag{2.54}$$

where t_i are (gauge-dependent) coefficients that depend on the parameters in (2.1), which makes the term proportional to the equations of motion non-covariant. Therefore, when we compute the divergent part of a Feynman diagrams with only graviton external legs, we would get also the terms proportional to the equations of motion. In practice, this mean that those terms always need to be taken into account when computing the divergent part of diagrams in Stelle theory and, in general, the results obtained in this way are not covariant. In other words, the functional G_n cannot be written as a function of $g_{\mu\nu}$ only but it depends also on $h_{\mu\nu}$ in an independent way. Only after removing those terms, we can safely conclude that any other one-particle irreducible diagrams would generate divergent terms that sum together to give a covariant result. Another way to see this is to consider the wave functions renormalization of $h_{\mu\nu}$ and its trace as independent. This depends on the gauge choice since the wave function renormalization are unphysical. For example, the background field method, which can be seen as a gauge choice, takes care of terms like those in σX_n . If we do not choose a particular gauge or a method that preserves general covariance at each steps, we can remove the σ -exact term by means of the transformation

$$\Phi^{i'} = \Phi^i - \frac{\delta X_n(\Phi, K)}{\delta K_i'}, \quad K_i' = K_i - \frac{\delta X_n(\Phi, K)}{\delta \Phi^i}. \tag{2.55}$$

which is canonical, i.e. preserves the antiparentheses. In components, it reads

$$h'_{\mu\nu} = h_{\mu\nu} - \Delta g_{\mu\nu}, \quad C'^\sigma = C^\rho - \Delta C^\rho, \quad \bar{C}'^\tau = \bar{C}^\tau, \quad B'^\nu = B^\nu, \tag{2.56}$$

while the sources transform as

$$K_g'^{\mu\nu} = K_g^{\mu\nu} + \tilde{K}_g^{\alpha\beta} \frac{\delta \Delta g_{\alpha\beta}}{\delta h_{\mu\nu}} - K_\sigma^C \frac{\delta \Delta C^\sigma}{\delta h_{\mu\nu}}, \quad K_\sigma'^C = K_\sigma^C - K_\alpha^C \frac{\delta \Delta C^\alpha}{\delta C^\sigma}, \tag{2.57}$$

$$K_\tau'^{\bar{C}} = K_\tau^{\bar{C}}, \quad K_\tau'^B = K_\tau^B.$$

We obtain

$$\Gamma_{\text{div}}^{(n)}(\Phi', K') = \Gamma_{\text{div}}^{(n)}(\Phi, K) - \sigma X_n(\Phi, K) = \int \sqrt{-g} \left[2\Lambda_n + \zeta_n R + \frac{\alpha_n}{2} C^2 - \frac{\xi_n}{6} R^2 \right], \tag{2.58}$$

where all the terms of dimension 0, 2 and 4 have been included as possible counterterms and the parameters with the subscript n denote the renormalization of the associated quantities.

We stress that we used the fact that the divergent terms can be organized as (2.45) in any gauge. Additional details can be found in [32] and [29]. However, we could choose a gauge fixing and a parametrization of the graviton field [11] such that $\Delta g_{\mu\nu} = \Delta C^\rho = 0$, so X_n is zero and the transformation (2.55) is trivial. In that case, no terms proportional to the equations of motion appear, the counterterms are gauge invariant and there is no need to use the theorem that shows (2.45) to prove renormalizability in that specific setting.

3 Unitarity and cutting identities

In this section we recall how unitarity is formulated diagrammatically by using the so-called *cutting equations*. Unitarity is the condition

$$SS^\dagger = \mathbb{1} \tag{3.1}$$

on the S matrix, which can be rewritten in terms of its nontrivial part T as

$$-i(T - T^\dagger) = TT^\dagger. \tag{3.2}$$

The matrix elements of Eq. (3.2) are obtained by choosing initial and final states, $|i\rangle, |f\rangle$ and by inserting the completeness relation between T and T^\dagger on the right-hand side of (3.2)

$$-i\langle f|(T - T^\dagger)|i\rangle = \sum_{|n\rangle} \int d\Pi_n \langle f|T|n\rangle \langle n|T^\dagger|i\rangle, \tag{3.3}$$

where the sum runs over all the possible final states and the integral symbolically denote the integration over the phase space of those final states. The matrix element $\langle f|T|i\rangle$ is associated to the amplitude \mathcal{A}_{fi} up to a δ function for the conservation of four momentum. If we write Eq. (3.2) in terms of the amplitudes for an elastic scattering, i.e. $|f\rangle = |i\rangle$, we obtain

$$2\text{Im}\mathcal{A}_{ii}^{(f)}(s) = \sum_n \int d\Pi_n |\mathcal{A}_{in}|^2 = 2\Phi\sigma_{\text{tot}}(s), \tag{3.4}$$

where $\mathcal{A}^{(f)}$ is the forward amplitude, s is the center-of-mass energy squared, Φ is the flux factor of initial particles and σ_{tot} is the total cross sections. This has a more familiar form known from the optical theorem in nonrelativistic scattering and this is why

also (3.2) is named like that. However, Eq. (3.2) is more general, since it is valid for any type of process and on its right-hand side typically involves quantities that are not cross sections. Since quantum field theory is defined perturbatively by means of Feynman diagrams, Eq. (3.2) can be expanded diagrammatically, obtaining a set of equations order by order. There is a more powerful set of identities that are valid diagram by diagram: the cutting equations, which hold in any local quantum field theory.³ From now on we refer to them as *cutting identities*. However, only if the theory satisfies additional requirements the cutting identities imply the optical theorem.

The first step to derive the cutting identities is the *largest-time equation*, which is their coordinate-space version. For the purpose of this section we consider only the scalar ϕ^3 theory given by the action (2.26) without the higher-derivative terms. The generalization to fermions, gauge theories and gravity involves some caveats related to gauge invariance, which we do not address here. More details can be found in [33, 34]. First, we recall that the Feynman propagator

$$D_{ij}(x) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ipx}, \quad x \equiv x_i - x_j \tag{3.5}$$

can be written as

$$D_{ij}(x) = \theta(x^0)D_{ij}^+(x) + \theta(-x^0)D_{ij}^-(x), \tag{3.6}$$

where

$$D_{ij}^\pm(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{e^{\mp i\omega(\mathbf{p})x^0} e^{\pm i\mathbf{p}\cdot\mathbf{x}}}{2\omega(\mathbf{p})}, \quad \omega(\mathbf{p}) \equiv \sqrt{\mathbf{p}^2 + m^2},$$

θ is the Heaviside step function and the bold symbols represent the space components of vectors. Then, consider a generic Feynman diagram represented by a function $F(x_1, \dots, x_n)$, where each spacetime point x_i is associated to a vertex. For example, a one-loop three-point function is given by

$$F(x_1, x_2, x_3) = (-i\lambda)^3 D_{12}D_{23}D_{31}, \tag{3.7}$$

where λ indicates a generic coupling constant. We introduce another function where one or more x_i are marked by a hat. The new function is obtained from F by means of the following substitutions

$$D_{ij}(\hat{x}_i - x_j) \rightarrow D_{ij}^+, \quad D_{ij}(x_i - \hat{x}_j) \rightarrow D_{ij}^-, \quad D_{ij}(\hat{x}_i - \hat{x}_j) \rightarrow (D_{ij})^*. \tag{3.8}$$

Moreover, each vertex associated with a marked point must be substituted with its complex conjugate. Using the example above we can have

$$F(\hat{x}_1, \hat{x}_2, x_3) = (-i\lambda)^3 (D_{12})^* D_{23}^+ D_{31}^-. \tag{3.9}$$

Now, suppose that a time component x_i^0 for some i is larger than all the others. Then any diagram where x_i is not marked is equal to minus the same diagram in which x_i is marked. This follows straightforwardly from the definition of marked points and from (3.6). For example, in the case of the three-point function above, assuming that x_1^0 is the largest time component, we can write

$$F(x_1, x_2, x_3) = (-i\lambda)^3 D_{12}D_{23}D_{31} = (-i\lambda)^3 D_{12}^+ D_{23} D_{31}^- = -F(\hat{x}_1, x_2, x_3), \tag{3.10}$$

$$F(x_1, x_2, \hat{x}_3) = -(-i\lambda)^3 D_{12}D_{23}^- D_{31}^+ = -(-i\lambda)^3 D_{12}^+ D_{23}^- D_{31}^+ = -F(\hat{x}_1, x_2, \hat{x}_3) \tag{3.11}$$

and so on. Therefore, for any time configuration we have a set of identities that hold by construction, as long as it is possible to determine which time component is the largest, i.e. the vertices are local, and they have distinct times. Coincident points can generate contact terms which do not spoil the result [34]. We can write the largest-time equation in a compact form without specifying which time component is the largest as

$$\sum_m F_m(x_1, \dots, \hat{x}_i, \dots, \hat{x}_j, \dots, x_n) = 0, \tag{3.12}$$

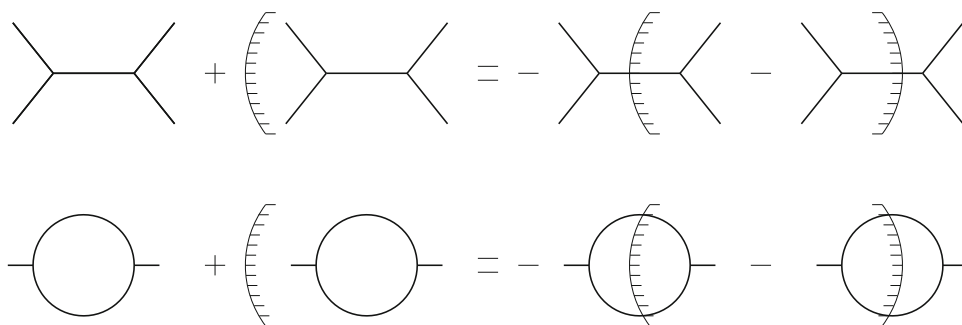
where the sum runs over all the possible ways of marking the points (including the cases where all and no vertices are marked) and the F_m are the diagram with marked vertices. Then we can take the Fourier transform of (3.12). The Fourier transform of D^\pm can be written in the form

$$\tilde{D}^\pm(p) = 2\pi\theta(p^0)\rho(p^2), \tag{3.13}$$

where $\rho(p^2)$ is a distribution such that the Fourier transform of the propagator is

$$\tilde{D}(p) = \int_0^\infty \frac{ds}{2\pi} \frac{i\rho(s)}{p^2 - s + i\epsilon}. \tag{3.14}$$

Fig. 1 On the top, the cutting identity for a tree-level diagram. On the bottom, the cutting identity for the bubble diagram



In the case of a standard scalar field theory we have

$$\rho(p^2) = \delta(p^2 - m^2), \quad \tilde{D}(p^2) = \frac{i}{p^2 - m^2 + i\epsilon}, \quad \tilde{D}^\pm = 2\pi\theta(\pm p^0)\delta(p^2 - m^2). \tag{3.15}$$

In order to have a better graphical view, we introduce cut diagrams, instead of marked ones, where an internal line is cut if it connects a marked vertex with an unmarked one. The cut is represented by a continuous line and has a “shaded region” on the side where the marked vertex lies, which represents the energy flow given by the θ -functions in (3.13). In practice, the cut diagrams are obtained from the main diagram by means of the substitution

$$\frac{i}{p^2 - m^2 + i\epsilon} \rightarrow 2\pi\theta(\pm p^0)\delta(p^2 - m^2) \tag{3.16}$$

for each cut propagator and by complex conjugating everything that lies on the shadowed region. Finally, a diagram where all the vertices lie in the shadowed region correspond to the complex conjugate of the diagram with no cut. With these definitions we can write the cutting identities

$$G + G^* = - \sum_{\text{cuts}} G_c, \tag{3.17}$$

where G is the diagram in momentum space and G_c are the cut diagrams. Simple examples of the cutting identities are given by those in Fig. 1. The one on the top reads

$$\begin{aligned} -\lambda^2 \left(\frac{i}{p^2 - m^2 + i\epsilon} + \frac{-i}{p^2 - m^2 - i\epsilon} \right) &= -\lambda^2 2\pi\theta(p^0)\delta(p^2 - m^2) - \lambda^2 2\pi\theta(-p^0)\delta(p^2 - m^2) \\ &= -\lambda^2 2\pi\delta(p^2 - m^2). \end{aligned} \tag{3.18}$$

The identity on the bottom in Fig. 1 is shown by considering equal masses m for the particles in the loop. In that case the bubble diagram reads

$$\mathcal{B}(p^2) = -\frac{i\lambda^2}{2(4\pi)^2} \int_0^1 dx \ln \left(\frac{-xp^2(1-x) + m^2 - i\epsilon}{\mu^2} \right) + \mathcal{B}_{\text{div}}, \tag{3.19}$$

where x is a Feynman parameter, μ is the renormalization scale and \mathcal{B}_{div} is the divergent part of the diagram evaluated in dimensional regularization, which might contain also some finite constant depending on the scheme. Therefore,

$$\mathcal{B} + \mathcal{B}^* = -\frac{\lambda^2}{16\pi} \theta(p^2 - 4m^2) \sqrt{1 - \frac{4m^2}{p^2}}. \tag{3.20}$$

Note that \mathcal{B}_{div} cancels in (3.20), since it is always purely imaginary. The cut diagrams are

$$\mathcal{B}_{12} = \lambda^2 \int \frac{d^4q}{(2\pi)^2} \theta(p^0 + q^0) \delta((p+q)^2 - m^2) \theta(-q^0) \delta(q^2 - m^2) \tag{3.21}$$

and analogous expression for \mathcal{B}_{21} , where the subscripts indicate the energy flow. It is easy to check the identity by direct computation and show that

$$\mathcal{B} + \mathcal{B}^* = -\mathcal{B}_{12} - \mathcal{B}_{21}. \tag{3.22}$$

In this case, where the scalar field propagator has a positive residue, the cutting identities are equivalent to the optical theorem. For example, remembering that $G = iT$, the right-hand side of (3.20) is equal to the right-hand side of (3.3) when the initial state is

³ Here locality means that it is possible to build consistent Feynman rules that involve local vertices.

a single particle with center-of-mass energy squared p^2 and the final state is two particles of mass m^2 . Note that the θ -function in (3.20) gives the correct kinematical condition $p^2 > 4m^2$ for the decay to happen. However, it is not always true that (3.17) implies (3.2). In fact, we did not assume any positivity property on the distribution ρ in (3.13) to prove (3.17), which, instead, is crucial for the optical theorem. The simplest example is the case of a ghost. In that case we have $\rho = -\delta(p^2 - m^2)$ and the right-hand side of (3.18) does not have the correct sign to match the right-hand side of (3.3).

In general we can say that the cutting identities implies a modified version of the optical theorem called *pseudounitariness equation*, which, written in a compact form, reads

$$-i(T - T^\dagger) = TCT^\dagger, \quad C = \text{diag}(1, \dots, 1, -1, \dots, -1, \dots), \tag{3.23}$$

where the minus ones in the matrix C depend on the presence of ghosts in the theory. Therefore, the task of determine whether a theory is unitary or not reduces to find a way to deal with the matrix C . For example, in gauge theories the matrix C is not the identity, due to the presence of Faddeev-Popov ghosts as well as longitudinal and temporal component of the gauge fields. However, thanks to the BRST symmetry it is possible to project the Fock space onto a subspace which is generated only by creation operators of the physical fields, i.e. the transverse gauge bosons. This operation is consistent with unitarity and the degrees of freedom that are projected out in this way are not generated back by loop corrections. In the next section we show that it is possible to achieve the same type of projection without the help of a symmetry by changing the quantization prescription for the degrees of freedom that we want to remove from the spectrum. However, there is a crucial difference in the case of the fakeon prescription: unlike the BRST case, purely virtual particles are not unphysical, so they can still contribute to the interactions between other standard degrees of freedom and give physical effects.

4 Spectral identities and unitarity with fakeons

In this section we show how to turn a standard degree of freedom into a purely virtual particle. This result is achieved by means of a set of operations that we call the *fakeon prescription*. There are a few different equivalent ways of implementing the fakeon prescription. Here we present what we believe to be the clearest one, i.e. via the so-called *threshold decomposition*. The amplitudes are obtained by using the Feynman prescription for every propagator in the theory and only at the end an additional step is performed. In this way it is clear which properties of the standard Feynman prescription survive the fakeon prescription and which get modified. Roughly speaking, what we want to achieve with the fakeon prescription is to remove all the parts of the amplitudes that are related to the (would-be-purely-virtual) particle being on-shell. For example, the real part of the bubble diagram (3.19) tells whether the particle in the external line can decay into the (on-shell) particles in the loop, as shown in (3.20). Once those pieces are subtracted, we perform a projection at the level of the Fock space by choosing to work in a subspace where the particles that we want to make purely virtual are not external legs. Combining the two operations (removing pieces of the amplitudes and projecting onto a subspace) guarantees that quantum corrections will not generate back the degrees of freedom that we remove from the set of particle that can appear on shell.

In order to identify the subtractions, we decompose the amplitudes as sums of terms that are associated to single thresholds. In Feynman diagrams thresholds tell which kinematic configurations allow for the virtual particles to become on shell. To obtain the threshold decomposition it is sufficient to integrate each Feynman diagram over the loop energies. The integral over space momenta is postponed to after the decomposition is performed.

We consider one-loop diagrams, the generalization to higher loops can be found in [35]. Moreover, we derive the decomposition for a scalar theory, since all the crucial points are related to the singularities and branch cuts of the amplitudes, which are not modified by the presence of nontrivial numerators. For a one-loop diagram with N legs, we define the *skeleton diagram* as

$$G_N^s = \int \frac{dq^0}{2\pi} \prod_{a=1}^N \frac{2\omega_a}{(q^0 + k_a^0)^2 - \omega_a^2 + i\epsilon_a}, \quad \omega_a = \sqrt{(\mathbf{q} - \mathbf{k}_a)^2 + m_a^2}, \tag{4.1}$$

where k_a are the external momenta.⁴ With this definition, a standard Feynman diagram is

$$G_N = \int \frac{d^{D-1}\mathbf{q}}{(2\pi)^{D-1}} \left(\prod_{a=1}^N \frac{1}{2\omega_a} \right) G_N^s. \tag{4.2}$$

The skeleton diagrams satisfy the identity [35]

$$G^s + (G^s)^* = - \sum_{\text{cuts}} G_c^s, \tag{4.3}$$

⁴ In order to have more symmetric formulas we have introduced one momentum for each external leg, which is redundant.

where G_c^s are the cut skeleton diagrams. Moreover, each diagram in (4.3) can be decomposed into a sum of independent terms, which means that (4.3) gives a set of identities that holds independently. The decomposition works as follows. First, perform the integral over the loop energies by means of the residue theorem. After this operation, the result has the form

$$G_N^s = \sum_i c'_i \prod_{j=1}^N \frac{1}{D'_{ij}}, \tag{4.4}$$

where the coefficients c'_i can depend on the spatial external momenta and the denominators D' are linear combinations of external energies k_a^0 and frequencies $\omega'_a = \omega_a - i\epsilon_a$. It is always possible to manipulate the above expression such that the denominators contain only sums of frequencies, i.e.

$$G_N^s = \sum_i c_i \prod_{j=1}^N \frac{1}{D_{ij}}, \quad D_{ij} = P^0 - \sum_a \omega'_a, \tag{4.5}$$

where P^0 is a combination of external energies k_a^0 . This means that the true physical nonanalyticities are always associated to thresholds which Lorentz invariant expressions have the form

$$P^2 \geq \left(\sum_i m_i \right)^2, \tag{4.6}$$

where P^2 is some invariant build with external momenta and m_i are internal masses. Thresholds that contains differences of masses are called *pseudothresholds* and they need to disappear from the amplitudes in order to have a physically consistent theory. In fact, pseudethresholds are responsible for instabilities. For example, if the pseudothreshold

$$P^2 \geq (m_1 - m_2)^2, \tag{4.7}$$

is present, then a particle with squared mass $P^2 = m^2$ would be allowed to decay into two heavier particles with masses $m_1, m_2 > m$ as long as their difference satisfies the bound.

The absence of pseudothresholds is nontrivial and depends on the choice of the prescription. In fact, the location of the poles in the loop-energy complex plane might change with different prescriptions and the integration contour used in the residue theorem would enclose different poles. For example, choosing the Feynman prescription for each degrees of freedom fixes the relative sign between ω_a and $i\epsilon_a$ to be the same in every ω'_a . In general, mixing the prescription could change that. An example is given by mixing Feynman and anti-Feynman prescriptions in the same diagram by choosing the former for some degree of freedom and the latter for others. In that case, some of the physical thresholds are switched with pseudothresholds [36]. From what follows, it is clear that mixing the Feynman prescription and the fakeon prescription does not lead to instabilities and the pseudothresholds still cancel.

Once a skeleton diagram is reduced in the form (4.5) we apply the formula

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x + i\epsilon} = \mathcal{P} \frac{1}{x} - i\pi \delta(x), \tag{4.8}$$

to each term, where the limit is understood in the sense of distributions and \mathcal{P} is the Cauchy principal value. At this level each threshold, or at least what would be a threshold once we finalize the integral over the space components of loop momenta, is associated with a delta function. It is important to separate the terms that contain independent thresholds. For this purpose we introduce a few definitions in order to rewrite the skeleton diagram. We define

$$\mathcal{P}^{ab} = \mathcal{P} \frac{1}{e_a - e_b - \omega_a - \omega_b}, \quad \mathcal{Q}^{ab} = \mathcal{P} \frac{2\omega_b}{(e_a - e_b - \omega_a)^2 - \omega_b^2}, \quad \Delta^{ab} = \pi \delta(e_a - e_b - \omega_a - \omega_b), \tag{4.9}$$

where $e_a \equiv k_a^0$ and the subscripts a, b etc... label the internal legs. For example the skeleton bubble and skeleton triangle diagrams are

$$B^s = -i\mathcal{P}_2 - \Delta^{12} - \Delta^{21} \tag{4.10}$$

$$C^s = -i\mathcal{P}_3 + \sum_{\text{perms}} \left[-\Delta^{ab} \mathcal{Q}^{ac} + \frac{i}{2} \Delta^{ab} (\Delta^{ac} + \Delta^{cb}) \right], \tag{4.11}$$

respectively, where

$$\mathcal{P}_2 = \mathcal{P}^{12} + \mathcal{P}^{21}, \quad \mathcal{P}_3 = \mathcal{P}^{12} \mathcal{P}^{13} + \text{cycl} + (e \rightarrow -e). \tag{4.12}$$

In general the quantity \mathcal{P}_n is a sum of products of $n - 1$ different \mathcal{P}^{ab} and always contains the ultraviolet divergences of the associated diagram.

Each Feynman diagram, as well as the correspondent cut diagrams, can be decomposed in this way (see [35] for a general strategy). It is useful to show these results in a table. The case of the bubble diagram is depicted in Table 1, where each column below

Table 1 Threshold decomposition of the bubble diagram and its cut diagrams

Terms	Diag.			
\mathcal{P}_2	$-i$	i	0	0
Δ^{12}	-1	-1	2	0
Δ^{21}	-1	-1	0	2

Table 2 Threshold decomposition of the triangle diagram and its cut diagrams

Terms	Diag.							
\mathcal{P}_3	$-i$	i	0	0	0	0	0	0
$\Delta^{12}\mathcal{Q}^{13}$	-1	-1	2	0	0	0	0	0
$\Delta^{23}\mathcal{Q}^{21}$	-1	-1	0	2	0	0	0	0
$\Delta^{31}\mathcal{Q}^{32}$	-1	-1	0	0	2	0	0	0
$\Delta^{21}\mathcal{Q}^{23}$	-1	-1	0	0	0	2	0	0
$\Delta^{32}\mathcal{Q}^{31}$	-1	-1	0	0	0	0	2	0
$\Delta^{13}\mathcal{Q}^{12}$	-1	-1	0	0	0	0	0	2
$\Delta^{12}\Delta^{13}$	i	$-i$	$2i$	0	0	0	0	$-2i$
$\Delta^{23}\Delta^{21}$	i	$-i$	0	$2i$	0	$-2i$	0	0
$\Delta^{31}\Delta^{32}$	i	$-i$	0	0	$2i$	0	$-2i$	0
$\Delta^{21}\Delta^{31}$	i	$-i$	0	0	$2i$	$-2i$	0	0
$\Delta^{32}\Delta^{12}$	i	$-i$	$2i$	0	0	0	$-2i$	0
$\Delta^{13}\Delta^{23}$	i	$-i$	0	$2i$	0	0	0	$-2i$

the diagrams shows the coefficients of the decomposition that multiply the terms in the first column. The sum of the diagrams in the table vanishes due to the cutting identities. However, it is clear from the table that each row cancel independently. This new set of identities are called *spectral identities*. In general, the independent terms that compose a skeleton diagram can be separated by the number of Δ that they contain. Indeed, in more complicated diagrams different products of Δ 's appear. The threshold decomposition of a L -loop diagram with N legs $G_N^{(L)}$ and its cut diagrams $G_N^{(L,n)}$ can be schematically written as

$$G_N^{(L)} = -i\mathcal{P}_N^{(L)} + \sum_j \sum_m c_j \mathcal{O}_{j\Delta,m} \tag{4.13}$$

$$G_N^{(L,n)} = \sum_j \sum_m c_j^{(n)} \mathcal{O}_{j\Delta,m}, \tag{4.14}$$

where $\mathcal{O}_{j\Delta,m}$ are terms that contain j number of Δ 's, m labels the number of terms with the same j , n labels the number of cut diagrams and $c_j, c_j^{(n)}$ are constant coefficients. The spectral identities then read

$$c_j + c_j^* + \sum_n c_j^{(n)} = 0, \quad \forall j. \tag{4.15}$$

Moreover, a few other properties can be derived. First, by construction, the coefficients with even number of Δ 's are purely imaginary, while those with an odd number are real, i.e.

$$\text{Re}[c_{2r}] = \text{Re}[c_{2r}^{(m)}] = 0, \quad \text{Im}[c_{2r+1}] = \text{Im}[c_{2r+1}^{(m)}] = 0, \quad \forall r, m. \tag{4.16}$$





Then, from this property and the spectral identities it follows that

$$\sum_m c_{2r} = 0, \quad \forall r. \tag{4.17}$$

These properties can be better appreciated in the case of the triangle diagram shown in Table 2.

The spectral identities (4.15) and the threshold decomposition allow us to reduce the optical theorem to a set of algebraic equations. To summarize, the sum of the spectral identities of a given diagram gives the cutting identities for that diagram, while the sum of

Table 3 The threshold decomposition of a triangle diagram and its cut diagrams where particle 1 is purely virtual

Terms	Diag.			
				
\mathcal{P}_3	$-i$	i	0	0
$\Delta^{23} \mathcal{Q}^{21}$	-1	-1	2	0
$\Delta^{32} \mathcal{Q}^{31}$	-1	-1	0	2

the cutting identities implies the optical theorem. As mentioned in Sect. 3, the last implication is true only if some conditions are satisfied, for example, if all the degrees of freedom have positive residue. In the case of gauge theories it is necessary to include Faddeev-Popov ghosts in order to obtain unitarity in the Fock subspace where they are projected away, together with the longitudinal and temporal components of the gauge fields. On the other hand, theories with ghosts satisfy the cutting identities but violate the optical theorem. However, thanks to the threshold decomposition and the spectral identities we know how to modify the amplitudes and obtain a unitary theory in the subspace where some degrees of freedom are removed from the states that can be on shell. This is achieved by simply setting to zero all the Δ 's that contain at least one frequency associated to the particles that we want to remove from the spectrum, i.e. $\Delta^{ab} = 0$ if the leg a and/or b is purely virtual. Indeed it is easy to see from Tables 1 and 2 that if we cancel any row we would just remove one of the spectral identities, which hold independently.

For example, Table 2 reduces to Table 3 if particle 1 is purely virtual. Note that, after setting to zero every $\Delta^{1,a}$ also some of cut diagrams vanishes (those where particle-1 leg is cut). In the end the surviving rows still sum to zero. Therefore, when ghosts are present, if we remove the rows that are responsible for the violation of the optical theorem, we get a modified set of cutting identities which imply unitarity, once we project onto the subspace where the ghosts are not external states. In other words, the fakeon prescription allows us to consistently remove the ghosts (or any degrees of freedom we decide to) from the physical spectrum without relying on a symmetry and without changing the properties under renormalization (we never cancel the row that contains \mathcal{P}_n). Note that this removal is a true one and it is understood at any energy. This makes purely virtual particles radically different from resonances or unstable particles, for which always exists a Lorentz frame where they can be long lived and therefore detectable as every other particle.

In practice, the operation just described can be implemented by simply changing (4.8) into

$$\frac{1}{x + i\epsilon} \rightarrow \mathcal{P} \frac{1}{x} - i\tau\pi\delta(x), \tag{4.18}$$

where $\tau = 0$ if x contains at least one fakeon frequency and $\tau = 1$ otherwise. Then we can safely consider only diagrams where the purely virtual particles do not appear in the external legs. This is the fakeon prescription, which can be applied to both ghosts and standard particles and is consistent at any loop order [35, 37].

A given amplitude \mathcal{A}_f in theories with fakeons, which is obtained using (4.18), can be written as

$$\mathcal{A}_f = \mathcal{A} - \Delta_f \mathcal{A}, \tag{4.19}$$

where \mathcal{A} is a standard amplitude obtained using the Feynman prescription and $\Delta_f \mathcal{A}$ is a functions obtained by summing all the terms that need to be set to zero by the fakeon prescription. This subtraction has been explicitly calculated in [38] for the case of one-loop bubble, triangle and box diagrams, which are the most relevant in particle physics phenomenology. The modified functions exhibit new nonanalyticities and singularities that can be used to discriminate models with fakeons from models without. It is even possible to explore the possibility that fakeons might exist in general, regardless the problem of ghosts in quantum gravity. For example, in [39] a inert-doublet model where the new scalars are turned into fakeons has been considered as an extension of the standard model and compared with the standard inert-doublet model. The results show that in some portion of the parameters space the two models can be quite different. In particular, it was shown that the contribution to the decay of the Higgs boson into two photons differs in the two cases due to the modifications of the form (4.19). Another possibility is to consider whether one of the particles in the standard model can be purely virtual or not. The only particle that cannot be ruled out from being a fakeon by the present experimental data is the Higgs boson [40]. This possibility is more interesting because there is no freedom in the parameters space and therefore we can look for some energy domains where the differences between the two cases (standard Higgs versus purely virtual Higgs) are relevant. This will be published in a forthcoming paper [41].

The fakeon prescription opens the way for a new understanding of model building and particle physics. However, its main application remains quantum gravity. Although in principle quantum gravity with purely virtual particles can be discriminated from Stelle gravity or Einstein theory in terms of scattering amplitudes, it is quite challenging from the practical point of view, given the energies in play. Instead, a good arena where to test quantum gravity is inflationary cosmology. In the next section we show how to obtain an important prediction from quantum gravity in that context.

5 Quantum gravity with purely virtual particles and cosmology

In this section we write the action (2.1) in an alternative form, in order to make clear the presence of the ghost and show how to apply the fakeon prescription. Then we proceed with the derivation of observables in inflationary cosmology and derive a prediction for the tensor-to-scalar ratio.

The procedure to rewrite the action is a composition of a Weyl transformation and a metric field redefinition, together with the introduction of scalar and spin-2 auxiliary fields ϕ and $\chi_{\mu\nu}$, respectively. The details can be found in [42]. The result of this procedure is that the action (2.1) can be rewritten as

$$S_{\text{HD}}(g, \phi, \chi) = \tilde{S}_{\text{HE}}(g) + S_\chi(g, \chi) + S_\phi(g + \psi, \phi), \tag{5.1}$$

where

$$\tilde{S}_{\text{HE}}(g) = -\frac{1}{2\kappa^2} \int \sqrt{-g} (2\tilde{\Lambda}_C + \tilde{\zeta} R), \quad \frac{\tilde{\Lambda}_C}{\Lambda_C} = \left(1 + \frac{2}{3} \frac{(\alpha + 2\xi)\Lambda_C}{\zeta^2}\right), \quad \tilde{\zeta} = \zeta \frac{\tilde{\Lambda}_C}{\Lambda_C}, \tag{5.2}$$

$$S_\phi(g, \phi) = \frac{3\hat{\zeta}}{4} \int \sqrt{-g} \left[\nabla_\mu \phi \nabla^\mu \phi - \frac{m_\phi^2}{\kappa^2} (1 - e^{\kappa\phi})^2 \right], \quad \hat{\zeta} = \zeta \left(1 + \frac{4}{3} \frac{\xi \Lambda_C}{\zeta^2}\right), \tag{5.3}$$

and $S_\chi(g, \chi)$ is the covariantized Pauli–Fierz action with the “wrong” sign for the kinetic term plus nonminimal couplings and infinite interactions between g and χ , i.e.

$$S_\chi(g, \chi) = -\frac{\tilde{\zeta}}{\kappa^2} S_{\text{PF}}(g, \chi, m_\chi^2) - \frac{\tilde{\zeta}}{2\kappa^2} \int \sqrt{-g} R^{\mu\nu} (\chi^\sigma_\sigma \chi_{\mu\nu} - 2\chi_{\mu\rho} \chi^\rho_\nu) + S_\chi^{(>2)}(g, \chi), \tag{5.4}$$

$$S_{\text{PF}}(g, \chi, m_\chi^2) = \frac{1}{2} \int \sqrt{-g} \left[\nabla_\rho \chi_{\mu\nu} \nabla^\rho \chi^{\mu\nu} - \nabla_\rho \chi^\mu_\mu \nabla^\rho \chi^\nu_\nu + 2\nabla_\mu \chi^{\mu\nu} \nabla_\nu \chi^\rho_\rho - 2\nabla_\mu \chi^{\rho\nu} \nabla_\rho \chi^\mu_\nu - m_\chi^2 (\chi_{\mu\nu} \chi^{\mu\nu} - \chi^\mu_\mu \chi^\nu_\nu) \right] \tag{5.5}$$

where $S_\chi^{(>2)}(g, \chi)$ contains terms that are at least cubic in χ . They can be derived from the Einstein–Hilbert action as

$$S_\chi(g, \chi) = \tilde{S}_{\text{HE}}(g + \psi) - \tilde{S}_{\text{HE}}(g) - 2 \int \chi_{\mu\nu} \frac{\delta \tilde{S}_{\text{HE}}(g + \psi)}{\delta g_{\mu\nu}} + \frac{\tilde{\zeta}^2}{2\alpha\kappa^2} \int \sqrt{-g} (\chi_{\mu\nu} \chi^{\mu\nu} - \chi^\mu_\mu \chi^\nu_\nu) \Big|_{g \rightarrow g + \psi}, \tag{5.6}$$

and

$$\psi_{\mu\nu} = 2\chi_{\mu\nu} + \chi_{\mu\nu} \chi^\rho_\rho - 2\chi_{\mu\rho} \chi^\rho_\nu. \tag{5.7}$$

Finally, the squared masses of ϕ and χ read

$$m_\phi^2 = \frac{\zeta}{\xi}, \quad m_\chi^2 = \frac{\tilde{\zeta}}{\alpha}. \tag{5.8}$$

We highlight that the tilde and hat quantities are useful to deal with a nonvanishing cosmological constant and derive the correct change of variables that makes only m_χ modified by Λ_C . In the case where the cosmological constant is negligible, all the tilde and hat quantities are equal to the usual ones and it is convenient to choose

$$\psi_{\mu\nu} = 2\chi_{\mu\nu}. \tag{5.9}$$

For details, see [42].

Now that all the degrees of freedom are manifestly represented by fields, it is clear how to apply the fakeon prescription: in every diagram we set to zero all the Δ ’s that contain a frequency associated to χ . Then we restrict only to diagrams with external $h_{\mu\nu}$ and/or ϕ legs. In general we can couple any type of matter Φ to quantum gravity and the interactions in the variables (5.1) are obtained by means of the substitution

$$S_m(g, \Phi) \rightarrow S_m(g e^{\kappa\phi} + \psi e^{\kappa\phi}, \Phi), \tag{5.10}$$

where S_m is the action of matter.

We stress that the action (2.1) [or (5.1)] is an interim one. This means that it is not the action that we would obtain in the classical limit. Roughly speaking, the reason is that the fakeon is a purely quantum object and does not have a classical counterpart. Therefore, the fakeon prescription and all its physical effects cannot be derived from a classical action. We first need to start from a quantum theory with the desired properties and then perform the classical limit. We call the true classical action the *projected action* and the interim one the *unprojected action*. Then, one could wonder why we do not start directly from the projected action, where the fields χ are already projected away, and quantize it with standard techniques. There are various reasons for this. First, it is very hard to obtain it explicitly, since the fakeon prescription is derived from perturbative quantum field theory, the whole procedure is also perturbative and it is not known at the moment how to implement it at a nonperturbative level. Moreover, the projected nonperturbative action

is expected to be nonlocal and rather cumbersome. The good of using action (2.1) is that it is local and we can define Feynman rules as usual (although we eventually modified the results as explained in Sect. 3). We could think of this in a reversed way as follows. Assume that a nonperturbative projected action of (5.1) exists and it is known explicitly. This action, $S_{\text{np}}^{\text{proj}}(g, \phi)$, would be dependent only on the metric and the scalar field ϕ (and on other physical fields, if present). Moreover, it would be of a form where renormalizability is not manifest and it is complicated (or even impossible) to define Feynman rules. The fakeon prescription tells us that it is possible to introduce a new field χ and define a new action $S(g, \phi, \chi)$, which is physically equivalent to $S_{\text{np}}^{\text{proj}}$, provided that the fakeon prescription is used for χ . The new action is (5.1) and such that

$$S(g, \phi, \chi(g, \phi)) = S_{\text{np}}^{\text{proj}}(g, \phi), \tag{5.11}$$

where $\chi(g, \phi)$ is obtained from the nonperturbative version of the fakeon prescription. The advantage of $S(g, \phi, \chi)$ is that it is local. Moreover, being also equivalent to (2.1), it can be proved to be renormalizable, as explained in Sect. 2. At the present it is possible to obtain only the perturbative version of projected action by studying the classical limit of the fakeon prescription, which is enough to work out predictions. Some explicit examples for simpler models have been obtained in [43]. The projected action in the case of gravity has been derived at quadratic level in the perturbations around inflationary background at several orders in the slow-roll expansion [44, 45] as we review below.

5.1 Inflation

The theory of quantum gravity with purely virtual particles can be tested in the context of inflationary cosmology. In fact, the scalar degree of freedom introduced by the R^2 term can be viewed as the inflaton and used to explain the anisotropies of the cosmic microwave background [17]. Since our action contains also the C^2 term, we need to treat it in a fashion similar to what is explained in Sect. 3 in the case of scattering amplitudes. For this reason, it is necessary to understand the fakeon prescription in curved spacetime. As explained below, this task is simplified by the fact that in cosmology we do not need to go as far as computing loop corrections. Therefore, we can work with the classical limit of the fakeon prescription [44].

It is more convenient to use the action

$$S(g, \phi) = - \frac{M_{\text{pl}}^2}{16\pi^2} \int \sqrt{-g} \left(R + \frac{1}{2m_\chi^2} C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} \right) + \frac{1}{2} \int \sqrt{-g} [\nabla_\mu \phi \nabla^\mu \phi - 2V(\phi)], \tag{5.12}$$

$$V(\phi) = \frac{m_\phi^2}{2\hat{\kappa}^2} (1 - e^{\hat{\kappa}\phi})^2, \quad \hat{\kappa} = M_{\text{pl}}^{-1} \sqrt{16\pi/3}, \tag{5.13}$$

where we have explicitly introduced the Planck mass $M_{\text{pl}}^2 = 8\pi^2 \zeta / \kappa^2$, the fakeon mass $m_\chi^2 = \zeta / \alpha$ and the inflaton mass $m_\phi^2 = \zeta / \xi$, and we have canonically normalized the field ϕ . Moreover, we set the cosmological constant to zero since it is unimportant for our purposes. Alternatively, it is possible to use directly the action (2.1). Here we consider only the action (5.12). The two possibilities are physically equivalent and they are connected by perturbative redefinitions of the quantities [44].

Before to proceed we fix some notation. We expand the action (5.12) around the Friedmann-Lemaître-Robertson-Walker (FLRW) background

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \quad \bar{g}_{\mu\nu} = \text{diag}(1, -a^2, -a^2, -a^2) \tag{5.14}$$

up to the quadratic order in the perturbations $\delta g_{\mu\nu}$, where $a(t)$ is the scale factor. We anticipate to the reader that in the end only the tensor and scalar perturbations propagate and no vector or additional tensors/scalars are present due to the fakeon projection. Therefore, instead of going through the details for the two cases separately, we show the general procedure which is valid for both (specifying the differences where necessary) and then present the power spectra and the spectral indices.

We work in space Fourier transform and label the space momentum as \mathbf{k} and its modulus as $|\mathbf{k}| = k$. We define the slow-roll parameter

$$\varepsilon \equiv - \frac{\dot{H}}{H^2}, \quad H = \frac{\dot{a}}{a} \tag{5.15}$$

and express every quantities as power series in ε , up to overall, non-polynomial factors. In particular, it is easy to show that

$$\frac{d^n \varepsilon}{dt^n} = H^n \mathcal{O}(\varepsilon^{\frac{n+2}{2}}) \tag{5.16}$$

so other slow-roll parameters that are typically defined in the literature, such as $\eta \equiv 2\varepsilon - \frac{\dot{\varepsilon}}{2H\varepsilon}$, are also series in ε . In general, inflationary models contain one independent slow-roll parameter for each field that participates to inflation. It is also useful to derive the expansions

$$H = \frac{m_\phi}{2} \left(1 - \frac{\sqrt{3\varepsilon}}{2} + \frac{7\varepsilon}{12} + \mathcal{O}(\varepsilon^{3/2}) \right),$$

$$v \equiv -aH\tau = 1 + \varepsilon + \mathcal{O}(\varepsilon^{3/2}),$$

where τ is the conformal time

$$\tau = - \int_t^{+\infty} \frac{dt'}{a(t')}, \tag{5.17}$$

with the initial condition chosen to have $\tau = -1/(aH)$ in the de Sitter limit $\varepsilon \rightarrow 0$. Moreover, it is convenient to define

$$\lambda \equiv \frac{\hat{k}\dot{\phi}}{2H} = \sqrt{-\frac{\dot{H}}{3H^2}} = \sqrt{\frac{\varepsilon}{3}}, \tag{5.18}$$

since every quantity is an expansion in $\sqrt{\varepsilon}$ (and eventually also in k). The parameter λ is small during inflation and can be viewed as a ‘‘coupling constants’’ from which the whole process of cosmic inflation is interpreted as a sort of ‘‘renormalization-group flow’’ [46], in analogy with particle physics. This is a mathematical correspondence between the quantities in inflation and in perturbative quantum field theory and it is useful to systematize the computations by exporting the techniques of the latter to the former. This idea is generalized for several single-field inflationary models [47], where potentials can be classified and even ruled out. The case of double field inflation is studied in [48]. For future use we write the background field equations in terms of λ . The Friedmann equations are

$$\dot{H} = -\frac{3\hat{k}^2}{4}\dot{\phi}^2, \quad H^2 = \frac{\hat{k}^2}{4}(\dot{\phi}^2 + 2V), \quad \ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0. \tag{5.19}$$

Then, by combining Eq. (5.19), it is easy to show that the parameter λ satisfies

$$\frac{\dot{\lambda}}{H} = -3\lambda(1 - \lambda^2) - \frac{\hat{k}}{2H^2} \frac{dV}{d\phi}. \tag{5.20}$$

Finally, we denote with $u_{\mathbf{k}}(t)$ the space Fourier transform of a general perturbation and the power spectrum of \mathcal{P}_u is defined as

$$\langle u_{\mathbf{k}}(\tau)u_{\mathbf{k}'}(\tau) \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') \frac{2\pi^2}{k^3} \mathcal{P}_u, \quad \mathcal{P}_u = \frac{k^3}{2\pi^2} |u_{\mathbf{k}}|^2. \tag{5.21}$$

Moreover, replacing $|\tau|$ by $1/k_*$, where k_* is a reference scale, we write

$$\ln \mathcal{P}_u(k) = \ln A_u + n_u \ln \frac{k}{k_*}, \tag{5.22}$$

where A_u and n_u are called amplitude and spectral index, respectively. From now on we omit the subscript \mathbf{k} in the perturbations and simply write $u(t)$. Terms like $u\dot{u}$ are understood as either $u_{\mathbf{k}}\dot{u}_{-\mathbf{k}}$ or $u_{-\mathbf{k}}\dot{u}_{\mathbf{k}}$.

In the case of tensor perturbations, the quadratic action is of the form

$$S(u) = \frac{M_{\text{Pl}}^2}{8\pi} \int dt a(t)^3 \left[f(t)\dot{u}^2 - g(t)u^2 - \frac{1}{m_\chi^2} \ddot{u}^2 \right], \tag{5.23}$$

where f and g are time-dependent functions. Note that, because of the C^2 term, there is a higher-derivative term, as expected. At this level, it is convenient to remove it with a field redefinition and introduce explicitly an additional perturbation. In order to do this we consider the extended action

$$S'(u, B) = S(u) + \Delta S, \quad \Delta S = \frac{M_{\text{Pl}}^2}{8\pi m_\chi^2} \int a^3 (B - \ddot{u} - \tilde{f}\dot{u} - \tilde{g}u)^2, \tag{5.24}$$

where B is an auxiliary field and \tilde{f}, \tilde{g} are functions to be determined. The two actions coincide once we substitute B with the solution of its own equation of motion $B(u)$, i.e.

$$S'(u, B(u)) = S(u). \tag{5.25}$$

Finally, we perform the field redefinitions

$$u = U + bV, \quad B = V + cU, \tag{5.26}$$

where b and c are other functions to be determined. In the case of scalars, there are no higher-derivative terms but some fields are not dynamical and can be removed algebraically by using their field equations.

After these procedures the action in both cases is reduced in the form

$$S'(U, V) = \frac{1}{2} \int dt Z(\dot{U}^2 - \omega^2 U^2 - \dot{V}^2 + \Omega^2 V^2 + 2\sigma UV), \tag{5.27}$$

where Z , ω , Ω and σ are time-dependent functions that are different for tensor and scalar perturbations. It is now explicit that the field V is problematic, due to the different sign of its kinetic term. Therefore, we need to quantize it as purely virtual and remove it from the spectrum. As mentioned above, we do not need to compute loops and the tree-level version of the fakeon prescription is enough. In order to obtain it we note that formula (4.18) with $\tau = 0$ can be written as

$$\frac{1}{2} \left(\frac{1}{x + i\epsilon} + \frac{1}{x - i\epsilon} \right) = \mathcal{P} \frac{1}{x} \tag{5.28}$$

and apply it to the fakeon perturbations by averaging their retarded and advanced Green functions. The result of this operation is called *fakeon Green function*. More explicitly, the equation of motion for V is

$$\Sigma V \equiv \left(\frac{d^2}{dt^2} + \frac{\dot{Z}}{Z} \frac{d}{dt} + \Omega^2 \right) V = -\sigma U \tag{5.29}$$

and the solution obtained with the fakeon Green function is

$$V(t) = (G_f * F)(t), \quad F \equiv -\sigma U, \tag{5.30}$$

where G_f is the fakeon Green function of the operator Σ and “*” is the convolution. The result for G_f in de Sitter space is [44]

$$G_f(t, t') = \frac{i\pi \operatorname{sgn}(t - t') e^{-3H(t-t')/2}}{4H \sinh(n_\chi \pi)} \left[J_{in_\chi}(\check{k}) J_{-in_\chi}(\check{k}') - J_{in_\chi}(\check{k}') J_{-in_\chi}(\check{k}) \right], \tag{5.31}$$

where J_n denotes the Bessel function of the first kind and

$$n_\chi = \sqrt{\frac{m_\chi^2}{H^2} - \frac{1}{4}}, \quad \check{k} = \frac{k}{a(t)H}, \quad \check{k}' = \frac{k}{a(t')H}. \tag{5.32}$$

In this way we can write V as a function of U

$$V(U) = -G_f * (\sigma U) \tag{5.33}$$

and plug it back in the action (5.27), obtaining the projected action

$$S^{\text{proj}}(U) = S'(U, V(U)). \tag{5.34}$$

This is the classical version of fakeon prescription. It is worth note a few properties. First, the function σ is of order λ^2 [44] and, by (5.33), so is $V(U)$, and at the leading order in the slow-roll expansion the fakeon prescription gives $V = 0$. Therefore, the nonlocal term $\sigma U V(U)$ in (5.34) is $\mathcal{O}(\lambda^4)$, which means that the projected action is unaffected by $V(U)$ up to the order λ^3 included. This simplifies the computations, which can be pushed to higher-order without including the nonlocal contributions [45]. However, the change of variables (5.26) tells us that we need $V(U)$ for the power spectra, since the true physical variable is $u(U, V)$. Moreover, the power spectra are computed in the so-called *superhorizon limit* $|k\tau| \rightarrow 0$, so having $V(U)$ in that limit is enough. This further simplifies its derivation. The function $V(U)$ up to the order λ^3 is of the form

$$V(U) = \lambda^2 (v_1 + v_2 \lambda) U + v_3 \lambda^3 \dot{U} + \mathcal{O}(\lambda^4), \tag{5.35}$$

where the coefficients v_i depend on m_χ and m_ϕ and can be found in [45].⁵

5.2 Power spectra

At this stage we need only to compute the solution to the equation of motion for U and then derive the power spectrum for u . In order to do it, we first change variable to a rescaled conformal time $\eta \equiv -\bar{k}\tau$, where $\bar{k} = k(1 + \mathcal{O}(\lambda))$ is due to the fakeon prescription, so $\bar{k} = k$ if the C^2 term is absent. Then we perform an additional field redefinition to put the projected action in the Mukhanov-Sasaki form

$$S_w^{\text{proj}}(w) = \frac{1}{2} \int d\eta \left[\left(\frac{dw}{d\eta} \right)^2 - w^2 + \left(\frac{2 + \sigma_w}{\eta^2} \right) w^2 \right], \tag{5.36}$$

where w is the new variable, the $'$ denotes the derivative with respect to η and $\sigma_w = \mathcal{O}(\lambda)$ is a power series in λ which encodes the deviations from scale-invariant power spectrum. The solution of the associated equation of motion can be derived by imposing the usual Bunch-Davies condition for the field w , which in these variables reads

$$w(\eta) \simeq \frac{e^{i\eta}}{\sqrt{2}}, \quad \text{for } \eta \rightarrow \infty. \tag{5.37}$$

⁵ Note that in [45] the quantity (5.18) is labelled as α , while here we have used the letter λ to avoid confusion with the parameter in front of the C^2 term.

The solution is expanded in powers of λ as

$$w(\eta) = w_0(\eta) + \sum_{n=1}^{\infty} \lambda^n w_n(\eta), \tag{5.38}$$

where

$$w_0(\eta) = \frac{i(1 - i\eta)}{\eta\sqrt{2}} e^{i\eta} \tag{5.39}$$

and $w_{n>0}$ are other functions that depend on the type of perturbations (tensor or scalar), since also σ_u depends on that. In particular, σ_u is $\mathcal{O}(\lambda)$ and $\mathcal{O}(\lambda^2)$, for scalar and tensor perturbations, respectively.

Finally, the power spectrum is obtained from $|u|^2$ by tracing back all the field redefinitions and change of variables starting from the solution $w(\eta)$, so we have

$$u = U(w) + bV(U(w)), \quad w = w(\bar{k}|\tau|) \tag{5.40}$$

and the power spectrum is

$$\mathcal{P}_u(k) = \frac{k^3}{2\pi^2} |u(k)|^2, \tag{5.41}$$

It is important to highlight that in the superhorizon limit the dependence of \mathcal{P}_u on τ drops and that the one on k is all encoded in $\lambda_k \equiv \lambda(1/k)$, where λ is obtained from the background equations of motion (5.20) as a function of τ and then set $\tau = 1/k$. The results for the power spectra for tensor and scalar perturbations are [44, 45]

$$\mathcal{P}_t(k) = \frac{4m_\phi^2 \zeta}{\pi M_{\text{Pl}}^2} \left[1 + 3\zeta\lambda_k + \lambda_k^2 \left(6\zeta\gamma_M + \frac{47}{4}\zeta^2 + \frac{11}{8} \frac{m_\phi^2}{m_\chi^2} \zeta \right) + \mathcal{O}(\lambda_k^3) \right], \quad \zeta \equiv \frac{1}{\left(1 + \frac{m_\phi^2}{2m_\chi^2} \right)}, \tag{5.42}$$

$$\mathcal{P}_s(k) = \frac{m_\phi^2}{12\pi M_{\text{Pl}}^2 \lambda_k^2} \left[1 + \lambda_k(5 - 4\gamma_M) + \lambda_k^2 \left(4\gamma_M^2 - \frac{40}{3}\gamma_M + \frac{7}{3}\pi^2 - \frac{67}{12} - \frac{m_\phi^2}{2m_\chi^2} F\left(\frac{m_\phi^2}{m_\chi^2}\right) \right) + \mathcal{O}(\lambda_k^3) \right], \tag{5.43}$$

respectively, where $\gamma_M = \gamma_E + \ln 2$, γ_E being the Euler–Mascheroni constant and F is a function which can be obtained recursively as a series up to arbitrary orders [45]. The spectral indices are given by

$$n_u - \theta = \frac{d \ln \mathcal{P}_u(\lambda_k)}{d \ln k} = -\beta(\lambda_k) \frac{\partial \ln \mathcal{P}_u}{\partial \lambda_k}, \quad \beta(\lambda_k) = -\frac{d\lambda_k}{d \ln k}, \tag{5.44}$$

where $\theta = 0, 1$ for tensor and scalar perturbations, respectively. The *beta function* β is a power series in λ_k and it is obtained from

$$\frac{d\lambda}{d \ln|\tau|} = -\frac{1}{v} \dot{\lambda} \tag{5.45}$$

by using the expansions (5.16) and (5.17). We report its expression up to the order λ^4

$$\beta(\lambda) = -2\lambda^2 \left[1 + \frac{5}{6}\lambda + \frac{25}{9}\lambda^2 + \mathcal{O}(\lambda^3) \right] \tag{5.46}$$

but can be derived up to any order. In the renormalization-group-flow analogy mentioned above the quantity (5.46) plays the role of the beta function in quantum field theory and a necessary condition for the background to be asymptotically de Sitter in the infinite past is that (5.46) be at least quadratic and negative (in analogy with asymptotic freedom).

Finally, the spectral indices read

$$n_t = -6\zeta\lambda_k^2 + \mathcal{O}(\lambda_k^3), \quad n_s - 1 = -4\lambda_k \left[1 - \lambda_k \left(\frac{5}{3} - 2\gamma_M \right) \right] + \mathcal{O}(\lambda_k^3). \tag{5.47}$$

Note that in the case of scalar perturbations the contributions due to the fakeon (the m_χ dependence) start to appear at the order λ_k^2 for \mathcal{P}_s and at the order λ_k^3 for n_s (not shown here). This means that up to those orders the predictions of scalar perturbations are indistinguishable from those of the Starobinsky model. On the other hand, the predictions for the tensor perturbations get modified already at the leading order.

5.3 Consistency condition

The study of the fakeon prescription on curved spacetime leads to other nontrivial consequences. The whole procedure cannot be applied in the case of tachyons and specific conditions must be imposed. In flat spacetime, the so-called *no-tachyon conditions* typically constrains the parameters in the action. For example, in the case of the quantum gravity action (2.1) we must have $\alpha, \xi > 0$ in order to avoid tachyons, i.e. $m_\phi^2, m_\chi^2 > 0$. In FLRW spacetime we need to impose that the mass squared of the field V be positive, which can be read from (5.29) after the redefinition that cancels the term with a single derivative, i.e.

$$V \rightarrow \frac{V}{\sqrt{Z}}. \tag{5.48}$$

Then the effective mass is

$$m(t)^2 = \Omega^2 + \frac{\dot{Z}^2}{4Z^2} - \frac{\ddot{Z}}{2Z}, \tag{5.49}$$

where the functions Ω and Z are different for each perturbations. However, for tensor, vector and scalar perturbations it is of the form

$$m(t)^2 = m_\chi^2 - \frac{H^2}{4} + \frac{k^2}{a^2} + \mathcal{O}(\epsilon, k^4). \tag{5.50}$$

It is enough to derive the no-tachyon condition in de Sitter spacetime, since we are expanding perturbatively around it. Moreover, we can further simplify the expression (5.50) by taking the superhorizon limit $k/(aH) \rightarrow 0$. Then, recalling that $H(\epsilon = 0) = m_\phi/2$ and by imposing $m(t)^2|_{k/(aH) \rightarrow 0} > 0$ at $\epsilon = 0$ we get the consistency condition

$$m_\chi > \frac{m_\phi}{4}. \tag{5.51}$$

It is possible to impose a stronger bound by requiring that $m(t)^2 > 0$ for every k . However, the positivity of a time-dependent function is not reparametrization invariant. This can be shown by considering the most general transformation $t \rightarrow t'(t)$, $V(t) \rightarrow \hat{V}(t)$ that leaves the kinetic term invariant but changes the mass $m(t)^2$ into $M(t')^2$. Then, $m(t)^2 > 0$ does not necessarily imply $M(t')^2 > 0$. On the other hand, if m^2 is time independent and positive then the most general transformation that leaves M^2 t' -independent also leaves it positive [44]. Imposing the no-tachyon condition for every k gives the same bound (5.51) in the case of tensor and vector perturbations and a stronger one in the case of scalar ones. However, in what follows we consider only the no-tachyon condition in the superhorizon limit and conclude that the presence of negative mass squared in some time interval is not necessarily a lack of consistency.

The bound (5.51) it is important for phenomenological reasons, since, by combining it with experimental data, it narrows the allowed window for the tensor-to-scalar ratio to be less than an order of magnitude (see below).

5.4 Predictions

The best experimental data used to test inflationary models are those given by the Planck collaboration [49] combined with those of the BICEP/Keck collaboration [50]. In order to compare the data with the theoretical calculations we define the (dynamical) tensor-to-scalar ratio

$$r(k) = \frac{\mathcal{P}_t(k)}{\mathcal{P}_s(k)} \tag{5.52}$$

so that the usual tensor-to-scalar ratio is $r(k)$ at a reference scale k_* . From the results above we obtain

$$r(k) = 48\zeta\lambda_k^2 [1 + \lambda_k(3\zeta - 5 + 4\gamma_M)] + \mathcal{O}(\lambda_k^3). \tag{5.53}$$

The best data available at the moment give

$$n_s(k_*) = 0.9649 \pm 0.0042, \quad \ln(10^{10}\mathcal{P}_s(k_*)) = 3.044 \pm 0.014, \quad r(k_*) < 0.035, \tag{5.54}$$

where $k_* = 0.05 \text{ Mpc}^{-1}$. From the measurement of $n_s(k_*)$ we can extract the value of $\lambda_* \equiv \lambda_{k_*}$, while from $\mathcal{P}_s(k_*)$ we extract m_ϕ . The results are

$$\lambda_* = 0.0087 \pm 0.0010, \quad m_\phi = (2.99 \pm 0.36) \times 10^{13} \text{ GeV}. \tag{5.55}$$

Then, knowing the allowed values for λ_* and using the bound (5.51), we can plot the tensor-to-scalar ratio as a function of λ and restrict its values between the two curves in the cases $m_\chi = m_\phi/4$ and $m_\chi \rightarrow \infty$. The result up to two-sigma level is shown in the left panel of Fig. 2, while on the right panel we show it as a function of the number of e-folds N

$$N = \int_{t_i}^{t_f} H(t') dt', \tag{5.56}$$

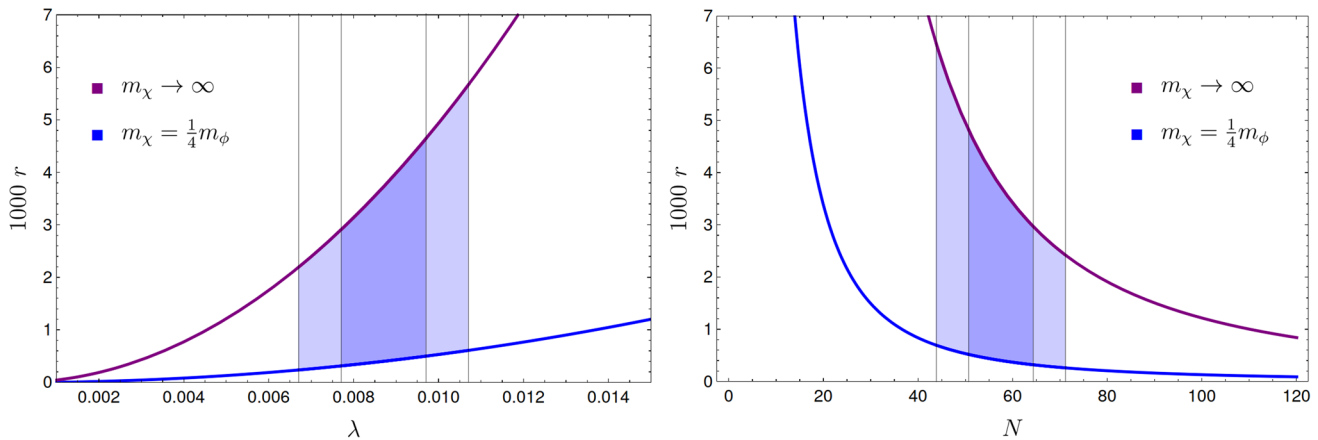


Fig. 2 On the left panel, the tensor-to-scalar ratio as a function of the parameter λ . On the right panel the same quantity is plotted as a function of the number of e-folds. The purple lines represent r in the Starobinsky model, while the blue lines represent r when the consistency condition is saturated. The darker regions indicate the one-sigma level, while the lighter ones the two-sigma level

which is often used in the literature to express the results. In general, it is more useful to use the variable λ , since the results can be expressed as power series of it (up to overall terms), while N it is not a perturbative quantity. This follows from the relation between N and λ

$$N = \int_{\lambda}^{\frac{1}{\sqrt{3}}} \frac{H(\lambda')}{\dot{\lambda}(\lambda')} d\lambda' = \int_{\lambda}^{\frac{1}{\sqrt{3}}} \frac{d\lambda'}{2\lambda'^2} \left[1 - \frac{5}{6}\lambda' + \mathcal{O}(\lambda'^0) \right] = \frac{1}{2\lambda} + \frac{5}{12} \ln \lambda + \mathcal{O}(\lambda^0), \tag{5.57}$$

From (5.57) we see that the $N(\lambda)$ is not a power series. The plots in Fig. 2 show that in the theory of quantum gravity with purely virtual particles, where the field ϕ plays the role of the inflaton, the tensor-to-scalar ratio is confined in a window that is around an order of magnitude. For concreteness, if we take $N = 60$ the tensor-to-scalar ratio is

$$0.37 \lesssim 1000 r \lesssim 3.41. \tag{5.58}$$

Future experiments, such as LiteBIRD [51], might be able to test this result, if the so-called B -modes are detected within the expected sensitivity ($\delta r < 0.001$ for LiteBIRD). We highlight that by measuring one new quantity, such as $r(k_*)$ or $n_t(k_*)$, the parameter m_χ would be fixed and every other potential prediction would be a precision test of the theory.

6 Conclusions

We have reviewed an approach to quantum gravity in the framework of quantum field theory. The theory is built by requiring the same guiding principles that have led to the standard model, i.e. locality, renormalizability and unitarity, and by being as conservative as possible in accommodating those requirements. Renormalizability already singles out a unique action for quantum gravity, which, besides the massless graviton, contains a scalar that can be viewed as the inflaton, a cosmological constant and a massive spin-2 particle. The latter, if quantized by means of standard techniques, leads to a violation of unitarity. However, this degree of freedom is necessary to achieve renormalizability and cannot be removed without losing that property. In order to restore unitarity without renouncing to renormalizability we use a different quantization procedure for the massive spin-2 field, the fakeon prescription. Such procedure is very general and in principle can be applied to any degree of freedom. Effectively, the fakeon prescription amounts to compute every amplitude using the Feynman prescription as usual and then subtract certain functions, whose role is to remove the on-shell parts of the degrees of freedom that we want to quantize in this way. The outcome is that those particles become purely virtual and cannot appear on shell. In the case of quantum gravity, this is crucial since the possibility of having an on-shell spin-2 ghost violates unitarity. Using the fakeon prescription we can remove the ghost from the spectrum of the theory without losing it from the possible virtual states. In this way we obtain a renormalizable and unitary theory of quantum gravity. We have shown that a good arena to test this theory is inflationary cosmology and shown that the consistency of the fakeon prescription in that context leads to a pretty sharp prediction for the tensor-to-scalar ratio, which could be tested in future experiments that measure the polarization of the cosmic microwave background.

When compared to other approaches to quantum gravity, the fakeon idea has several advantages. First it is perturbative and does not add more complications from the computational point of view. The effort in computing Feynman diagrams in the theory of quantum gravity with fakeons is comparable to that required for the standard model. Moreover, by reconciling renormalizability and unitarity in quantum gravity, the fakeon approach makes unnecessary to invoke nonperturbative approaches to renormalization, such as asymptotic safety. Furthermore, the whole procedure truly removes the ghost degrees of freedom from the theory, unlike

the approach [27], which proves unitarity in Stelle theory only in situations where the unstable ghost has decayed. In this regards, it is worth to highlight that ghosts can be removed by means of the fakeon prescription even if they are stable.

Finally, we mention some future directions in this research line. From the theoretical point of view it is important to understand some aspects of the theory of quantum gravity, for example its perturbative validity. In fact, the fakeon prescription applied to Stelle gravity introduces two power countings, so to speak. On the one hand the renormalization sector is perturbative up to arbitrary energies, since the power counting is that of Stelle gravity. On the other hand some quantities, such as absorptive parts of the amplitudes, obey the power counting of general relativity. An on-going work is to understand whether semi-nonperturbative techniques, combined with the fakeon prescription, can improve the nonrenormalizable behaviors of scattering amplitudes. On a more phenomenological side, it would be interesting to study the role of the spin-2 fakeon in post-inflationary eras to see how cosmology is impacted. Furthermore, the results of [38] need to be explored more deeply, to uncover other new effects in particle physics that could be directly related to the presence of fakeons.

Data Availability Statement No data associated in the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. D. Anselmi, M. Piva, A new formulation of Lee-Wick quantum field theory. *JHEP* **06**, 066 (2017). [arXiv:1703.04584](https://arxiv.org/abs/1703.04584)
2. D. Anselmi, On the quantum field theory of the gravitational interactions. *JHEP* **06**, 086 (2017). [arXiv:1704.07728](https://arxiv.org/abs/1704.07728)
3. G.'t Hooft, M.J.G. Veltman, One loop divergencies in the theory of gravitation. *Ann. Inst. H. Poincaré Phys. Theor. A* **20**, 69–94 (1974)
4. M.H. Goroff, A. Sagnotti, The ultraviolet behavior of Einstein gravity. *Nucl. Phys. B* **266**, 709–736 (1986)
5. E. Fermi, Tentativo di una teoria dell'emissione dei raggi beta. *Ric. Sci.* **4**, 491–495 (1933)
6. D. Anselmi, Infinite reduction of couplings in non-renormalizable quantum field theory. *JHEP* **08**, 029 (2005). [arXiv:hep-th/0503131](https://arxiv.org/abs/hep-th/0503131)
7. D. Anselmi, Absence of higher derivatives in the renormalization of propagators in quantum field theories with infinitely many couplings. *Class. Quant. Grav.* **20**, 2355–2378 (2003). [arXiv:hep-th/0212013](https://arxiv.org/abs/hep-th/0212013)
8. D. Anselmi, Properties of the classical action of quantum gravity. *JHEP* **05**, 028 (2013). [arXiv:1302.7100](https://arxiv.org/abs/1302.7100)
9. S. Weinberg, *Ultraviolet divergences in quantum theories of gravitation*, pp. 790–831, (1980)
10. O. Lauscher, M. Reuter, Ultraviolet fixed point and generalized flow equation of quantum gravity. *Phys. Rev. D* **65**, 025013 (2002). [arXiv:hep-th/0108040](https://arxiv.org/abs/hep-th/0108040)
11. K.S. Stelle, Renormalization of higher derivative quantum gravity. *Phys. Rev. D* **16**, 953–969 (1977)
12. J. Julve, M. Tonin, Quantum gravity with higher derivative terms. *Nuovo Cim. B* **46**, 137–152 (1978)
13. E.S. Fradkin, A.A. Tseytlin, Renormalizable asymptotically free quantum theory of gravity. *Nucl. Phys. B* **201**, 469–491 (1982)
14. S. Kawasaki, T. Kimura, A possible mechanism of ghost confinement in a renormalizable quantum gravity. *Prog. Theor. Phys.* **65**, 1767 (1981)
15. E.T. Tomboulis, Unitarity in higher derivative quantum gravity. *Phys. Rev. Lett.* **52**, 1173 (1984)
16. I.G. Avramidi, A.O. Barvinsky, Asymptotic freedom in higher derivative quantum gravity. *Phys. Lett. B* **159**, 269–274 (1985)
17. A.A. Starobinsky, A new type of isotropic cosmological models without singularity. *Phys. Lett. B* **91**, 99–102 (1980)
18. T.D. Lee, G.C. Wick, Negative metric and the unitarity of the S matrix. *Nucl. Phys. B* **9**, 209–243 (1969)
19. T.D. Lee, G.C. Wick, Finite theory of quantum electrodynamics. *Phys. Rev. D* **2**, 1033–1048 (1970)
20. L. Modesto, I.L. Shapiro, Superrenormalizable quantum gravity with complex ghosts. *Phys. Lett. B* **755**, 279–284 (2016). [arXiv:1512.07600](https://arxiv.org/abs/1512.07600)
21. L. Modesto, Super-renormalizable or finite Lee-Wick quantum gravity. *Nucl. Phys. B* **909**, 584–606 (2016). [arXiv:1602.02421](https://arxiv.org/abs/1602.02421)
22. C.A. Nelson, E.C.G. Sudarshan, Quantum field theories with shadow states. i. soluble models. *Phys. Rev. D* **6**, 3658–3678 (1972)
23. E.C.G. Sudarshan, C.A. Nelson, Quantum field theories with shadow states. ii. low-energy pion-nucleon scattering. *Phys. Rev. D* **6**, 3678–3688 (1972)
24. C.A. Nelson, Physical unitarization of indefinite-metric theories by shadow state summation, bjoeken scaling and light quarks. *Lett. Nuovo Cim.* **4S2**, 913–918 (1972)
25. Y.V. Kuzmin, The convergent nonlocal gravitation. *Sov. J. Nucl. Phys.* **50**, 1011–1014 (1989). (In Russian)
26. L. Modesto, L. Rachwał, Nonlocal quantum gravity: a review. *Int. J. Mod. Phys. D* **26**(11), 1730020 (2017)
27. J.F. Donoghue, G. Menezes, Unitarity, stability and loops of unstable ghosts. *Phys. Rev. D* **100**(10), 105006 (2019). [arXiv:1908.02416](https://arxiv.org/abs/1908.02416)
28. M.J.G. Veltman, Unitarity and causality in a renormalizable field theory with unstable particles. *Physica* **29**, 186–207 (1963)
29. D. Anselmi, M. Piva, The ultraviolet behavior of quantum gravity. *JHEP* **05**, 027 (2018). [arXiv:1803.07777](https://arxiv.org/abs/1803.07777)
30. H. Kluberg-Stern, J.B. Zuber, Renormalization of nonabelian gauge theories in a background field gauge. 1. green functions. *Phys. Rev. D* **12**, 482–488 (1975)
31. H. Kluberg-Stern, J.B. Zuber, Renormalization of nonabelian gauge theories in a background field gauge. 2. Gauge invariant operators. *Phys. Rev. D* **12**, 3159–3180 (1975)
32. D. Anselmi, Background field method and the cohomology of renormalization. *Phys. Rev. D* **93**(6), 065034 (2016). [arXiv:1511.01244](https://arxiv.org/abs/1511.01244)
33. G.'t Hooft, An algorithm for the poles at dimension four in the dimensional regularization procedure. *Nucl. Phys. B* **62**, 444–460 (1973)
34. D. Anselmi, Aspects of perturbative unitarity. *Phys. Rev. D* **94**, 025028 (2016). [arXiv:1606.06348](https://arxiv.org/abs/1606.06348)
35. D. Anselmi, *Diagrammar of physical and fake particles and spectral optical theorem*, [arXiv:2109.06889](https://arxiv.org/abs/2109.06889)
36. D. Anselmi, The quest for purely virtual quanta: fakeons versus Feynman-Wheeler particles. *JHEP* **03**, 142 (2020). [arXiv:2001.01942](https://arxiv.org/abs/2001.01942)
37. D. Anselmi, Fakeons and Lee-Wick models. *JHEP* **02**, 141 (2018). [arXiv:1801.00915](https://arxiv.org/abs/1801.00915)
38. A. Melis and M. Piva, *One-Loop Integrals for Purely Virtual Particles*, [arXiv:2209.05547](https://arxiv.org/abs/2209.05547)
39. D. Anselmi, K. Kannike, C. Marzo, L. Marzola, A. Melis, K. Müürsepp, M. Piva, and M. Raidal, *Phenomenology of a Fake Inert Doublet Model*, [arXiv:2104.02071](https://arxiv.org/abs/2104.02071)
40. D. Anselmi, On the nature of the Higgs boson. *Mod. Phys. Lett. A* **34**(16), 1950123 (2019). [arXiv:1811.02600](https://arxiv.org/abs/1811.02600)

41. A. Melis and M. Piva *In preparation*
42. D. Anselmi, M. Piva, Quantum gravity fakeons and microcausality. *JHEP* **11**, 021 (2018). [arXiv:1806.03605](#)
43. D. Anselmi, Fakeons and the classicization of quantum gravity: the FLRW metric. *JHEP* **04**, 061 (2019). [arXiv:1901.09273](#)
44. D. Anselmi, E. Bianchi, M. Piva, Predictions of quantum gravity in inflationary cosmology: effects of the Weyl-squared term. *JHEP* **07**, 211 (2020). [arXiv:2005.10293](#)
45. D. Anselmi, High-order corrections to inflationary perturbation spectra in quantum gravity. *JCAP* **02**, 029 (2021). [arXiv:2010.04739](#)
46. D. Anselmi, Cosmic inflation as a renormalization-group flow: the running of power spectra in quantum gravity. *JCAP* **01**, 048 (2021). [arXiv:2007.15023](#)
47. D. Anselmi, F. Fruzza, M. Piva, Renormalization-group techniques for single-field inflation in primordial cosmology and quantum gravity. *Class. Quant. Grav.* **38**(22), 225011 (2021). [arXiv:2103.01653](#)
48. D. Anselmi, Perturbation spectra and renormalization-group techniques in double-field inflation and quantum gravity cosmology. *JCAP* **07**, 037 (2021). [arXiv:2105.05864](#)
49. Y. Planck Collaboration, Akrami, et al., Planck 2018 results: X-Constraints on inflation. *Astron. Astrophys.* **641**, 10 (2020). [arXiv:1807.06211](#)
50. BICEP, Keck Collaboration, P. A. R. Ade, et al., Improved constraints on primordial gravitational waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 observing season. *Phys. Rev. Lett.* **127**(15), 151301 (2021). [arXiv:2110.00483](#)
51. M. Hazumi et al., LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection. *J. Low Temp. Phys.* **194**(5–6), 443–452 (2019)