On the Phenomenology of Evolving Black Holes, Dark Energy and Bouncing Cosmology

Cyril Renevey

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : PHYS - Physique
Spécialité : Physique Théorique
Unité de recherche : Laboratoire de Physique Subatomique et Cosmologie

Sur la phénoménologie des trous noirs en évolution, l'énergie noire et les cosmologies en rebond

On the Phenomenology of Evolving Black Holes, Dark Energy and Bouncing Cosmology

Présentée par :

Cyril RENEVEY

Direction de thèse :

Aurélien BARRAU — Directeur de thèse
PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes

Rapporteurs :

CAMILLE BONVIN
ASSOCIATE PROFESSOR, UNIVERSITE DE GENEVE
ALEXANDRE ARBEY
MAITRE DE CONFERENCES HDR, UNIVERSITE LYON 1 - CLAUDE BERNARD

Thèse soutenue publiquement le **29 septembre 2023**, devant le jury composé de :

CAMILLE BONVIN ASSOCIATE PROFESSOR, UNIVERSITE DE GENEVE	Rapporteure
ALEXANDRE ARBEY MAITRE DE CONFERENCES HDR, UNIVERSITE LYON 1 - CLAUDE BERNARD	Rapporteur
LUCAS LOMBRISER ASSISTANT PROFESSOR, UNIVERSITE DE GENEVE	Examinateur
RICHARD TAILLET PROFESSEUR DES UNIVERSITES, UNIVERSITE DE CHAMBERY	Examinateur
GILLES HENRI PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES	Président

Pour mon Grand-père, qui a été une immense source d'inspiration et un modèle d'admiration. Merci Pipix pour ces moments passés en ta companie, pour l'intéret que tu as porté à mes études et pour l'avenir que tu m'as offert.

Acknowledgments

Cela fait maintenant plus de 10 ans depuis que j'ai suivi mon premier cours de physique à l'EPFL. Tout au long de ce parcours, j'ai rencontré de nombreuses personnes exceptionnelles qui m'ont aidé dans mes études et mon bien-être, la plupart d'entre elles devenant mes amis les plus proches. J'ai eu la chance de passer de longues heures à travailler sur nos séries d'exercices et de terminer nos journées à Satellite avec Vincent, Clément et Yves. J'ai passé des moments inoubliables pendant ces premières années exigeantes à Lausanne et nous n'avons jamais oublié de profiter de la vie en cours de route. Mon échange à Stockholm a été un tournant dans ma vie, J'ai pu construire un groupe d'amis, Antoine, Alexis, Etienne, Gaétan et aussi Mélanie, Delphine, Soline, avec qui j'ai passé des moments inoubliables, que ce soit en Suède, en Suisse, en France et même à Madagascar!

It was also during this incredible year that I met the most important person with whom I will spend the rest of my life. Thank you Pie for your love and support, for these countless hours in bars and restaurants discussing about everything and nothing, for these cosy Netflix evening together in Grenoble and these unexpected hikes with Toffee boy. Thank you for following me in this little town that is now our home. I am having the greatest time of my life and it is all thanks to you! I can't wait for the next chapters of our book to be written.

Je suis également très chanceux de pouvoir compter sur ma famille pour embellir ces années. Merci à toi Maman pour toutes ces soirées passées ensemble et pour les nombreuses visites à Grenoble. Merci pour ta confiance et ton amour, d'être toujours là pour moi et pour être tout simplement la meilleure maman du monde. un grand merci également à Hervé pour les petits coup de mains, les excellentes bouteilles de vin et les visites. J'ai aussi eu la chance d'avoir des visites de ma soeurette préférée, on a passé de super moment à Grenoble et aussi lors de mes passages au pays! Merci Marraine et Henri pour vos encouragements et félicitations, merci également pour les repas en famille dans le jardin et ces délicieux knöpfli! Merci Papa pour ton support tout au long de mes études, ces beaux moments passés ensemble à Bourguillon et ailleurs, tes encouragements et ta fierté. J'espère que l'on pourra reconstruire une relation père-fils pleine de partage et bonheur. And of course, a special thanks to you Tage and Helena for the great moments that we spent together driving around France, from Nice to Champagne, Stockholm and Finland. Thank you for making me feel welcome and comfortable in the Erikson-Pohjaranta family!

Bien sur, rien n'aurait été possible sans une équipe de choc au LPSC! Merci à toi Aurélien pour la confiance que tu m'as donné et pour avoir été un directeur de thèse extra, toujours à mon écoute et prêt à m'aider en cas de besoin. J'ai pu m'épanouir pleinement grâce à toi, aux projets proposés et aux nombreuses discussions. Plus généralement, merci beaucoup beaucoup à toute l'équipe théorie du LPSC, Killian, Rémy, Fernando, Mohamed, Tim, Christopher, Sabine et Jérémie, entre autre, pour tous les moments de convivialité passés ensemble. J'ai passé trois années inoubliables grâce à vous tous!

Contents

Introduction

The theory of General Relativity and the ΛCDM model are widely regarded as the most successful frameworks for understanding gravitational interaction and cosmological observations. However, their development was not without challenges. Albert Einstein's groundbreaking contributions, starting with the discovery of Special Relativity in 1905 and later General Relativity in 1915, revolutionized our understanding of space and time. Einstein showed that space and time are not separate entities but form a dynamic entity called space-time, influenced by the matter within it. Despite its significant advancements, GR initially faced skepticism, even after resolving the problem of Mercury's perihelion advance and confirming the deflection of light during a solar eclipse in 1919, due to uncertainties surrounding Arthur S. Eddington's observations.

One remarkable consequence of General Relativity is its first solution to the Einstein field equations, which suggested the existence of black holes. Karl Schwarzschild discovered this solution, now bearing his name, in 1916, just a year after the theory's publication. The Schwarzschild metric describes a non-rotating, uncharged, spherically symmetric black hole and introduced the concept of an event horizon—the boundary beyond which no information can escape from the black hole's interior. The theoretical discovery of black holes sparked extensive investigations into their properties and behaviours. Observational evidence has since confirmed their existence, solidifying black holes' status as captivating celestial entities that continue to intrigue scientists and inspire further inquiry.

Even though they appear as very simple objects in the Schwarzschild approach, black holes have not yet revealed all their secrets. Their complexity was first put forward when Hawking radiation was discovered in a first attempt to mix quantum field theory and gravity in a semi-classical description. We had long known that black holes could grow by accreting surrounding matter and together with their newly predicted ability to shrink, it makes them intriguing dynamical objects. Unfortunately, the Schwarzschild solution assumes a static horizon and faces challenges in rapidly evolving conditions, when surrounded by a very dense fluid for example. In this thesis, we will examine the evolving behaviour of the event horizon of a Schwarzschild black hole bathed in a test fluid at thermal equilibrium. The limitations of such a model will be investigated as well.

One could try to relax the static assumption, at the cost of more complex calculations. Two approaches for that purpose have been attempted. One can make an educated guess on the form of the metric and later match the corresponding energy-momentum tensor through the Einstein field equations, an approach followed to design the McVittie and the generalized McVittie metrics. The latter solution being closer to realistic phenomena as it describes black holes surrounded by an imperfect fluid. However, in the context of cosmology black holes are expected to be embedded in a perfect fluid instead. One could then rather keep the most general isotropic metric and try to solve the Einstein field equations for such a

perfect fluid. In this work, we will quickly describe the generalized McVittie solution and we will study in more depth the second approach, initiated by Nielsen and Visser, to describe dynamical black holes.

While the first direct detection of gravitational waves from a black hole binary has been made only recently by LIGO, their existence was already confirmed in 1974 using the decaying orbit of the Hulse-Taylor binary system of pulsars. As opposed to pulsars, black holes' mass is expected to decrease through Hawking radiation and the resulting effect is an expansion of the orbit of a binary system. In a recent paper included in this thesis, we studied in depth the competition between the loss of energy through gravitational waves emission and the loss of mass not only from Hawking radiation, but also from phantom dark energy, a cosmological fluid violating the equivalence principle.

The study of cosmology in the framework of General Relativity also began shortly after the discovery of the Einstein field equations. Friedmann developed in the 1920s the first equations for a homogeneous and isotropic Universe. Lemaître used these equations to propose an expanding Universe, later confirmed by Hubble's observation in 1929. Lemaître also predicted that the universe was very hot and dense at early times and the discovery of the big bang nucleosynthesis brought the big bang model on the front scene. Further observations, such as the cosmic microwave background in the sixties, reinforced the status of General Relativity and the Big Bang model. However, the cosmological journey faced its first hurdles in the mid-1930s with the discovery of dark matter by Zwicky, later confirmed by Vera Rubin in 1970. The discovery of accelerated expansion in 1998 led to another revolution in the cosmology community and the ΛCDM model was born. Combining data from the cosmic microwave background and other independent cosmological surveys, scientists determined that the known Universe consists of approximately 68.9% dark energy and 31.1% matter, with dark matter being the primary component. This suggests that we have only understood about 5% of the observable Universe, despite the current model's success.

The best model that we have so far in order to describe dark energy is the cosmological constant. However this model is not without theoretical flaws. For a start, the cosmological constant is very small and an extreme fine-tuning is needed. Moreover, while a constant contribution to the Einstein field equations fits the experimental observations relatively well, the constraints on the dynamics of dark energy remain loose. These considerations led cosmologists to study dynamical approaches to dark energy and a natural candidate for that purpose is a scalar field. Such a choice is motivated by several factors. It naturally satisfies Lorentz invariance and the isotropic principle used in cosmology. In addition, a scalar field naturally emerges in theories with extra dimensions such as the Dvali Gabadadze Porrati (DGP) model or string theory, making scalar-tensor theories compelling extensions to General Relativity. Although string theory can be described as a scalar-tensor theory in the low energy effective approach, it has proven so far impossible to construct such an effective theory that could sustain a long period of de Sitter expansion. This led to the postulation of the de Sitter conjecture by Vafa, which restricts the form of the scalar potential in order to be consistent with string theory. While the conjecture is already in slight conflict with current constraints on the dynamics of dark energy, future surveys, such as the Euclid satellite or the Vera Rubin observatory, are expected to exacerbate the tension. A detailed analysis of the expected constraints on the de Sitter conjecture from the predicted precision of future surveys is conducted in this work.

Scalar-tensor theories derived from string theory are not the only candidates that are explored. The cubic Galileon theory for example, the effective theory describing the DGP model or ghost-free massive gravity in the decoupling limit, is a well studied candidate. It belongs to the larger set of all non-minimally coupled scalar-tensor theories gathered under the general Horndeski theory. An interesting property of the cubic Galileon theory, shared with other viable Horndeski models, is its ability to hide the effect of the scalar field on astrophysical scales through a screening mechanism, making the theories exhibiting such a mechanism consistent with solar system tests of General Relativity. However, screening mechanisms are non-linear behaviours of a theory, which makes them difficult to deal with in a post-Newtonian approach, often used for astrophysical tests or gravitational waves emissions from a binary system. The recent development of an effective approach to screening mechanisms led by Gabadadze, McManus and others, allowed the derivation of a post-Newtonian expansion for screened Horndeski theories. These results in turn made it possible to describe the waveform generated by a binary system of compact objects. The calculations of the waveform, using Will's post-Newtonian approach to gravitational waves, in the case of Vainshtein and Chameleon screening are exposed in this thesis. The now numerous detections of binary mergers by the LIGO/VIRGO collaboration, together with the calculated waveform, open a new perspective to test screened Horndeski theories.

Beyond the lack of understanding of dark matter and dark energy, the ΛCDM model faces other conceptual issues. For one thing, the current concordance model of cosmology is plagued with a big bang singularity at the origin of the universe, as proven by Hawking and Penrose. Moreover, the universe appears to be extremely flat and regions of space, that have never been in causal contact, happened to have the same temperature. An inflationary period of exponential growth at the beginning of the universe naturally resolves both problems and is so far supported by experimental evidence in the cosmic microwave background. Even though inflation is very effective at flattening the spatial sections, it does not tell if the universe is closed or infinite. In the former case of a positively curved universe, one can actually show that the big bang singularity is naturally replaced by a bounce, without the need of quantum gravity or exotic physics. Under this reasonable postulate, one could hope that the prediction of inflation for the cosmic microwave background are altered by the curvature bounce. In this thesis we will study in details the pre-bounce universe and the bounce behaviour as well as the scalar and tensor primordial power spectra resulting from this bouncing scenario.

Of course, there is no guarantee that our universe is closed and other solutions to the big bang singularity should be explored. At very early times when the energy density of the universe approaches the Planck scale, quantum gravity effects are expected to kick in. The canonical approach to quantizing gravity, loop quantum gravity, predicts that there exists a smallest area, resulting in the cosmological consequence of modified Friedmann equations at early times. Loop quantum cosmology, the theory described by these modified equations, also resolves the big bang singularity with a bounce and is also so far consistent with the cosmic microwave background power spectrum. However, as is common in quantum mechanics, several ambiguities might alter the quantum corrections to the Friedmann equations as predicted by the theory. Therefore, it is important to take these ambiguities into consideration by studying more general modified Friedmann equations. In a series of papers that are included to this thesis, we scrutinized the consistency of the predictions of loop quantum cosmology, when taking different form of quantum corrections to the Friedmann

equations.

This thesis is organized as follow. In chapter 1, we introduce the fundamental principles of General Relativity as well as the most famous solutions to the theory, namely the Friedmann-Lemaître-Robertson-Walker (FLRW) metric and the Schwarzschild metric. The limitations of the Schwarzschild solution for black holes immersed in a thermal bath are exposed and potential extensions to dynamical black holes using the generalized McVittie metric or Painlevé-Gullstrand coordinates are given. It is also in Chap. 1 that the competition between the effect of the loss of mass and loss of energy due to radiation for binary black holes is studied. The main successes on the current ΛCDM model to cosmology are described at the beginning of chapter 2. In this chapter, we also introduce the concept of Quintessence, before deriving the predicted constraints from Euclid and the Vera Rubin observatory on the de Sitter conjecture. Later on in the same chapter, we derive the gravitational waveform sourced by a binary system for screened Horndeski models, after a short introduction on Horndeski theories and screening mechanism. The third and last chapter, is dedicated to early universe cosmology and bouncing scenarii. We first show how one can generate an inflationary period using a single scalar field and we describe cosmological perturbation theory. We then discuss the curvature bounce and its potential imprints in the cosmic microwave background. Later on in Chap. 3, we give a brief introduction to loop quantum gravity and its cosmological consequences. The last chapter ends with the study of the effect of generalized holonomy corrections on the predictions of loop quantum cosmology. We finally end the thesis with a conclusive note.

Throughout the thesis, we adopt the signature $(-,+,+,+)$ for the metric. Greek indices run from 0 to 3 and Latin indices from 1 to 3. A comma indicates a partial derivative, $A_{,\mu} \equiv \partial_\mu A$, while a semicolon denotes a covariant derivative associated with the Levi-Civita connection, $A_{\mu;\nu} \equiv \nabla_\nu A_\mu$. Bold symbols $\mathbf{x}$ represent three-dimensional vectors. Symmetrization of indices follows the notation $T_{(\mu_1\ldots\mu_n)} \equiv \frac{1}{n!}\sum T_{\sigma_i(\mu_1\ldots\mu_n)}$, where the sum is made on the set of all permutations of $\{\mu_1,\ldots,\mu_n\}$ and similarly for anti-symmetrization. Units are such that $c = G = \hbar = 1$, unless otherwise stated.

Chapter 1

The theory of gravity

1.1 General Relativity

Since the discovery of special relativity in 1905 [1], Albert Einstein was puzzled by the inconsistency of Newton's theory of universal gravitation with his newly born relativistic mechanics. In particular, the gravitational force between two bodies was thought to act instantly, which was conflicting with the postulate of finite speed of "causality". Moreover, Einstein realized that there are no local experiments that can be done to differentiate between a freely falling observer subjected to gravity and an observer floating in empty space. This hypothesis is often called the Einstein equivalence principle (EEP). To formalize these two ideas, pseudo-Riemannian geometry comes in handy. While the Lorentzian metric $g^{\mu\nu}$ solves the problem of faster than light travel, the property of differentiable manifold to be locally flat satisfies the EEP.

Once we understood that spacetime should be described by a curved manifold, we can start to build physics on top of it. The first step is to describe the path $x^\mu(\lambda)$ of a free point-like object of mass m embedded in such a spacetime. Its action can be defined as

$$S[x] = -m \int \sqrt{-g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}} \, \mathrm{d}\lambda \,, \tag{1.1}$$

and using the principle of least action we find that free particles follow the geodesic equation

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu{}_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0, \tag{1.2}$$

where $\Gamma^\mu{}_{\alpha\beta}$ is the connection associated to the metric $g_{\mu\nu}$. In order for the connection to be uniquely defined, we are required to add two additional conditions: metric compatibility $\nabla_\alpha g_{\mu\nu} = 0$, the covariant derivative of the metric is zero and $\Gamma^\mu{}_{\alpha\beta} = \Gamma^\mu{}_{\beta\alpha}$, the connection is symmetric in its lower indices, or in other word, the manifold is torsion-free. These imply that we can write

$$\Gamma^\sigma{}_{\mu\nu} = \frac{1}{2} g^{\sigma\alpha} (\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\nu\alpha} - \partial_\alpha g_{\mu\nu}) \,, \tag{1.3}$$

which are commonly called the *Christoffel symbols*. If one compares the equation of motion (1.2) for a massive particle in curved spacetime to the flat case $\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} = 0$, one can deduce that to move from the former case to the latter, we can promote partial derivatives to covariant

ones, namely $\partial_\mu \to \nabla_\mu$. This idea is sometimes called the *covariance principle*, but it is in fact a direct consequence of the EEP.

Now that we learned gravity is described by curved spacetime and that we discovered the principle of covariance, how do we relate the curvature tensor $R^\sigma{}_{\rho\mu\nu}$ to matter, as it is done in Newton's theory, where gravity is related to the mass of the objects? Let us first remind ourselves what we mean by curvature. It is defined as the shift of a vector when parallel transported along a closed curve. In terms of the Christoffel symbols, the Riemann or curvature tensor can be written as

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma^\rho{}_{\nu\sigma} - \partial_\nu \Gamma^\rho{}_{\mu\sigma} + \Gamma^\rho{}_{\mu\lambda}\Gamma^\lambda{}_{\nu\sigma} - \Gamma^\rho{}_{\nu\lambda}\Gamma^\lambda{}_{\mu\sigma}\,. \tag{1.4}$$

Then, one can specify how we describe the matter sector. As it was understood a decade before GR was developed, the concept of mass and energy are completely related in special relativity and one can expect energy to also affect the curvature of spacetime. Therefore, it is natural to consider the energy-momentum tensor $T^{\mu\nu}$ to be the quantity describing matter. This tensor is symmetric and in the flat case, it is conserved, i.e. $\partial_\mu T^{\mu\nu} = 0$. Using the covariance principle, we deduce that the energy-momentum tensor needs to satisfy $\nabla_\mu T^{\mu\nu} = 0$ and thus, it must be related to a spacetime quantity satisfying the same condition as well as being symmetric. We have a handful of straightforward choices such as the metric itself or the Ricci tensor $R_{\mu\nu} := R^\sigma{}_{\sigma\mu\nu}$. However the $g_{\mu\nu} = \kappa T_{\mu\nu}$ would lead to instantaneous gravitational influence or other inconsistencies and $R_{\mu\nu} = \kappa T_{\mu\nu}$ to $\nabla_\mu T = \partial_\mu T = 0$ using Bianchi identities, meaning T is constant through spacetime (this would also be the case with the metric). An other tensor satisfying the required criteria is the Einstein tensor $G_{\mu\nu} := R_{\mu\nu} - g_{\mu\nu}R/2$, with $R = R^\alpha{}_\alpha$ being the Ricci scalar. Remarkably, by choosing the relation

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi\, T_{\mu\nu}\,, \tag{1.5}$$

one can recover Newtonian gravity in the weak field limit and the precession of Mercury is correctly predicted using this new theory. This was in a nutshell the historical discovery of General Relativity (GR) by Albert Einstein and the field equations (1.5) now bear his name [2].

At the same time than the discovery of GR, David Hilbert addressed the question of gravity using the approach of symmetries and Lagrangian mechanics [3]. He found that the diffeomorphism invariant action

$$S_{\mathrm{EH}}[g] := \frac{1}{2\kappa} \int \sqrt{-g} R \,\mathrm{d}^4 x\,, \tag{1.6}$$

where $\kappa = 8\pi$, leads to the Einstein field equations in vacuum by applying the principle of least action. To take matter into account, one simply add the Einstein-Hilbert action (1.6) with the relevant matter action to get the full theory. More recently, David Lovelock showed [4] that the most general diffeomorphism and Lorentz invariant action for a tensor field with field equations of degree two can be written as

$$S_{\mathrm{EH}}[g] := \frac{1}{2\kappa} \int \sqrt{-g}(R - 2\Lambda) \,\mathrm{d}^4 x\,, \tag{1.7}$$

where Λ a constant, named the *cosmological constant*, and the factor -2 in front of it is included by convention. The derived field equation is simply

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu} \,, \tag{1.8}$$

and it still converges to Newtonian gravity in the weak field limit, if Λ is small.

It is interesting to note at this point that the choice of describing gravity using the curvature tensor, while keeping the torsion and *non-metricity* $\nabla_\rho g_{\mu\nu}$ to be zero is historical and often considered the most elegant. However, one can construct equivalent theories (up to boundary terms in the action) using only non-zero torsion (while keeping curvature and non-metricity to be zero) [5] or only non-zero non-metricity [6]. In addition, GR can be viewed as a gauge theory with the gauge group being the general linear group $GL(4, \mathbb{R})$. The fact that GR can be described as a geometrical phenomenon and not the other gauge interactions is directly related to the equivalence principle. For a detailed description of the different interpretations of GR, we refer to Ref. [7] and a detailed description of General Relativity as well as related topics can be found in Refs. [8–12].

1.2 Homogeneity and isotropy: the FLRW universe

While the Einstein field equations (1.8) look elegant at first, their complexity is hidden behind clever notations. Exact solutions to these equations are hard to come by and only exist for highly symmetric systems. Fortunately for cosmologists, one such system is our observable universe. From a close perspective, the space around us looks clumpy with most of the matter concentrated in stars and galaxies, while the rest seem to be empty space. However, the distribution of galaxies on very large scales, more than 60 or 70 Mpc, can be safely assumed homogeneous with high precision [13]. In a similar way, the universe looks the same independently of the line of sight, hence it can be regarded as isotropic [14]. The assumptions of homogeneity and isotropy of space make up together the *cosmological principle*. Thankfully, under this fundamental principle the Einstein field equations simplify tremendously and one can study the evolution of the universe. In this section, we introduce the fundamental elements of modern cosmology and discuss the physical interpretations of different approaches. In the following we make extensive use of the lecture notes from Daniel Baumann [15–17] as well as the book by Sean Carroll [9].

1.2.1 The metric and coordinate systems

In this part, we would like to understand how the degrees of freedom of the metric $g_{\mu\nu}$ reduce under the symmetries of the cosmological principle. In General Relativity, it is common practice to display the components of the metric using the squared line element $\mathrm{d}s^2$, namely

$$\mathrm{d}s^2 := g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \,, \tag{1.9}$$

where $\mathrm{d}x^\mu$ are the differential of the coordinate functions. We will sometimes refer to $\mathrm{d}s^2$ as being the metric for simplicity. Since our spacetime manifold is only homogeneous and isotropic in space and not in time, this means that one can foliate the manifold into spacelike

hypersurfaces, where each spatial slice satisfies the symmetries. In other words, one can say that spacetime has a topology $\mathbb{R} \times \Sigma$, where $\mathbb{R}$ represents the time direction, while Σ is the remaining homogeneous and isotropic three-manifold. Under such conditions, the metric becomes

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j\,, \tag{1.10}$$

where t is the time coordinate, $a(t)$ is a function of time that is called the *scale factor* and γ_{ij} is the three-metric of Σ. The remaining three-metric carries six degrees of freedom, while the space of symmetries of the spatial manifold is six dimensional (in short, three translations for homogeneity and three rotations for isotropy), meaning that Σ is *maximally symmetric*. This has the useful consequences that γ_{ij} must have no remaining degree of freedom and the three-Ricci scalar must be equal to a constant, let us say $6k$ (this choice of the constant will be justified later on). Using the fact that Σ is spherically symmetric, one can write the three-metric using the spherical coordinates (r, θ, ϕ)

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = e^{2f(r)}\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) := e^{2f(r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.11}$$

and the function $f(r)$ is found using the constancy of the Ricci scalar. The result gives us the well-known *Friedmann-Lemaitre-Robertson-Walker (FLRW) metric*

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \frac{a^2(t)}{1-kr^2}\mathrm{d}r^2 + a^2(t)r^2\mathrm{d}\Omega^2\,. \tag{1.12}$$

Starting from the ten degrees of freedom contained in the metric $g_{\mu\nu}$, we reduced them to only the scale factor $a(t)$ and the constant k, which is known as the *curvature parameter*.

The resulting FLRW metric (1.12) retains a residual symmetry, it is invariant under the re-scaling

$$a \to \lambda a, \quad r \to r/\lambda, \quad k \to \lambda^2 k\,, \tag{1.13}$$

hence one can choose to have a dimensionless scale factor that can be set to one at a particular instant. Such a choice, which we adopt in the work, would make the curvature constant of dimension $[\text{length}]^{-2}$. The radial coordinate r is what is called a *comoving coordinate*. as one can deduce from the FLRW metric (assuming $k = 0$), physical quantities are calculated using the physical coordinate $r_\varphi = a(t)r$. Figure 1.1 shows the difference between comoving and physical coordinates. It is also interesting to look at the physical velocity of an object

$$v_\varphi = \frac{\mathrm{d}(a(t)r)}{\mathrm{d}t} = a(t)\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\mathrm{d}a}{\mathrm{d}t}r = v_{\mathrm{pec}} + HR\,, \tag{1.14}$$

with $R = ar$ the comoving radius. In other word, the physical velocity of an object can be decomposed into a peculiar velocity v_{pec} and the effect of the expansion of the universe, the Hubble flow, described by the *Hubble constant* $H = \dot{a}/a$, with the dot notation standing for the time derivative.

Moreover, the topology of the spatial section Σ depends on the sign of k. If $k = 0$, the three-metric γ_{ij} is in fact the euclidean metric δ_{ij} written in terms of the spherical coordinates. This means that space is globally flat and we usually associate it with the

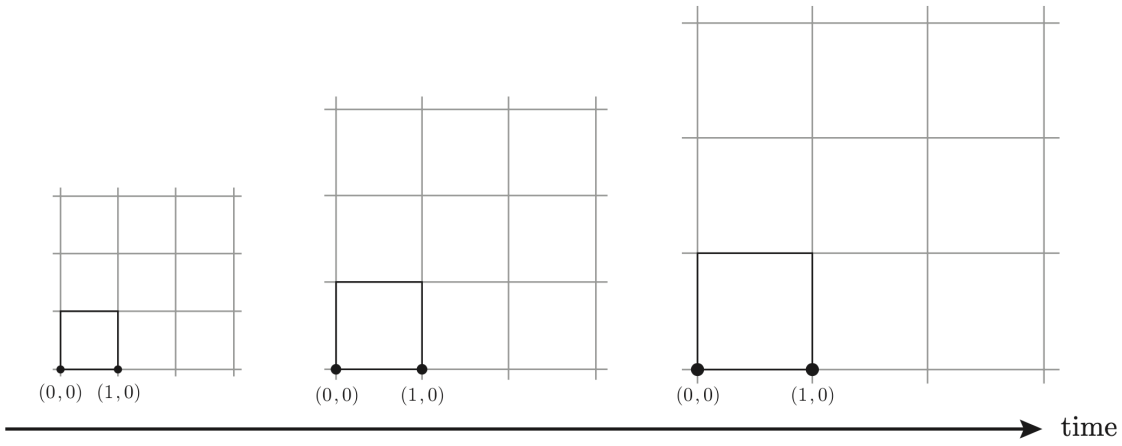


Figure 1.1: Evolution of the comoving and physical distance in an expanding universe. The comoving distance between the two black points is constant and equal to one, while the physical distance between them increases with the expansion. Retrieved from [17].

topology $\Sigma = \mathbb{R}^3$, even though this is not the only choice. In the case where the curvature parameter is positive, space is closed and its topology is the sphere, $\Sigma = \mathbb{S}^3$. To see this, one can change the radial coordinate $\sqrt{k}r \to \sin(\sqrt{k}\chi)$ to get the three-metric

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = \mathrm{d}\chi^2 + k^{-1}\sin^2(\sqrt{k}\chi)\,\mathrm{d}\Omega^2\,, \tag{1.15}$$

which is the metric describing the sphere. Finally when $k < 0$, we use a similar trick as the closed space case and make the coordinate change $\sqrt{k}r \to \sinh(\sqrt{k}\chi)$ to get the three-metric

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = \mathrm{d}\chi^2 + k^{-1}\sinh^2(\sqrt{k}\chi)\,\mathrm{d}\Omega^2\,, \tag{1.16}$$

describing a three-dimensional space with constant negative curvature. Such a space is often called *open* because it necessarily extends to infinity and can thought off having the shape of a saddle.

Once the coordinate system has been chosen, it is useful to move on and choose an observer as well. In general relativity, we define an *observer* as a time-like future pointing worldline γ on the space-time manifold, together with an orthonormal basis e^μ_α which includes the tangent vector field V^μ, otherwise also named *velocity* vector field, to the worldline. A particular and interesting choice that can be made is the *Eulerian observer*. It is defined as the observer which is at rest on spatial slices [18, 19]. This means that the velocity field of the Eulerian observer is aligned with the normal vector n^μ to the spatial slices of constant time coordinate and can be written as

$$V_\mu = (-\sqrt{-g^{00}}, 0, 0, 0)\,. \tag{1.17}$$

Using the coordinate system leading to the FLRW metric (1.12), we get the velocity field

$$V_\mu = (-1, 0, 0, 0) \implies V^\mu = (1, 0, 0, 0)\,. \tag{1.18}$$

As an example, we arbitrarily give other vectors that could complete the orthonormal basis

defining the observer

$$e^{\mu}_{\alpha} = (V^{\mu}, S^{\mu}, \theta^{\mu}, \phi^{\mu}) \,, \tag{1.19}$$

$$\text{with} \quad S^{\mu} = (0, \sqrt{g^{11}}, 0, 0) \,, \quad \theta^{\mu} = (0, 0, \sqrt{g^{22}}, 0) \,, \quad \phi^{\mu} = (0, 0, 0, \sqrt{g^{33}}) \,. \tag{1.20}$$

The Eulerian observer plays a particular role in cosmology as it is the frame in which the fluid describing the energy content of the universe is at rest. This can be understood from the assumption that the universe is homogeneous and isotropic. An other equivalent way to define the cosmological principle is to say that the universe appears isotropic from any given point in space-time and a fluid in motion with respect to the observer at rest would violate this assumption. It will be shown later on in Sec. 1.4.3 that assuming isotropy with respect to a single point together with the fluid at rest from the point of view of the Eulerian observer is in fact equivalent to the cosmological principle. While this statement could appear trivial in the context of cosmology, it is less straight forward when one wants to add a black hole in the central point. The above equivalence basically means that one cannot have a black hole surrounded by an Eulerian fluid except if the fluid describes a cosmological constant.

An other point that is often heard in cosmology is the fact that galaxies are not moving away from each other, but it is the space between them that expands. This statement can be understood from the velocity field of the fluid (1.18), which shows the fluid at rest, and the expansion of the spatial slices $\mathrm{d}s^2_3 := a^2(t)/(1 - kr^2)\mathrm{d}r^2 + a^2(t)r^2\mathrm{d}\Omega^2$. However, as we will see in Sec. 1.2.4 this point of view is merely a coordinate artifact. Moreover, as it was already discussed by Lemaître and others [20, 21], if the universe was composed of an expanding large gas cloud siting in empty space, we would not be able to differentiate it from an FLRW universe, provided that the gas cloud we live in is large enough. This somewhat contradicts the common saying that the big bang does not look simply like an explosion. From today's perspective, we could very well assume that it is the case and as for the expansion of space, it is a matter of interpretation.

1.2.2 Dynamics

So far, we have only discussed the geometrical part of the Einstein field equations and informally introduced how we model the energy content of the universe as a fluid. One of the prime candidates to model this energy fluid is *perfect fluids*. They have the property of being isotropic in their rest frame, which naturally satisfies the cosmological principle. The energy-momentum tensor of a perfect fluid can be written as

$$T_{\mu\nu} = (p + \rho)U_{\mu}U_{\nu} + p g_{\mu\nu} \,, \tag{1.21}$$

where p and ρ are the pressure and the density of the fluid, respectively, and U_{μ} is the four-velocity of the fluid. In the case of a comoving fluid $U_{\mu} = (-1, 0, 0, 0)$, the energy-momentum tensor of a perfect fluid reduces to

$$T^{\mu}{}_{\nu} = \mathrm{diag}(-\rho, p, p, p). \tag{1.22}$$

As we have seen earlier in the historical approach to GR, the conservation of the energy-momentum tensor plays an important role. In the context of cosmology, one can apply the

concept of conservation to find the *continuity equation* of the fluid. This reads

$$\nabla_\mu T^\mu{}_0 = \partial_\mu T^\mu{}_0 + \Gamma^\mu_{\mu\nu} T^\nu{}_0 - \Gamma^\nu_{\mu 0} T^\mu{}_\nu = 0 \tag{1.23}$$

$$\implies \dot\rho + 3\frac{\dot a}{a}(\rho + p) := \dot\rho + 3\frac{\dot a}{a}(1+w)\rho = 0 \tag{1.24}$$

$$\implies \frac{\dot\rho}{\rho} = -3(1+w)\frac{\dot a}{a}\,, \tag{1.25}$$

where we used the *equation of state* $p = w\rho$ in the last line.

Our universe is filled with different types of perfect fluid, satisfying different equations of state. We summarize here the list of fluids contained by the universe as accepted by the current standard model of cosmology.

Dust This type of matter is constituted by the stars or galaxies, for example, and is characterized by its collisionless property. This implies that the pressure of the fluid can be neglected and the equation of state reduces to

$$\frac{\dot\rho}{\rho} = -3\frac{\dot a}{a} \implies \rho \propto a^{-3}\,. \tag{1.26}$$

As expected, this means that the density of dust is diluted with the expansion. There is an overwhelming amount of clues that the dust content of the universe is dominated by some invisible matter, called *dark matter*, and the study of this dark sector is one of the biggest puzzle of modern physics [22].

Radiation One cannot neglect the pressure of electromagnetic radiation or gases of particles with relativistic velocities. This can be seen using the energy-momentum tensor of the electromagnetic field, i.e.

$$T^{\mu\nu} = \frac{1}{4\pi}\left(F^{\mu\alpha}F^\nu{}_\alpha - \frac{1}{4}g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}\right) \implies T^\mu{}_\mu = 0\,, \tag{1.27}$$

but as a perfect fluid, the tensor also satisfy Eq. (1.21). Therefore one can relate the density and the pressure as

$$T^\mu{}_\mu = -\rho + 3p = 0 \implies p = \frac{1}{3}\rho\,. \tag{1.28}$$

The continuity equation further leads us to $\rho \propto a^{-4}$, meaning that the energy density of radiation decreases faster with the expansion than the density of dust, due to the loss of energy of individual photons through *redshift*.

Dark energy While the cosmological constant is included in the Einstein field equations (1.8) as an additional term, it can be considered as an other form of energy. It also behaves as a perfect fluid and its equation of state can be easily derived as

$$T_{\mu\nu} = -\frac{\Lambda}{8\pi}g_{\mu\nu} \implies \rho = -p = \frac{\Lambda}{8\pi}\,, \tag{1.29}$$

meaning that $w = -1$. A fluid with such an equation of state has the particularity that its energy stays constant with the expansion of space, $\rho \propto a^0$. One could wonder, if such a constant is needed in our theory and it seems to be the case in today's cosmological paradigm.

There are numerous evidences that the expansion of our universe is accelerating [23–25] and the cosmological constant is the best model we have so far to explain this behaviour.

Next we would like to understand how a coupled system, composed of the gravitational field and a perfect fluid, evolves. For that purpose, we need to reduce the Einstein field equations using the symmetries at hand. This is simply done by substituting what we know about perfect fluids and the FLRW metric. Using the equations for the 00-component and the ij-component, we can find

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3p) \tag{1.30}$$

$$\left(\frac{\dot{a}}{a}\right)^2 := H^2 = \frac{8\pi}{3}\rho - \frac{k}{a^2}\,, \tag{1.31}$$

the *Raychaudhuri* and *Friedmann equations*, respectively. Together with the continuity equation (1.25), the system of three differential equations for the two functions $\rho(t)$ and $a(t)$ is not linearly independent and one can discard one of the equations. This is useful in the case where one considers several fluids composing the matter sector. In this case, we can use the additive property of the density to rewrite the Friedmann equation as

$$H^2 = \frac{8\pi}{3}\left(\rho_{\text{dust}} + \rho_{\text{rad}} + \rho_\Lambda\right) - \frac{k^2}{a^2}\,, \tag{1.32}$$

while each individual perfect fluid still satisfies its continuity equation. Therefore, one can further improve the Friedmann equation as

$$H^2 = \frac{8\pi}{3}\left(\rho_{\text{dust},0}\frac{a_0^3}{a^3} + \rho_{\text{rad},0}\frac{a_0^4}{a^4} + \rho_\Lambda\right) - \frac{k^2}{a^2}\,, \tag{1.33}$$

with $\rho_{\text{dust},0}$, $\rho_{\text{rad},0}$ and a_0 being the initial conditions at time t_0 of the relevant variables. This last equation only depends on the scale factor and can be solved using the initial condition $a(t_0) = a_0$. Another version of the Friedmann equation that can often be seen in the literature involves the density parameters Ω_m, Ω_r, Ω_Λ and Ω_k of dust, radiation, the cosmological constant and the curvature, respectively. This equation is dimensionless and is obtained by dividing Eq. (1.33) by the initial condition H_0^2 of Hubble constant squared at some instant t_0, giving

$$\frac{H^2}{H_0^2} = \left(\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k\right)\,, \quad \text{with } \Omega_i := \frac{8\pi\rho_{i,0}}{3H_0^2}\,, \tag{1.34}$$

where we also set $a_0 = 1$ and defined $\rho_{k,0} = -3k/(8\pi)$.

1.2.3 Cosmological horizons

In today's cosmological paradigm, the universe started at a finite past time and its expansion is accelerating. This means that there are patches of the universe that either didn't have time or will never reach us. The surface that separates the region are called *horizons*. There are different types of horizons depending on what is meant by "observable" and if we work with global or local horizons. This section is largely inspired by the book of Faraoni [26].

Particle horizon This is often the first type of horizons that is encountered in cosmology textbooks since its interpretation is relatively straightforward. It is the largest distance R_P from which light had time to reach us since the beginning of the universe. The light that travelled the furthest before reaching us was following radial null geodesics, which means that the line element with $k = 0$ becomes

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\mathrm{d}r^2 = 0\,, \tag{1.35}$$

giving us the infinitesimal comoving radius $\mathrm{d}r = -\mathrm{d}t/a(t)$. We can integrate these quantities between the emission of the light ray at $r = R_P$ at the time of the big bang t_{bb} and its detection here $r = 0$ and today at time t. In physical coordinates, this gives

$$a(t)\int_{R_P}^{0}\mathrm{d}r = -a(t)\int_{t_{bb}}^{t}\frac{\mathrm{d}t'}{a(t')} \implies R_P(t) = a(t)\int_{t_{bb}}^{t}\frac{\mathrm{d}t'}{a(t')}\,. \tag{1.36}$$

We found the proper particle horizon R_P.

Event horizon An other question that can arise naturally in cosmology is to wonder how far can we ever send signals to. Or in an other perspective, at which distance lies the furthest object that one can ever observe from earth. This is the cosmological definition of event horizons. At first, one can think that in an infinite amount of time one will be able to observe everything, but this is not the case because of the accelerating nature of the expansion. In physical coordinates, the proper event horizon for a spatially flat universe is

$$R_E(t) = a(t)\int_{t}^{t_{max}}\frac{\mathrm{d}t'}{a(t')}\,, \tag{1.37}$$

where t_{max} is the time of the end of the universe and is $t_{max} = \infty$ in most cases. For fluids with $w > -1$, one can show that the comoving event horizon is

$$R_E(t) = t^{2/(3(w+1))}\left[\frac{3(w+1)}{3w+1}t'^{(3w+1)/(3(w+1))}\right]_{t'=t}^{t'=\infty}\,. \tag{1.38}$$

we see that for $w > -1/3$ we have $R_E \to \infty$ meaning that there are no event horizons, while R_E is finite for $w < -1/3$. This latter inequality coincides with the condition required for an accelerated expansion $\ddot{a} > 0$.

Event horizons are also very important in the context of black hole physics. In this context, we also introduce the more formal definition of an event horizon. It is defined as the connected component of the boundary of the causal past of future null infinity. In other words, it is the boundary of all the null geodesics that reaches infinity. One can also say that an event horizon arises from the null geodesics that fails to reach infinity.

Hubble horizon Due to the expansion of the universe, the further away an object is the faster it moves away from us. One can therefore wonder where lies the limit for which objects moves at the speed of light with respect to us. We can deduce this distance using the physical velocity of objects (1.14), assuming no peculiar velocity and $k = 0$. We find

that the Hubble horizon is

$$R_H = \frac{1}{H(t)}\,. \tag{1.39}$$

Apparent horizon We found the Hubble horizon by fixing the velocity of comoving objects to the speed of light. However, this approach is not appropriate for more complicated system and it is not mathematically well defined. One can treat this type of local horizon with the more formal concept of apparent horizons. A future apparent horizon is defined as surface where future-pointing and outgoing geodesics stop propagating outward of the horizon [26,27]. This is usually quoted as the surface from which light cannot escape. Let l^μ and n^μ be the two future-pointing null vectors normal to this surface. Let's assume without loss of generality that l^μ is pointing in the radial direction with increasing r, while n^μ is pointing in the opposite direction, towards decreasing r. what we mean by that if we define the radial vector in the FLRW coordinate system (1.71)

$$S^\mu = (0, 1, 0, 0), \tag{1.40}$$

we have $l^\mu S_\mu > 0$ and $n^\mu S_\mu < 0$. We further define the *expansions* θ_l and θ_n of the normal vectors l^μ and n^μ respectively

$$\theta_l = h^{\mu\nu}\nabla_\mu l_\nu \quad \text{and} \quad \theta_n = h^{\mu\nu}\nabla_\mu n_\nu\,, \tag{1.41}$$

$$\text{with} \quad h^{\mu\nu} = g^{\mu\nu} + \frac{l^\mu n^\nu + l^\nu n^\mu}{-l^\mu n_\mu}\,. \tag{1.42}$$

$h_{\mu\nu}$ is the induced metric on the surface of normal vectors l^μ and n^μ. In short, when the expansion of a vector is negative, it means that the geodesic tangent to this vector is evolving inward, towards decreasing r, while a positive expansion leads to an outward propagation of the geodesic. When both expansions of the normal vectors are negative, null geodesics are trapped and drifting toward $r = 0$. One therefore define a cosmological apparent horizon as the limiting case, which is the marginally outer trapped surface for which we have

$$\theta_l > 0 \quad \text{and} \quad \theta_n = 0\,. \tag{1.43}$$

These conditions mean that the geodesic tangent to the vector n^μ stops propagating inward, while the geodesic associated to the vector l^μ still propagate outward.

Coming back to the FLRW metric (1.12), the two future pointing normal vectors are

$$l^\mu = (1, \sqrt{(1-kr^2)/a^2}, 0, 0) \quad \text{and} \quad n^\mu = (1, -\sqrt{(1-kr^2)/a^2}, 0, 0)\,, \tag{1.44}$$

giving us the expansions

$$\theta_l = 2\left(\sqrt{\frac{1-kr^2}{a^2r^2}} + H\right) \quad \text{and} \quad \theta_n = 2\left(-\sqrt{\frac{1-kr^2}{a^2r^2}} + H\right)\,. \tag{1.45}$$

If one chooses the comoving radius

$$R_A = \sqrt{H^2 + k/a^2}^{-1}\,, \tag{1.46}$$

we can show that we have indeed $\theta_l > 0$ and $\theta_n = 0$. R_A is therefore the apparent horizon and we recover the Hubble horizon for $k = 0$.

In the case of black hole, we would like that null geodesics are drifting towards $r \to 0$, hence both expansions being negative. The limiting case is then

$$\theta_n < 0 \quad \text{and} \quad \theta_l = 0, \tag{1.47}$$

which is defined as a black hole apparent horizon, the geodesic tangent to the outgoing vector l^μ stops propagating outwards.

A very interesting property of apparent horizons is that they always lie inside an event horizon, provided that the null energy condition (NEC), $T^{\mu\nu}l_\mu l_\nu \geq 0$, for l^μ a null vector, is satisfied [26]. This means that the existence of an apparent horizon is sufficient to show that there exists an event horizon. In particular, from Penrose and Hawking singularity theorems [28, 29] we conclude that there is always a singularity included in an apparent horizon.

1.2.4 Cosmology in Painlevé-Gullstrand coordinates

One of the foundational principle of General Relativity is the invariance under general coordinate transformation. In choosing the FLRW metric (1.12) we fixed the coordinate system and therefore the gauge, but we are free to change the coordinates as we like. So far, we have used the comoving radial coordinate r and we often moved to the proper or physical coordinate $R = a(t)r$ to describe physical properties. One could instead directly start by writing the metric in terms of R. For the following, we assume a spatially flat space-time, $k = 0$. Using the product rule for derivatives, we find the infinitesimal *Schwarzschild radial coordinate* $\mathrm{d}R$, often called this way due to its connection with the radial coordinate used in the description of black holes with the Schwarzschild metric,

$$\mathrm{d}R = a(t)\mathrm{d}r + H(t)R\ \mathrm{d}t\,. \tag{1.48}$$

Substituting $\mathrm{d}r$ in the FLRW metric we find the metric in *Painlevé-Gullstrand* coordinates [19]

$$\mathrm{d}s^2 = -\left(1 - H^2(t)R^2\right)\mathrm{d}t^2 - 2H(t)R\ \mathrm{d}R\mathrm{d}t + \mathrm{d}R^2 + R^2\mathrm{d}\Omega^2\,. \tag{1.49}$$

Under these new coordinates, the metric is no more diagonal, which is a rather unusual feature. Moreover, the spatial section

$$\mathrm{d}s_3^2 = \mathrm{d}R^2 + R^2\mathrm{d}\Omega^2\,, \tag{1.50}$$

is now non-expanding. This is somewhat in conflict with the interpretation that galaxies move away from each other due to the expansion of space, which is the case with the FLRW coordinates. To see how and why galaxies recede, we need to look at the velocity field of the matter fluid. The fluid moves as an Eulerian or comoving observer with the covariant

velocity $V_\mu = (-1, 0, 0, 0)$, which becomes $V^\mu = (1, H(t)R, 0, 0)$ in the contravariant basis, because of the non-diagonal term. This means that a galaxy with no peculiar velocity is moving in this coordinate system. For the readers familiar with the 3+1 decomposition and the ADM formalism [30], the coordinate transformation (1.48) transfers the dynamics from the lapse function to the shift vector.

1.3 The Schwarzschild solution and its limitations

A common simplifying assumption used in Newtonian gravity consists of considering a dominating and spherical central mass with orbiting test objects with no gravitational influence. Moreover, such a system lies in vacuum outside of the central mass and has a static gravitational field, which makes the calculations easier. One can adopt similar assumptions, spherical symmetry and vacuum, in the context of General Relativity in order to describe the gravitational field outside of spherical objects. Historically, the solution to the Einstein field equations with these symmetries was the first analytic solution that has been found [31]. Beyond its use in planetary systems or other spherically symmetric models, the *Schwarzschild* solution has led to the surprising prediction of black holes. These bizarre objects, from which light cannot escape, has recently been directly imaged for the first time [32] strongly supporting their existence. In the context of cosmology, in which the universe is filled with a perfect fluid, one could try to add a test fluid surrounding a black hole, but with no gravitational influence. After introducing the Schwarzschild solution, we will look at the consequences of adding a test fluid around a black hole.

1.3.1 Static and spherically symmetric space-time

Without entering in too many details, one can always reduce a spherically symmetric metric into the form [9]

$$\mathrm{d}s^2 = -e^{2\alpha(t,r)}\mathrm{d}t^2 + e^{2\beta(t,r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.51}$$

where α and β are the remaining dynamical variables at play. In short, the spherical symmetry fixes the induced metric on the spherical surfaces to be $\mathrm{d}s_2^2 = r^2\mathrm{d}\Omega^2$, leaving only three remaining variables in the metric. An additional reparametrization of the time coordinate reduces the system to two dynamical functions as in the metric (1.51). One further chooses an exponential form for these arbitrary functions so as to have a Lorentzian metric. Finally, α and β cannot depend on the angular coordinates, which would otherwise break the spherical symmetry. For simplicity in the notation, the radial coordinate is labelled r, but it is in fact what was called the Schwarzschild radial coordinate earlier in the chapter.

Next we postulate that outside of the spherically symmetric object we have a perfect vacuum, hence $T^{\mu\nu} = 0$. One can therefore derive the Einstein field equations and we get

$$1 - e^{2\beta} = 2r\beta' \tag{1.52}$$

$$1 - e^{2\beta} = -2r\alpha' \tag{1.53}$$

$$\dot{\beta} = 0\,. \tag{1.54}$$

Firstly, we see that β cannot depend on time from the third equation and secondly we have

that $\alpha' = -\beta'$ using the two first equalities. Furthermore, the first equation can be solved analytically leading to the solutions

$$\alpha(t,r) = C(t) + \frac{1}{2}\ln\left(1+\frac{\mu}{r}\right), \tag{1.55}$$

$$\beta(t,r) = -\frac{1}{2}\ln\left(1+\frac{\mu}{r}\right), \tag{1.56}$$

where $C(t)$ and μ are integration constants. In fact, $C(t)$ can simply be absorbed in the time coordinate with the coordinate transformation $e^{C(t)}\mathrm{d}t \to \mathrm{d}t$. In the weak field limit, we need to recover the Newtonian gravitational potential and therefore make the choice $\mu = GM$, with M the mass of the central object. The final result is the Schwarzschild metric

$$\mathrm{d}s^2 = -\left(1-\frac{2M}{r}\right)\mathrm{d}t^2 + \left(1-\frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.57}$$

where we set again $G = 1$. Interestingly, in the vacuum solution space-time is static, even though we haven't required the source to be static. This means that, as long as the source stays spherically symmetric, the latter can be evolving in time without affecting the outside gravitational field. This is known as *Birkhoff's theorem* and in particular, this also implies that the Schwarzschild metric can very well describe the gravitational field outside a collapsing star or a supernova for example. A striking property of the Schwarzschild metric (1.57) is its singular behaviour at $r \to 2M$. In fact, there lies an event horizon at the *Schwarzschild radius* $r_S = 2M$.

As in the homogeneous and isotropic solution to the Einstein field equations, we can change the coordinate system, in particular the time coordinate, to express the Schwarzschild metric (1.57) in the Painlevé-Gullstrand form. Using the new time coordinate

$$\tilde{t} = t - 2M\left(-2y + \ln\left(\frac{y+1}{y-1}\right)\right), \quad y = \sqrt{\frac{r}{2M}}\,, \tag{1.58}$$

we find the metric

$$\mathrm{d}s^2 = -\left(1-\frac{2M}{r}\right)\mathrm{d}\tilde{t}^2 + 2\sqrt{\frac{2M}{r}}\mathrm{d}\tilde{t}\mathrm{d}r + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,. \tag{1.59}$$

For illustrative purposes, one can look for the apparent horizon associated to this metric. The future-pointing light-like vectors that are needed are

$$l^\mu = \left(1, 1-\sqrt{\frac{2M}{r}}, 0, 0\right) \quad \text{and} \quad n^\mu = \left(1, -1-\sqrt{\frac{2M}{r}}, 0, 0\right), \tag{1.60}$$

from which we derive the expansions

$$\theta_l = \frac{2}{r}\left(1-\sqrt{\frac{2M}{r}}\right) \quad \text{and} \quad \theta_n = -\frac{2}{r}\left(1+\sqrt{\frac{2M}{r}}\right). \tag{1.61}$$

therefore we conclude that there exists an apparent horizon for $r = 2M$ and it coincides

with the event horizon.

1.3.2 Schwarzschild black hole with a test fluid

The spherically symmetric vacuum solution is undoubtedly of crucial importance for planetary systems and to describe with a reasonable accuracy the behaviour of black holes. However, the surrounding of such objects is not perfectly empty and one could ask how one can add a fluid in the vicinity of a black hole. In the case where the energy density around the central object is small, one can make the assumption that such a fluid has negligible influence on the gravitational field relative to the effect of the central mass. Hence, one can keep the Schwarzschild solution as a good approximation. Not all perfect fluids can be used with these assumptions. A fluid made of bosonic matter for example is not well suited because it is strongly self-interacting and would form an accretion disk, which can break spherical symmetry or the negligible back-reaction of the test fluid. Dark matter on the other hand, is usually assumed to be interaction-free or very weakly interacting [22] and radiation doesn't interact with itself, hence they could both be of interest in describing a central black hole surrounded by a test fluid.

A black hole with mass M bathed in a radiation fluid at some possibly varying temperature $T(t)$ will grow in size and mass as it absorbs matter. To measure the energy flow into the black hole's horizon and thus the change in mass, we use the viewpoint of an observer far away in flat space-time. To this observer, the black hole and its gravitational field appear as a sphere with a radius of $b = \sqrt{27}M$. b is in fact the impact parameter of the black hole for photons at infinity. In addition to the physical cross-section, one also need to take into account the scattering cross-section of photons with larger wavelength for the calculation of the inward flux. One last ingredient that one can add to fully describe the overall energy flow is the *Hawking radiation* [29]. It has been shown that black holes radiate in a similar way as a black body and its temperature is inversely proportional to its mass. One needs to take into account this energy loss when computing the overall energy flow across the horizon. We are now fully equipped to describe the evolution of the mass of a black hole surrounded by a radiation fluid that is evolving in time as it is the case in an expanding or contracting universe.

Catastrophic fate of Schwarzschild black holes in a thermal bath

Aurélien Barrau, Killian Martineau, and Cyril Renevey
Laboratoire de Physique Subatomique et de Cosmologie, Univ. Grenoble-Alpes, CNRS/IN2P3 53, avenue des Martyrs, 38026 Grenoble cedex, France

(Received 31 January 2022; accepted 25 June 2022; published 13 July 2022)

Using the Schwarzschild metric as a rudimentary toy model, we pedagogically revisit the curious prediction that the mass of a classical black hole in a constant temperature thermal bath diverges in a finite amount of time. We study in detail how this instability behaves if the temperature of the bath is allowed to vary with time and conclude that whatever the background behavior (but for a zero-measure subspace of the initial conditions), the black hole mass either diverges or vanishes in a finite time if the Hawking radiation is taken into account. The competition between both effects is subtle and not entirely governed by the hierarchy of the relevant temperatures. This instability is also shown to be reached before the background singularity in a contracting universe, which has implications for bouncing models. The results are generalized to spaces with extra dimensions, and the main conclusions are shown to remain true. The limitations of the model are reviewed, both from the point of view of the dynamical black hole horizon and from the point of view of the background space expansion. Comparisons with other approaches are suggested and possible developments are underlined.

DOI: 10.1103/PhysRevD.106.023509

I. INTRODUCTION

This work does not contain fundamentally new results. It basically aims at pedagogically "rediscovering" in a wider way a strange behavior of black holes absorbing a continuous flux of energy and at exploring interesting situations relevant for cosmology. Most of our points are established using the Schwarzschild metric in a Bondi-Hoyle-Lyttleton basic approximation scheme, which can only be considered as a very rough toy model. Interesting features can be guessed while neglecting both the background space expansion and the truly dynamical nature of the black hole horizon but extreme care should be taken when considering the results literally: technical and conceptual issues should be addressed before firm conclusions are drawn. Our aim is therefore not to make any strong claim but to encourage further studies so as to clarify interesting features and to attract the attention of the unfamiliar reader to some strange situations.

A classical black hole in a thermal bath obviously grows by absorbing the surrounding radiation. If the temperature of the bath is constant, it might be naively expected that its mass tends to infinity after an infinite amount of time. We show that, without any exotic assumption, the mass actually diverges at *finite* time. This has first been noticed in [1] (translated in [2]). Many works were devoted to the so-called self-similar solution, to the investigation of different equations of state, to the existence of a Friedmann or quasi-Friedmann asymptotic behavior, and to the separate universe issue [3–17]. Although those points are of unquestionable importance, we do not here deal with these subtleties and mostly focus on some strange consequences of a naïve Schwarzschild-based analysis. Our aim is not to derive reliable conclusions—as a static metric is used beyond its regime of validity—but to underline some (maybe) surprising situations that could deserve a closer look and seem, to the best of our knowledge, quite ignored by the community.

We show that the pathological behavior depends neither on the kind of radiation nor on the initial black hole mass. If the background varies with time, the situation becomes intricate. Depending both on the speed of the space dilation and on the initial conditions, the mass of the black hole can either diverge in a finite amount of time or tend to a finite asymptotic value. The "natural" expectation ($M \to \infty$ for $t \to \infty$) actually happens only for a subset of zero measure in the parameter space (which has already received a great deal of attention in the literature and will not be our focus).

We compare the critical black hole mass with the Hubble mass and show that different hierarchies have to be considered. When the Hawking evaporation is also taken into account, the mass can vanish for a part of the possible initial states. This shows that, under the strong simplifications performed, in all cases, the black hole is unstable and its mass either diverges or vanishes in a finite amount of time. The relative value between the initial temperature of the bath and the initial temperature of the black hole does *not* determine the long run behavior.

Importantly, in contracting cosmological solutions, the black hole instability is always reached before the background singularity. Some consequences for bouncing

models are drawn, underlying that this catastrophic behavior can happen before the Planck energy is reached. The competition between the absorption and the evaporation is quite involved.

We study how extra dimensions would influence the picture and conclude that the diverging behavior remains true although the singularity is reached later when $D > 4$.

Finally, some comparisons with analog black holes are suggested, and links with chemistry and statistical physics are pointed out. The limitations of the toy model used in this work are discussed in detail and possible developments are underlined, focusing on some relevant generalized metrics. We stress that the curious results—some already well known, others quite new—shown in this work are more to be understood as a strangeness of the Schwarzschild solution pushed beyond what it was intended for than as realistic physical effects. Still, they can be useful to discover situations deserving deeper investigations.

II. ABSORPTION BY A BLACK HOLE IN A THERMAL BATH AT CONSTANT TEMPERATURE

To fix ideas, let us begin by considering a black hole in a constant temperature thermal bath of photons, disregarding the Hawking evaporation. Throughout the article we use Planck units except as otherwise stated. The energy density of the equilibrium distribution is given by

$$\rho = \frac{g}{2\pi^2}\int_0^\infty \frac{E}{e^{\frac{E-\mu}{T}}-1}E^2 dE, \tag{1}$$

where g is the number of internal degrees of freedom, E is the energy, and μ is the chemical potential assumed to be negligible in the following. This leads to the standard relation

$$\rho = \frac{\pi^2}{15}T^4. \tag{2}$$

The radius of a static[1] nonspinning and uncharged black hole of mass M is $2M$. Photons will fall in the black hole if they approach its center at a distance smaller that $3M$ in Schwarzschild coordinates. This corresponds to an impact parameter at infinity $b = \sqrt{27}M$. Assuming that the wavelength of the background radiation is much smaller than the size of the black hole, that is, assuming that the optical limit holds, the effective cross section is simply given by $108M^2$ (the factor π being lost by averaging over the isotropic distribution). The mass evolution therefore immediately reads

$$\frac{dM}{dt} = \frac{36}{5}\pi^2 T^4 M^2. \tag{3}$$

[1]Once again, we emphasize that this hypothesis is not fulfilled in our calculation, hence the "toy model" qualification.

The energy density given by Eq. (2) and entering Eq. (3) is defined at infinity. The equation is trivially integrated in

$$\frac{1}{M_i} - \frac{1}{M} = \frac{36}{5}\pi^2 T^4 t, \tag{4}$$

where M_i is the initial mass of the black hole and t is time elapsed since its formation. Interestingly (and maybe surprisingly), the mass does not diverge in the limit $t \to \infty$ but at *finite* time:

$$t_d = \frac{5}{36\pi^2 T^4 M_i}, \tag{5}$$

such that

$$\lim_{t\to t_d} M = \infty. \tag{6}$$

This basically means that any black hole in a thermal bath, described at this level of approximation, is unstable. (Alternatively, this also means, with the same limitations, that any steady-state cosmological scenario—relying on an infinitely old universe and trying to account for a blackbody radiation usually assumed to be at constant temperature—is basically incompatible with the existence of black holes in the usual sense.)

This result remains true if the initial mass of the black hole is such that the wavelength of the surrounding radiation is much larger than the Schwarzschild radius. In this case, the scattering cross section is given by [18]

$$\sigma = \frac{64\pi M^4 E^2}{3}. \tag{7}$$

This new behavior is due to the fact that the incoming wave cannot be approximated to a point particle anymore. The additional E^2 factor shows that the greybody factor is, as expected, suppressed in the limit $ME \to 0$. In this regime, and assuming $E \sim T$, the evolution equation becomes

$$\frac{dM}{dt} = \frac{256}{45}\pi^2 T^6 M^4. \tag{8}$$

Calling λ the mean wavelength of the thermal radiation, we define the equilibrium time (corresponding to the transition between a black hole smaller than the typical photons to a black hole larger than the typical photons) t_e such that $M(t_e) \sim \lambda \sim 1/T$. This defines the change of regime between the optical cross section $\propto M^2$ and the low-energy cross section $\propto M^4E^2$. Neglecting the oscillations (which would not change the order of magnitude), t_e can be estimated to be given by integrating Eq. (8):

$$t_e = \frac{15}{256\pi^2 T^6}(M_i^{-3} - T^3), \tag{9}$$

with $M_i < 1/T$ by hypothesis.

After the time t_e, the black holes are in the usual regime and the remaining time before divergence is

$$\Delta t = \frac{5}{36\pi^2 T^3}. \tag{10}$$

The full time between formation and divergence is therefore in this case

$$t_d = \frac{5}{36\pi^2 T^3} + \frac{15}{256\pi^2 T^6}(M_i^{-3} - T^3) \tag{11}$$

$$= \frac{186}{2304\pi^2 T^3} + \frac{15}{256\pi^2 T^6 M_i^3}, \tag{12}$$

which is again finite for any value of M_i and T. The dependence upon the initial mass is, as expected, stronger than before as the absorption at the beginning is highly suppressed. This phenomenon, however, does not prevent the mass divergence at finite time.

If the thermal bath is made of relativistic fermions instead of photons, the result is mostly the same in the case of a black hole larger that the mean wavelength of the radiation. The energy density of the bath is simply modified by a factor of $7g/16$ and

$$t_d = \frac{20}{63\pi^2 g T^4 M_i}. \tag{13}$$

The "small mass regime," however, strongly differs from the case of photons. The scattering cross section for fermions in the $E \to 0$ limit is [18]

$$\sigma = 2\pi M^2. \tag{14}$$

The same reasoning as previously leads to

$$t_e = \frac{30}{7\pi^2 g T^4}(M_i^{-1} - T). \tag{15}$$

And the full divergence time reads

$$t_d = \frac{20}{63\pi^2 g T^3} + \frac{30}{7\pi^2 g T^4}(M_i^{-1} - T) \tag{16}$$

$$= \frac{30}{7\pi^2 g T^4 M_i} - \frac{256}{63\pi^2 g T^3}. \tag{17}$$

This shows that, in all cases, the mass of a classical black hole in a constant temperature bath tends to infinity in a finite amount of time.

This unusual behavior is, of course, entirely rooted in the specific mass-radius relation of black holes. A ball of standard matter with a cross section proportional to $M^{2/3}$ would grow gently as $M \propto t^3$ and would never experience any singularity.

III. ABSORPTION IN A THERMAL BATH AT DECREASING TEMPERATURE

Although the mass divergence in finite time might come, at first sight, as a physical surprise, it is a mathematically obvious consequence of having $dM/dt \propto M^\delta$ with $\delta > 1$. It is now worth considering the case of a classical black hole immersed in a thermal bath with decreasing temperature. Let us assume that

$$T = T_0\left(\frac{t}{t_0}\right)^\alpha. \tag{18}$$

The constants T_0 and t_0 could be absorbed in a single parameter but keeping both of them helps the physical intuition. The exponent α is negative (otherwise the divergence is just trivially amplified). The evolution equation (focusing on the case of an initial mass larger than the radiation mean wavelength) is

$$\frac{dM}{dt} = kT^4M^2 = kT_0^4\left(\frac{t}{t_0}\right)^{4\alpha}M^2, \tag{19}$$

with $k = 36\pi^2/5$ for photons. This leads to

$$\frac{1}{M} = \frac{1}{M_i} - \frac{kT_0^4}{(4\alpha+1)t_0^{4\alpha}}(t^{4\alpha+1} - t_i^{4\alpha+1}), \tag{20}$$

where t_i is the formation time of the black hole and M_i its corresponding mass. Calling $\beta = 4\alpha + 1$, the mass diverges $(1/M = 0)$ at time

$$t_d = e^{\frac{1}{\beta}\ln\left(\frac{\beta t_0^{\beta-1}}{kM_iT_0^4} + t_i^\beta\right)}. \tag{21}$$

This will actually happen if the argument of the logarithm is positive. If $\beta > 0$, this is always true. Otherwise stated, if the cooling of the universe in slow enough ($T \propto t^\alpha$ with $\alpha > -1/4$), the mass divergence at finite time happens whatever the initial conditions. On the other hand, if $\beta < 0$, the divergence requires $M_i > M_c$ with

$$M_c = -\frac{\beta t_0^{\beta-1} t_i^{-\beta}}{kT_0^4} \tag{22}$$

$$= \left(\frac{8}{3(1+w)} - 1\right)\frac{t_0^{-\frac{8}{3(1+w)}}}{kT_0^4}t_i^{\left(\frac{8}{3(1+w)-1}\right)}. \tag{23}$$

In a cosmological setting, $T \propto a^{-1}$—with a the scale factor–that is, $T \propto t^{-\frac{2}{3(1+w)}}$ with $w = p/\rho$ the equation of

state parameter for the dominant fluid. The condition $\beta > 0$, or equivalently $\alpha > -1/4$, translates into $w > 5/3$. This value, greater than one, corresponds to "superstiff" matter. Although quite exotic, this behavior can be encountered in realistic models such as Horava-Lifshitz gravity (see [19]). In this case, even if the temperature is decreasing, the black hole mass diverges whatever its initial value. If $\alpha < -1/4$, the divergence is associated with the condition given by Eq. (23). The important point is that there always exists an initial mass beyond which the divergence does occur. In this sense, part of the parameter space is unavoidably unstable.

It is interesting to compare M_c with the Hubble mass M_H. At time t_i, the latter is of order $M_H \sim t_i$. The condition $M_c < M_H$ is always fulfilled when $\beta > 0$, which means that the considered black hole can form without any causality issue. On the other hand, if $\beta < 0$, the condition reads

$$t_i^{-\beta-1} < \frac{kT_0^4}{-\beta t_0^{\beta-1}}. \tag{24}$$

Let us define the critical crossing time

$$t_{cH} = \left(\frac{kT_0^4}{-\beta t_0^{\beta-1}}\right)^{\frac{-1}{\beta+1}}. \tag{25}$$

If $-\beta - 1 > 0$ (that is, $\alpha < 1/2$ or $w < 1/3$), the condition translates into $t_i < t_{cH}$, whereas if $-\beta - 1 < 0$ (that is, $\alpha > 1/2$ or $w > 1/3$), the condition translates into $t_i > t_{cH}$. For content with an equation of state softer than radiation, unstable (with diverging mass) black holes can be causally formed early in the history of the Universe, while for a background equation of state stiffer than radiation, unstable black holes can form late in the cosmological history. If $w = 1/3$ exactly, the condition simply reads $-\beta t_0^{\beta-1} < kT_0^4$. In all cases a part of the parameter space leads to diverging black holes. The detailed investigation of hierarchy between the horizon of the black hole and the cosmological horizon (particle horizon before the inflation, Hubble horizon after the inflation) has been extensively studied, e.g., in [3,6,11], and we will not repeat the analysis here.

Interestingly, the naively expected behavior ($M \to \infty$ for $t \to \infty$), which is also the one that has generated some interest from the point of view of general relativity, happens only for a zero measure parameter space. In the general case where the instability is avoided, the mass tends to a finite asymptotic value M_∞ in the remote future:

$$M_\infty = \left(\frac{1}{M_i} + \frac{kT_0^4 t_i^\beta}{\beta t_0^{\beta-1}}\right)^{-1}. \tag{26}$$

This is illustrated in Fig. 1. Depending on the initial mass and on the speed at which the temperature of the thermal bath decreases, the evolution of the black hole mass corresponds to one of the two cases displayed.

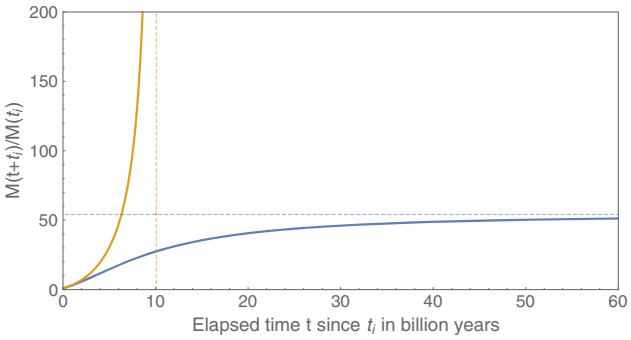


FIG. 1. Evolution of the mass of a black hole in a perfect fluid with $w = 0$. The blue curve corresponds to a black hole with an initial mass slightly lower than the critical mass M_c, whereas the yellow curve corresponds to a mass slightly higher. Dashed lines represent asymptotic behaviors. No "in-between" dynamics is possible.

IV. SWITCHING ON THE EVAPORATION

Obviously, the full picture requires one to also take into account the Hawking evaporation [20] which is mandatory to have a consistent thermodynamical understanding (although the truly dynamical horizon is still ignored). The evolution equation now reads

$$\frac{dM}{dt} = kT^4M^2 - \gamma M^{-2}. \tag{27}$$

In principle, the γ parameter depends on the mass M as the number of degrees of freedom available increases with the black hole temperature $T_{\rm BH} = 1/(8\pi M)$. New channels are opened each time the temperature becomes higher than the rest mass of a given particle species. For this study it is clearly sufficient to assume γ to be constant [21]. Its numerical value can be straightforwardly calculated by integrating the Hawking spectrum (multiplied by the energy Q of the emitted particle):

$$\frac{dM_{\rm evap}}{dt} = -\int \frac{\Gamma}{2\pi}\left(e^{\frac{Q}{T_{\rm BH}}} - (-1)^{2s}\right)^{-1} Q dQ, \tag{28}$$

where s is the spin of the particle and $\Gamma = Q^2\sigma/\pi$ is the greybody factor.

The Hawking evaporation has an immediate consequence. If the black hole happens to be in an initial state where the evaporation dominates over the absorption, it automatically remains in this regime: the mass decreases and the M^{-2} term becomes more and more important with respect to the absorption one. The situation considered at the end of the first section is therefore purely academic: in practice, if the size of the black hole is smaller than the wavelength of the surrounding radiation, the black hole

behavior is dominated by the evaporation. Its mass vanishes in a time

$$t_{\text{evap}} = \frac{M_i^3}{3\gamma}. \tag{29}$$

This means that, in a thermal bath, a black hole is anyway unstable: its mass either diverges or reaches zero[2] in a finite amount of time. The Hawking time scales as M_i^3 and the divergence time scales as M_i^{-1}.

The reciprocal of the previous statement is, however, not true in an expanding universe. If the absorption initially dominates, this does not mean that it will remain so forever. Actually, it *cannot* remain true in most cases. If the cooling of the Universe is "usual" (e.g., associated with an equation of state $w = 1/3$ or $w = 0$) and the initial black hole mass is such that the divergence is avoided ($M_i < M_c$), the evaporation will inevitably dominate at some point. In the absorption-dominated regime, the mass of the black hole increases and its temperature decreases. However, as previously shown, the mass necessarily tends to a finite value in the remote future. This means that, as far as absorption is concerned, $dM/dt \to 0$. On the other hand, the mass variation due to the evaporation has a constant asymptotic value $dM/dt = -\gamma M_\infty^{-2}$. After some time t_* this latter term will dominate. This basically means that the naive idea according to which, in a dynamical thermal bath, a black hole either absorbs radiation or evaporates depending only on the respective (initial) values of the bath and of the black hole temperatures is wrong. If the initial temperature of the black hole is smaller than the initial temperature of the background, the black hole temperature will first decrease (it will grow by absorption), in accordance with the usual view. However (if the dynamics is not the diverging one), after a finite time t_*, the temperature will start to increase again (the black hole will shrink by evaporation). Once the evaporation dominates, the behavior will not reverse until the disappearance of the black hole. The minimum temperature reached by the black hole is

$$T_{\min} = \frac{1}{8\pi}\left(\frac{1}{M_i} + \frac{kT_0^4 t_i^\beta}{\beta t_0^{\beta-1}}\right). \tag{30}$$

The transition time t_* is such that

$$\sqrt{\frac{\gamma}{k}}\frac{t_0^{2\alpha}}{T_0^2 t_*^{2\alpha}}\left(\frac{1}{M_i} - \frac{kT_0^4}{(4\alpha+1)t_0^{4\alpha}}(t_*^{4\alpha+1} - t_i^{4\alpha+1})\right) \tag{31}$$

$$= (M_i^3 - 3\gamma(t_* - t_i))^{\frac{1}{3}}. \tag{32}$$

If the dynamics of the background spacetime is, however, such that the black hole mass should diverge due to absorption, the situation is quite subtle. For most of the parameter space, the black hole just grows and the evaporation does not play any role. But it can be shown that for highly tuned initial conditions the second time derivative of the mass can vanish and the evaporation can overcome the growth. This corresponds to a very particular case worth being mathematically pointed out but most probably without any phenomenological consequences.

This shows why black holes in a thermal bath are *unstable*. The mass either diverges or vanishes in a finite amount of time. Figure 2 illustrates the typical behavior in the case where the absorption initially dominates without being diverging: the mass first increases quite fast, then remains close to its asymptotic value for most of the evolution, and then decreases until it completely vanishes. The very highly tuned case where the mass would diverge without evaporation but where the dynamics is finally overcome by the Hawking effect is exhibited in Fig. 3.

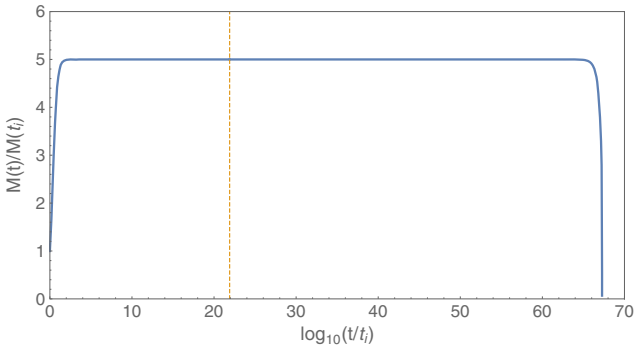


FIG. 2. Evolution of the mass of a black hole in a perfect fluid with $w = 0$ taking into account Hawking radiation. The blue curve corresponds to a black hole with an initial mass slightly lower than the critical mass M_c, whereas the yellow dashed line corresponds to the time when evaporation starts to dominate.

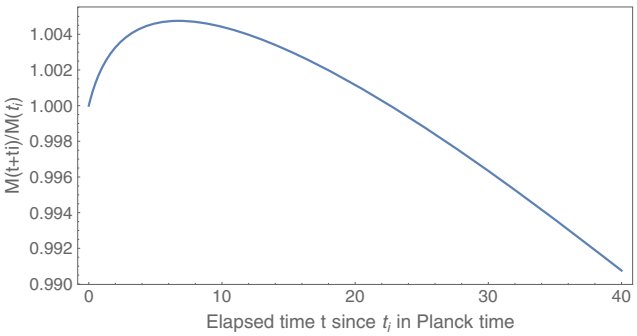


FIG. 3. Very special case where the absorption initially dominates in an expanding space—in a regime such that the mass should diverge without evaporation—but where the Hawking effects still finally overcome the evolution.

[2]There are countless arguments and models in quantum or extended gravity to avoid the naked singularity—see references in [22]—but this is not the point we wish to make here.

Let us summarize:

(i) If the initial black hole temperature is higher than the background temperature, the black hole simply evaporates and vanishes.
(ii) If the initial black hole temperature is smaller than the background temperature and $\alpha > -1/4$ (or $\alpha < -1/4$ and $M > M_c$), the black hole mass generically diverges (however, there exists a tiny set of parameters—corresponding to initial conditions highly tuned close to the critical values—where the evaporation finally dominates).
(iii) If the initial black hole temperature is smaller than the background temperature (and $\alpha < -1/4$ and $M < M_c$), the black hole mass vanishes after having reached a "plateau" where it stayed for the vast majority of its lifetime.

The important feature is that the infinite or vanishing mass is always reached in a finite amount of time. Let us mention, although this is not the point of this study, that even if some kind of remnant or relic is formed at the end of the evaporation, they are anyway usually expected to disappear at some point [23].

V. THE DE SITTER CASE

When the scale factor expands exponentially, the temperature evolution can be written as

$$T = T_0 e^{-\kappa t}. \tag{33}$$

The mass evolution (when the black hole initial mass is such that the absorption dominates over evaporation) then reads

$$\frac{1}{M} = \frac{1}{M_i} + \frac{kT_0^4}{4\kappa}(e^{-4\kappa t} - e^{-4\kappa t_i}). \tag{34}$$

As expected, no divergence occurs in this case and the mass always reaches its asymptotic value,

$$M_\infty = \left(\frac{1}{M_i} - \frac{kT_0^4}{\kappa} e^{-4\kappa t_i}\right)^{-1}. \tag{35}$$

Once again, the evaporation unavoidably dominates at some time t_* such that

$$\sqrt{\frac{\gamma}{k}} kT_0^{-2} e^{2\kappa t_*}\left(\frac{1}{M_i} + \frac{kT_0^4}{4\kappa}(e^{-4\kappa t_*} - e^{-4\kappa t_i})\right) = \tag{36}$$

$$(M_i^3 - 3\alpha(t_* - t_i))^{\frac{1}{3}}. \tag{37}$$

The de Sitter horizon is endowed with a temperature

$$T_{\mathrm{dS}} = \frac{1}{2\pi}\sqrt{\frac{\Lambda}{3}}, \tag{38}$$

where Λ is the cosmological constant. This, however, does not change the picture in any noticeable way as the black hole temperature always remains higher than the de Sitter temperature.

VI. THE CONTRACTING UNIVERSE CATASTROPHE

Quite a lot of theories beyond general relativity predict a cosmological bounce instead of the big bang singularity (see e.g., [24–26] for reviews). This is even possible in general relativity without exotic matter [27,28]. In such models, the Universe was contracting before entering the current expanding branch. If space was, before the bounce, filled with black holes and radiation,[3] the catastrophic growth of black hole masses that we have established in a constant-temperature bath will even be worsened.

It makes sense to evaluate the time taken by a black hole to reach its absorption singularity. Contracting spaces are known to exhibit some paradoxes (see [29,30]), and it is interesting to compare how long it takes for the black hole to become unstable when compared to the time required for the background radiation to reach the Planck density (triggering quantum gravity effects). If the contracting branch is filled with relativistic matter, the dynamics of the black hole reads

$$\frac{1}{M_i} - \frac{1}{M} = kT_0^4 t_0^2\left(\frac{1}{t_i} - \frac{1}{t}\right), \tag{39}$$

where t is now negative and the conventions are the same as previously. (Contrary to what is sometimes believed, if a cosmological variable scales as t^q in an expanding universe, it will *not* behave as t^{-q} in a contracting one.) Although this expression is formally the same as in the expanding universe, it does induce, due to the fact that t is now negative, a far *faster* divergence, all the other parameters being the same, as it can be seen in Fig. 4. The background energy density varies as $\rho \propto (-t)^{-2}$, and Fig. 4 shows that the black hole mass divergence can (depending on the initial conditions) be reached before the energy density becomes Planckian. This means that the singularity resolution provided by the bounce in quantum gravity models (see e.g., [31]) might not solve this specific problem which should be seriously considered. To be more illustrative, let us once again consider a prebounce universe similar to our expanding one, i.e., such that the temperature was $T \sim 3 \times 10^3$ K at time $t \sim -3 \times 10^5$ years before the bounce. Then a stellar mass black hole with an initial mass $M \sim 10\ \mathrm{M}_\odot$ would diverge at time $t_d \sim -10^{-5}$ s

[3]There are *no* motivations for this assumption often made in the framework of bouncing models. This is only justified by the desire of studying a symmetric situation that might be the less unjustified guess. It remains hazardous from the causality point of view.

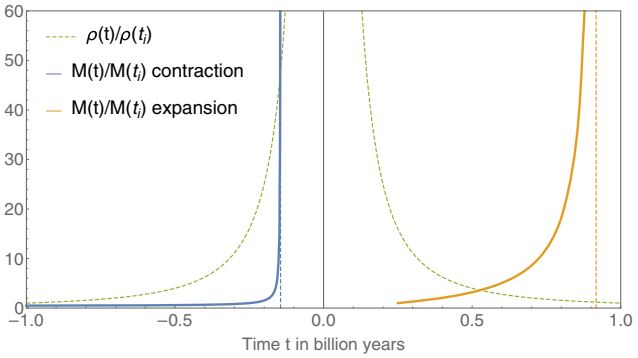


FIG. 4. The blue curve in the left panel and the yellow curve in the right panel represent the time evolution of black holes with the same initial masses in a radiation-dominated universe which is, respectively, contracting (left) and expanding (right). The dashed green curves correspond to the evolution of the background density.

when the temperature was $T(t_d) \sim 10^{12}$ K, well before the Planck era.

This also leads to an important remark. Let us consider the contracting solution to the Friedmann equation

$$H = -\sqrt{\frac{8\pi G\rho}{3}}, \tag{40}$$

where H is the Hubble parameter. This describes a universe which reaches a singularity at $t = 0$. It is, however, clear from Eq. (39) that the point $1/M = 0$ will inevitably be reached before $t = 0$, whatever the initial conditions. The black hole instability precedes the big crunch singularity. As long as there is a black hole in a contracting universe filled with radiation, this phenomenon will take place before the breakdown of the smooth background evolution.

In a contracting space, if the absorption initially dominates, the evaporation obviously never plays any role and the mass diverges. What, however, happens if the evaporation is dominant at the formation time? One might naïvely expect that a severe competition takes place between absorption (which increases at a given mass due to the increase of the temperature) and evaporation (which increases due to the decrease of the mass). This, however, is not the case in general. One can straightforwardly show that the absorption overcomes the evaporation if the background temperature satisfies

$$T_{\text{back}} > \left(\frac{\gamma}{k}\right)^{\frac{1}{4}} \frac{1}{M}. \tag{41}$$

The background temperature, however, diverges at $t = 0$. This means that—once again except for a tiny and mostly irrelevant part of the parameter space corresponding to fine-tuned initial conditions—the Hawking evaporation remains dominant if the Hawking time is smaller than the time available before the crunch (or the bounce).

VII. EXTRA DIMENSIONS

Gravity with more than three spatial dimensions is far richer. This is the case for mathematical reasons: the rotation group $SO(D)$ has Cartan subgroup $U(1)^N$ with $N = E[D/2]$. This is also the case for physical reasons: the radial falloff of the Newtonian potential scales as r^{D-2}, whereas the centrifugal barrier does not depend on D. Let us consider a spacetime with D spatial dimensions and define

$$\mu = \frac{16\pi M}{(D-1)\Omega_{D-1}}, \tag{42}$$

where $\Omega_{d-1} = 2\pi^{D/2}/\Gamma(\frac{D}{2})$ is the area of a unit $(D-1)$-sphere. The generalized Schwarzschild metric reads [32,33]

$$ds^2 = -\left(1 - \frac{\mu}{r^{D-2}}\right)dt^2 + \frac{dr^2}{1 - \frac{\mu}{r^{D-2}}} + r^2 d\Omega_{D-1}^2. \tag{43}$$

It is straightforward to show that, in this framework, the time evolution of the black hole mass becomes (in a constant temperature background)

$$\frac{dM}{dt} \propto M^{\frac{D-1}{D-2}}. \tag{44}$$

This diverges at finite time if $(D-1)/(D-2) > 1$, which is always true. However, the higher the number of extra dimensions, the less drastic the divergence becomes. Care should be taken when using Planck units with $D > 3$. The full expression of the radius of the horizon is

$$R_D = \frac{1}{\sqrt{\pi} M_*}\left(\frac{M}{M_*}\right)^{\frac{1}{D-2}}\left(\frac{8\Gamma(\frac{D}{2})}{D-1}\right)^{\frac{1}{D-2}}. \tag{45}$$

In this expression M_* is the fundamental Planck scale and *not* the usual $(3+1)$-dimensional one. The Hawking temperature of the black hole reads

$$T_{\text{BH}_D} = \frac{D-2}{4\pi R_D}. \tag{46}$$

The behavior is qualitatively the same as in the four-dimensional case. Whatever the number of extra dimensions, the black hole mass diverges at finite time in a static thermal bath. If, however, the temperature of the bath decreases with time and the black hole is not too large at the initial time, the black hole mass will also tend to an asymptotic value. The dynamics will then, at some point, be dominated by the evaporation and the mass will vanish.

VIII. ANALOGIES

The simple approach we have chosen here is obviously oversimplified. The dynamics of Eq. (3) reads, for the radius of the black hole,

$$\frac{dR}{dt} = \frac{72}{5}\pi^2 T^4 R^2. \tag{47}$$

This means that $dR/dt \to \infty$ as the critical time is approached. This is also true in the other backgrounds considered as far as the instability occurs. Is this physical? We will come back to this question at a more fundamental level in the next sections, and we remain here at the level of the simple consistency of the toy model. Obviously, if a particle of total energy ϵ falls into a large black hole of radius R, it does not make sense to assume that the horizon jumps instantaneously to a perfect sphere whose radius is simply increased by 2ϵ. The intricate relaxation procedure through quasinormal modes should, in principle, be considered to accurately describe the time evolution of the shape of the horizon. For obvious causality reasons the horizon cannot globally vary in shape "faster than light" as a consequence of a local deformation. The situation considered in this work is, however, different. One deals with a nearly continuous background of ingoing energy.

The very notion of a event horizon is, in nature, global and teleological (see e.g., [34]). It is *the boundary of future null infinity*. This, however, does not prevent the change of the horizon shape, at $t + dt$, due to an incoming flux of energy on a black hole preexisting at time t, to be *locally* determined. Associated conceptual and technical considerations were studied, e.g., in [35,36], following [37]. The growth of the horizon is locally driven by the global past history *and* the recent local events. There is, then, no reason to discard a "faster than light" expansion of the horizon. This is no stranger than a wave crest moving arbitrarily fast: no matter or information is transported. The region from which (in the black hole case) light rays can escape to infinity might, without violating causality, shrink faster than light. This eventually happens when a black hole is fed by a continuous and homogeneous flux of energy.

Let us consider for example analog black holes (see e.g., [38]). The position of the horizon is determined both by nonlinear effects and by the environment, whereas the speed of sound is an entirely linear property. The speed of the first can therefore be higher than the speed of the second. Concretely, this would clearly be possible with a two-component Bose-Einstein condensate. The sound speed for the two species, respectively c_1 and c_2, are *a priori* different. If they do interact, the position of the horizon for the first will depend on the density of the second (and the other way around). As there are nonlinear solutions—such as grey solitons—moving close to sound speed, there can be, from the viewpoint of the first species, a horizon moving at a speed close to c_2, which can be greater than c_1.

With the obvious limitations of the metric used, the growth of the horizon at an arbitrary high speed is not inconsistent in the considered framework. This, by the way, also happens when considering the Hawking evaporation where $dR/dt \to \infty$ when $M \to 0$. In this case, one could argue that the semiclassical formula should be modified by quantum gravity effects in this regime but the divergence of the speed is not the reason why a better description should be searched for. The phenomenon underlined in this work has no reason not to be possibly real. Although a kind of "bubble nucleation" (at the speed of light) similar to a tunneling effect from a false vacuum to the true vacuum (see e.g., [39]) would also make sense, the considered process can be faster and even more catastrophic.

Finally, it is important to notice that the phenomenon exhibited in this article is not unique in physics. One could first think about the disorder correlator. Functional renormalization shows that its curvature explodes at a finite scale ℓ_c (developing a cusp at the origin). This can be interpreted through shocks and avalanches [40].

Closer to our situation is probably the case of cubic autocatalysis (that is, of order $n = 2$). In such chemical processes, one of the reaction products is also a catalyst for the same reaction. This reads $2A + B \to 3A$. If the rate of the reverse reaction is vanishing (the rate of the forward reaction being r) and if the concentration x_B of the species B is kept constant, the number of A particles evolves (in the large N limit) as $\partial N/\partial t = r x_b N^2$. This is the same formal equation as the one we have considered in this work for a constant temperature bath. In a way, the black hole corresponds to species A.

IX. THE BLACK HOLE AS AN UNDERDENSITY

Among others, a curious aspect of the situation considered in this article is the following. When it grows by absorption of the surrounding radiation, a black hole might end being *less* dense that the medium in which it is embedded. Obviously, the Schwarzschild solution, which is a vacuum solution to Einstein's equation, is not appropriate to handle this situation. The very meaning of the Arnowitt Deser Misner (ADM) mass defining the families of solutions is ill-defined. It is nevertheless worth considering a simple thought experiment to understand whether it is possible to make sense out of this. An interesting case is the one of thick shells collapsing toward a preexisting black hole [41]. The exact solution was first derived using comoving coordinates,

$$ds^2 = d\tau^2 - e^{\tilde{\omega}(R,\tau)} dR^2 - e^{\omega(R,\tau)}(d\theta^2 + \sin^2\theta d\phi^2), \tag{48}$$

where there is only one nonvanishing component in the stress-energy tensor. Using the machinery developed by Oppenheimer [42] and appropriate matching conditions,

the result can be transformed to the usual coordinates and is easy to interpret. If several shells collapse on the black hole, one can show that all the incoming matter but the outermost layer of the last one do cross the horizon [41] (exact solutions for null fluid collapse were obtained in [43]). This does not rely on any assumption about the respective densities of the black hole and of the shells. The latter can be denser than the former (the mean density of a black hole is anyway not a physically relevant quantity). It should be kept in mind that we do not consider here accretion of dust, which would be severely impacted, but absorption of radiation whose penetration in the black hole is (mostly) not driven by gravitational effects.

Obviously, a black hole does not form in a static homogeneous space without a triggered gravitational instability. Otherwise, one would be led to the absurd conclusion that black holes spontaneously appear in any large enough static space as a region of size R falls inside its own gravitational radius as long as $R > \sqrt{3/(8\pi\rho)}$. This is clearly wrong as this analysis would only be correct for the vacuum case. However, once the black hole is formed, there is no reason to discard the possibility that its mean density becomes smaller than the average one of the surrounding medium. Although unusual, this situation is not impossible.

X. FULL DYNAMICS

Two crucial dynamical aspects, beyond the scope of this article, should be accounted for before any reliable conclusion can be drawn: the expansion (or contraction) of the Universe beyond its consequence on the radiation density and the evolution of the black hole horizon itself.

Many works are devoted to both aspects but it remains hard to get a clear and noncontroversial picture. Black holes in a radiation-dominated universe were, in particular, studied in detail in [44]. The metric obtained is

$$ds^2 = \left(1 - \frac{r_g}{r} - \frac{1}{2} + \frac{t}{2\sqrt{t^2 - r^3}}\right) dt^2 \tag{49}$$

$$- \left(\left(1 - \frac{r_g}{r}\right)^{-1} - \frac{1}{2} + \frac{t}{2\sqrt{t^2 - r^3}}\right) dr^2, \tag{50}$$

in curvature coordinates. This line element develops a curvature singularity at $r = r_g$ and is strictly valid only for $r_g < r < t$, that is, between the black hole horizon and the cosmological horizon. The basic behavior, building on the analysis of [45], is, however, in agreement with the picture drawn in this article. Many situations beyond the radiation-dominated background were also considered (see e.g., [46]—pushed further in [47]—which is very useful for deriving generalized metrics in a nonvacuum environment). Some interesting hints can already be found with the Vaidya metric, describing the absorption (or emission) of null dust. The very meaning of the associated horizons is, however, still debated [48]. The most promising approach to deal with dynamical black holes in the framework we consider is probably the one advocated in [49]. Event horizons are indeed not well suited for evolving black holes, and quite a lot can be inferred from the evolution and the apparent horizon in Painlevé-Gullstrand coordinates (that are regular at the horizon).

Although the expansion (or contraction) of the Universe is not the most important aspect for the points we are making here, it would in principle make sense to generalize this study by considering metrics found e.g., in [50,51]. In particular, when the mass depends on time in a flat Friedmann-Lemaître-Roberston-Walker (FLRW) background, the line element—generalizing the McVittie results—reads [52]

$$ds^2 = -\frac{\left[1 - \frac{M(t)}{2a(t)r}\right]^2}{\left[1 + \frac{M(t)}{2a(t)r}\right]^2} dt^2 + a^2(t)\left[1 + \frac{M(t)}{2a(t)r}\right]^4 \times (dr^2 + r^2 d\Omega^2). \tag{51}$$

Interesting paths are also suggested in [53,54].

Two specifically interesting situations were recently considered in the case of evaporating black holes. The first one relies on the Thakurta metric [55]. The Thakurta spacetime is a generalized McVittie black hole with accretion. This solution approximately corresponds (at distances such that $M \ll R \ll 1/H$) to a Newtonian point particle with growing mass, the accretion rate being proportional to the Hubble rate. Important consequences were e.g., investigated in [56], deeply changing the LIGO bounds on primordial black holes. Several concerns were raised: the energy flux required for this specific mass growth seems nonphysical [57] and neither an event horizon nor a trapping horizon seems allowed [58]. Answers were provided in [57], adopting the foliation associated with the Kodama time. Other counterarguments were exhibited in [59]. Anyway, the peculiar mass evolution, $\dot{M} = HM$, associated with this spacetime is very different from the one of our study, as even the asymptotic behaviors do not coincide [3,17,44,45].

The second one is based on the Sultana-Dyer spacetime [60]. It is a Petrov type D metric describing a black hole embedded in a spatially flat FLRW universe with scale factor $a(t) \propto t^{2/3}$, generated by mapping the Schwarzschild timelike Killing field ξ^a into a conformal Killing field (for $\xi^a \nabla_a \Omega \neq 0$) [61]. In this framework, an exact model for evaporating primordial black holes in the cosmological spacetime was developed in [62], showing potentially important deviations with respect to the usual behavior. This approach, however, assumes a matter-dominated universe which not only is different from the one considered in this study but also is incompatible with the specific kind of accretion assumed here.

The situation is not fully clear. The "correct" horizon to consider highly depends on the physical properties being investigated. Apparent horizons are usually advocated for phenomenological purposes but their foliation-dependence problem remains mostly unsolved [63] (even though arguments in [57] show that a preferred foliation of the Thakurta spacetime does exist). There is no consensus [63], and the question of horizons in a dynamical framework is still an open one (interesting points are made in [64]).

In this work, we do *not* pretend to give a fully general exact and analytical solution for growing black holes in a radiation-dominated dynamical universe. This goal is still beyond available models. We simply focus on some specific features of the Schwarzschild metric in a photon background. This is obviously an oversimplification but it exhibits intriguing behaviors that deserve future investigations. As one can guess from the generalized McVittie metric (where $M(t)$ is the Hawking-Hayward quasilocal mass), it is probable that, at late times, the rate of growth of the black hole mass becomes comparable to the Hubble rate and the black hole becomes comoving [50]. The situation in which such a result is derived is, however, different from the one studied in this article and no firm conclusion can be reached.

XI. CONCLUSION

The behavior of black holes in a thermal bath is counterintuitive when first encountered. This is simply grounded in the very special mass/radius relation of black holes. For standard matter, the area scales as $dA \propto M^{-1/3}dM$, whereas for black holes, $dA \propto MdM$. This is the key point. For usual matter, the area variation induced by an incoming mass dM decreases with the mass M of the star (or whatever), whereas for black holes, it *increases* with M: the larger the black hole, the larger the area variation induced by the absorption of a quantum of given energy. This is the straightforward cause of the supercritical mass behavior. It is, however, mandatory to underline that all those "strange" features, when taken literally, might very well be artifacts of the Schwarzschild metric used beyond its domain of validity more than real physical effects. More work is needed and the aim of this article is mostly to invite the interested readers to investigate in more detail the curious behaviors presented here.

Within those (strong) limitations, we have established that for very simple reasons, the evolution of a black hole in a thermal bath is catastrophic. If the bath is at constant temperature, the black hole mass inevitably diverges at finite time (ignoring, of course, cosmological horizon issues). If the temperature of the bath increases—e.g., in a contracting space—the phenomenon is (obviously) even faster and the black hole singularity is reached before the background singularity. If the temperature decreases—e.g., in an expanding space—the mass can either diverge or vanish, in finite time once again. For tuned initial conditions a (questionable) self-similar solution is possible. In the vanishing case, the black hole spends most of its life on a long "plateau."

The popular expression "black hole bomb" [65] is strengthened and acquires a wider meaning. As a possible nongravitational development of this work, it would be interesting to investigate how the studied phenomenon can be viewed as a phase transition. Analogies with an appropriate Ising model should be fruitful.

ACKNOWLEDGMENTS

We thank Vivien Lecomte, Florent Michel, Antoine Rignon-Bret, and Léonard Ferdinand for very interesting remarks and helpful discussions.

[1] Y. B. Zel'dovich and I. D. Novikov, Astron. Zh. **43**, 758 (1966).
[2] Y. B. Zel'dovich and I. D. Novikov, Sov. Astron. **10**, 802 (1967).
[3] B. J. Carr and S. W. Hawking, Mon. Not. R. Astron. Soc. **168**, 399 (1974).
[4] B. J. Carr, Astrophys. J. **201**, 1 (1975).
[5] G. V. Bicknell and R. N. Henriksen, Astrophys. J. **225**, 237 (1978).
[6] B. J. Carr and A. Yahil, Astrophys. J. **360**, 330 (1990).
[7] H. Maeda, J.-i. Koga, and K.-i. Maeda, Phys. Rev. D **66**, 087501 (2002).
[8] D. Lin, B. J. Carr, and S. M. Fall, Mon. Not. R. Astron. Soc. **177**, 51 (1976).
[9] S. Hacyan, Astrophys. J. **229**, 42 (1979).
[10] M. S. Madsen, Classical Quantum Gravity **5**, 627 (1988).
[11] T. Harada and B. J. Carr, Phys. Rev. D **71**, 104009 (2005).
[12] T. Harada, H. Maeda, and B. J. Carr, Phys. Rev. D **74**, 024024 (2006).
[13] H. Maeda, T. Harada, and B. J. Carr, Phys. Rev. D **77**, 024023 (2008).
[14] M. Kyo, T. Harada, and H. Maeda, Phys. Rev. D **77**, 124036 (2008).
[15] B. J. Carr, T. Harada, and H. Maeda, Classical Quantum Gravity **27**, 183101 (2010).
[16] M. Kopp, S. Hofmann, and J. Weller, Phys. Rev. D **83**, 124025 (2011).
[17] B. J. Carr and T. Harada, Phys. Rev. D **91**, 084048 (2015).
[18] J. H. MacGibbon and B. R. Webber, Phys. Rev. D **41**, 3052 (1990).
[19] S. Kalyana Rama, Phys. Rev. D **79**, 124031 (2009).

[20] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975); **46**, 206(E) (1976).
[21] F. Halzen, E. Zas, J. H. MacGibbon, and T. C. Weekes, Nature (London) **353**, 807 (1991).
[22] A. Barrau, K. Martineau, F. Moulin, and J.-F. Ngono, Phys. Rev. D **100**, 123505 (2019).
[23] S. B. Giddings, Phys. Rev. D **46**, 1347 (1992).
[24] D. Battefeld and P. Peter, Phys. Rep. **571**, 1 (2015).
[25] M. Lilley and P. Peter, C. R. Phys. **16**, 1038 (2015).
[26] R. Brandenberger and P. Peter, Found. Phys. **47**, 797 (2017).
[27] A. Barrau, Eur. Phys. J. C **80**, 579 (2020).
[28] C. Renevey, A. Barrau, K. Martineau, and S. Touati, J. Cosmol. Astropart. Phys. 01 (**2021**) 018.
[29] A. Barrau, K. Martineau, and F. Moulin, Phys. Rev. D **96**, 123520 (2017).
[30] A. Barrau and L. Linsefors, J. Cosmol. Astropart. Phys. 12 (2014) 037.
[31] A. Ashtekar and P. Singh, Classical Quantum Gravity **28**, 213001 (2011).
[32] R. Emparan and H. S. Reall, Living Rev. Relativity **11**, 6 (2008).
[33] P. Kanti, Int. J. Mod. Phys. A **19**, 4899 (2004).
[34] S. Bhattacharya and S. Shankaranarayanan, Mod. Phys. Lett. A **32**, 1750186 (2017).
[35] E. Kopteva, P. Jalůvková, I. Bormotova, and Z. Stuchlík, Astrophys. J. **866**, 98 (2018).
[36] S.-X. Zhao and S.-N. Zhang, Chin. Phys. C **42**, 085101 (2018).
[37] W. Israel, Nuovo Cimento B **44S10**, 1 (1966); **48**, 463(E) (1967).
[38] X.-H. Ge and S.-J. Sin, J. High Energy Phys. 06 (2010) 087.
[39] Y. Chen and A. Vilenkin, J. Cosmol. Astropart. Phys. 03 (2021) 035.
[40] K. J. Wiese, arXiv:2102.01215.
[41] Y. Liu and S. N. Zhang, Phys. Lett. B **679**, 88 (2009).
[42] J. R. Oppenheimer and H. Snyder, Phys. Rev. **56**, 455 (1939).
[43] V. Husain, Phys. Rev. D **53**, R1759 (1996).
[44] E. Babichev, V. Dokuchaev, and Y. N. Eroshenko, Astron. Lett. **44**, 491 (2018).
[45] E. Babichev, V. Dokuchaev, and Y. Eroshenko, Phys. Rev. Lett. **93**, 021102 (2004).
[46] V. V. Kiselev, Classical Quantum Gravity **20**, 1187 (2003).
[47] Y. Heydarzade, H. Moradpour, and F. Darabi, Can. J. Phys. **95**, 1253 (2017).
[48] A. B. Nielsen, Galaxies **2**, 62 (2014).
[49] A. B. Nielsen and M. Visser, Classical Quantum Gravity **23**, 4637 (2006).
[50] V. Faraoni and A. Jacques, Phys. Rev. D **76**, 063510 (2007).
[51] C. Gao, X. Chen, V. Faraoni, and Y.-G. Shen, Phys. Rev. D **78**, 024008 (2008).
[52] V. Faraoni, C. Gao, X. Chen, and Y.-G. Shen, Phys. Lett. B **671**, 7 (2009).
[53] E. Abdalla, N. Afshordi, M. Fontanini, D. C. Guariento, and E. Papantonopoulos, Phys. Rev. D **89**, 104018 (2014).
[54] D. C. Guariento, M. Fontanini, A. M. da Silva, and E. Abdalla, in *Proceedings of the 13th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories* (2015), pp. 1398–1400.
[55] S. Thakurta, Indian J. Phys. **55B**, 304 (1981).
[56] C. Boehm, A. Kobakhidze, C. A. J. O'hare, Z. S. C. Picker, and M. Sakellariadou, J. Cosmol. Astropart. Phys. 03 (2021) 078.
[57] A. Kobakhidze and Z. S. C. Picker, arXiv:2112.13921.
[58] T. Harada, H. Maeda, and T. Sato, arXiv:2106.06651.
[59] C. Boehm, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker, and M. Sakellariadou, arXiv:2105.14908.
[60] J. Sultana and C. C. Dyer, Gen. Relativ. Gravit. **37**, 1347 (2005).
[61] V. Faraoni, Phys. Rev. D **80**, 044013 (2009).
[62] S. Xavier, A. Sunny, and S. Shankaranarayanan, arXiv:2110.14379.
[63] V. Faraoni, Universe **4**, 109 (2018).
[64] L. Vanzo, G. Acquaviva, and R. Di Criscienzo, Classical Quantum Gravity **28**, 183001 (2011).
[65] W. H. Press and S. A. Teukolsky, Nature (London) **238**, 211 (1972).

1.4 Beyond Schwarzschild for cosmological black holes

As we have seen in the previous section, the Schwarzschild solution is not suited to describe the evolution of a black hole horizon when it is evolving too quickly. This was of course expected since a Schwarzschild black hole is supposedly static. To go beyond, we need to relax the static or vacuum assumption for the metric. Several approaches have been explored [33] and we will discuss two of them that have recently been put forward. The *generalized McVittie* metric [34], which describes a central mass in an imperfect fluid and has the advantage of being relatively easy to work with. We will also discuss an other more general approach developed in Ref. [35], which is more complicated, but allows us to work with perfect fluids.

1.4.1 The case of the generalized McVittie metric

A common metric that is used to describe cosmological black holes is called the generalized McVittie (gMcVittie) metric [33, 34, 36–38]

$$\mathrm{d}s^2 = -\frac{B^2(T,\bar{R})}{A^2(T,\bar{R})}\mathrm{d}T^2 + a^2(T)A^4(T,\bar{R})\left(\mathrm{d}\bar{R}^2 + \bar{R}^2\mathrm{d}\Omega\right) , \tag{1.62}$$

where

$$A(T,\bar{R}) = 1 + \frac{M(T)}{2\bar{R}} , \quad B(T,\bar{R}) = 1 - \frac{M(T)}{2\bar{R}} , \tag{1.63}$$

and is directly inspired from the older McVittie metric [39]. While the latter metric suffers from physical and conceptual problems [40], it was shown that the gMcVittie metric can in fact describe a spherically symmetric mass with an apparent horizon bathed in an imperfect fluid. This metric contains two arbitrary functions $a(T)$ and $M(T)$, which depend only on the time coordinates. According to [38] and reference therein, the domain of validity of the gMcVittie metric should be black holes embedded in an imperfect fluid, whose stress-energy tensor is defined as

$$T_{\mu\nu} = \big(\rho(T,R) + p(T,R)\big)\, U_\mu U_\nu + p(T,R) g_{\mu\nu} + f_\mu(T,R)U_\nu + f_\nu(T,R)U_\mu , \tag{1.64}$$

where U_μ is the velocity field of the fluid and one can interpret ρ, p and f_μ as the density, pressure and energy flux, respectively, of the fluid as measured by the observer in the reference frame of the fluid. The energy flux must satisfy $f_\mu U^\mu = 0$. Before moving further, we note that for large $\bar{R}$ the gMcVittie metric (1.62) reduces to the FLRW metric and the function $a(T)$ can be interpreted as the scale factor. While the perfect fluid, often used in cosmology, is a special case of the fluid described above with $f_\mu = 0$, the gMcVittie metric can only account for perfect fluids satisfying $p = -\rho$, representing a cosmological constant [33]. This is the case even if one considers that the fluid is not at rest with respect to the Eulerian observer. In this section, we quickly review from [34] how the black hole and cosmological horizons evolve. We only look at the case where the energy flux is of the form $f^\mu = (0, q(T,R), 0, 0)$ and the velocity field is following the Eulerian observer $U_\mu = (-1,0,0,0)\sqrt{-g_{00}}$.

First, we assume that we recover the FLRW behaviour far from the black hole, which

means we can set

$$a(T) \propto \sqrt{T}\,, \tag{1.65}$$

we chose to have the big bang singularity at $T = 0$. This comes from the fact that we recover the FLRW metric for $\bar{R} \to \infty$ and $a(T)$ only depends on the time coordinate. Furthermore, the stress-energy tensor of radiation needs to be traceless [41], which gives us the usual equation of state $p(T, R) = \rho(T, R)/3$ even though we are considering an imperfect fluid. The Einstein field equations give us the behaviour of the density and the energy flux as well as the field equation for $M(T)$, namely

$$q(T, \bar{R}) = -\frac{\dot{M}a + M\dot{a}}{4\pi \bar{R}^2 a^3 A^6} \tag{1.66}$$

$$\rho(T, \bar{R}) = \frac{3A^2}{B^2}\left(\frac{\dot{a}}{a} + \frac{\dot{M}}{\bar{R}A}\right)^2 \tag{1.67}$$

$$\frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{\dot{a}}{a} + \frac{\dot{M}}{\bar{R}A}\right) + \left(\frac{\dot{a}}{a} + \frac{\dot{M}}{\bar{R}A}\right)\left(2\left(\frac{\dot{a}}{a} + \frac{\dot{M}}{\bar{R}A}\right) + \frac{\dot{M}}{\bar{R}AB}\right) = 0\,. \tag{1.68}$$

Since both M and a only depend on T, the only solution that can satisfy (1.68) for all $\bar{R}$ is $M(T) = M_0$, with M_0 constant. Introducing this solution in (1.68) gives the the Raychaudhury equation for the FLRW universe. This was expected since $a(T)$ was interpreted as being the usual scale factor. This result does not mean that the central black hole has a constant mass. In fact, it was discussed in [36] that the relevant quantity describing the mass of the black hole is $\tilde{M}(T) = m(T)a(T) = M_0 a(T)$. Finally, it was shown in [34] that one can find the apparent horizons $\bar{R}_H(T)$ of the system with the equations

$$1 - \frac{2M_0 a(T)}{\bar{R}_H(T)} \pm \frac{\dot{a}(T)}{a(T)}\bar{R}_H(T) = 0\,, \tag{1.69}$$

where the $\pm$ sign corresponds to the contracting or expanding cases, respectively. For illustrative purposes, we show the evolution of the horizons for the contracting and expanding branches on Fig. 1.2, with $M_0 = 0.2$ and $a(T) = \sqrt{T}$. We see that in the contracting phase both horizons shrink, except right before their merger, where the black hole horizon slightly grows. In the case of an expanding universe, we have the exact opposite behaviour.

1.4.2 Dynamical apparent horizons in Painlevé-Gullstrand coordinates

The Schwarzschild solution being unsuitable to describe dynamical horizons and the FLRW metric unfitted for inhomogeneous models, we begin by simply assuming spherical symmetry. We start again with the most general Lorentzian metric then reads [10, 41]:

$$\mathrm{d}s^2 = g_{\tilde{t}\tilde{t}}(\tilde{t}, r)\mathrm{d}\tilde{t}^2 + g_{\tilde{t}r}(\tilde{t}, r)\mathrm{d}\tilde{t}\mathrm{d}r + g_{rr}(\tilde{t}, r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.70}$$

with arbitrary functions $g_{\tilde{t}\tilde{t}}(\tilde{t}, r) < 0$, $g_{\tilde{t}r}(\tilde{t}, r)$ and $g_{rr}(\tilde{t}, r) > 0$, together with $\mathrm{d}\Omega^2 = \mathrm{d}\phi^2 + \sin^2(\phi)\mathrm{d}\theta^2$. In order to rewrite the metric in Painlevé-Gullstrand coordinates, we need to change the time coordinate, $\tilde{t} \to t(\tilde{t}, r)$, such that the coefficient in front of $\mathrm{d}r^2$ becomes

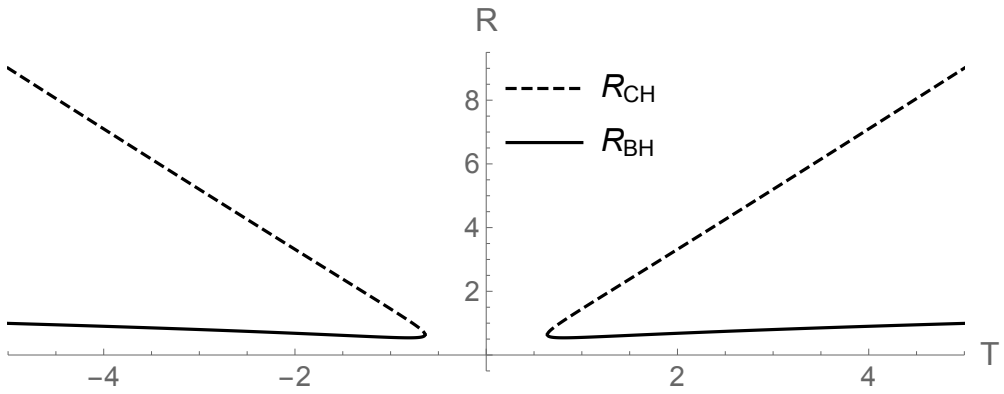


Figure 1.2: Evolution of a black hole R_{BH} and cosmological $_{CH}$ horizon in a contracting and expanding universe, as predicted by the generalized McVittie approach. The black hole horizon can increase or decrease with time while the cosmological horizon is monotonic.

unity. For better comparison with previous works, in particular with Ref. [35], we rewrite the metric under the form

$$\mathrm{d}s^2 = -\left(c^2(t,r) - v^2(t,r)\right)\mathrm{d}t^2 + 2v(t,r)\mathrm{d}r\mathrm{d}t + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,. \tag{1.71}$$

In the following, we will refer to the metric (1.71) as the PG metric. Note that the radial coordinate r is what we called the Schwarzschild radial coordinate. This metric can be constructed with the functions $c(\tilde{t},r)$, $v(\tilde{t},r)$ and $t(\tilde{t},r)$ satisfying the system

$$g_{\tilde{t}\tilde{t}} = (c^2 - v^2)\left(\frac{\partial t}{\partial \tilde{t}}\right)^2, \tag{1.72}$$

$$g_{\tilde{t}r} = 2\left(v + (c^2 - v^2)\frac{\partial t}{\partial r}\right)\frac{\partial t}{\partial \tilde{t}}, \tag{1.73}$$

$$g_{rr} = 1 - 2v\frac{\partial t}{\partial r} + v^2\left(2\frac{\partial t}{\partial r} - 1\right)\frac{\partial t}{\partial r} + c^2\left(1 - 2\frac{\partial t}{\partial r}\right)\frac{\partial t}{\partial r}, \tag{1.74}$$

which is determined up to boundary conditions for $t(\tilde{t},r)$.

In the effort to describe evolving black hole horizons, it proved useful to substitute

$$v(t,r) = \pm c(t,r)\sqrt{2m(t,r)/r}, \tag{1.75}$$

where $m(t,r)$ is a newly defined positive function, as it is the case in Ref. [35]. In the latter reference, the $+1$ convention in $v(t,r)$ was chosen for the calculations. In Ref. [35], $c(t,r)$ was assumed to be positive at all times, without loss of generality. If $c(t,r)$ was to change sign, the metric would become either degenerate if $v(t,r) = 0$ or Euclidean. It was shown that the function $m(t,r)$ corresponds to the Hernandez-Misner mass function [42] as well as to the Hawking-Israel quasi-local mass, meaning that $m(t,r)$ represents the mass included in the radius r at time t. The function $r_H(t)$ satisfying $2m(t,r_H) = r_H$ was proven to be the apparent horizon of a black hole, analogous to the Schwarzschild case. Furthermore, the black hole apparent horizon was shown to be only growing with time if the null energy condition (NEC) was fulfilled, i.e. $T_{\mu\nu}\tilde{l}^\mu\tilde{l}^\nu > 0$ with $T_{\mu\nu}$ the stress-energy tensor and $\tilde{l}^\mu$ an

arbitrary light-like vector field. However, the conclusion that has been drawn might not be as straight forward as it seems. Let us start by briefly reviewing how it was proven that the black hole horizon cannot shrink. The evolution equation of the apparent horizon, as derived in Ref. [35], is

$$\dot{r}_H = \frac{8\pi r_H^2 c(t, r_H) T_{\mu\nu} \tilde{l}^\mu \tilde{l}^\nu}{1 - 2m'(t, r_H)}\,, \tag{1.76}$$

where $\tilde{l}^\mu = (c^{-1}, 1 - \sqrt{2m/r}, 0, 0)$ is a light-like vector field. This equation is obtained from the equality $G_{\mu\nu}\tilde{l}^\mu\tilde{l}^\nu = 8\pi T_{\mu\nu}\tilde{l}^\mu\tilde{l}^\nu$ and using the derivative of the horizon equation $2m(t, r_H) = r_H$. One can show that the denominator in Eq. (1.76) is positive and the NEC guaranties $T_{\mu\nu}\tilde{l}^\mu\tilde{l}^\nu > 0$ leading to a growing apparent horizon. A first problematic fact is that, by fixing the sign of $v(t, r)$ one fixes the general behaviour of the universe (either expanding or contracting). This can be understood by considering the homogeneous limit of the PG metric with $v(t, r)$ as in Eq. (1.75), which is obtained with the limit $m(t, r) \to H^2(t)r^3/2$, where $H(t)$ is the Hubble parameter, and $c(t, r) \to 1$. The Raychaudhuri equation then reads

$$\pm\frac{\mathrm{d}\big|H(t)\big|}{\mathrm{d}t} = 3H^2(t) + 8\pi p(t) = 8\pi\big(p(t) + \rho(t)\big)\,, \tag{1.77}$$

where p is the pressure and ρ the density. The absolute value comes from the square root of $m(t, r)$ in Eq. (1.75), while the $\pm$ sign originates from the sign of $v(t, r)$. We recall that in spatially flat FLRW cosmology, the universe can either exclusively grow or exclusively shrink if the NEC is fulfilled, meaning $p > -\rho$. This behaviour can be mathematically described using the Hubble horizon $|H(t)|^{-1}$, which grows or shrinks in the same way the universe does. From the Raychaudhuri equation (1.77), we can directly see using the last equality that the evolution of the Hubble horizon is monotonic when the NEC is satisfied. Moreover, the $\pm$ sign tells us if the horizon is shrinking ($+1$ case) or growing (-1 case). Hence it can be argued that the proof as in Ref. [35] of growing black hole horizon is restricted to contracting universe scenarios. Moreover, when one tries to take the other sign convention, meaning $v = -c\sqrt{2m/r}$, we find that there is no black hole apparent horizon anymore. A last but crucial objection that can be raised about the choice of fixing $v(t, r)$ as in Eq. (1.75), when describing the universe as a whole, is the lack of cosmological apparent horizon when $v(t, r) > 0$ and the lack of black hole apparent horizon when $v < 0$.

To see the failure of describing both apparent horizons by fixing the sign of v, we first choose the future-pointing null vectors that are orthogonal to a spherical surface, namely

$$l^\mu = (1, 1 - v, 0, 0)/c \quad \text{and} \quad n^\mu = (1, -1 - v, 0, 0)/c. \tag{1.78}$$

One can verify that they are indeed null vectors and $l^\mu S_\mu = 1$ and $n^\mu S_\mu = -1$, where $S^\mu = (0, 1, 0, 0)$. The expansions associated to these vectors are

$$\theta_l = \frac{2}{cr}(c - v) \quad \text{and} \quad \theta_n = -\frac{2}{cr}(c + v). \tag{1.79}$$

From these alone we can see that if one replaces $v \to c\sqrt{2m/r}$, we cannot have a cosmological apparent horizon, since the equation $1 + \sqrt{2m/r} = 0$ admits no solution. Conversely,

with $v \to -c\sqrt{2m/r}$, we do not have a black hole apparent horizon. Moreover, from the expansions (1.79) we see that $v(t,r)$ must change sign to describe both apparent horizons that should be part of our model.

Even though we expressed some criticism towards the conclusion that were drawn using Eq. (1.76) for $v(t,r)$, we believe that the method used is of great use and we arrive at the same conclusion using the more general metric (1.71). Let us follow a similar reasoning as in Ref. [35] to show that a black hole apparent horizon cannot shrink. First, the horizon equation for a black hole becomes

$$v(t,r_H) = c(t,r_H) \implies \dot{v} + v'\dot{r}_H = \dot{c} + c'\dot{r}_H, \tag{1.80}$$

where dot and prime stand for the partial derivative with respect to t and r, respectively. Then, using $G_{\mu\nu}l^\mu l^\nu = 8\pi T_{\mu\nu}l^\mu l^\nu$ evaluated at the horizon $r_H(t)$ as well as the derivative of the horizon equation above, we find

$$\dot{r}_H(t) = \frac{8\pi c^2(t,r_H) r_H(t) T_{\mu\nu}l^\mu l^\nu}{2\big(c'(t,r_H) - v'(t,r_H)\big)}. \tag{1.81}$$

We are left to show that the denominator of Eq. (1.81) is indeed positive. Since $r_H(t)$ describes a black hole apparent horizon, we want that $\theta_l < 0$ for $r < r_H(t)$, hence

$$\left.\frac{\partial\theta_l}{\partial r}\right|_{r=r_H} > 0 \implies \frac{2\big(c'(t,r_H) - v'(t,r_H)\big)}{c(t,r_H)r_H} > 0, \tag{1.82}$$

which proves what was needed.

The fact that the apparent black hole horizon must always grow is independent of the form of the stress-energy tensor, as long as the system is spherically symmetric. This appears to be in contradiction with the generalized McVittie approach, in which the black hole horizon decreases right before the merger with the cosmological horizon. The potential solution to this issue might come from the definition of the apparent horizon, which depends on the foliation of the 3-surface defining the marginal surfaces. Nonetheless, further investigations should be pursued.

1.4.3 Perfect fluid moving as an Eulerian observer

Let us now focus on a perfect fluid moving as an Eulerian observer with velocity field $U_\mu = (-1,0,0,0)$. The stress-energy tensor therefore takes the usual form

$$T^{\mu\nu} = (p(t,r) + \rho(t,r))U^\mu U^\nu + p(t,r)g^{\mu\nu}\,. \tag{1.83}$$

Note that the density and pressure defined here are different from the case of the gMcVittie case, due to the change of coordinates. The full tetrad basis that defines the observer can be chosen to be

$$e_a{}^\mu = \left(U^\mu, S^\mu, \hat{\theta}^\mu, \hat{\phi}^\mu\right)\,, \tag{1.84}$$

with

$$U^\mu = (c^{-1}, v/c, 0, 0)\,, \quad S^\mu = (0, 1, 0, 0)\,, \tag{1.85}$$

$$\hat{\theta}^\mu = (0, 0, r^{-1}, 0)\,, \quad \hat{\phi}^\mu = (0, 0, 0, r^{-1}\sin^{-1}\theta)\,. \tag{1.86}$$

Since the observer is in the rest frame of the fluid, the projection of the stress-energy tensor onto the tetrad basis is diagonal, meaning that the fluid is isotropic in its rest frame, as it should for a perfect fluid. Let us now derive the Einstein's field equations for the system. Projected on the tetrad basis, the independent equations for $r > 0$ are

$$8\pi r^2\rho = \frac{v}{c^2}\left(v + 2rv'\right)\,, \tag{1.87}$$

$$\frac{v}{r}\frac{c'}{c} = 0\,, \tag{1.88}$$

$$4\pi c^2 rp = \dot{v} - \frac{v^2}{2r} - v\frac{\dot{c}}{c} - vv'\,, \tag{1.89}$$

$$8\pi p = -\frac{1}{c^2}\left(r\frac{\dot{c}}{c}v' + v\frac{\dot{c}}{c} + v(2v' + rv'')\right) + c\left(\dot{v} + v\dot{v}' - rv'^2\right)\,, \tag{1.90}$$

where we have kept implicit the dependence of c, v, ρ and p on t and r. The $\theta\theta$ and $\phi\phi$-component equations are the same. We have used $c' = 0$ to simplify the other field equations. The continuity equations, $\nabla_\mu T^{\mu\nu} = 0$, are not independent from the field equations, but they offer a new perspective to the system. In particular, one obtains:

$$\nabla_\mu T^{\mu\nu} S_\nu = p' = 0\,, \tag{1.91}$$

the pressure does not depend on the radial component. Moreover, we fix the continuity equation such that $p = w\rho$ and therefore the density does not depend on r either. This implies that Eq. (1.87) is now analytically solvable and we get

$$v(t, r) = \pm\sqrt{\frac{8\pi}{3}c^2(t)r^2\rho(t) + \frac{\bar{v}(t)}{r}}\,, \tag{1.92}$$

with $\bar{v}(t)$ an integration constant. With this solution at our disposal, we can look at the second continuity equation, which gives us

$$\nabla_\mu T^{\mu\nu} V_\nu = 0 \implies \frac{\dot{\rho}(t)}{\rho(t)} = \pm(1+w)\frac{3\bar{v}(t) + 16\pi c^2(t)r^3\rho(t)}{2r^2\sqrt{\bar{v}(t)/r + 8\pi c^2(t)r^2\rho(t)/3}}\,. \tag{1.93}$$

The LHS of this equation does not depend on time, so should the RHS. This is possible only if $w = -1$ or if $\bar{v}(t) = 0$. Let's analyze the latter first. With this new condition and further defining $H^2(t) := 8\pi\rho(t)/3$, the continuity equation becomes

$$\frac{\dot{\rho}(t)}{c(t)} = \pm 3(1+w)\big|H(t)\big|\rho(t)\,. \tag{1.94}$$

All the Einstein field equations with these additional conditions give the same equation as this one. $c(t)$ can be absorbed in the time coordinate with the redefinition $c(t)\mathrm{d}t \to \mathrm{d}t$. We recall that in a FLRW universe with no spatial curvature and non-negative cosmological

constant, the scale factor is monotonic, therefore the Hubble constant keeps the same sign. This means that Eq. (1.94) is equivalent to the continuity equation of an FLRW universe and the $\pm$ sign decides if we are in an expanding or contracting branch. This can also be seen if we write the metric using the solution for $v(t)$ (1.92) and the condition $\bar{v} = 0$, i.e.

$$\mathrm{d}s^2 = -c^2(t,r)\left(1 - H^2(t)r^2\right)\mathrm{d}t^2 \pm 2c(t)\big|H(t)\big|r\mathrm{d}r\mathrm{d}t + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.95}$$

where we can again redefine the time coordinate to get rid of the function $c(t)$. This metric is the FLRW solution in the Painlevé-Gullstrand form. Therefore we conclude that assuming a perfect fluid with $w > -1$ is moving as an Eulerian observer is equivalent to homogeneity for an isotropic spacetime and therefore cannot describe a cosmological black hole.

We are left with the case $w = -1$. The equations are then

$$\dot{\rho}(t) = 0 \quad \text{and} \quad 2\dot{\bar{v}}(t)c(t) = \bar{v}(t)\dot{c}(t)\,. \tag{1.96}$$

We recover the continuity equation of de-Sitter space-time and we find that $\bar{v}(t) = 2mc^2(t)$, with m an integration constant. Once again, we can use the solution (1.92) and a redefinition of the time coordinate to find the metric

$$\mathrm{d}s^2 = -\left(1 - H^2(t)r^2 + \frac{2m}{r}\right)\mathrm{d}t^2 \pm 2\sqrt{H^2(t)r^2 + \frac{2m}{r}}\mathrm{d}r\mathrm{d}t + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,,. \tag{1.97}$$

This metric is known as the *Kottler or Schwarzschild-de Sitter metric* and describe a black hole immersed in a de-Sitter space-time.

1.5 Competition between GWs emissions and mass loss in black hole binaries

One of the most striking evidence for the existence of black holes has been the observation of the gravitational waves (GWs) that are emitted when a binary system coalesces [43]. The loss of energy due to the emission of these waves shrinks the orbits of binary black holes. Moreover, as we have seen earlier in Sec. 1.3.2 Hawking radiation decreases the mass of black holes and we expect such a mass loss to have the opposite effect on binary black holes of the one from GWs. While the hawking radiation should be very small for the solar mass black holes that are observed with current GWs' observatories, we could still wonder, as an academic exercise, what is the behaviour of a binary system, in circular orbit, when both effects are considered. Hawking radiation might not be the only phenomenon that can decrease the mass of a black hole. If the equation of state parameter of dark energy can become smaller than -1, a fluid commonly called *phantom energy*, then the mass of a black hole immersed in such a fluid could decrease [44]. In the following paper, we study the competition between the effect of the loss of mass due to Hawking radiation or phantom energy and the loss of energy from GWs radiation.

Competitive effects between gravitational radiation and mass variation for two-body systems in circular orbits

Baptiste Blachier,[1,2,*] Aurélien Barrau,[1,†] Killian Martineau,[1,‡] and Cyril Renevey[1,§]

[1] *Laboratoire de Physique Subatomique et de Cosmologie, Univ. Grenoble-Alpes, CNRS/IN2P3 53, avenue des Martyrs, 38026 Grenoble cedex, France*

[2] *Department of Physics, École Normale Supérieure de Lyon, 69364 Lyon, France*

(Dated: June 16, 2023)

This work provides, at lower order, general analytical solutions for the orbital separation, merging time, and orbital frequency of binary systems emitting gravitational waves while being submitted to mass variations. Specific features, depending on the exponent of the mass derivative, are investigated in details. Two phenomenologically interesting cases are explicitly considered : i) binaries formed by two light primordial black holes submitted to Hawking evaporation and ii) bodies driven by a Bondi accretion of phantom dark energy. It is shown that three different regimes arise, including an intricate non-monotonic behaviour of the system. We study subtle imprints that could be associated with those phenomena. A careful analysis of the conditions of validity of the different hypotheses performed is finally carried out.

I. INTRODUCTION

Newtonian orbits are stable for objects with constant masses. This is obviously not the case anymore if the considered masses become time-dependent, whatever the reason for this variation. This also becomes wrong in General Relativity (GR) when the system radiates energy through gravitational waves. The very existence of this energy has been intensively debated, even quite recently [1], as one might wonder if objects following geodesics – thus being "force-free" – can dissipate energy. It is however well settled since the works performed by Penrose, Bondi, and Sachs in the 1960s, although the conceptual arguments are somehow subtle and the technicalities quite involved [2]. This work aims at investigating – at lowest order – the detailed evolution of the orbits of systems submitted both to the emission of gravitational waves (GWs) and to mass variations.

One of the most striking sources of GWs is the coalescence of black holes (BHs), which has been exhaustively studied theoretically and has proved to be experimentally fruitful since most of the GW signals detected so far by LIGO and Virgo collaborations come from this phenomenon (see, *e.g.*, [3, 4]). Although the most simple and straightforward cases are those of binaries with constant masses, inspiral binaries with time-varying masses and their consequences on the emitted gravitational radiation have also been studied in the literature [5–15]. It is well-known that the emission of GWs tend to make the two bodies of the binary come closer to each other, until they ultimately coalesce [16]. However, the time-varying mass can have the converse effect, depending on whether the bodies gain or lose mass. If they gain mass, for instance through matter accretion, they inspiral, thus enhancing the effect of the emission of GWs. On the other hand, if they lose mass, they outspiral: in that case, mass loss produces an antagonistic effect to the one of gravitational radiation.

In this note, we focus on presenting a comprehensive view of the dynamics induced by the competitive effects of gravitational radiation and mass loss, using Newtonian analysis and treating the binary system as Keplerian, as well as the characteristic features of the imprint it produces on the emitted GWs. In section II, we review how one can theoretically deal with the coupling between mass variation and the backreaction produced by the emission of gravitational waves. However, unlike previous works – see Refs. [7–12] for accreting bodies or Refs. [13–15] for evaporating BHs in the context of braneworld models and extra-dimensions scenarios – we present a generic analytic solution of the differential equation satisfied by the orbital separation when the mass variation is taken into account. This leads in turn to analytic solutions for the frequency and strain of the emitted GWs.

For cosmology and phenomenology, we focus specifically on two situations that are driven by such a competition. In section III, we investigate the case of inspiral binaries of light primordial black holes (PBHs) submitted to Hawking evaporation. Due to their low mass, they are good candidates as sources of high-frequency GWs [17] – in addition to standing as plausible dark matter candidates and valuable probes of the early universe.

The second case of interest is the one of binaries of BHs accreting phantom energy (*i.e.* violating energy dominance conditions) which makes their mass *decrease* [18]. Although the accretion of a fluid onto a black hole is a long-standing problem in astrophysics [19], the study of this so-called phantom energy has proved particularly relevant in the context of dark energy [20]. Assuming that the latter can be described by a perfect fluid of density ρ and pressure p, an equation of state parameter $w = p/\rho$

* baptiste.blachier@ens-lyon.fr
† barrau@lpsc.in2p3.fr
‡ martineau@lpsc.in2p3.fr
§ renevey@lpsc.in2p3.fr

smaller than -1 would make dark energy behave as phantom energy: accretion would decrease the mass. Coining down the value of w is a hot spot in modern cosmology [21] and doing so by examining small-scale systems like BHs binaries was proposed in [7, 22] and questioned in [8]. In section IV, we refine the analysis and draw a clear conclusion.

Beyond the specific examples chosen for their physical relevance, this study aims at offering an exhaustive classification of possible behaviors of binary systems emitting gravitational waves and submitted to mass variation effects (either accretion or radiation) described by a (positive or negative) power law. These generic results are presented in section V

Eventually in section VI, we show how accounting for a time-varying mass changes the formal expression of the radiated gravitational power and in section VII we discuss how the results of section V are changed when considering non-identical masses presenting a high hierarchy.

II. COUPLING MASS VARIATION AND GWS EMISSION

The two-body problem with variable mass, for point-like entities, was extensively studied in Ref. [23, 24]. When the variation of mass is isotropic, the equation of motion is given by

$$\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = -\frac{Gm_{\rm tot}(t)}{R^3}\boldsymbol{r}, \tag{1}$$

where $\boldsymbol{r} = \boldsymbol{r}_2 – \boldsymbol{r}_1$, thus $R \equiv |\boldsymbol{r}|$ is the orbital separation between the two bodies of respective mass m_1 and m_2, the total mass being $m_{\rm tot}(t) \equiv m_1(t) + m_2(t)$. For orbital separations much larger than the Schwarzschild radii of the black holes, $R \gg R_{\rm s} = 2Gm/c^2$, we can describe the binary system using Keplerian dynamics. If $\mu \equiv m_1 m_2/(m_1+m_2)$ stands for the reduced mass, the orbital angular momentum of the system, for a circular orbit, whose generic expression is given by

$$J_{\rm orb} = \mu\sqrt{Gm_{\rm tot}R}, \tag{2}$$

is conserved, even in the case of variable masses [23]. If we consider that the two bodies have the same mass $m(t)$, then the conservation of the orbital angular momentum $\dot{J}_{\rm orb} = 0$ provides

$$\dot{R} = -3\frac{\dot{m}}{m}R, \tag{3}$$

showing that in the case of mass loss ($\dot{m} < 0$) the two bodies outspiral, *i.e.* they drive away from each other, which is consistent with Refs. [14, 15, 23, 25]. From now on, we shall always assume, unless otherwise stated, that the two bodies have same mass $m(t)$. This is the most interesting situation and it is phenomenologically sound – at least as a first approximation – if the mass spectrum of PBHs is nearly monochromatic (see [26] for a review of formation mechanisms). Only in section VII, will we drop this assumption in order to examine how certain results presented in the following are modified when the binary system is made of bodies of different masses, but presenting a strong hierarchy (typically one mass dominates over the other).

The mass loss, in addition to separating the two bodies also induces a change in the orbital energy $E_{\rm orb} = -Gm^2/(2R)$, which is the total energy of the system, sum of kinetic and potential energy of the orbit. For variable masses, the variation of orbital energy $-\frac{\mathrm{d}E_{\rm orb}}{\mathrm{d}t}$ must be equal to the power associated to mass loss $P_{\rm ml}$, thus leading to

$$P_{\rm ml} = \frac{5}{2}\frac{G\dot{m}m}{R}. \tag{4}$$

On the other hand, the emission of GWs costs energy which is taken from the orbital energy of the system, carried away at a rate [16]

$$P_{\rm gw}(t) = \frac{64}{5}\frac{G^4}{c^5}\frac{m^5(t)}{R^5}. \tag{5}$$

In fact, m having an explicit time-dependence, it affects the form of the power radiated by GWs thus modifying Eq. (5), as already noted in [8], through several corrective terms. We discuss this point in further detail in section V, and show how the choice of the form (5) for $P_{\rm gw}$ is connected to the condition of circularity of the orbit.

To consider the concomitant effect of gravitational radiation and mass variation, one can use the conservation of energy and write [8]

$$-\frac{\mathrm{d}E_{\rm orbit}}{\mathrm{d}t} = P_{\rm ml} + P_{\rm gw}, \tag{6}$$

with $P_{\rm ml}$ and $P_{\rm gw}$ respectively given by Eqs. (4) and (5). Taking the derivative of the orbital energy leads to

$$\dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m^3}{R^3} - 3\frac{\dot{m}}{m}R, \tag{7}$$

which is the general differential equation relevant for our problem.

This agrees with the result of Ref. [8] – up to the corrected prefactor 3 in the last term. This deserves a brief specific discussion. Equation (3) holds only if the masses are identical. There are two other limit cases allowing for simple formulas (as can immediately be seen from the expression of the angular momentum). They correspond to a strong mass hierarchy with asymmetrical losses. We specifically investigate those cases in the last section of this work. We will show there that if the varying mass is the small one, the prefactor 2, inappropriately used in [8] indeed appears. Hence the probable origin of the mistake.

If the functions $m(t)$ and $\dot{m}(t)$ are explicitly known, this differential equation, although non-linear, can be integrated since it is a Bernoulli differential equation of

the form $y' + P(t)y = Q(t)y^n$ with $n = -3$ whose general solution is known. In the following two sections, we shall focus on mass losses described by rates of the form $\dot{m} \propto -1/m^2$ (see section III) and $\dot{m} \propto -m^2$ (see section IV). More generic results shall be found in section V.

Throughout all this work, we call "merging" or "coalescence" the situation corresponding to the precise vanishing of the orbital separation ($R = 0$), and not to the "contact" of horizons ($R = 4Gm/c^2$). For most observables this makes nearly no difference at all, but it is important when considering the end of the process.

III. INSPIRAL BINARIES WITH EVAPORATING BHS

Because of their small masses, primordial black holes can be particularly sensitive to the Hawking evaporation process (see [27, 28] for pioneering works and [26] for a recent review). Quantum field theory in curved spacetimes predicts that black holes evaporate with a temperature [29, 30]

$$T = \frac{\hbar c^3}{8\pi G k_{\rm B} m}, \tag{8}$$

for a BH of mass m, $\hbar$ being the Planck constant, $k_{\rm B}$ the Boltzmann constant, c the speed of light, and G the gravitational constant. Even though for typical solar-mass BHs, Hawking evaporation is too weak to play any significant role in the dynamics of binaries, for two-body systems made of PBHs, it could produce non-negligible effects on the overall dynamics, and leave a possible imprint on the emitted gravitational waves. The Hawking process leads to a mass loss at the rate [27, 28]

$$\dot{m} = -\frac{\alpha_{\rm H}}{m^2}, \tag{9}$$

where $\alpha_{\rm H}$ accounts for the degrees of freedom of each particle contributing to the evaporation. For simplicity, we shall assume that $\alpha_{\rm H}$ is a constant, which is a good approximation. In this case, the differential equation is separable and integrates into

$$m(t) = m_0 \left(1 - \frac{t}{t_{\rm ev}}\right)^{1/3}, \tag{10}$$

where m_0 represents the initial mass of the BH whereas $t_{\rm ev} \equiv m_0^3/(3\alpha_{\rm H})$ corresponds to the time of evaporation, *i.e.* the typical time it takes for the BH to evaporate completely. The other time scale of the problem under scrutiny is the *time of coalescence* which accounts for the limited life duration of the binary system. However, contrary to $t_{\rm ev}$ which only depends on constants fixed during the initial setup (namely the initial mass m_0), the time of coalescence has an explicit dependence on the mass m (and not only on the initial mass m_0). Otherwise stated, if $t_{\rm ev}$ is not to be altered by the dynamics of the system, the time of coalescence will necessarily be affected by the mass loss. Consequently, we will call $t_{\rm cc}$ the time of coalescence of the binary system *if* the two BHs were of constant mass m_0, with an initial orbital separation R_0, *i.e.* [16]

$$t_{\rm cc} = \frac{5}{512}\frac{c^5 R_0^4}{G^3 m_0^3}, \tag{11}$$

while $t_{\rm coal}$ will denote the real time of coalescence, that is to say the one taking into account the mass loss, which is likely to be different from $t_{\rm cc}$ and yet to be determined.

A. Evolution of the orbital separation

Specifying (7) to the case of Hawking evaporation provides

$$\dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m^3}{R^3} + \frac{3\alpha_{\rm H}}{m^3}R. \tag{12}$$

Evaluating the above at an initial time $t_0 = 0$ gives

$$\left.\frac{\dot{R}}{R}\right|_{t_0=0} = -\frac{1}{4t_{\rm cc}} + \frac{1}{t_{\rm ev}}. \tag{13}$$

If one assumes that the system is prepared – which boils down to giving ourselves m_0 and an initial separation R_0 – such that $t_{\rm ev} < 4t_{\rm cc}$, then the evaporation process initially dominates. Since the evaporation increases R and since the GW part is in m^3/R^3 – see Eq. (12) – when m diminishes and R increases, both these variables contribute to the dwindling of this term, that will thus never be able to grow again to counterbalance the outspiraling effect. In conclusion, when $t_{\rm ev} < 4t_{\rm cc}$, the system initially outspirals and continues to do so until the PBHs eventually evaporate completely, leading to the disappearance of the binary.

In the other case, no definite conclusion can be stated by the sole inspection of Eq. (12). When $4t_{\rm cc} < t_{\rm ev}$, we are in a regime where, initially, the radiation of GWs dominates, thus R decreases. But as R and m diminish, there might be a counterbalancing effect due to the term of relative mass loss. Using the explicit expression (10) and solving the Bernoulli differential equation provides

$$R(t) = R_0 \left(\frac{t_{\rm ev}}{t_{\rm ev} - t}\right)\left(1 + \frac{1}{6}\frac{t_{\rm ev}}{t_{\rm cc}}\left[\left(1 - \frac{t}{t_{\rm ev}}\right)^6 - 1\right]\right)^{1/4}. \tag{14}$$

When $t_{\rm cc} \to \infty$, the second bracket boils down to unity and we recover the solution where no gravitational waves are emitted with the sole effect of mass loss leading to an outspiralling dynamics. Equation (14) is the exact analytical solution to the problem under consideration.

The equation $R(t) = 0$ has for solution

$$t_{\rm coal} = t_{\rm ev}\left(1 - \left[1 - 6\frac{t_{\rm cc}}{t_{\rm ev}}\right]^{1/6}\right), \tag{15}$$

which corresponds to the real time of coalescence. Mathematically, this solution is well-defined only if $1 - 6t_{\rm cc}/t_{\rm ev} > 0$ which leads to the condition

$$t_{\rm ev} > 6t_{\rm cc}. \tag{16}$$

If Eq. (16) is satisfied, the equation $R(t) = 0$ always admits a solution and thus the two PBHs do coalesce, in a similar fashion to what is observed in the case of constant masses. However, as shown in Fig. 1, because of the mass loss which renders the system less tightly bound, $t_{\rm coal}$ is always larger than $t_{\rm cc}$.

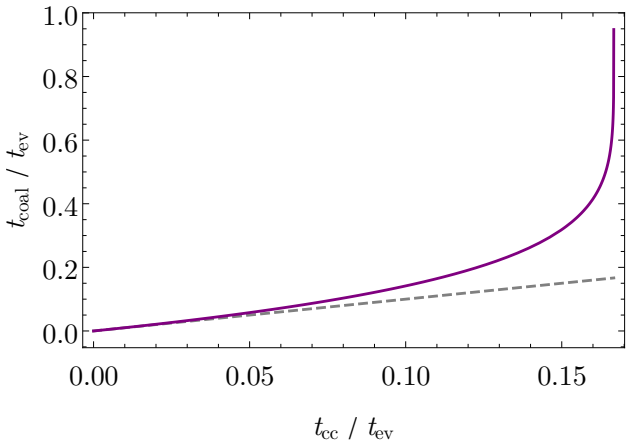


FIG. 1. Effective time of coalescence $t_{\rm coal}$ as given by Eq. (15) divided by the time of evaporation $t_{\rm ev}$ (purple curve) in function of the ratio $t_{\rm cc}/t_{\rm ev}$. At $t_{\rm cc}/t_{\rm ev} = 1/6$, this function is finite equal to 1, and for higher values, it ceases to be mathematically well-defined. The grey dashed line is the identity function.

This immediately raises the question of what happens in the small window of parameters for which $4t_{\rm cc} < t_{\rm ev} < 6t_{\rm cc}$. Equation (13) indicates that we are then in a regime where the emission of gravitational waves *initially* dominates, thus leading, in a first step, to the decrease of the orbital separation. However, since the condition (16) is not fulfilled, the equation $R(t) = 0$ does not admit any solution, which indicates that the system will never merge. This is due to the evaporation term that gets progressively more important and eventually dominates the GW emission. Once the evaporation process leads, there is no coming back and it ultimately makes the two PBHs outspiral, until they completely disappear. This is an interesting and highly non-trivial case where the orbital separation evolution is non-monotonic.

To sum up, there are three distinct regimes (see Fig. 2) :

- If $t_{\rm ev} < 4t_{\rm cc}$, the system outspirals due to the domination of the evaporation process and R increases with a final divergence corresponding to the full evaporation of the two PBHs;
- If $4t_{\rm cc} < t_{\rm ev} < 6t_{\rm cc}$, the two PBHs come closer together in a first step due to the emission of GWs and then, as the evaporation takes the lead back, they ultimately outspiral as in the first case; and
- If $t_{\rm ev} > 6t_{\rm cc}$ the emission of GWs dominates entirely, leading to the merger of the system, in a similar fashion to what is observed in the case of constant masses. The only real influence of the mass loss is to increase the time of coalescence.

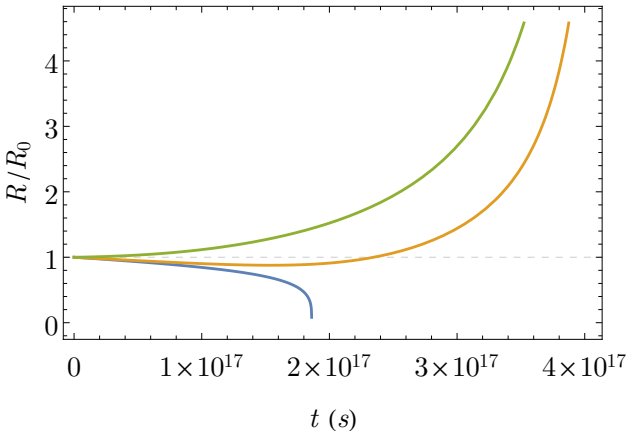


FIG. 2. Evolution of the orbital separation R (normalized to the initial orbital separation R_0) between two evaporating black holes with initial mass $m_0 = 10^{12}$ kg, as a function of time (in seconds). Three distinct regimes arise : for $t_{\rm ev} < 4t_{\rm cc}$ (green curve) the binary outspirals whereas for $t_{\rm ev} > 6t_{\rm cc}$ (blue curve) it inspirals following a trend similar to the standard case of constant masses. For $4t_{\rm cc} < t_{\rm ev} < t_{\rm cc}$ (orange curve), one has an intermediate regime mixing the two behaviours, with an unusual non-monotonic evolution.

The above inequalities can be conveniently expressed as conditions on the initial orbital separation R_0 for a given initial mass m_0 by using Eq. (11). If one defines

$$R_1 \equiv \left(\frac{256}{45}\frac{G^3}{c^5\alpha_{\rm H}}\right)^{1/4} m_0^{3/2}, \tag{17}$$

and

$$R_2 \equiv \left(\frac{128}{15}\frac{G^3}{c^5\alpha_{\rm H}}\right)^{1/4} m_0^{3/2}, \tag{18}$$

the ouspiralling regime is reached when $R_0 > R_2$, the inspiralling one when $R_0 < R_1$, while the intermediate regime corresponds to $R_1 < R_0 < R_2$. This might be relevant when evaluating the merging rate of PBHs [31–33].

B. Analysis of the frequency

Since the system is assumed to be Keplerian at every step in its evolution, using Kepler's third law $\omega^2 =$

$2Gm/R^3$ along with Eq. (14) provides an analytical expression for the orbital frequency ω:

$$\omega(t) = \omega_0 \left(1 - \frac{t}{t_{\rm ev}}\right)^{\frac{5}{3}} \left(1 + \frac{1}{6}\frac{t_{\rm ev}}{t_{\rm cc}}\left[\left(1 - \frac{t}{t_{\rm ev}}\right)^6 - 1\right]\right)^{-\frac{3}{8}} \tag{19}$$

with $\omega_0^2 \equiv 2Gm_0/R_0^3$ the initial orbital frequency. As for the orbital separation, it admits three distinct regimes (see Fig. 3). The frequency of gravitational waves is simply twice the orbital frequency.

As expected, when $R \to \infty$ in the outspiralling case, $\omega \to 0$. This basically means that the two bodies are not sufficiently tightly bound to be considered anymore as a binary system, hence the very notion of orbital frequency becomes irrelevant.

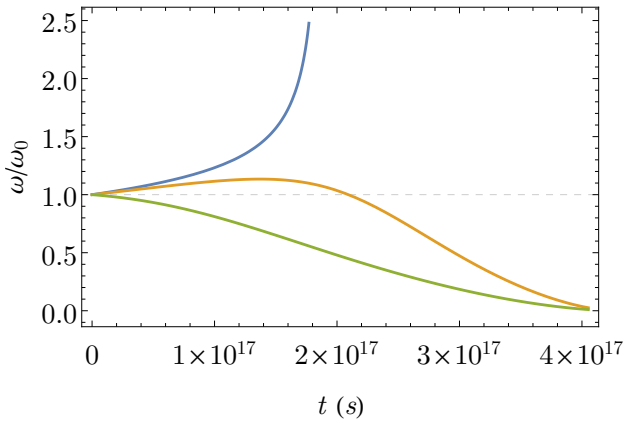


FIG. 3. Evolution of the orbital frequency ω (normalized to the initial orbital frequency ω_0) between two evaporating black holes with initial mass $m_0 = 10^{12}\,\mathrm{kg}$, as a function of time (in seconds). The color of the three different curves correspond to the ones used in Fig. 2.

When the system actually merges (blue curve in Fig. 3), we recover the familiar chirping trend. To determine the chirping frequency, we place ourselves at the innermost stable circular orbit (ISCO) which corresponds to an orbital separation $R_{\rm ISCO} = 12Gm_{\rm ISCO}/c^2$ where $m_{\rm ISCO}$ is the time-varying mass evaluated at $t = t_{\rm ISCO}$. Thus one can solve the equation $R(t) = R_{\rm ISCO}$ using Eq. (14), find the corresponding time $t_{\rm ISCO}$ and then evaluate $\omega(t_{\rm ISCO})$ with Eq. (19). Since $R(t) = R_{\rm ISCO}$ is an implicit equation that doesn't admit any simple analytical solution, this should be done numerically.

Nonetheless, two important features can be noticed. First, as for the standard case, $t_{\rm ISCO}$ is very close to the effective time of coalescence $t_{\rm coal}$. Solving $R(t) = R_{\rm ISCO}$ for various values of m_0 (typically from $10^9\,\mathrm{kg}$ to $10^{20}\,\mathrm{kg}$) shows that, if γ is the order of magnitude of the initial mass, *i.e.* if $m_0 \simeq 10^\gamma$ kg, then $t_{\rm coal} - t_{\rm ISCO} \simeq 10^{\gamma-34}$ s. For PBHS, one can thus legitimately approximate $t_{\rm ISCO} \sim t_{\rm coal}$. Conveniently, we readily have an explicit expression for $t_{\rm coal}$, see Eq. (15). However, by construction, $\omega(t_{\rm coal}) = \infty$. Although tiny in itself, the difference between $t_{\rm ISCO}$ and $t_{\rm coal}$ is, in principle, relevant to determine the final frequency.

Analytically, the expression (19) is not of great use and one should simply go back to Kepler's third law. In fact, whenever the binary system actually merges, the relative variation of mass during the process is small ; more specifically, we shall demonstrate that it is perturbatively small at ISCO. Let us introduce

$$\varepsilon \equiv \frac{m_0 - m_{\rm ISCO}}{m_0} = 1 - \left(1 - \frac{t_{\rm ISCO}}{t_{\rm ev}}\right)^{1/3}, \quad \varepsilon > 0, \tag{20}$$

where the last equality comes from Eq. (10). Using that $t_{\rm ISCO} \sim t_{\rm coal}$ and Eq. (15), we are led to

$$\frac{t_{\rm ISCO}}{t_{\rm ev}} \sim \frac{t_{\rm coal}}{t_{\rm ev}} = 1 - \left[1 - 6\frac{t_{\rm cc}}{t_{\rm ev}}\right]^{1/6}. \tag{21}$$

Looking at Fig. 1, unless we put ourselves right at the saturated condition $t_{\rm ev} = 6t_{\rm cc}$ – this case is treated at the end of the discussion – we have $t_{\rm ISCO}/t_{\rm ev} \lesssim 0.4$ which leads to $\varepsilon \lesssim 0.2 \ll 1$, *i.e.* the variation of mass is always small whenever the binary system merges. Consequently, one can rely on the usual expression for the frequency at ISCO, but evaluated for bodies that have lost a fraction ε of their initial masses such that $m(t_{\rm ISCO}) = m_0(1-\varepsilon)$. If $f^{\rm H}_{\rm ISCO}$ (resp. $f^{\rm cc}_{\rm ISCO}$) denotes the frequency at ISCO for the evaporating binary system (resp. for a binary system formed of two *constant* masses m_0), then

$$\frac{f^{\rm H}_{\rm ISCO}}{f^{\rm cc}_{\rm ISCO}} = \frac{1}{1-\varepsilon} \sim 1 + \varepsilon > 1, \tag{22}$$

hence for an evaporating binary system, we expect (slightly) higher frequencies. For numerical estimates, plugging Eq. (20), supplemented by Eq. (21), into Eq. (22), one is led to:

$$\begin{aligned} f^{\rm H}_{\rm ISCO} &\simeq \frac{1}{12\pi\sqrt{6}}\left(1 - \frac{t_{\rm coal}}{t_{\rm ev}}\right)^{-1/3}\frac{c^3}{Gm} \\ &\simeq 2200\,\mathrm{Hz}\left(1 - \frac{t_{\rm coal}}{t_{\rm ev}}\right)^{-1/3}\frac{M_\odot}{m}, \end{aligned} \tag{23}$$

where $M_\odot$ is the stellar mass. As for the strain amplitude at the ISCO, it is basically given by the usual formula [34]:

$$h_{\rm max} \approx \frac{2}{D}\left(\frac{GM_{\rm c}}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm ISCO}}{c}\right)^{2/3}, \tag{24}$$

where $M_{\rm c}$ is the chirp mass – simply given by $M_{\rm c} = m/2^{1/5}$ in the case of bodies of identical masses – and D is the distance to the observer. Applying the same reasoning than for the frequency associated with the ISCO, one easily obtains that $h^{\rm H}_{\rm max}/h^{\rm cc}_{\rm max} = 1 - \varepsilon < 1$, *i.e.* using again Eqs. (20) and (21):

$$h^{\rm H}_{\rm max} \simeq \left(1 - \frac{t_{\rm coal}}{t_{\rm ev}}\right)^{1/3} h^{\rm cc}_{\rm max}. \tag{25}$$

Let us now examine the specific case for which $t_{\rm ev} = 6t_{\rm cc}$. According to Fig. (1) – or equivalently Eq. (15) – it means that the system merges concomitantly to the full evaporation of the two BHs, *i.e.* $t_{\rm coal} = t_{\rm ev}$. Furthermore, plugging the condition $t_{\rm ev} = 6t_{\rm cc}$ into Eq. (14) provides

$$R(t) = R_0\sqrt{1-\frac{t}{t_{\rm ev}}}, \tag{26}$$

which proves that in that case, the system does actually merge. On the other hand, using Eq. (19), it is easy to see that the frequency is simply given by

$$\omega(t) = \omega_0\left(1-\frac{t}{t_{\rm ev}}\right)^{-7/12}. \tag{27}$$

We recover that at coalescence, the frequency formally diverges. The analysis of the frequency reached at the ISCO can however be here conducted in a fully analytical way. Solving $R(t) = R_{\rm ISCO}$ indeed leads to

$$t_{\rm ISCO} = t_{\rm ev}\left[1-\left(\frac{12Gm_0}{R_0c^2}\right)^6\right]. \tag{28}$$

One can plug the above expression into Eq. (27) and use the explicit form of ω_0, as well as the condition $t_{\rm ev} = 6t_{\rm cc}$, which enables to express the initial orbital separation R_0 as a function of the initial mass m_0 and some constants (or vice versa), to obtain that the frequency associated at the ISCO is given by:

$$\omega(t_{\rm ISCO}) = \left(\frac{c^3}{G}\right)^{3/2}\frac{1}{\sqrt{34992\alpha_{\rm H}}}. \tag{29}$$

Nicely, this formula does depend neither on m_0, nor on R_0. As the ISCO is reached very late in the process (one should keep in mind that, in this case, the mass decreases during the inspiral and vanishes at merging), the frequency is huge, not far from the Planck frequency. This is obviously a purely academic situations but it enlightens the behaviour of the system in the most extreme (merging) case. Interestingly, one should notice that the binary system always reaches its ISCO, which was not *a priori* obvious, as $R_{\rm ISCO} = 12Gm/c^2$ decreases as R decreases. The two curves ($R(t)$ and $R_{\rm ISCO}(t)$) however necessarily intersect each other.

IV. BONDI ACCRETION OF PHANTOM DARK ENERGY

We now turn to the study of binaries formed of two black holes accreting phantom dark energy. As in Ref. [7, 8], to describe the accretion of matter from interstellar medium by a compact object, we shall use the standard Bondi approximation corresponding to a rate of mass change $\dot{m} \propto -m^2$, which integrates into

$$m(t) = \frac{m_0}{1+t/\tau}, \tag{30}$$

the typical time of evolution τ reading, in the context of dark energy [8, 18],

$$\tau = \frac{3\times10^{40}\,{\rm s}}{|1+w|}\left(\frac{\rho_{\rm d}}{\rho_c}\right)^{-1}\frac{M_\odot}{m_0}, \tag{31}$$

where $M_\odot$ is the Solar mass, $w = p/\rho$ is the equation of state parameter, $\rho_{\rm c} \sim 10^{-26}\,{\rm kg/m^3}$ is the critical density of the universe, and $\rho_{\rm d}(\infty)$ is the dark energy density (which is of the same order of magnitude as shown by observations [35]). For numerical estimates, we assume in the following $1 + w \sim -0.1$.

Several remarks are in order at this point. First, let us emphasize that we focus here on phantom energy ($w < -1$), that is on "anti-accretion" (accretion inducing a mass decrease) as the case $w > -1$ leads to a standard accretion during which the mass increases. Although interesting in itself, this situation does not bring any significant new features to the evolution of the orbital separation: both the effects of gravitational waves and of accretion play in the same direction. The case of a pure cosmological constant, that is $w = -1$, leads to no mass variation at all.

Second, it should be emphasized that in the case of a standard accretion, as it will be discussed later in this article, Eq. (30) would lead to a mass divergence in a *finite* amount of time. Although not directly related to this work, this opens interesting phenomenological features [36]. The situation is, in a sense, formally close – although reversed – to the one of the Hawking evaporation. The deep reason for this is quite simple. In the case of Hawking evaporation, as in the case of standard accretion, the mass variation gets amplified by the evolution it generates. When it evaporates, a BH becomes smaller and smaller, hence hotter and hotter. The phenomenon diverges in finite time. Exactly as what happens for standard accretion: the more a BH absorbs usual matter (with $w > -1$), the larger the cross section gets and the faster its mass grows, leading to a divergence in finite time. Both cases are fundamentally unstable and the relevant question is basically to wonder if the coalescence (if any) happens before or after the singularity associated with the mass variation. On the other hand, the case of phantom energy accretion is stable, hence the regular behaviour of Eq. (30), whatever the value of t. This is because, when $w < -1$, as the anti-accretion takes place, the BH gets smaller and smaller and, therefore, sees its cross section and, consequently its mass loss rate, *decrease* with time. It is a negative feedback whereas the diverging cases correspond to positive feedbacks.

A. Evolution of the orbital separation

Focusing on the case of phantom dark energy, the same method as used in section III can be applied. The differ-

ential equation now reads

$$\dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m^3}{R^3} + \frac{3}{\tau}\frac{m}{m_0}R, \tag{32}$$

which clearly shows that, depending on an initial setup favouring either inspiral or outspiral, there is always one of the two terms which presents a competitive behaviour between the time evolution of the mass and its orbital separation counterpart (in addition to the obvious competition between the two terms themselves, simply due to the opposite sign). Explicitly: if R decreases, the evolution of the amplitude of the first term is not obvious as m also decreases, making the fate of m^3/R^3 *a priori* non-trivial. If, the other way round, R increases, the evolution of the amplitude of the second term is now not obvious as m still decreases, making the evolution of mR possibly intricate.

At an initial time $t_0 = 0$, one has

$$\left.\frac{\dot{R}}{R}\right|_{t_0=0} = -\frac{1}{4t_{\rm cc}} + \frac{3}{\tau}, \tag{33}$$

and the integration of Eq. (32) provides

$$R(t) = R_0\left(1+\frac{t}{\tau}\right)^3\left(1+\frac{1}{14}\frac{\tau}{t_{\rm cc}}\left[\left(1+\frac{t}{\tau}\right)^{-14}-1\right]\right)^{1/4}. \tag{34}$$

Solving the equation $R(t) = 0$ gives the following effective time of coalescence

$$t_{\rm coal} = \tau\left(\left[1-14\frac{t_{\rm cc}}{\tau}\right]^{-1/14}-1\right), \tag{35}$$

which is well-defined mathematically only for $\tau > 14t_{\rm cc}$. As a consequence, the system outspirals for $\tau < 12t_{\rm cc}$, inspirals when $\tau > 14t_{\rm cc}$ and for $12t_{\rm cc} < \tau < 14t_{\rm cc}$ we recover this so-called intermediate regime beginning with an initial inspiral but with an ultimate outspiral due to the time-varying mass. Contrary to the case of Hawking evaporation though, the outspiralling dynamics does not lead to any divergence of the orbital separation at finite time, as it can be seen in Fig. 4. As the evolution law for the mass – Eq. (30) – does not present any pole, the orbital radius gently tends to infinity for $t \to \infty$.

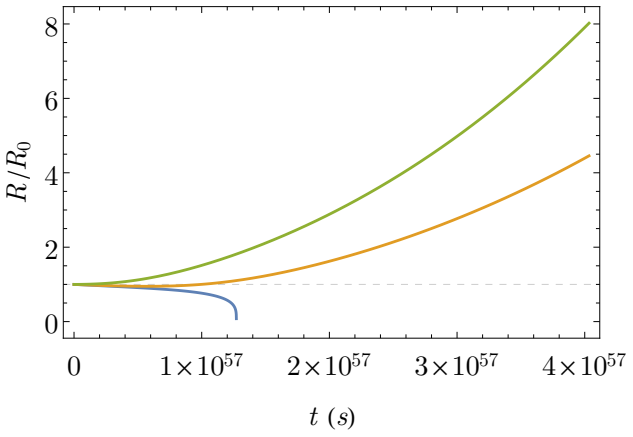


FIG. 4. Evolution of the orbital separation R (normalized to the initial orbital separation R_0) between two black holes submitted to a Bondi-type accretion of phantom dark energy with initial mass $m_0 = 10^{12}\,\mathrm{kg}$, as a function of time (in seconds). We recover the outspiralling dynamics for $\tau < 12t_{\rm cc}$ (green curve), the inspiralling one for $\tau > 14t_{\rm cc}$ (blue curve), and the intermediate non-monotonic regime for $12t_{\rm cc} < \tau < 14t_{\rm cc}$ (orange curve).

B. Evolution of the frequency

It is straightforward, using Kepler's third law, to obtain an analytical expression for the orbital frequency. Its evolution is shown in Fig. 5. The ISCO analysis performed in the case of the Hawking evaporation is still applicable. One can indeed demonstrate that $t_{\rm ISCO} \sim t_{\rm coal}$ and plotting the curve $t_{\rm coal}/\tau$ using Eq. (4) shows that the previously introduced ε parameter is such that $\varepsilon \lesssim 0.23$. Then an approximate formula for the frequency associated to the ISCO when the system is in the inspiralling regime is

$$\begin{aligned} f^{\rm B}_{\rm ISCO} &\simeq \frac{1}{12\pi\sqrt{6}}\left(1+\frac{t_{\rm coal}}{\tau}\right)\frac{c^3}{Gm} \\ &\simeq 2200\,\mathrm{Hz}\left(1+\frac{t_{\rm coal}}{\tau}\right)\frac{M_\odot}{m}, \end{aligned} \tag{36}$$

with the time of coalescence $t_{\rm coal}$ being given by Eq. (35). Again, the correction brought by the time-varying mass is meager. Carrying out the same analysis, but with a *non*-phantom dark energy, *i.e.* with $w > -1$, shows, as expected, that the binary system always merges. In addition, because the mass growth enhances the inspiralling effect, it does so with a shorter time of coalescence. However, as we shall show later on, for all ranges of mass, the difference between these two times of coalescence is too small in practice to constitute a reliable experimental criteria to differentiate the case $w > -1$ from the case $w < -1$.

C. Strain - observational distance relation

We now consider the possibility, in the light of our results and in the same vain as in Ref. [8], to use observations of a binary system to measure local dark energy properties. To get an order of magnitude one can safely assume that, for the accretion of dark energy not to be negligible, the two terms entering the differential equation (32) must be of the same order of magnitude. This implies that the two bodies should be then separated by

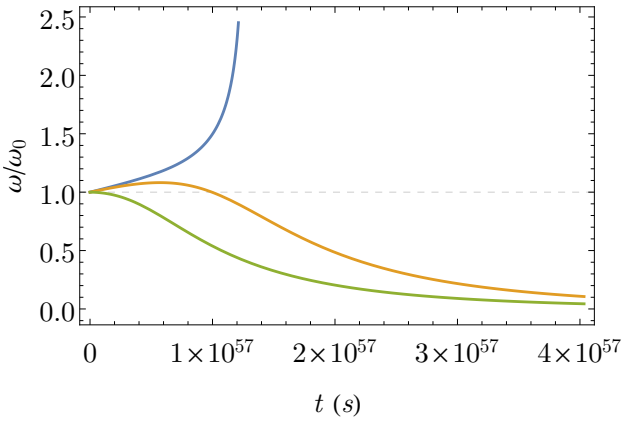


FIG. 5. Evolution of the orbital frequency ω (normalized by the initial orbital frequency ω_0) between two black holes submitted to Bondi-type accretion of phantom dark energy with initial mass $m_0 = 10^{12}\,\mathrm{kg}$, as a function of time (in seconds). The color of the three different curves correspond to the ones used in Fig. 4.

(at least) a distance

$$R \sim \left(\frac{128}{15}\frac{G^3}{c^5}\tau m_0\right)^{1/4}\sqrt{m}. \tag{37}$$

Using Kepler's third law and the above relation to express the frequency solely in function of the mass, and injecting the result in the strain formula, Eq. (24), shows that, given a minimum detectable strain amplitude $h_{\min}$, and demanding $h > h_{\min}$, constrains the distance to the observer D to be less than

$$D_{\max} \approx 2.9\times10^{-2}\,\mathrm{pc}\left(\frac{m}{M_\odot}\right)^{3/2}\left(\frac{h_{\min}}{10^{-23}}\right)^{-1}, \tag{38}$$

where the normalizing value 10^{-23} corresponds to the order of magnitude of the minimum measurable strain amplitude currently reached by gravitational wave detectors (see *e.g.* [37]). This expression depends only on the mass as the orbital separation is chosen so that the anti-accretion term plays a significant role. It should be stressed that it is very different – not only in prefactors but also in the mass functional dependence – from the usual formula encountered in gravitational wave physics, $D < 1.6(m/M_\odot)/(h\times10^{23})$ Gpc, as the latter assumes that the system is seen at the merging time. This is far from being the case for the situation we consider here.

Beyond the distance $D_{\max}$, even if the radial change due to the accretion of dark energy were big when compared to the one due to the emission of GWs, the strain produced would be too small to be detected. For solar-mass BHs, with $h_{\min} \sim 10^{-23}$, the corresponding order of magnitude is approximately the size of the solar system. However, for supermassive black holes with $m = 10^8 M_\odot$, Eq. (38) gives $D_{\max} \sim 29\,\mathrm{Gpc}$ which is more than the cosmological horizon radius corresponding to the limit of the observable universe. It therefore seems that wherever a binary system of $10^8 M_\odot$ BHs, seen early enough in the inspiralling process, exists, the accretion of phantom dark energy could be measured on Earth. This is however not that simple.

Whenever phantom dark energy dominates over gravitational radiation, the binary system is in an outspiralling regime, the orbital frequency thus decreases (see the orange and green curves in Fig. 5). This is an appealing feature as it never happens in standard cases corresponding to a dynamics driven only by the emission of GWs. It could be a "smoking gun" for an effect beyond the standard model (BSM). For the strain to be high enough to be detected without requiring the system to violate obvious bounds (as it would, *e.g.*, be the case if one assumes the existence solar mass BHs within the solar system), the best candidates are supermassive BHs. Equation (37) shows that they should be separated by a distance of order $R \sim 10^{18}\,\mathrm{m}$, corresponding to a time of merging above the Hubble time. However, although the strain would actually be measurable and the sign of the frequency drift would be favourable (that is negative and, therefore, distinct from any usual phenomenon), the amplitude of the *variation* of frequency – which stands as the relevant parameter – would be ridiculously small. Fixing orders of magnitude by assuming constant masses, one would have $|\dot{f}/f| \sim 10^7 f^{8/3}$ which leads to $|\dot{f}/f| \sim 10^{-30}$ for $m = 10^8 M_\odot$. By inspection of Fig. 5, it is obvious that taking into account the variation of the mass and the precise form of the frequency would only worsen the effect. In addition, the absolute value of the frequency would also be way below anything measurable in the next decades.

It should be noticed that, in the context of binaries of evaporating BHs, the question of the maximum distance of observation was already considered in Ref. [38]: BHs undergoing the Hawking effect (with masses typically below $M_* \equiv 10^{12}\,\mathrm{kg}$) would have fully evaporated in a time smaller than the age of the Universe. There is no chance to observe them in our necessarily confined scope of observation. In this case, the phenomenological relevance of our work is not about the detection of a single event but, rather, about the statistical features of the expected merging rate.

V. GENERIC RESULTS

Beyond the examples of (potential) phenomenological interest given in the previous sections, we now focus on offering an exhaustive classification of the possible behaviours of binaries emitting GWs submitted to mass variation effects described by either a positive or a negative power law. The analytical solutions we give – not previously known in the relevant literature to the best of our knowledge – can be used in many different situations.

We consider mass loss rates of the form:

$$\begin{aligned}
&(i) \quad \dot{m} = -\frac{\beta}{m^{k-1}}, k > 0,\\
&(ii) \quad \dot{m} = -\beta m^{k+1}, k > 0,\\
&(iii) \quad \dot{m} = \beta m^{k+1}, k > 0,\\
&(iv) \quad \dot{m} = \frac{\beta}{m^{k-1}}, k > 0,
\end{aligned} \tag{39}$$

with $\beta > 0$ in all cases. For now, we ignore the cases for which the above differential equations do not lead to a mass evolution following a power law – that is corresponding to $\dot{m} \propto m$ which implies an exponential trend. Those specific types of evolution are dealt with at the end of this section.

Let us comment about the physical signification of some particular cases. Hawking radiation corresponds to (i) with $k = 3$. A constant mass loss rate – as it would be the case for a star during most of its life – corresponds to (i) with $k = 1$. A standard Bondi accretion by a black hole – that is to say a spherically symmetric, adiabatic, steady-flow gas accretion – corresponds to (iii) with $k = 1$. A Bondi anti-accretion – that is with a phatom fluid – corresponds to (ii) with $k = 1$. A naive mass loss rate proportional to the area of the compact (non-black hole) object corresponds to (ii) with $k = 1/3$. A hypothetical situation where the accretion rate of a BH would be solely driven by its surface gravity corresponds to (iv) with $k = 2$.

Each of these differential equations for β constant are separable and easily integrate into :

$$\begin{aligned}
&(i) \quad m(t) = m_0 \left(1 - \frac{t}{t_{\rm ev}}\right)^{1/k}, \quad t_{\rm ev} = \frac{m_0^k}{k\beta} > 0,\\
&(ii) \quad m(t) = m_0 \left(1 + \frac{t}{\tau}\right)^{-1/k}, \quad \tau = \frac{1}{m_0^k k\beta} > 0,\\
&(iii) \quad m(t) = m_0 \left(1 - \frac{t}{\tau_*}\right)^{-1/k}, \quad \tau_* = \frac{1}{m_0^k k\beta} > 0,\\
&(iv) \quad m(t) = m_0 \left(1 + \frac{t}{\tilde{\tau}}\right)^{1/k}, \quad \tilde{\tau} = \frac{m_0^k}{k\beta} > 0.
\end{aligned} \tag{40}$$

It should be noticed that (ii) and (iv) exhibit behaviours qualitatively different from (i) and (iii). In the latter two cases, there is a critical point in the evolution of the mass: for (i) it corresponds to $t = t_{\rm ev}$, where the mass suddenly plunges to zero whereas for (iii) it corresponds to the divergence at finite time $t = \tau_*$. As previously explained, those cases are physically related to unstable processes driven by positive feedbacks. Conversely, the cases (ii) and (iv) correspond, respectively, to a smooth decrease and a smooth increase of the mass. They are related to stable processes with negative feedbacks. Then, τ and $\tilde{\tau}$ can be understood as typical time scales on which the considered body has lost or gained a significant amount of mass.

The differential equations satisfied by the orbital separation R are respectively given by :

$$\begin{aligned}
&(i) \quad \dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m_0^3}{R^3}\left(1 - \frac{t}{t_{\rm ev}}\right)^{3/k} + \frac{3}{k}\frac{R}{t_{\rm ev} - t},\\
&(ii) \quad \dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m_0^3}{R^3}\left(1 + \frac{t}{\tau}\right)^{-3/k} + \frac{3}{k}\frac{R}{\tau + t},\\
&(iii) \quad \dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m_0^3}{R^3}\left(1 - \frac{t}{\tau_*}\right)^{-3/k} - \frac{3}{k}\frac{R}{\tau_* - t},\\
&(iv) \quad \dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m_0^3}{R^3}\left(1 + \frac{t}{\tilde{\tau}}\right)^{3/k} - \frac{3}{k}\frac{R}{\tilde{\tau} + t},
\end{aligned} \tag{41}$$

which are all Bernoulli equations. The solutions to these four differential equations are given by :

$$\begin{aligned}
&(i) \quad R(t) = R_0 \left(\frac{t_{\rm ev}}{t_{\rm ev} - t}\right)^{3/k} \times \\
&\qquad \left(1 + \frac{k}{15 + k}\frac{t_{\rm ev}}{t_{\rm cc}}\left[\left(1 - \frac{t}{t_{\rm ev}}\right)^{\frac{15+k}{k}} - 1\right]\right)^{1/4},\\
&(ii) \quad R(t) = R_0 \left(1 + \frac{t}{\tau}\right)^{3/k} \times \\
&\qquad \left(1 - \frac{k}{k - 15}\frac{\tau}{t_{\rm cc}}\left[\left(1 + \frac{t}{\tau}\right)^{\frac{k-15}{k}} - 1\right]\right)^{1/4},\\
&(iii) \quad R(t) = R_0 \left(1 - \frac{t}{\tau_*}\right)^{3/k} \times \\
&\qquad \left(1 + \frac{k}{k - 15}\frac{\tau_*}{t_{\rm cc}}\left[\left(\frac{\tau_*}{\tau_* - t}\right)^{\frac{15-k}{k}} - 1\right]\right)^{1/4},\\
&(iv) \quad R(t) = R_0 \left(\frac{\tilde{\tau}}{t + \tilde{\tau}}\right)^{3/k} \times \\
&\qquad \left(1 - \frac{k}{15 + k}\frac{\tilde{\tau}}{t_{\rm cc}}\left[\left(1 + \frac{t}{\tilde{\tau}}\right)^{\frac{15+k}{k}} - 1\right]\right)^{1/4}.
\end{aligned} \tag{42}$$

The times of coalescence are given by

$$\begin{aligned}
&(i) \quad t_{\rm coal} = t_{\rm ev}\left(1 - \left[1 - \frac{15 + k}{k}\frac{t_{\rm cc}}{t_{\rm ev}}\right]^{\frac{k}{15+k}}\right),\\
&(ii) \quad t_{\rm coal} = \tau\left(\left[1 - \frac{15 - k}{k}\frac{t_{\rm cc}}{t_{\rm ev}}\right]^{\frac{k}{k-15}} - 1\right),\\
&(iii) \quad t_{\rm coal} = \tau_*\left(1 - \left[1 + \frac{15 - k}{k}\frac{t_{\rm cc}}{\tau_*}\right]^{\frac{k}{15-k}}\right) \text{for } k < 15,\\
&\qquad t_{\rm coal} = \tau_* \text{ for } k \geq 15,\\
&(iv) \quad t_{\rm coal} = \tilde{\tau}\left(\left[1 + \frac{15 + k}{k}\frac{t_{\rm cc}}{\tilde{\tau}}\right]^{\frac{k}{15+k}} - 1\right).
\end{aligned} \tag{43}$$

The different situations are deeply nonequivalent.

- In the case (i), the system outspirals whenever $t_{\rm ev} < \frac{12}{k} t_{\rm cc}$. It first inspirals and, then, outspirals, for $\frac{12}{k} t_{\rm cc} < t_{\rm ev} < \frac{15+k}{k} t_{\rm cc}$. It only inspirals when $t_{\rm ev} > \frac{15+k}{k} t_{\rm cc}$.
- The case (ii) has to be dived into two sub-cases:
  - If $k < 3$, the system behaves as previously, but with different bounds. It outspirals for $\tau < \frac{12}{k} t_{\rm cc}$. It first inspirals and, then, outspirals, for $\frac{12}{k} t_{\rm cc} < \tau < \frac{15-k}{k} t_{\rm cc}$. It only inspirals when $\tau > \frac{15-k}{k} t_{\rm cc}$.
  - If, however, $k > 3$, the system outspirals for $\tau < \frac{15-k}{k} t_{\rm cc}$. It first outspirals and, then, inspirals for $\frac{15-k}{k} t_{\rm cc} < \tau < \frac{12}{k} t_{\rm cc}$. It inspirals with a monotonic behaviour for $\tau > \frac{12}{k} t_{\rm cc}$. This new behaviour was not encountered in the phenomenological cases presented before. It is illustrated in Fig. 6 with $k = 4$ (which would correspond to a mass loss rate proportional to T^4, with the temperature T proportional to $m^{5/4}$ as expected in some stellar models).
  - If $k = 3$, there is no intermediate regime. The system outspirals for $4t_{\rm cc} < \tau$ and inspirals for $4t_{\rm cc} > \tau$.
- In the cases (iii) and (iv), there is only one regime corresponding to a pure inspiralling behavior leading to the coalescence of the binary system. This was expected since a gain of mass enhances the effect due to gravitational waves. In the case (iii), the mass diverges in a finite amount of time but the system coalesces anyway, either before (when $k < 15$) or precisely at this time (when $k \geq 15$). This is not trivial and this results from the fact that the very rapidly growing mass also speeds up the inspiral.

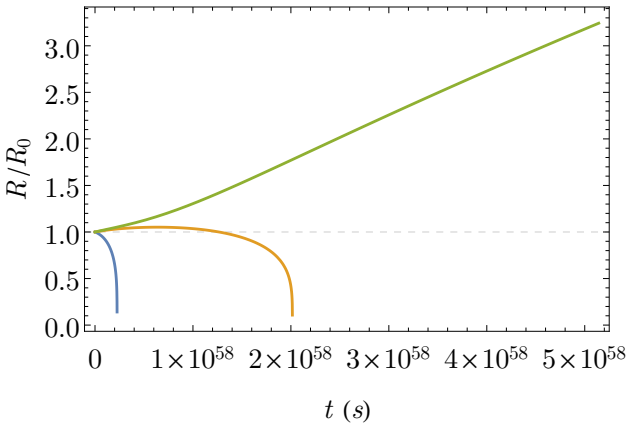
The richness of the solutions encountered lies in the asymmetrical role played by the different terms entering the differential equation: the mass variation can either tend to decrease or increase the orbital separation (depending on the sign of $\dot{m}$) and can either be smooth of catastrophic (depending on the exponent of m) but the GW term always plays in the same direction, although with a strength coupled to the mass loss.

Let us now turn to the case where $\dot{m} = \pm\beta m, \beta > 0$ which leads to a mass evolution exponentially decreasing or increasing, depending on the sign. The above differential equation integrates into

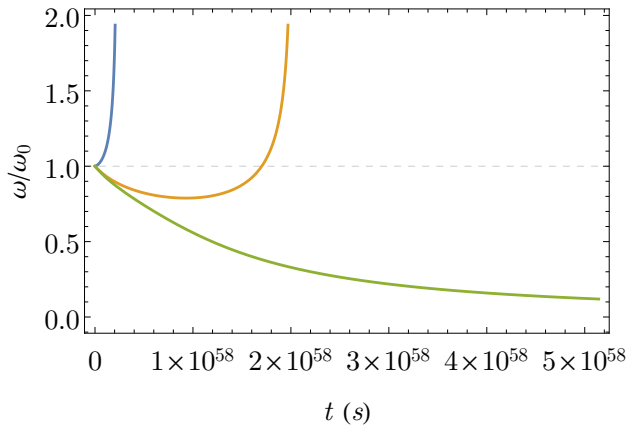
$$m_\pm(t) = m_0 e^{\pm t/t_*}, \quad t_* = \beta^{-1}. \tag{44}$$

The differential equation for the orbital separation simply reads

$$\dot{R} = -\frac{128}{5}\frac{G^3}{c^5}\frac{m_0^3}{R^3} e^{\pm t/t_*} \pm \frac{3}{t_*} R \tag{45}$$



(a) Orbital separation in function of time (in seconds)



(b) Frequency in function of time (in seconds)

FIG. 6. Illustration of the behaviour for the case (ii) with $k > 3$. The plot corresponds to $k = 4$, with initial mass $m_0 = 10^{12}$ kg. Panel (a) shows the orbital separation R (normalized to the initial orbital separation R_0) while panel (b) displays the corresponding orbital frequency. For $\tau < \frac{11}{4} t_{\rm cc}$ the binary system outspirals (green curve) whereas for $\tau > 3t_{\rm cc}$ it inspirals (blue curve). The regime $\frac{11}{4} t_{\rm cc} < \tau < 3t_{\rm cc}$ presents a non-monotonic evolution with an initial outspiral followed by an inspiral when the gravitational radiation takes the lead and makes the binary coalesce (orange curve).

and integrates into these two solutions (again, depending on the initially chosen sign):

$$R_\pm(t) = R_0 e^{\pm 3t/t_*} \left(1 \pm \frac{1}{15}\frac{t_*}{t_{\rm cc}} e^{\mp 15t/t_*}\right)^{1/4}. \tag{46}$$

In the case R_+, The dynamics is always outspiralling. The exponential accretion overwhelms any effect coming from gravitational radiation. It can be clearly seen analytically by taking the limit $t \to \infty$ in Eq. (46) for which the term in brackets vanishes and at large t, the overall dynamics is dictated by e^{3t/t_*}. Furthermore, it is straightforward to notice that the equation $R(t) = 0$ does not admit any mathematically sound solution. Conversely, for R_-, the binary system always merges with a

time of coalescence given by

$$t_{\rm coal} = \frac{t_*}{15}\ln\left(\frac{15t_{\rm cc}}{t_*}\right). \tag{47}$$

As it could have been expected, in both cases, the decreasing or increasing exponential evolution of the mass completely overwhelms the effects coming from gravitational radiation. Although mathematically rigorous, this case of exponential evolution of the mass might suffer from the fact that the formal expression of the radiated power by gravitational waves should be modified (beyond the simple accounting for the $m(t)$ term) when one allows a time-dependent mass, notably including terms related to the rate of variation of the mass. This point is discussed in the following section.

VI. RADIATED GW POWER AND CIRCULARITY CONDITION

In section II, we have taken into account the radiated power of emitted gravitational waves by substituting in the standard formula a varying mass. One should however be careful about the fact that the standard formula for $P_{\rm gw}$ is derived with the assumption of a constant mass. For circular orbits and time-varying masses, with orbital frequency ω, the second mass moments $M^{ij} = \mu x^i(t)x^j(t)$ read:

$$\begin{aligned} M_{11}(t) &= m(t)R^2\frac{1-\cos(\omega t)}{4},\\ M_{22}(t) &= m(t)R^2\frac{1+\cos(\omega t)}{4},\\ M_{12}(t) &= -\frac{1}{4}m(t)R^2\sin(\omega t). \end{aligned} \tag{48}$$

since for identical masses, the reduced mass is given by $\mu = m/2$. The total radiated power due to the emission of gravitational waves involves the third temporal derivatives of the second mass moments (48) since it is given, in the quadrupole approximation, by [16]

$$P_{\rm gw} = \frac{G}{5c^5}\langle \dddot{M}_{ij}\dddot{M}_{ij} - \frac{1}{3}(\dddot{M}_{kk})^2\rangle, \tag{49}$$

where the average $\langle\cdot\rangle$ is in the time domain over several characteristic periods of the gravitational waves. Using Eq. (48), we get

$$P_{\rm gw} = \frac{GR^4}{10c^5}\Big\langle \frac{\dddot{m}^2}{3} + \frac{3}{2}\left(\frac{3}{2}\ddot{m}^2 - \dddot{m}\dot{m}\right)\omega^2 + \frac{3}{2}\left(\frac{3}{2}\dot{m}^2 - \ddot{m}m\right)\omega^4 - \dddot{m}m\omega^3\sin(\omega t)\cos(\omega t) + \frac{1}{4}m^2\omega^6\Big\rangle. \tag{50}$$

If only the last term in the average in Eq. (50) is taken into account, one recovers the formula (5). All the other terms involve derivatives (up to order three) of the mass and are thus the direct consequence of allowing time-varying masses. If $m^{(n)} \equiv \frac{{\rm d}^n m}{{\rm d}t^n}$ denotes the n-th derivative of the mass, with n a positive integer, then these corrective terms can be neglected if

$$\forall n, \quad |m^{(n)}(t)| \ll m\omega^n, \tag{51}$$

which imposes that the mass of the bodies is slowly varying in comparison to the typical orbital evolution of the binary system. Then, over the temporal window on which the average is performed, that is to say on a few characteristic periods of the gravitational waves, the mass remains nearly constant equal to $m(t)$.

The second assumption used to compute the radiated GW power thanks to Eqs. (48) and (49) relies on the fact that the orbit remains circular. In the standard case of bodies with constant mass, the circularity is ensured by demanding $|\dot{\omega}| \ll \omega^2$ [16]. We investigate if this condition is changed by the variability of the mass. By taking the derivative of Kepler's third law, one has

$$\dot{R} = -\frac{2}{9}(\omega R)\frac{\dot{\omega}}{\omega^2} + \frac{1}{9}\frac{\dot{m}}{m}R. \tag{52}$$

Making use of the triangle inequality

$$|\dot{R}| \leq \frac{2}{3}(\omega R)\frac{|\dot{\omega}|}{\omega^2} + \frac{1}{3}\frac{|\dot{m}|}{m}R. \tag{53}$$

Since $|\dot{R}| \equiv v_\parallel$ is the radial velocity of the system, while $\omega R \equiv v_\perp$ represents its tangential velocity, Eq. (53) shows that the orbits are (quasi)-circular, *i.e.* $v_\parallel \ll v_\perp$, if $|\dot{\omega}| \ll \omega^2$ *and* $|\dot{m}| \ll m\omega$. This last condition exactly corresponds to the case $n = 1$ of Eq. (51). Under those two conditions, the approximation of a circular orbit with slowly varying radius is still applicable, even in the case of a variable mass system.

In conclusion, the *slowly varying mass condition*, encapsulated by the conditions (51), which enables to simplify the radiated GW power (50) to Eq. (5), is linked to the condition of (quasi-)circularity of the orbital trajectories during the inspiral phase.

To conclude this section, one should emphasize that the generic results of section V, although perfectly correct from a mathematical point of view, might suffer some lack of physical relevance in certain cases. Indeed, for mass evolutions presenting a stiff variation, the derivatives $\dot{m}, \ddot{m}, \dddot{m}$ cannot be legitimately neglected in the radiated power $P_{\rm gw}$ and Eq. (50) should be used instead, thus providing supplementary terms to the differential

equation on R. Investigating the precise effect of these additional terms would constitute a work in itself, as it might as well, as explained just above, have some consequences on the form of the orbital trajectories. Let us however note that, for instance, in the context of BHs submitted to Hawking evaporation, the slowly varying mass assumption is accurately satisfied – except for the higly fine-tuned case for which $t_{\rm coal} = t_{\rm ev}$ which might have an influence on the result of Eq. (29).

VII. CASE OF A HIGH HIERARCHY OF MASS

Henceforth, we have examined the case of a binary system constituted by two identical masses $m(t)$, following the same rate of variation. We now turn to the case where one of the bodies forming the binary has constant mass M whereas the other body has a varying mass $m(t)$, assuming that the masses obey the hierarchy $\forall t, M \gg m(t)$. In this case, the orbital angular momentum reads

$$J_{\rm orb} \approx m\sqrt{GMR}, \tag{54}$$

and its conservation leads to

$$\frac{\dot{R}}{R} = -2\frac{\dot{m}}{m}. \tag{55}$$

As Kepler's third law involves the *sum* of the masses of the two bodies, it now simply reads $\omega^2 \approx GM/R^3$. Furthermore, the orbital energy is

$$E_{\rm orbit} = -\frac{GmM}{2R}, \tag{56}$$

so that the power radiated by the mass loss is

$$P_{\rm ml} = \frac{3}{2}\frac{G\dot{m}M}{R}. \tag{57}$$

One must also be careful about the expression of $P_{\rm gw}$ in this case. We established in the previous analysis that, in full generality, one must take into account the variable mass through the computations of the derivatives of the second mass moments. It is still the case for a high hierarchy because the second mass moments M_{ij} depend on the reduced mass $\mu \approx m$ for $M \gg m$.

Since Kepler's third law shows that $\omega^2 \propto M$, the circularity condition of the orbit is satisfied only by demanding $|\dot{\omega}| \ll \omega^2$, as in the standard case of constant masses. Consequently, if one still wants to get rid of the higher derivatives terms of m in the expression of $P_{\rm gw}$ (*i.e.* $\dot{m}, \ddot{m}$ and $\dddot{m}$), it boils down to an additional assumption put by hand, which has – contrary to the case of identical evaporating masses – no link whatsoever with the circularity of the orbit. Under this hypothesis, one can use

$$P_{\rm gw} = \frac{32}{5}\frac{c^5}{G}\left(\frac{GM_c\omega}{c^3}\right)^{10/3} \approx \frac{32}{5}\frac{G^4}{c^5}\frac{M^3m^2}{R^5}. \tag{58}$$

The generic differential equation satisfied by the orbital separation (*i.e.* the counterpart of Eq. (7)) is consequently

$$\dot{R} = -\frac{64}{5}\frac{G^3}{c^5}\frac{M^2m}{R^3} - 2\frac{\dot{m}}{m}R, \tag{59}$$

which is again a Bernoulli differential equation which can be analytically integrated given an explicit expression of the mass. If one takes the same mass evolutions as the one given in Eq. (40), then Eqs. (42) and (43) are still valid up to the following systematic substitution (in the exponents as well as in some prefactors, the various signs being unchanged)

$$\frac{3}{k} \longrightarrow \frac{2}{k} \quad \text{and} \quad \frac{k}{15 \pm k} \longrightarrow \frac{k}{9 \pm k}. \tag{60}$$

In addition, the typical time $t_{\rm cc}$ now reads

$$t_{\rm cc} \equiv \frac{5}{256}\frac{c^5R_0^4}{G^3m_{\rm tot}^2\mu_0} \approx \frac{5}{256}\frac{c^5R_0^4}{G^3M^2m_0}. \tag{61}$$

One can also investigate the converse situation for which it is the biggest mass that varies in time *i.e.* $\forall t, M(t) \gg m$ and m is constant. Let us note that for Hawking radiation for instance, this situation is physically irrelevant since bigger BHs are less sensitive to the effect. For completeness, however, we shall treat all the cases. The conservation of the orbital angular momentum, using Eq. (54), leads this time to

$$\frac{\dot{R}}{R} = -\frac{\dot{M}}{M}. \tag{62}$$

In the expression of the orbital energy, and thus in the power radiated by the variation of mass, m and M play symmetric roles, so Eq. (56) is unchanged and we have

$$P_{\rm ml} = \frac{G\dot{M}m}{R}. \tag{63}$$

The conservation of the energy leads to the following (Bernoulli) differential equation for the orbital separation

$$\dot{R} = -\frac{64}{5}\frac{G^3}{c^5}\frac{M^2m}{R^3} - \frac{\dot{M}}{M}R. \tag{64}$$

If one takes the same mass evolutions as the ones given by Eq. (40), then Eqs. (42) and (43) are still valid up to the following systematic substitution (in the exponents as well as in some prefactors, the various signs being unchanged)

$$\frac{3}{k} \longrightarrow \frac{1}{k} \quad \text{and} \quad \frac{k}{15 \pm k} \longrightarrow \frac{k}{6 \pm k} \tag{65}$$

and the typical time $t_{\rm cc}$ is still given by Eq. (61).

VIII. CONCLUSION

Somehow surprisingly, even at lowest order, the dynamics of binary systems with varying mass is highly non-trivial when the emission of gravitational waves is taken into account.

In this work, we have investigated in details the evolution of binaries composed of evaporating primordial black holes with equal masses. We have shown that, depending on the initial conditions, there exist three different regimes. Both the orbital separation and the frequency of emitted gravitational waves are studied under the assumption of circularity. The critical case corresponding to an evaporation time exactly equal to the coalescence time was also considered.

Then, we focused on the case of a Bondi accretion of phantom dark energy. We took this opportunity to correct a mistake made in the literature and to settle a controversy about possible observations of this effect. This required to consider the strain, the frequency and the frequency drift of emitted gravitational waves.

Building on those results, a general study of all possible (power law) cases, including both mass losses and mass gains, was proposed. We gave full analytical solutions and, interestingly, suggested a taxonomy based only on variables easy to determine *a priori* – that is the typical variation time-scale and the time the coalescence would take with constant masses (not the actual coalescence time). We have shown that quite a few different situations appear, depending on initial conditions, on the sign of the mass variation, and on the regular or "explosive" nature of this variation. The existence of new non-monotonic regimes is underlined.

Finally, we have carefully stated the domain of validity of the different assumptions performed throughout the work and considered also the case of a high mass hierarchy between the inspiralling (or outspiralling) black holes.

In the future, it would be welcome to extend this work to the case of eccentric orbits and to investigate deeper the possible phenomenological consequences.

[1] P. M. Duerr, Stud. Hist. Phil. Sci. B **65**, 25 (2019).
[2] H. Gomes and C. Rovelli, (2023), arXiv:2303.14064 [physics.hist-ph].
[3] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **11**, 021053 (2021), arXiv:2010.14527 [gr-qc].
[4] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), (2021), arXiv:2111.03606 [gr-qc].
[5] M. Holgado and P. Ricker, The Astrophysical Journal **882**, 39 (2019).
[6] E. Barausse, V. Cardoso, and P. Pani, Phys. Rev. D **89**, 104059 (2014), arXiv:1404.7149 [gr-qc].
[7] L. Mersini-Houghton and A. Kelleher, Nuclear Physics B - Proceedings Supplements **194**, 272 (2009), new Horizons for Modern Cosmology.
[8] J. Enander and E. Mörtsell, Physics Letters B **683**, 7 (2010).
[9] S. M. O'Neill, M. C. Miller, T. Bogdanovic, C. S. Reynolds, and J. Schnittman, Astrophys. J. **700**, 859 (2009), arXiv:0812.4874 [astro-ph].
[10] C. F. B. Macedo, P. Pani, V. Cardoso, and L. C. B. Crispino, Astrophys. J. **774**, 48 (2013), arXiv:1302.2646 [gr-qc].
[11] A. Sarkar, K. Rajesh Nayak, and A. S. Majumdar, Phys. Rev. D **100**, 103514 (2019), arXiv:1904.13261 [astro-ph.CO].
[12] A. Sarkar, A. Ali, K. R. Nayak, and A. S. Majumdar, Phys. Rev. D **107**, 084038 (2023), arXiv:2210.12502 [gr-qc].
[13] K. Yagi, N. Tanahashi, and T. Tanaka, Phys. Rev. D **83**, 084036 (2011), arXiv:1101.4997 [gr-qc].
[14] J. H. Simonetti, M. Kavic, D. Minic, U. Surani, and V. Vijayan, Astrophys. J. Lett. **737**, L28 (2011), arXiv:1010.5245 [astro-ph.HE].
[15] C. Yuan, R. Brito, and V. Cardoso, Phys. Rev. D **104**, 124024 (2021), arXiv:2107.14244 [gr-qc].
[16] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments* (Oxford University Press, 2007).
[17] N. Aggarwal *et al.*, Living Rev. Rel. **24**, 4 (2021), arXiv:2011.12414 [gr-qc].
[18] E. Babichev, V. Dokuchaev, and Y. Eroshenko, Phys. Rev. Lett. **93**, 021102 (2004), arXiv:gr-qc/0402089.
[19] H. Bondi, Monthly Notices of the Royal Astronomical Society **112**, 195 (1952), https://academic.oup.com/mnras/article-pdf/112/2/195/9073555/mnras112-0195.pdf.
[20] E. Babichev, V. Dokuchaev, and Y. Eroshenko, J. Exp. Theor. Phys. **100**, 528 (2005), arXiv:astro-ph/0505618.
[21] S. M. Carroll, M. Hoffman, and M. Trodden, Phys. Rev. D **68**, 023509 (2003), arXiv:astro-ph/0301273.
[22] X. He, B. Wang, S.-F. Wu, and C.-Y. Lin, Phys. Lett. B **673**, 156 (2009), arXiv:0901.0034 [gr-qc].
[23] J. D. Hadjidemetriou, Icarus **5**, 34 (1966).
[24] F. Verhulst, Celestial Mechanics **11**, 95 (1975).
[25] S. T. McWilliams, Phys. Rev. Lett. **104**, 141601 (2010), arXiv:0912.4744 [gr-qc].
[26] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, Rept. Prog. Phys. **84**, 116902 (2021), arXiv:2002.12778 [astro-ph.CO].
[27] J. H. MacGibbon, Phys. Rev. D **44**, 376 (1991).
[28] F. Halzen, E. Zas, J. H. MacGibbon, and T. C. Weekes, Nature **353**, 807 (1991).
[29] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975), [Erratum: Commun.Math.Phys. 46, 206 (1976)].
[30] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*, Cambridge Monographs on Mathematical Physics (Cambridge Univ. Press, Cambridge, UK, 1984).
[31] B. Kocsis, T. Suyama, T. Tanaka, and S. Yokoyama, Astrophys. J. **854**, 41 (2018), arXiv:1709.09007 [astro-ph.CO].
[32] M. Raidal, C. Spethmann, V. Vaskonen, and H. Veermäe, JCAP **02**, 018 (2019), arXiv:1812.01930 [astro-ph.CO].

[33] A. D. Gow, C. T. Byrnes, A. Hall, and J. A. Peacock, JCAP **01**, 031 (2020), arXiv:1911.12685 [astro-ph.CO].
[34] J. M. Antelis, J. M. Hernández, and C. Moreno, J. Phys. Conf. Ser. **1030**, 012005 (2018).
[35] N. Aghanim *et al.* (Planck), Astron. Astrophys. **641**, A6 (2020), [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv:1807.06209 [astro-ph.CO].
[36] A. Barrau, K. Martineau, and C. Renevey, Phys. Rev. D **106**, 023509 (2022), arXiv:2203.13297 [gr-qc].
[37] F. Acernese *et al.* (Virgo), (2022), arXiv:2210.15633 [gr-qc].
[38] A. Barrau, J. García-Bellido, T. Grenet, and K. Martineau, (2023), arXiv:2303.06006 [gr-qc].

Chapter 2

Testing exotic models of dark energy

2.1 The standard cosmological model

It has been almost one hundred years since we discovered that the universe is expanding [45], giving rise to the *big bang model*. The big bang model is a widely held theory of the evolution of the universe. It proposes that the universe emerged from a state of extremely high temperature and density, the so-called big bang that occurred 13.8 billion years ago. The nucleosynthesis process that took place within roughly the first 10 minutes after the big bang is called big bang nucleosynthesis (BBN) or primordial nucleosynthesis. At this early epoch, the universe was dense and hot enough to allow for nuclear reactions to take place, producing the lightest elements such as hydrogen, helium and lithium [46].

After BBN, the universe was filled with a hot plasma of particles and radiation. As it expanded and cooled, protons and electrons combined to form neutral hydrogen atoms. This process is called recombination and it occurred about 380,000 years after the big bang [47,48]. It released photons that have now been redshifted to microwave frequencies, forming the cosmic microwave background radiation [49], which provides us with a snapshot of the universe when it was just a few hundred thousand years old.

Over the next few hundred million and billion years, gravity caused matter in the universe to clump together, forming the first structures, such as galaxies, and clusters of galaxies. However, baryons are not sufficient to explain the large scale structures, the rotation curve of galaxies or the gravitational lensing from galaxy clusters that we observe today [50]. These observations requires the existence of additional non-relativistic matter called *dark matter* [50].

In the late 1990s, observations of distant supernovae revealed that the expansion of the universe is accelerating, rather than slowing down as expected [23]. This unexpected acceleration is thought to be caused by a cosmological constant or more generally dark energy, which makes up about 68% of the energy density of the universe.

In this section we will introduce further the necessary ingredients of the standard model of cosmology: BBN, recombination, dark matter and dark energy, as well as discuss their observational signatures. Refs. [50–58] were the primary sources for this short introduction.

2.1.1 From nucleosynthesis to recombination

Today, the observations tell us that matter is much more abundant than radiation, the latter being dominated by the cosmic microwave background. However, this picture was reversed in

the early universe, because radiation density, which evolves as $\rho \propto a^{-4}$, is more diluted than matter, which scales as $\rho \propto a^{-3}$, with the expansion. It is during the radiation dominated era that the BBN took place, while the cosmic microwave background was emitted around its end. To better understand the reactions that took place at early times, we start be introducing thermodynamic equilibrium and decoupling.

The early universe was characterized by a high-density and high-temperature state. Particle interactions during this period were mainly described by the reaction rate denoted as Γ. If the reaction rate was much larger than the Hubble rate H, particles could remain in thermodynamic equilibrium. Under these conditions, particles could be treated as perfect Fermi-Dirac and Bose-Einstein gases with a distribution given by

$$F_i(E,T) := \frac{g_i}{2\pi^3}\frac{1}{e^{(E-\mu_i)/T} \pm 1} \tag{2.1}$$

where E is the energy of a single particle, T is the temperature, g_i is the degeneracy factor and μ_i is the chemical potential. In general, different types of particles can have different temperatures. But all the species that interact with photons will have the same temperature T, the temperature of radiation, which at this epoch is also the temperature of the universe. Using the distribution (2.1), one can find the macrostate variables such as the number density or the density, namely

$$N_i = \int F_i(\mathbf{p},T)\mathrm{d}^3\mathbf{p}\,, \quad \rho_i = \int F_i(\mathbf{p},T)E(\mathbf{p})\mathrm{d}^3\mathbf{p}\,. \tag{2.2}$$

In particular, we used in the paper in section 1.3.2 the following relation between energy density and temperature,

$$\rho(T) = g_*\frac{\pi^2}{30}T^4\,. \tag{2.3}$$

If the cross-section of a reaction varies with temperature as $\sigma \sim T^x$ (for example $x = 2$ for the electroweak interaction), then the reaction rate, which is proportional to both the number density of particles N and the cross-section, varies as $\Gamma \sim N\sigma \sim T^{x+3}$. On the other hand, during the radiation period of the universe the Hubble parameter varies as $H \sim T^2$. Hence, as long as $x+1 > 0$, there will always be a temperature below which the interaction becomes less efficient and decouples as the universe cools down. This decoupling process occurs because the interaction can no longer maintain the equilibrium between different particles due to cosmic expansion. This property is fundamental to the thermal history of the universe and the existence of *relics*, particles left over from early stages of the universe.

Big bang nucleosynthesis is one of the observational pillar of the standard model of cosmology, describing the early universe. It is the process by which light elements, such as helium, deuterium, and lithium, were synthesized during the first few minutes after the Big Bang. Its importance lies in its ability to constrain the baryonic density together with the abundance of light elements. The standard BBN scenario can be divided into three main steps:

Neutron-proton equilibrium At temperatures above 1 MeV, for $t < 1$ s, the weak interaction rates are rapid enough to maintain the neutron-proton equilibrium through reactions

such as $n \leftrightarrow p + e^- + \bar{\nu}_e$. As long as statistical equilibrium holds, the neutron to proton ratio is given by $n/p = \exp\big(-Q_{\rm np}/T\big)$, where $Q_{\rm np}$ is the neutron-proton mass difference.

Weak interaction freeze-out Around temperatures of 0.8 MeV for $t \sim 2$ s, the weak interaction rates become slower than the expansion rate of the Universe, $\Gamma \lesssim H$, and the neutron-proton ratio freezes out. The number of neutrons and protons changes only from the neutron β-decay until the temperature drops to about 0.1 MeV ($t \sim$ a few minutes), when the production of deuterium D is faster than its dissociation.

Synthesis of light elements For temperatures between 0.6 MeV and 0.05 MeV, or 3 s to 6 min after the big bang, the synthesis of light elements occurs, but only through two-body reactions. This is one reason why only an insignificant number of elements heavier than lithium were created. They require triple collisions to be formed such as the triple-alpha process producing carbon, but very high density and long period of time are needed, conditions only found in stars. Two-body reactions require the formation of deuterium through $p + n \to D$ and a negligible photo-dissociation. This happens roughly when $N_D \sim N_\gamma$, the number density of deuterium becomes significant compared to the number density of photons or equivalently when the energy of photons is lower than the binding energy of deuterium. In order to find the relative abundance of the elements D, H, ^{3}He, ^{4}He, ^{7}Li, etc., one needs to solve a system of coupled differential equations. Other observations fix the cosmological behaviour, therefore the time-dependent temperature and we have a good understanding of the nuclear cross-sections. This leaves only one free parameter to constrain, which is the baryon-photon ratio η, which is well constrained by the Planck satellite [24]. In Fig. 2.1, we show the comparison between the theoretical prediction of abundances of the light elements and the observational results. We note that there is a discrepancy for the abundance of ^{7}Li, which is one of the main mysteries of modern cosmology.

At temperatures higher than the hydrogen ionization energy, matter remains ionized, and photons are tightly coupled to electrons through Compton scattering. However, as the universe expands and cools, the formation of neutral atoms becomes favoured from a thermodynamical perspective. At this point, Compton scattering becomes less efficient, and radiation decouples from matter, leading to the formation of the CMB. This decoupling marks the point at which the universe becomes transparent to photons, and they can freely propagate in all directions. This process called *recombination* involves the transition of ionized hydrogen (protons and electrons) to neutral hydrogen atoms. The equilibrium between ionization and recombination is governed by the Saha equation,

$$\frac{X_{\rm e}^2}{1 - X_{\rm e}} = \left(\frac{m_{\rm e} T}{2\pi}\right)^{3/2} \frac{{\rm e}^{-E_{\rm I}/T}}{N_{\rm p} + N_{\rm H}}, \tag{2.4}$$

which relates the relative abundances $X_{\rm e}$ of electrons $N_{\rm e}$, protons $N_{\rm p}$, and neutral hydrogen atoms $N_{\rm H}$, namely

$$X_e = \frac{N_{\rm e}}{N_{\rm p} + N_{\rm H}}\,, \tag{2.5}$$

in terms of temperature T and the hydrogen ionization energy $E_I = m_p + m_e - m_H$, which is the difference between the mass of the constituents and the mass of the atom. Recombination

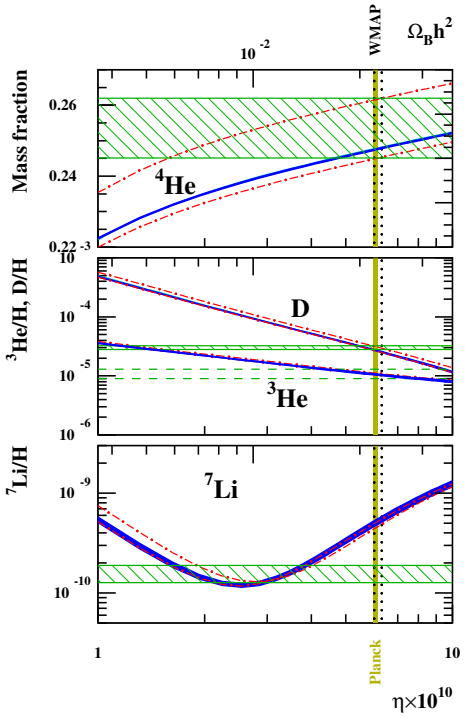


Figure 2.1: abundances of ^{4}He, D, ^{3}He, and ^{7}Li as a function of the baryon-photon ratio η. The blue line corresponds to the theoretical prediction. The vertical regions are the constraints on η derived from WMAP (dotted black) and Planck (yellow). The horizontal green shaded regions correspond to the observed abundances of the light elements. The red dot-dashed lines correspond to the range of values for the effective number of neutrino families N_{eff} obtained from Planck, with the extreme values being $N_{eff} = 3.02$ and $N_{eff} = 3.70$. Retrieved from [52].

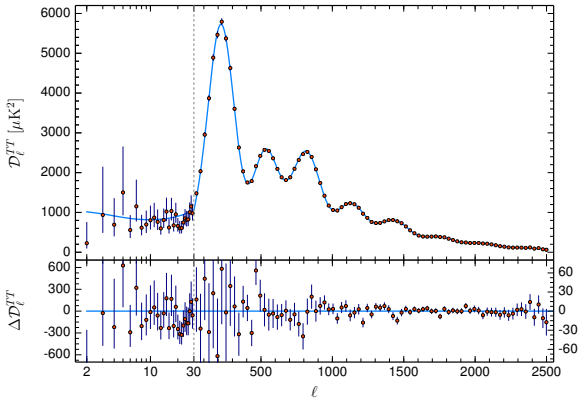


Figure 2.2: Power spectrum of the temperature anisotropies of the CMB. The red dots represent the experimental results with the standard deviation. The blue line follows the best fit from the Λ-CDM to the measurements. Retrieved from [24].

occurs when the temperature drops significantly below the hydrogen ionization energy ($T < E_\mathrm{I}$), leading to a sudden drop of the ionization fraction X_e. Using today's observations of the average temperature T_0 of the CMB and the photon number density $N_{\gamma 0}$, we can predict that recombination happened around redshift $z = 1200$. This is derived using the Saha equation with the relations $N_\mathrm{b} := N_\mathrm{p}+N_\mathrm{H} = \eta N_{\gamma 0}(1+z)^3$ and $T = T_0(1+z)$. The average temperature of the CMB, denoted as T_0, has been measured with precision by the COBE satellite, yielding a value of $T_0 = 2.725 \pm 0.001$ K at 2σ confidence level [59]. The observed spectrum of the CMB closely resembles a black body spectrum. The nearly perfect agreement with a black body spectrum indicates that the CMB has undergone thermalization, primarily through interactions with electrons. Any energy injection at redshifts lower than $z = 10^6$ would distort the Planck spectrum and can be constrained by observations.

After the removal of the monopole and dipole components from the cosmic microwave background (CMB), where the dipole arises from our local velocity relative to the CMB frame, small temperature anisotropies persist with a relative amplitude of approximately 10^{-5}. These anisotropies result from fluctuations in the radiation fluid during recombination, which have been redshifted due to the expansion of the universe. These anisotropies can be analyzed using spherical harmonics decomposition:

$$\frac{\delta T}{T}(\theta,\phi) = \sum_l \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\phi),, \tag{2.6}$$

where a_{lm} are the coefficients corresponding to the spherical harmonics $Y_{lm}(\theta,\phi)$. The angular power spectrum $C_l = \langle |a_{lm}|^2 \rangle$ is derived from these coefficients, which involves averaging the squared coefficients over m. Figure 2.2 displays the angular power spectrum $D_l = l(l+1)C_l$ as measured by the Planck satellite [24]. The blue line corresponds to the prediction from the Λ-CDM model described in this section with a quasi-invariant primordial power spectrum, which will be discussed in more details in the next chapter 3. The experimental results represented by the red dots show us the angular power spectrum as observed in the CMB. The large error bars at low multipole, or equivalently for large angles, originate from the lack of a sufficient amount of coefficients a_{lm} per multipole and could

never be alleviated. Understanding the origin and properties of these fluctuations is crucial in modern cosmology and will be a central part of the next chapter 3.

The current concordance model of cosmology, the Λ-CDM, is characterised by seven parameters that must be constrained experimentally. First, there are the density parameters in the Friedmann equation (1.34). The spatial curvature is assumed to be zero, hence there are three independent parameters that must be considered, radiation density Ω_r, matter density Ω_m and the Hubble parameter H_0. the cosmological constant can be recovered from these three parameters through the Friedmann equation. The model also distinguishes between baryonic matter Ω_b and cold dark matter Ω_c, raising the number of background parameters to 4. The photons from the CMB were affected by the ambient matter during the billion years long travel through Thomson scattering. This effect is summarized in the *optical depth parameter* τ. Finally, the initial conditions for the fluctuations that are observed in the CMB must be set. It appears that power spectrum of these primordial fluctuations is almost scale invariant. Hence, the primordial power spectrum can be describe using only two parameters: the *amplitude* A_s and the small *tilt* n_s of the primordial power spectrum. Together, these 7 parameters are enough to describe the current standard model of cosmology and the CMB allows us to constraint all of them, provided that some priors are given. This is the role of other independent experiments that are designed to measure one or several parameters, such as type Ia SNe for dark energy, galaxy lensing for dark matter or BBN for baryonic matter. The radiation parameter is constrained using the average temperature of the CMB, while the other parameters are constrained using the temperature anisotropies. The most recent constraints can be found in Ref. [24].

2.1.2 Main evidences for dark matter

The observation of the abundance of light elements as predicted from the BBN hypothesis was able to tightly constrain the baryon-photon ratio $\eta \sim 6 \cdot 10^{-10}$. If one extrapolates this value to today's density parameters using Friedmann's equations, we obtain that baryons should represent roughly 5% of the energy content of the universe, i.e. $\Omega_b \sim 0.05$. Using the angular power spectrum of the CMB displayed on Fig. 2.2, one can indeed confirm that baryons constitute 5% of the content, but one also needs the existence of some exotic matter, five times more abundant than baryons, that do not interact with radiation. This sort of non-interacting matter is necessary to explain the observed density fluctuations of the CMB. Hence, the known particles of the standard model, hereby baryons, are not sufficient to explain the observed CMB temperature anisotropies power spectrum. This important signature of dark matter is far from being the only evidence suggesting the existence of such exotic matter. Some attempts at modifying our theory of gravity to explain dark matter has been made, such as MOND theory [60, 61], but such modified theories struggle to explain the observed CMB as well as other measurements that will be described in this section and the community favours the matter hypothesis.

Historically, the first signs of dark matter has been found in 1933 by Fritz Zwicky [62]. He observed that the Coma galaxy cluster does not satisfy the virial theorem using its estimated mass and kinetic energy. Nowadays, galaxy clusters are great probes for dark matter, because their mass can be well estimated using various methods. First, one can study the velocity profile of galaxies within the cluster and derive the gravitational field or equivalently the mass distribution of the cluster. Its mass can also be extracted from

the gravitational lensing on distant sources. At last, the mass profile can also be found by measuring the gas cloud temperature and density using X-ray observatories. From the temperature, one can derive the pressure profile and therefore the mass of the cluster through the virial theorem, assuming hydrostatic equilibrium. In the other hand, the mass of baryons can be estimated from the luminosity of the cluster and observations lead to a 5-to-1 share of dark to baryons, agreeing with the CMB measurements. A particular cluster, called the bullet cluster, resulting from the collision of two clusters of galaxies has given strong evidence towards dark matter as opposed to modified gravity theories. Indeed, it has been observed that the distribution of dark matter of both clusters passed through each other without being affected, while baryons was slowed through interactions during the collision. The distributions of both type of matter do not overlap in this particular cluster, a fact that is hard to satisfy using modified gravity theories [63]. Furthermore, constraints on dark matter self-interactions could be inferred [64].

It was not before the 70s and the measurement of the rotation curve of the Andromeda galaxy by Vera Rubin [65], that dark matter was seriously studied. Since then, thousands of rotation curves of galaxies have been measured suggesting that there is ten times more dark matter than baryons in galaxies. These results somehow predicts more dark matter than CMB and cluster mass measurements, but the mass of galaxies only accounts for a small fraction of the total mass of matter in the universe. Therefore, we do not expect the amount of dark matter in galaxies to be representative of the distribution on cosmological scales. Recently, the discovery of dwarf galaxies without dark matter [66, 67] also disfavours modified gravity theories, due to the difficulty to explain such a phenomenon.

2.1.3 Late time universe and dark energy

We saw earlier that dark matter constitutes roughly 30% of the energy content of the universe, we have yet to talk about the remaining 70%. While the measurements by the Planck satellite have been a cornerstone of modern cosmology, the fit of the Λ-CDM parameters is not unique and it requires the input of priors. The value of the cosmological constant is no exception and other independent observations are necessary. Fortunately, since the discovery of the accelerated expansion in the late 90s [23], there has been numerous evidences of the existence of a positive cosmological constant.

At first, the accelerated expansion of the universe was observed using type Ia supernovae as standard candles to measure the expansion rate with respect to the distance. The expansion rate is measured using the redshift and the distance is deduced from the luminosity of the standard candle. In Fig. 2.3 we show the relation between the luminosity distance and the redshift for the Supercal compilation [68] of supernovae data and compare it to different model of expansion. Δ stands for the normalized difference in magnitude. We see that the data strongly favour a small cosmological constant, since the expansion was first decelerating when matter was dominating, but is now accelerating due to the dark energy contribution.

A very useful property of the type Ia supernovae is their ability to behave as a standard ruler allowing us to measure distances. In a cosmological setting, a very useful ruler was created at recombination in a process called *baryon acoustic oscillations (BAO)*. During the radiation era, the universe was filled with a photon-baryon plasma and dark matter. The energy content was very close to being homogeneous, but small fluctuations persisted

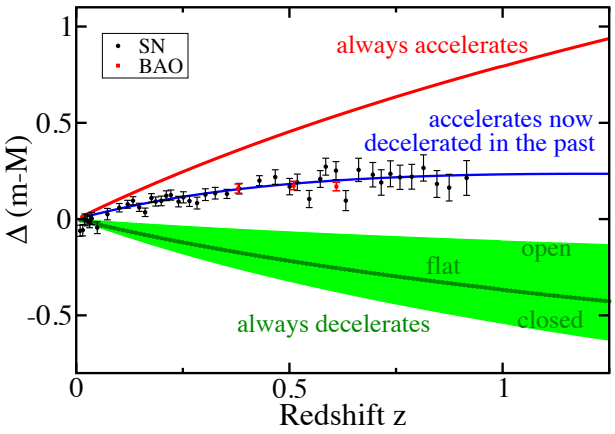


Figure 2.3: Evidence of past deceleration transitioning to present-day acceleration. The blue line represents a well-fitting model with late acceleration. In green, a range of matter-only models, with $0.3 \leq \Omega_m \leq 1.5$, covering open, flat, and closed geometries without dark energy. The red curve represents a constant acceleration model. Black data points are binned distance moduli from the Supercal compilation [68] of 870 supernovae, while three red data points represent distances inferred from recent BAO measurements [69]. Retrieved from [57].

and in particular fluctuations in the dark matter fluid. The latter being decoupled from the plasma, the fluctuations of dark matter grew from the action of gravity and attracted with it the baryon-photon plasma creating high density and high pressure regions. The pressure in these over-dense regions were such that sound waves travelling at the speed of sound $c_s \approx c/\sqrt{3}$ in the plasma were emitted outward. These sound waves travelled until recombination happened, when the speed of sound drastically dropped and effectively froze the over-dense spheres created by the wave. The size of the spheres r_{BAO} at recombination can be easily calculated from the distance the sound waves had time to travel in roughly 380000 years, namely

$$r_{BAO} = \int_0^{t_{rec}} \frac{c_s}{a(t)} \mathrm{d}t \sim 145\,\mathrm{Mpc}\,. \tag{2.7}$$

With time and the expansion of the universe, these over-dense regions grew and influenced the distribution of galaxies in the universe. Using galaxy catalogues such as the Sloan digital sky survey (SDSS) [70] listing galaxies and their redshift, one can map the distribution of galaxies and observe the over-dense spheres. Because we know the size of the sphere at recombination and one can measure their size as observed from earth today, we can derive the expansion history of the universe. In Fig. 2.3, the red dots represent the results from the measurements of these spheres and confirms the scenario of late time acceleration. Moreover, we show in Fig. 2.4 the over-density at comoving distance r_{BAO}.

The effect of dark energy is effective at large scales and it is natural to look for its contribution on the evolution of large scale structures. These structures, primarily composed of dark matter, distort the image we have of distant objects through gravitational lensing. In particular, they deform the shape of galaxies. The goal is to reconstruct the source of the lensing, i.e. the dark matter structures and derive the dark energy contribution. The distortion is measured using the *convergence* κ and the *complex shear* (γ_1, γ_2), which acts

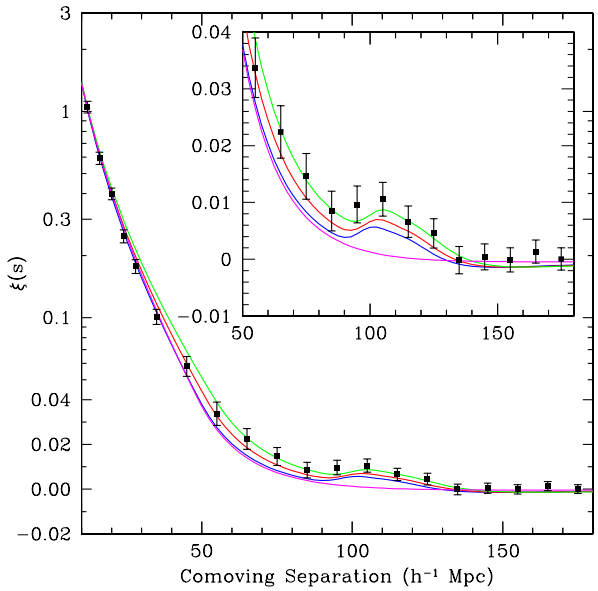


Figure 2.4: Large-scale redshift-space correlation function of the SDSS sample in black with the error bars. Models: $\Omega_m h^2 = 0.12$ (green), 0.13 (red), 0.14 (blue with peak); all with $\Omega_m h^2 = 0.024$, $n = 0.98$, and mild non-linear prescription. Magenta line: pure CDM model ($\Omega_m h^2 = 0.105$) without acoustic peak. Retrieved from [71].

as a coordinate transformation namely

$$\delta \mathbf{x}^S = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \delta \mathbf{x}^I \,, \tag{2.8}$$

where $\delta \mathbf{x}^{S,I}$ is the displacement vector in the plane of the source and the image, respectively. When studying gravitational lensing at large scale, we are mainly interested in *weak lensing*, which is characterized by κ, $\sqrt{\gamma_1^2+\gamma_2^2} \ll 1$. The main statistical signal that can be extracted from averaging over the shape of galaxies is the power spectrum of the shear, which turns out to be identical to the convergence power spectrum [72]

$$P_l^\kappa(z_s) = \int_0^{z_s} \frac{(1+z)^2}{d_L(z)H(z)} W^2(z) P\left(k = \frac{(1+z)^2 l}{d_L(z)}, z\right) \,, \tag{2.9}$$

where z_s is the redshift of the source, $d_L(z)$ the luminosity distance and $P(k,z)$ the matter power spectrum. $W(z)$ is the weight function that describes the efficiency of the lensing from galaxies in the foreground and is the highest for lenses lying at roughly half the distance between the source and the observer. Dark energy influences the shear power spectrum at different levels. Evidently, it affects the Hubble constant, but also the luminosity distance, the weight function and the matter power spectrum through the redshift evolution. On the left panel of Fig. 2.5, we show the resulting cosmic shear from a mass distribution that has been derived using a N-body simulation. On the right panel, we plot the shear power spectrum with the projected statistical errors from the future LSST, renamed Vera Rubin, observatory.

Dark energy is yet well described by a cosmological constant, but the constraints on its

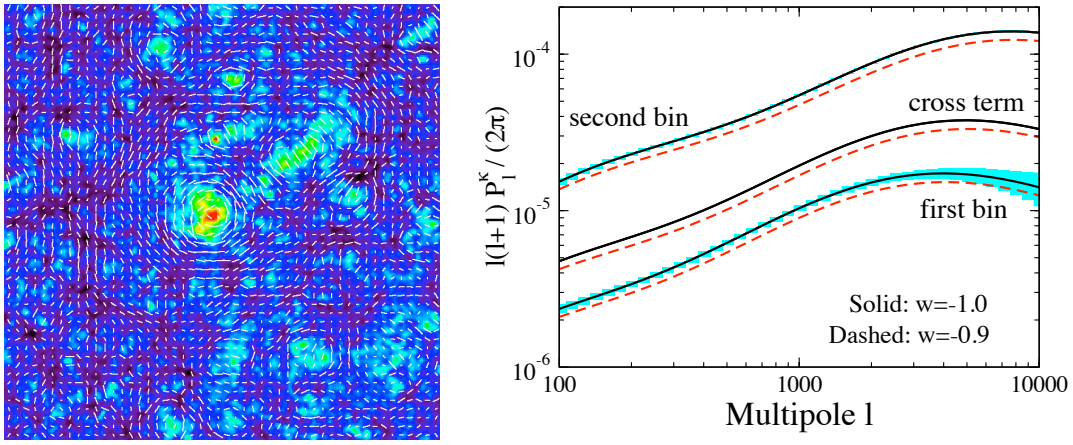


Figure 2.5: *Left panel:* Cosmic shear field (white lines) overlaying a projected mass distribution from a cosmological N-body simulation. Bright areas represent overdense regions, while dark areas represent underdense regions. *Right panel:* Angular power spectrum of cosmic shear with statistical errors expected for LSST in light blue. Shown for source galaxies in two redshift bins: $0 < z_s < 1$ (first bin) and $1 < z_s < 3$ (second bin). The cross-power spectrum between the bins (cross term) is also displayed. Retrieved from [57].

dynamics are still loose and the nature of dark energy remains under scrutiny. Current and future cosmological surveys measuring the CMB, supernovae or galaxies aim at constraining the dynamics of dark energy to confirm or reject the cosmological constant hypothesis. There are mainly two variables of interest when studying the accelerated expansion, the density parameter of dark energy $\Omega_\Lambda(a)$ and its equation of state parameter $w_\Lambda(a)$, both potentially evolving with time and therefore with the scale factor a. More precisely, cosmological surveys set constraints on the value of the density parameter today $\Omega_{\Lambda,0} \sim 0.7$, as well as the value of $w_\Lambda(a_0) = w_0$ and small variation around its current value $w_a = w'(a_0)$, with prime being the derivative with respect to a. According to the Λ-CDM model, we have $w_0 = -1$ and $w_a = 0$. Current constraints on $w_0 - w_a$ as well as $w_0 - \Omega_m$ (recall that $\Omega_m \sim 1 - \Omega_\Lambda$) using the previously mentioned experimental evidences are exposed in Fig. 2.6. We see indeed that the constraints on the dynamics of dark energy are loose and there is a need for new cosmological observations, such as galaxy surveys.

2.2 Dark energy as quintessence and future constraints

Even though the cosmological constant model is well fitted to describe the accelerated expansion of the universe, it still suffers from some theoretical issues. First and foremost, its existence and more importantly its very small value appears somewhat arbitrary and bizarrely fine-tuned. A more problematic issue emerges from an other side of fundamental physics and the description of subatomic particles using quantum field theory (QFT). QFT predicts that the energy density of the vacuum is non-zero and would remarkably behave as a positive cosmological constant in the context of cosmology [58]. The vacuum energy density originates from the creation of particle-antiparticle pairs and its strength is therefore related to the mass of the known elementary particles. However, rough estimates of

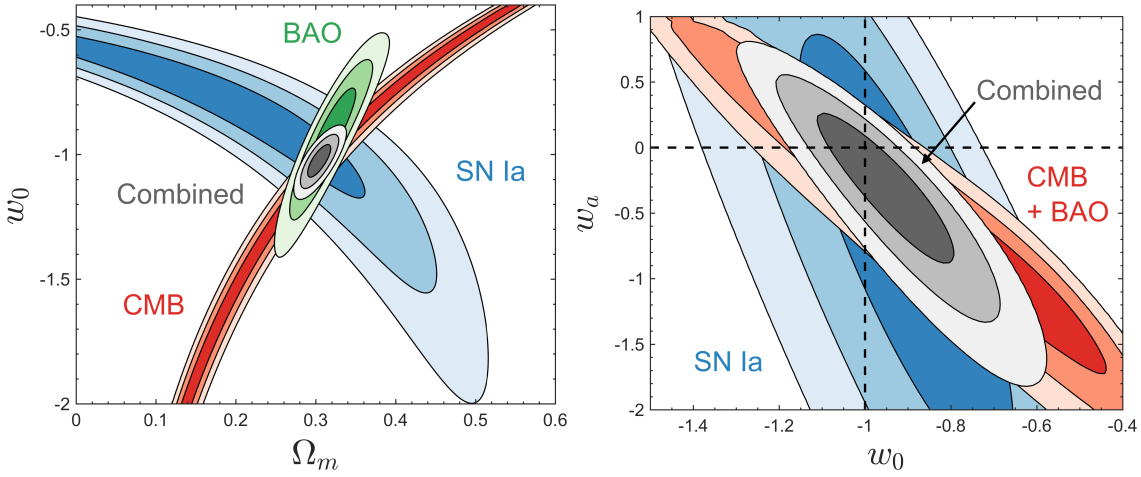


Figure 2.6: Cosmological parameter constraints from our analysis of three key probes: SNe Ia [73], BAO [69], and CMB [74]. *Left panel* displays constraints on Ω_m and w, while the *right panel* shows constraints on $w_0 - w_a$, marginalized over Ω_m. Contours represent likelihood regions of 1-σ, 2-σ and 3-σ. Retrieved from [57].

the cosmological constant derived from vacuum energy is off by over 50 orders of magnitude [75, 76]. Of course, one could claim that this vacuum energy is unphysical and QFT predictions are wrong, but it has been experimentally observed in the Lamb shift [77] or the Casimir effect [78] for example. The erroneous prediction for dark energy as vacuum energy is sometimes coined as the *cosmological constant problem*, even though the issue lies deeper in the QFT approach [58]. Through the renormalization procedure commonly used in QFT, one has some freedom in choosing the counter-terms that are used to get rid of infinities plaguing the theory. We can therefore make the cosmological constant as small as we like with finely tuned counter-terms in the one-loop correction, which would bring the cosmological constant problem back to a fine tuning problem. Unfortunately, the story does not end here. The counter-terms are not safe from higher order loop corrections, leading to an infinite tower of finely tuned counter-terms [58]. This tells us that the cosmological constant is very sensitive to UV physics and we do not have at present a UV-complete theory describing the standard model of particle physics.

A solution that is often brought forward consists in assuming some yet unknown symmetries that would render null the effect of vacuum energy on the expansion of space-time. But one is left to explain the observed accelerated expansion in other ways. Since the primary goal is to get rid of the cosmological constant and explain why the acceleration of the expansion happened late in the history of the universe, dark energy is made dynamical. The simplest field that we have at our disposal in order to make dark energy dynamical is the scalar field. It has the advantage of being simple and respects the cosmological principle, since it is naturally isotropic and can be made homogeneous. Of course, the random addition of a scalar degree of freedom feels somewhat arbitrary, but there are well-motivated arguments to postulate the existence of a scalar field. If the graviton carries a small mass for example, then the gravitational field propagate five degrees of freedom, including a scalar field. Moreover in the small graviton mass limit, the vector degrees of freedom are negligible and one is left with a scalar-tensor theory [7, 79, 80]. We will introduce the resulting scalar-tensor theory in Sec. 2.3.2. More generally, It was recently shown that [81–84] a scalar-field naturally emerges as a Goldstone boson when the time translation symmetry is broken.

Hence, studying scalar-tensor theories is part of the search of potential violations of the strong equivalence principle and the symmetry breaking of the time translation could explain the accelerated expansion. Such an approach to dark energy has been named *effective field theory (EFT) of dark energy.*

Quintessence is a scalar-tensor theory with minimal coupling between the scalar and tensor sector. It has gained considerable attention in theoretical cosmology thanks to its simplicity and ability to describe the accelerated expansion. The scalar field used in quintessence can of course be derived using the EFT approach to dark energy. Nevertheless, some suggested that in its low energy limit, string theory could give rise to a quintessence model [85, 86]. String theory is known to have an absurd number of different possible low energy EFT, called the *landscape* and it is no surprise that quintessence could emerge from such a vast theory. However, it has proven so far impossible to find an EFT compatible with a lasting period of accelerated expansion and this led, together with many other arguments, to the conjectured de-Sitter condition [85]

$$\frac{\mathrm{d}V/\mathrm{d}\phi}{V} \gtrsim \mathcal{O}(1)\,, \tag{2.10}$$

where $V(\phi)$ the potential of the scalar field, for a theory to be part of the landscape. Scalar-tensor theories that do not satisfy this condition (and other similar conditions that has been postulated [87]) are said to be part of the *swampland* of string theory. Such a condition on the behaviour of dark energy can in fact be measured experimentally and current results using the constraints on $w_0 - w_a$ lead to $\mathrm{d}V/\mathrm{d}\phi < 0.6V$, already in slight conflict with the de-Sitter conjecture.

In the coming years, several new missions with the possibility to study the behaviour of dark energy are expected to be launched. The square kilometer array (SKA) is an array of radio-telescopes, which is expected to observe a large amount of galaxies at low redshift. This will allow the refinement of the BAO constraints on the dynamics of dark energy [88]. Moreover, the future space-based Euclid and the ground-based Vera Rubin observatories will be able to catalog the shape and redshift of a lot of new galaxies, improving the constraints on the equation of state parameter from weak lensing [89,90]. The following paper pedagogically introduces the notion of swampland and quintessence, and studies the future constraints that we will be able to set on the de-Sitter conjecture using these future observatories.

The Astrophysical Journal, 912:99 (15pp), 2021 May 10 https://doi.org/10.3847/1538-4357/abedbe

The String Theory Swampland in the Euclid, Square Kilometer Array, and Vera Rubin Observatory Era

Aurlien Barrau, Cyril Renevey, and Killian Martineau
Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3 53, avenue des Martyrs, F-38026 Grenoble cedex, France
Received 2021 January 8; revised 2021 February 26; accepted 2021 March 9; published 2021 May 10

Abstract

This article aims to draw the attention of astronomers to the ability of future cosmological surveys to put constraints on string theory. The fact that "quantum gravity" might be constrained by large-scale astrophysical observations is a remarkable fact that has recently concentrated a great amount of interest. In this work, we focus on future observatories and investigate their capability to put string theory, which is sometimes said to be "unfalsifiable," under serious pressure. We show that the combined analysis of the Square Kilometer Array, Euclid, and the Vera Rubin observatory—together with Planck results—could substantially improve the current limits on the relevant string swampland parameter. In particular, our analysis leads to a nearly model-independent prospective upper bound on the quintessence potential, $|V'|/V < 0.16$, in strong contradiction of the so-called de Sitter conjecture. Some lines of improvements for the very long run are also drawn, together with generic prospective results, underscoring the efficiency of this approach. The conjectures used in this work are discussed pedagogically, together with the cosmological models chosen in the analysis.

Unified Astronomy Thesaurus concepts: Cosmology (343); Dark energy (351); Quintessence (1323)

1. Introduction

String theory might be the only available and viable candidate "theory of everything". It features many unique and appealing properties (see, e.g., Polchinski 2007; Danielsson 2001; Becker et al. 2006; Blumenhagen et al. 2013 for pedagogical introductions). Replacing point particles with one-dimensional quantum objects leads to a completely new paradigm that has been built over the last five decades. Among many others, an important consequence of this fundamental shift is the natural appearance of a massless spin-2 particle, the graviton. In a sense, (quantum) gravity is therefore an unavoidable prediction of string theory. The long history of string theory (see, e.g., Rickles 2014) went through several revolutions, from the discovery of superstrings—and the understanding that the theory might be capable of describing all elementary particles as well as all the interactions between them—to the unification of different versions of the model into the M-theory framework (see, e.g., Schwarz 1996). The ideas of string theory have far-reaching consequences in mathematical physics, cosmology, condensed matter physics, particle physics, nuclear physics, and black hole physics. More than a well-defined axiomatic model, string theory, in a broad sense, is a kind of rich and intricate framework (see, e.g., Tong 2009) made of many interconnected subfields.

In spite of its extraordinary mathematical elegance—beginning with the historical Green–Schwarz anomaly cancellation mechanism (Green & Schwarz 1984), which led to the first revolution—string theory raises questions about its falsifiability. One might provocatively argue that all the predictions made so far were somehow contradicted by observations. First, the number of dimensions required is not the one we know in nature (Lovelace & Lett 1971). Second, the world should be supersymmetric (Gliozzi et al. 1977). Third, the cosmological constant is expected to be negative (see, e.g., arguments in Witten,). Fourth, non-Gaussianities are to be observed in the cosmological microwave background (CMB; Lidsey & Seery 2007). Of course, ways out of those naïve expectations are well known, and it is not the purpose to describe them here in detail: compactified extra dimensions are possible, supersymmetry can be broken at a high-energy scale, the introduction of membranes as new fundamental objects might stabilize a positive cosmological constant, moduli fields might not contribute simultaneously to the inflation dynamics, etc. Without going into the quantitative arguments (the interested and unfamiliar reader can get a flavor of the main ideas in, e.g., Gubser 2010), it is fair to conclude that, at this stage, string theory is not in contradiction with observations. The main concern is different. Just the other way round, one might wonder if string theory is actually falsifiable (see, e.g., references in Smolin 2006; Rovelli 2013). It might be feared that the intrinsic richness of the paradigm is such that basically anything happening in nature could be accounted for by string theory. The very scientific nature of string theory could then be questioned, at least in a Popperian sense. Although quite a lot of possible "tests" were considered in the past (see, e.g., Kane 1998; Casadio & Harms 2000; Hewett et al. 2005; Kallosh & Linde 2007; Durrer & Hasenkamp 2011), the great novelty now emerging is that cosmological surveys might severely constrain—if not exclude—string theory. This work aims to derive new results on the limits that could be obtained in the next decade and to drive, pedagogically, the attention of astronomers on this somehow unexpected importance of the study of large-scale structures for fundamental high-energy physics.

As far as surveys are concerned, we focus on the following three main experiments: Euclid (Laureijs et al. 2011; Amendola et al. 2013, 2018), the Vera Rubin observatory (Abell et al. 2009; Alonso et al. 2018), and the Square Kilometer Array (SKA; Yahya et al. 2015; Maartens et al. 2015; Santos et al. 2015). Euclid is a European Space Agency mission using a Korsch-type space telescope. Its diameter is 1.2 m with a focal length of 24.5 m. The instruments are a visible-light camera (VIS) and a near-infrared camera/spectrometer (NISP). The Vera Rubin observatory, previously called LSST, is a ground-

based telescope whose diameter is 8.4 m with a focal length of 10.3 m. The focal plane is made of a 3.2 billion pixel matrix, and the redshifts will be determined photometrically. Both are scheduled for 2022. Finally, SKA is a radio telescope array project with stations extending out to a distance of 3000 km from a concentrated central core. It will be operated in several successive stages, and the fully completed array is expected around 2027. In the following, Euclid and LSST will often be mixed as they provide basically the same information in a complementary way.

We first explain, as intuitively as possible, the main ideas of the string theory swampland program that allow a connection between quantum gravity and very large-scale—ultra-low-energy—astronomical observations. This is rooted in the apparent impossibility of producing a real de Sitter space (positive cosmological constant) in string theory. According to the de Sitter conjecture, if we were to measure a scale factor behavior resembling too much a pure cosmological constant, the whole edifice might be in trouble. We then focus on quintessence models for dark energy and explain the specific potentials we choose together with the motivations underlying those choices to test the conjecture. The next section is devoted to the presentation of the results for each considered class. The aim is to answer this question: could the next generation of experiments show that we live in the so-called swampland, therefore excluding string theory (*if* the de Sitter conjecture, which will be described in detail, is correct)? We conclude by underscoring the limitations of the approach and possible developments to be expected for the future.

2. The String Theory Swampland

Although string theory is a tentative unification theory, it actually leads to a huge amount of false (metastable) vacua: the so-called landscape (see, e.g., Susskind 2003; Banks et al. 2004; Polchinski 2006; Arkani-Hamed et al. 2007). Technically, the tremendously large number of possibilities comes from choices of Calabi–Yau manifolds and of generalized magnetic fluxes over different homology cycles. More intuitively, this landscape represents a vast collection of effective low-energy theories. Otherwise stated, although the fundamental equations of string theory are often simple and elegant, their solutions are extraordinarily complex and diverse. This diversity is not necessarily a failure of physics. It may help in understanding some curious features of our universe. Just as Earth is a very specific—and anthropically favored—place in the observable universe, the observable universe could be a special sample in the "multiverse" (see, e.g., Stoeger et al. 2004; Garriga et al. 2006; Vilenkin 2007; Barrau 2007; Carr 2007; Hall & Nomura 2008). In principle, the idea of a multiverse, generated for example—but not necessarily—by inflation, filled with different low-energy landscape realizations of string theory is not outside of the usual scientific arena: probabilistic predictions can be made and tested (Freivogel 2011). Having a single sample at our disposal (our universe) makes things complicated but not radically different from the usual situation in physics where the amount of information available is always finite and incomplete (Carroll 2018). It has never been mandatory to check all the predictions of a theory to make it scientific. Of course, providing clear predictions in a string landscape multiverse is, to say the least, challenging: from the clear definition of a measure to the accurate estimate of the anthropic weight, many questions remain open (Stoeger et al. 2004). Still, the idea that string theory generates a landscape of solutions—which can be considered as a set of effective physical theories—remains central to the field. Although often associated with the multiverse, this wide variety of solutions can, of course, be also meaningfully considered in a single universe, which is the framework of this study.

The situation changed drastically with the emergence of the swampland idea (Vafa 2005; Ooguri & Vafa 2007). The swampland refers to the huge space of theories that seem compatible with (or possibly derived from) string theory but which, actually, are not. The landscape corresponds to the, more restricted, space of "possibly correct" theories, in the sense that they really emerge from string theory, which is here assumed to be the framework. The swampland corresponds to theories that are not consistent when considered in detail, although they appear to be correct at first sight. The landscape is a much smaller subset of the space of theories than the swampland: the island of consistent theories in a sea of incoherent proposals, even if those proposals were looking like viable candidates from a simple field theory point of view. In many works, the strategy consists in guiding the construction of effective field theories so that they belong to the landscape and not to the swampland. This is a valuable help in low-energy model building. In this article, we take a different view. We try to determine the correct description of the cosmological dynamics and show that it might actually correspond to a swampland theory. If the real world were to lie in the swampland, it might strongly suggest that either string theory itself or the swampland program is wrong.[1] This is a unique—although still intensively debated and clearly controversial—way of possibly falsifying a candidate theory of quantum gravity. Surprisingly, the detailed features of the acceleration of the universe are a privileged way to address this outstanding question.

Let us be slightly more specific. Excellent reviews of the swampland ideas can be found in Brennan et al. (2017), Palti (2019). We simply give here a flavor of the arguments to fix the ideas of the unfamiliar reader. In the space of theories, the swampland would be vastly larger than the landscape. Theories in the swampland are apparently fine. They can be used to make predictions and look like good physics. But, when looking more carefully, they cannot be consistently completed in the ultraviolet (that is, at high energy). This does not mean that they are not quantum theories of gravity, but rather, that they cannot appear consistently as a low-energy limit of a "string-theory-inspired" quantum gravity model. Otherwise stated: the coupling to gravity is problematic for those theories. One way of using this information is to discard such theories and focus on other models. However, as explained before, an interesting situation would be the one where our universe is described by a theory belonging to the swampland: the logic could then be reversed and the framework suggesting the theory to be discarded should be revised or abandoned. The swampland program is very wide and uses at the same time rigorous string theory arguments, general quantum gravitational ideas, and microphysics inputs.

Illustratively, one can consider an effective quantum field theory (EQFT) self-consistent up to a scale $E_{\rm self}$. If gravity is added to the game (with its own finite scale

[1] We will resist the temptation to consider the possibility that string theory is correct and that it is the real world that is wrong!

$E_{\rm Planck} = \sqrt{\hbar c^5/G}$), the new theory will exhibit a new limit energy scale $E_{\rm swamp}$. This is basically the energy above which the theory has to be modified if it is to become compatible with quantum gravity at very high energies. An interesting situation (Palti 2019) is one where $E_{\rm swamp} < E_{\rm self} < E_{\rm Planck}$. Even better, if $E_{\rm swamp}$ is smaller than any characteristic energy scale involved in the theory, it means that the whole theory lies in the swampland. This can be very helpful as a guide in a model-building approach.

At this stage, the swampland program is mostly an ensemble of conjectures. Some of them are very reliable and—to some extent—demonstrated. They are theorems (even if the hypotheses are sometimes stronger than one would like). Others are speculative and grounded in extrapolations (for example, from known vacua to all vacua). They rely on many different kinds of arguments, from very formal ones to qualitative *gedankenexperiments* involving black holes.

Among others, the following swampland conjectures are intensively being considered: the distance conjecture (Ooguri & Vafa 2007; Klaewer & Palti 2017; it is not possible to move too much in the field space), the weak gravity conjecture (Arkani-Hamed et al. 2007; Cheung & Remmen 2014; gravity is always the weakest force), the species scale conjecture (Veneziano 2002; Dvali et al. 2002; there is a bound on the cutoff scale that depends on the number of particles), the no global symmetry conjecture (Banks & Dixon 1988, 1988; there cannot be an exact global symmetry in a theory with a finite number of states), the completeness conjecture (Polchinski 2004; if there is gauge symmetry, the theory has to incorporate states with all possible charges), the emergence conjecture (Harlow 2016; Heidenreich et al. 2018; Grimm et al. 2018; Ooguri et al. 2019; the kinetic terms for all fields are emergent in the low-energy limit by integrating out states up to a finite scale), the non-supersymmetric AdS instability conjecture (Ooguri & Vafa 2017; non-supersymmetric anti–de Sitter space is unstable), the spin-2 conjecture (Klaewer et al. 2019; any theory with spin-2 massive fields has a cutoff scale[2]), etc. In this work, we shall focus on the so-called de Sitter conjecture (Obied et al. 2018; Andriot 2018; Garg & Krishnan 2019; Ooguri et al. 2019). It basically encodes the fact that it is very hard to build something resembling a positive cosmological constant in string theory. Let us describe the underlying idea.

The key point is that in the 11-dimensional supergravity theory arising as a low-energy limit of M-theory, there is no cosmological constant. What is usually described as a cosmological constant actually corresponds, in this framework, to the local minimum of a scalar potential. This results from the compactification of the higher-dimensional (supersymmetric) theory. A de Sitter space does not here refer to a "real" positive cosmological constant but, instead, to the (meta)stable state of a scalar field with a positive value of the potential. Anything looking like a positive cosmological constant is extremely hard to construct in string theory. The profound reasons behind this are, on the one hand, the instability of the so-called moduli[3] in the de Sitter vacuum and, on the other hand, the non-supersymmetric nature of the de Sitter space. The heart of the conjecture lies in the bet that the tremendous difficulties arising when trying to build a de Sitter vacuum are hints that such a state does not exist at all in the theory (Danielsson & Riet 2018).

This would not mean that the de Sitter–like behavior of the contemporary universe—that is, its accelerated expansion (see Riess et al. 1998 for the historical detection and Mortonson et al. 2013 for a brief and recent review)—is in radical conflict with the conjecture. Just as it was probably the case for the primordial inflation (see Senatore 2017 for an excellent introduction), it might be that the current acceleration is obtained dynamically by a scalar field rolling down a potential. In such a case, the conjecture might hold but would constrain the shape of the potential. Although they suffer from serious drawbacks and are obviously more complicated than a pure cosmological constant (Bianchi & Rovelli 2010), such quintessence models must not appear as too strange or unnatural. Indeed, if a true cosmological constant happens to become dominant in the cosmic dynamics, it will automatically remain so forever as it does not dilute, whereas the other fluids of the universe do. It is therefore impossible that inflation was driven by the cosmological constant and the introduction of something like a scalar field evolving in an appropriate potential is required. It makes sense to assume that the contemporary phase of acceleration, which resembles the primordial one, is due to the same cause. Quintessence models for dark energy are precisely focused on this idea (see Brax 2018) and receive strong support from both particle physics and EQFTs. This might provide a solution that pleases string theory and remains compatible with what we observe. For the sake of completeness, we should however mention that quintessence is *not* the only alternative to a cosmological constant.

However, it is not sufficient to choose a dynamical dark energy model instead of a pure cosmological constant to avoid falling in the swampland. The very interesting aspect of the de Sitter conjecture is that it provides a constraint that has to be satisfied by the potential. Let us consider the usual action for gravity and scalar field ϕ with a potential V:

$$S = \int d^4x\sqrt{-g}\left[R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \qquad (1)$$

where R is the Ricci scalar.[4] Unless otherwise stated, we use Planck units throughout this work. The original de Sitter conjecture (Obied et al. 2018) states that

$$\frac{|V'|}{V} > \lambda_c, \qquad (2)$$

where λ_c is a constant of order one, and the derivative is to be understood with respect to the field. Interestingly, recent studies set a clear numerical value on the large field limit (Bedroya & Vafa 2020; Andriot et al. 2020a): $\lambda_c > \sqrt{2/3} \approx 0.82$. It has already been shown (Agrawal et al. 2018) that current observations lead to $|V'|/V < 0.6$, which is in tension with the conjecture. We will show that the Euclid/Rubin/SKA generation of experiments should lead to a clear improvement of this limit and potentially show that the

[2] In a way, this simple fact about effective field theories significantly predates the entire Swampland program (Arkani-Hamed et al. 2003).

[3] Moduli fields are scalar fields with properties often associated with supersymmetric systems. They arise naturally in string theory.

[4] The forthcoming statements can be generalized to a set of scalar fields, introducing a supermetric G_{ij} in the field space, but we prefer to keep the notations simple at this stage.

actual behavior of the universe unambiguously does not satisfy the conjecture.

This simple version of the de Sitter conjecture is, however, not sufficient. The well-understood Standard Model of particle physics allows one to extrapolate the existence of other critical points of the potential. This argument was first made for the Standard Model Higgs potential (Denef et al. 2018) where some loopholes were pointed out in the argument. It was then elaborated on in Murayama et al. (2018) and extended in Choi et al. (2018) to consider the pion potential, leading to the firm conclusion that there is no way that a quintessence model coupled to the Standard Model can be consistent with the original de Sitter swampland conjecture that bounds only $|V'|/V$. Avoiding a critical point of the pion potential would require a large quintessence coupling in violation of stringent fifth-force constraints. These arguments led to a refined form of the bound so that the potential has tachyonic directions in the regions where V' is too small.[5]

The refined conjecture reads (Ooguri et al. 2019; Garg & Krishnan 2019)

$$\frac{|V'|}{V} > \lambda_c \quad \text{or} \quad \frac{V''}{V} < -\alpha_c, \tag{3}$$

where λ_c and α_C are both positive numbers of order 1. In addition to the previously given arguments, the second condition also comes from the fact that when it is satisfied, an instability develops, leading to a breakdown of the entropy-based argument for the first condition. In the following, we first consider the original conjecture and then comment on the refined version. It should also be mentioned that several other improvements are also being considered (Garg & Krishnan 2019; Andriot 2018; Ben-Dayan 2019; Dvali et al. 2019; Garg et al. 2019; Andriot & Roupec 2019).

Why are we to expect the derivative of the potential (normalized to the potential) not to be too small in string theory? The reasons are subtle (Ooguri et al. 2019; Garg & Krishnan 2019), and no obvious argument can be given. The de Sitter space is a highly nontrivial structure. It has a horizon and is endowed, analogously to what happens for black holes, with a temperature and entropy. This is expected not to be stable in quantum gravity. The de Sitter conjecture can be seen as resulting from the distance conjecture (Ooguri et al. 2019; Geng 2020). Very loosely speaking, this is because the distance traveled in the field space is linked to the dimension of the Hilbert space of the theory, which is itself (potentially) linked to the entropy of the considered de Sitter space. It can also be approached from thermodynamical considerations (Seo 2019): under several assumptions, it might be that when the number of degrees of freedom is enhanced as the modulus rolls down the potential, the bound on $|V'|/V$ becomes equivalent to the condition for the positive temperature phase. As mentioned before, the original de Sitter conjecture has been refined (Andriot 2018, 2019; Andriot & Roupec 2019), and there are now quite a lot of arguments supporting its validity although the accurate value of the bound λ_c is still debated. Let us introduce $\lambda \equiv -V'/V$ (which depends on ϕ, hence on time) and summarize: if a potential resembling too much a cosmological constant (that is $|\lambda| \ll 1$) were to be measured, it would hardly be compatible with string theory if the conjecture is correct.

In principle, studying the primordial inflation leads to more stringent constraints on λ than considering the contemporary acceleration of the universe. Although inflation is unquestionably part of the cosmological paradigm, it is, however, fair to note that other scenarios are also being considered: strictly speaking, inflation is not proven. Several alternatives are described in Durrer & Laukenmann (1996), Hollands & Wald (2002), Veneziano (2003), Brandenberger (2011), Creminelli et al. (2010), Popławski (2010), Wilson-Ewing (2013), Lilley & Peter (2015), and references therein. If a study aims to investigate how the actual behavior of the universe constrains —or even might exclude—string theory, it is mandatory to rely on established facts. For this reason, we focus on the late-time acceleration that is being directly observed.

3. Dark Energy Potentials

3.1. General Picture

The acceleration of the universe is now unquestionable. This has been observed using Type Ia supernovae (Perlmutter et al. 1999), using the cosmological microwave background (CMB) (Aghanim et al. 2020), and using baryon acoustic oscillations (BAOs; Blake et al. 2011). In a way, a pure cosmological constant—as a part of the full Einstein's equations according to Lovelock's theorem—works well to explain this acceleration (Bianchi & Rovelli 2010). If the quantum vacuum fluctuations are taken into account the question becomes merely the one of a suitable renormalization. This, however, raises many questions (see, e.g., Brax 2018) and the idea of quintessence is intensively studied to overcome most coincidence problems (see Zlatev et al. 1999; Martin 2008; Tsujikawa 2013 for reviews). Any model of quintessence must not only produce a scalar field potential that is sufficiently flat but also fine-tune away the cosmological constant. From the point of view of our approach, it is important to consider such models, as a pure cosmological constant lies trivially in the swampland if the de Sitter conjecture holds. To be conservative on a possible "exclusion" of string theory, it is necessary to focus on dynamical models.

It is not possible to derive limits on λ that are strictly model independent, in exactly the same way as it is not possible to obtain observational limits on the equation of state (EOS) of dark energy as a function of the redshift without assuming a given parameterization. One needs to consider a specific quintessence model that is the potential on which the scalar field is rolling and compare the cosmological trajectories with current—or future—data to derive limits on λ. We shall discuss the precise methodology in the next section, but it is worth emphasizing now the deep limitation of our approach. We will consider different classes of quintessence models that are the main paths discussed and conservatively assume that the bound to keep for the analysis is the less stringent one. It is, however, not possible to exclude that a future model might relax our limits. To minimize this risk, we investigate very different potentials corresponding to extremely different philosophies.

The second Friedmann equation, obtained by combining the first one with the trace of Einstein's equations reads, without a

[5] One should remain careful to avoid a possible loophole in the reasoning. If the conjecture is modified each time a model correctly describing our world contradicts it, there is no way it can be used to falsify the framework in which the conjecture is established. It is therefore mandatory to strengthen, in the near future, the theoretical ground on which the de Sitter conjecture is derived.

cosmological constant and for a flat space:

$$\frac{\ddot{a}}{a} = -\frac{1}{6}\sum_i(\rho_i + 3p_i), \tag{4}$$

where a is the scale factor and ρ_i and p_i are, respectively, the energy density and pressure for the fluid of type i. In a universe filled with matter (p_m=0), radiation ($p_R/\rho_R = 1/3$), and a scalar field ϕ (with energy density ρ_ϕ and pressure p_ϕ), this means

$$\frac{\ddot{a}}{a} = -\frac{1}{6}(\rho_m + 2\rho_R + \rho_\phi + 3p_\phi). \tag{5}$$

A dark energy quintessence model must therefore ensure that the scalar field dominates over the other components and exhibits a negative enough pressure to achieve $\ddot{a} > 0$. In Equation (5), the field pressure $p_\phi = \frac{1}{2}\dot{\phi}^2 - V$ is the only term that can be negative. In practice, this is realized by the domination of the potential energy over the kinetic energy of the field, hence the usual reference to flat enough potentials. The tricky part lies in the fact that this domination of the negative pressure of the scalar field over all other components has to happen very late in cosmological history. And, if possible, for a wide range of initial conditions. This set of constraints has basically led to three classes of relevant potentials that we have used to study the de Sitter conjecture in this work. For all of them, we define the EOS parameter $w \equiv p_\phi/\rho_\phi$. Intuitively speaking, in freezing models, the motion of the field gradually slows down because the potential becomes flat at low redshift. In the thawing model, the field was initially frozen due to the Hubble friction (as during inflation), and it started evolving when the Hubble rate became small enough. Thawing models have a value of w which began near –1 and increased with time, whereas freezing models have a value of w which decreased, usually (but not necessarily) toward –1 (Scherrer & Sen 2008; Chongchitnan & Efstathiou 2007; Duary & Banerjee 2019). In a way, freezing models more obviously answer the question of producing the required behavior of the universe but thawing models are easier to build and, in this sense, more natural.

Using the Friedmann and Klein–Gordon (K-G) equations, one can write first-order differential equations for the parameters w, $\Omega_\phi = \rho_\phi/(3H^2)$, and λ, with respect to the number of e-folds N:

$$\frac{w}{N} = (w-1)[3(1+w) - \lambda\sqrt{3(1+w)\Omega_\phi}], \tag{6}$$

$$\frac{\Omega_\phi}{N} = -3w\Omega_\phi(1-\Omega_\phi), \tag{7}$$

$$\frac{\lambda}{N} = -\sqrt{3(1+w)\Omega_\phi}\,(\Gamma - 1)\,\lambda^2, \tag{8}$$

where $\Gamma \equiv VV''/(V')^2$. Rewriting the equations of motion under the form of Equations (6)–(8) is very useful as we are ultimately interested in constraining $|\lambda|$ using constraints on $w_a \equiv -dw/dN|_{N=0}$ and $w_0 \equiv w|_{N=0}$ (with $a_0 = 1$). Furthermore, as observations show that $\Omega_{\phi,0} \sim 0.7$ (Aghanim et al. 2020), where the subscript 0 refers to contemporary values, we consider that physical trajectories have to fulfill $\Omega_{\phi,\mathrm{past}} \ll 1$ and $\Omega_{\phi,0} \sim 0.7$, which is easily implementable with Equations (6)–(8). In addition, these differential equations describe directly the parameters in which we are interested, as opposed to the Friedmann and K-G equations, thus improving the intuitive interpretations.

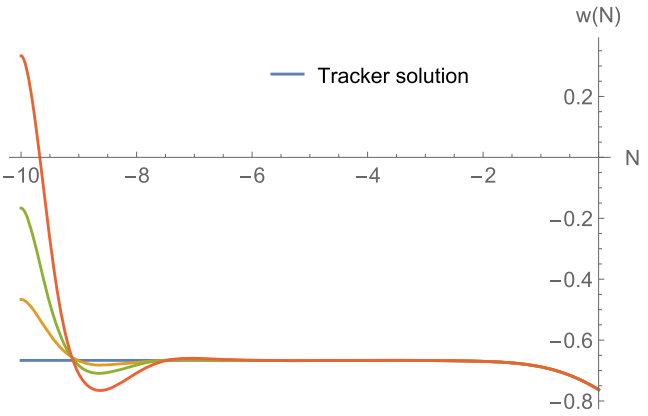


Figure 1. Stability of the solution for tracking freezing models. The blue curve represents the tracker trajectory, while the yellow, green, and orange curves represent trajectories starting with various initial conditions for $w(N_i)$.

3.2. Tracking Freezing Models

A very effective class of potentials fulfilling the previous conditions leading to tracking freezing solutions is the Ratra–Peebles potentials (Peebles & Ratra 1988; Ratra & Peebles 1988):

$$V(\phi) = M^{4+\alpha}\phi^{-\alpha}, \tag{9}$$

where M is a characteristic energy scale and $\alpha > 0$. The remarkable feature of those potentials lies in the fact that the scalar field "tracks" the background evolution. This means that the EOS parameter of the field, w, changes at the transition between radiation domination and matter domination. The field adapts itself to the scale factor behavior. The density of the field can be shown to decrease less rapidly than that of the surroundings, ensuring the late-time domination.

In addition to being a tracker, this solution is also an attractor. Impressively, the attractor solution is joined before present times for more than 100 orders of magnitude in the initial energy density of the field (Martin 2008). If α is large enough, it is possible that the energy scale M of the theory becomes high whereas the energy scale of the acceleration remains very small, which has far-reaching consequences for naturalness. This is somehow reminiscent of the so-called see-saw mechanism in particle physics. General considerations on tracking solutions are given in Steinhardt et al. (1999), where it is, in particular, shown that the very existence of a tracker depends on the behavior of the function $\Gamma(\phi)$, which needs to be greater than one at any time. In the case of the potential described by Equation (9), $\Gamma = 1 + 1/\alpha$, and the tracker condition is always satisfied.

The dynamics makes the system nearly insensitive to initial conditions. Whatever the reasonable choices made in the remote past, at the initial time N_i, trajectories converge to the tracker solution before $N = 0$. One aspect of this behavior is illustrated in Figure 1. In addition, because the parameter λ is directly related to w and dw/dN through Equation (6), the stability of the EOS parameter under changes in the initial conditions is reflected in the behavior of λ. It can be shown, using Equations (6)–(8), that when the EOS parameter w is

quasi-constant—that is, in the tracker regime—one has

$$w \approx w_i \coloneqq \frac{-2}{2+\alpha} \quad \text{and} \quad \lambda^2 \approx \lambda_i^2 \coloneqq 3\frac{1+w}{\Omega_\phi}. \tag{10}$$

When adding the requirement that $\Omega_{\phi,0} \approx 0.7$, the system is fully determined.

As a side effect, this raises a quite interesting epistemological question. In the limit $t \to \infty$, this potential leads to $w \to -1$ and $-V'/V \to 0$. This will inevitably violate the de Sitter conjecture at some point in the future. Does it, however, make sense to use a cosmological behavior that has not yet taken place to rule out a model? Or, at least, should we use the expected future dynamics in our sets of constraints? Although the answer is not obvious and deserves to be debated, we will remain conservative and resist the temptation to extrapolate beyond the contemporary epoch. First, because, strictly speaking, the future is not written. With the cosmic fluids considered so far, the evolution is obviously known but nothing prevents surprises (other fields, exotic matter, etc.). Second, because the de Sitter space might be unstable and "decay" before the conjecture is violated. Third, because it is a kind of "safe" principle to assume that physics is to describe what *does* exist and not what might or will exist.

3.3. Scaling Freezing Models

The so-called scaling freezing models typically rely on potentials of the form (Ferreira & Joyce 1998)

$$V(\phi) = V_0 e^{-\lambda\phi}, \tag{11}$$

where V_0 and λ are constants. The scaling solution itself corresponds to $\Omega_m/\Omega_\phi = \text{constant}$, where the Ω_k are the densities normalized to the critical density. It was shown (Copeland et al. 1998) to be realized at the fixed point

$$\frac{\dot\phi}{\sqrt{6}H} = \sqrt{\frac{3}{2}}\frac{1+w_m}{\lambda}, \tag{12}$$

$$\frac{\sqrt{V}}{\sqrt{3}H} = \sqrt{\frac{3(1-w_m^2)}{2\lambda^2}}. \tag{13}$$

Such a potential also exhibits a tracking-like solution that acts as an attractor for physical trajectories with different initial conditions. In this case, the function $-V'/V$ is constant and can be seen as a simple parameter. The system is solely described by Equations (6) and (7). Moreover, the stable tracking solution has a quasi-constant EOS parameter $w \approx w_i \approx -1$ and, as Ω_ϕ is initially set to have $\Omega_{\phi,0} \sim 0.7$, the system is fully determined.

More refined models were constructed (Sahni & Wang 2000; Albrecht & Skordis 2000), involving two exponential functions

$$V(\phi) = V_1 e^{-\lambda_1\phi} + V_2 e^{-\lambda_2\phi}, \tag{14}$$

where V_i and λ_i are constants. Constraints on the parameters have been derived (Chiba et al. 2013) for the model to be consistent with CMB and BAO data. Contrary to the previous potentials, neither λ nor Γ is independent of ϕ, and Equations (6)–(8) are ill defined. Fortunately, one can easily show that $\lambda(\phi)$ is bijective, which allows one to consider Γ as a function of λ without affecting the system. Once again, the attractor nature of the tracking solution also allows the set of initial conditions to be properly defined. In this case, the EOS parameter w is quasi-constant for $w \approx w_i \approx 0 \Longrightarrow \lambda \approx \lambda_i \coloneqq \sqrt{3/\Omega_\phi(N_i)}$, with $\Omega_\phi(N_i)$ being set so that $\Omega_{\phi,0} \sim 0.7$.

3.4. Thawing Models

In thawing models, the field is initially frozen with $w \approx -1$ and then begins to evolve. This is typically produced by potentials like (Frieman et al. 1995)

$$V(\phi) = V_0(1 + \cos(\sqrt{2}\ \phi/f)) \quad \text{or} \tag{15}$$

$$V(\phi) = V_0 \cos(\phi/f), \tag{16}$$

where V_0 and f are constants. Although the phenomenologies associated with those potentials are very close, we keep both expressions as some subtleties might differ. In this case, the evolution begins at late times but has to remain weak to account for data. Constraints were derived in Dutta & Scherrer (2008) and Gupta et al. (2012). The behavior of the thawing dynamics around the redshift interval $0.6 < z < 1$ is quite sensitive to deviations from a pure cosmological constant (Sen et al. 2010). The potentials we choose here are not the only possible ones but allow most of the dynamics for this kind of model to be caught.

To study the behavior of $w(N)$ with the potentials given by Equations (15) and (16), we use an analytical approximation from Dutta & Scherrer (2008) and Scherrer & Sen (2008), which reads

$$w(a) \approx -1 + (1+w_0)a^{3(K-1)}\mathcal{F}^2(a), \tag{17}$$

where

$$\mathcal{F}(a) \equiv \frac{(K-F(a))(F(a)+1)^K + (K+F(a))(F(a)-1)^K}{(K-\Omega_{\phi,0}^{-1/2})(\Omega_{\phi,0}^{-1/2}+1)^K + (K+\Omega_{\phi,0}^{-1/2})(\Omega_{\phi,0}^{-1/2}-1)^K} \tag{18}$$

with

$$K \equiv \sqrt{1+\frac{4}{3}c_i^2}, \tag{19}$$

$$F(a) \equiv \sqrt{1+(\Omega_{\phi,0}^{-1}-1)a^{-3}}. \tag{20}$$

In addition, one introduces $c_i^2 = -V''(\phi_i)/V(\phi_i)$. The approximation is valid when $w \approx -1$ and the scalar field is close to the top of its potential, which is suitable for the considered case. It is then easy to show that $c^2 = -\Gamma\lambda^2$. In the slow-roll regime, $\lambda(\phi)$ is bijective and can be inverted to get $\Gamma(\lambda)$, leading to a useful relation between c^2 and λ for the potentials of Equations (15) and (16):

$$c^2 = \frac{-\lambda^2}{2} + \frac{1}{f^2}, \tag{21}$$

$$c^2 = \frac{1}{f^2}, \tag{22}$$

respectively.

In the previously considered models, the parameter $|\lambda(N)|$ was found to be decreasing (or constant) with N through Equation (8). Using the contemporary constraints on the swampland conjecture was therefore the most efficient way to go. This is not the case for thawing models, where $|d\lambda/dN| > 0$ requires the past evolution of $\lambda(N)$ to be taken

into account. To obtain Equation (17), the de Sitter solution for $\Omega_\phi(N)$ was used, namely

$$\Omega_\phi(a) = \frac{\Omega_{\phi,0}\, a^3}{\Omega_{\phi,0}\, a^3 \;+\; 1 - \Omega_{\phi,0}}. \tag{23}$$

Approximations given by Equations (17) and (23), together with the ordinary differential equation (ODE), Equation (8), give the past evolution of $-V'/V$.

4. Results

Links between the swampland ideas and astronomical observations have already been studied in Agrawal et al. (2018), Raveri et al. (2019), Akrami et al. (2019), and Arjona & Nesseris (2021). In this work, however, we do not focus on existing data or question the validity of the de Sitter conjecture but, instead, we try to probe its exclusion power from the viewpoint of future surveys, as initiated in Heisenberg et al. (2018).

4.1. Experimental Projections

We consider, on the one hand, the Vera Rubin (formerly called LSST) observatory and the Euclid satellite, which we shall refer to as Large Optical Surveys (LOSs) and, on the other hand, the SKA radio-astronomy interferometric project. The aim is to investigate the limits they shall put on $|V'|/V$ and if the dramatic improvement in sensitivity might establish that we actually live in the Swampland (under the assumption that the de Sitter conjecture is correct), therefore suggesting that the string-inspired ideas behind this concept are incorrect.

Establishing in detail the sensitivity of future surveys to a given cosmological observable is tricky. In the specific case we are interested in, it is tempting to use as much information as possible and, in particular, to take into account the observational constraints obtained on the EOS parameter as a function of the redshift, $w(z)$, so as to compare the resulting curve with the expectations calculated for a given dark energy potential. This is the strategy followed in Zlatev et al. (1999). This, however, cannot be straightforwardly extrapolated to future experiments for which such detailed investigations are not yet available and, more importantly, this anyway relies on an assumed parameterization for the evolution of the EOS parameter. For those reasons, we prefer to, conservatively, use only the information on w_0 and w_a, assuming the usual form $w(a) = w_0 + (1-a)w_a$ (Chevallier & Polarski 2001), where a is the scale factor. We have checked that the results obtained using exclusively w_0 and w_a, with the exponential potential, are extremely close to those derived in Zlatev et al. (1999): the sensitivity remains practically the same.

The main difficulty when focusing on forecasts is to evaluate correctly the theoretical uncertainties on nonlinear scales, which become particularly relevant for the next generation of instruments that will probe the details of the growth of large-scale structures with extraordinary precision, up to redshifts of order 3. Billions of galaxies will be precisely located by the LOSs. The reionization era and the cosmic dawn, up to $z = 20$, are even expected to be probed by SKA. The accurate evaluation of the constraints put on dark energy crucially depends on the way small scales are taken into account. This is a highly complex problem that involves general relativistic corrections to the structure formation mechanisms (Tansella et al. 2018), the galaxy nonlinear bias (Jennings et al. 2016),

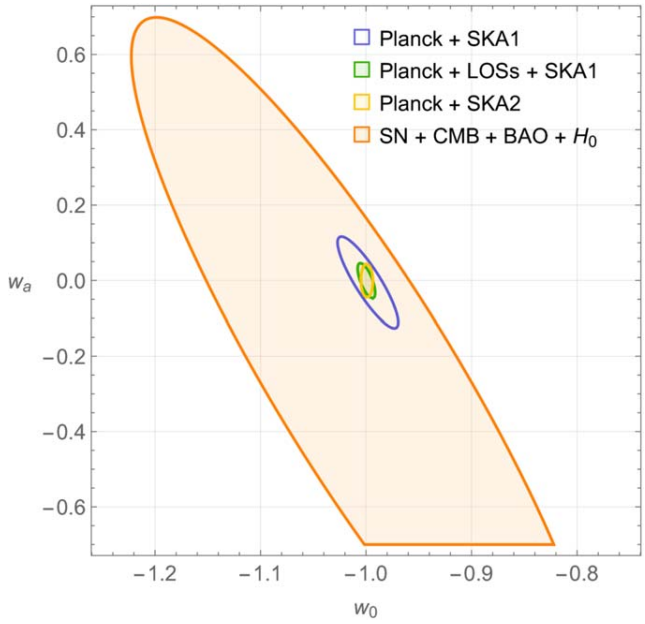


Figure 2. Comparison of the 95% CL constraints on w_0–w_a: current results (in orange) and expected improvements (in blue, green, and yellow).

the intrinsic alignment problem (Hilbert et al. 2017), the feedback of baryons (Schneider & Teyssier 2015), etc. The usual strategy has been to implement a cutoff scale below which data cannot be used. This obviously misses quite a lot of potentially relevant information, and it has been shown that nonlinear scales could be used in a controlled way (Baldauf et al. 2016). The key point lies in the correct estimation of the evolution of theoretical uncertainties at very high wavenumbers.

Instead of crude assumptions relying on either fully correlated or totally uncorrelated errors at small scales, a realistic numerical method to estimate uncertainties on the nonlinear spectrum was developed in Sprenger et al. (2019), following the ideas of Audren et al. (2013). A reasonable improvement of the modeling of nonlinear effects is obtained by taking into account the increase of numerical resources expected by the time the real data will be available. To this aim, a Bayesian Markov Chain Monte Carlo (MCMC) was used instead of the Fisher matrix formalism, subject to numerical instabilities. As massive neutrinos play an important role in the nonlinear growth of structures, in addition to w_0 and w_a, the total neutrino mass M_ν is also left as a free parameter. It is marginalized over in the (w_0, w_a) contour that we use. We basically consider three cases extracted from Sprenger et al. (2019): scenario 1 is "Planck+SKA1," scenario 2 is "Planck +LOSs+SKA1," and scenario 3 is "Planck+SKA2." The details of the SKA1 and SKA2 programs are given in Sprenger et al. (2019), Bull (2016), Harrison et al. (2016), and Bonaldi et al. (2016). Three probes are included: galaxy clustering, weak lensing (for LOSs and SKA), and HI intensity mapping (for SKA) at low redshift—probing the reionized universe. We draw in Figure 2 the comparison between current constraints on w_0 and w_a (Scolnic et al. 2018) and what is to be expected in the future. Throughout this study, we approximate the constraints by ellipses that match closely the numerical results.

Although the actual path might slightly differ in some cases (as will be detailed below), the methodology is as follows. To

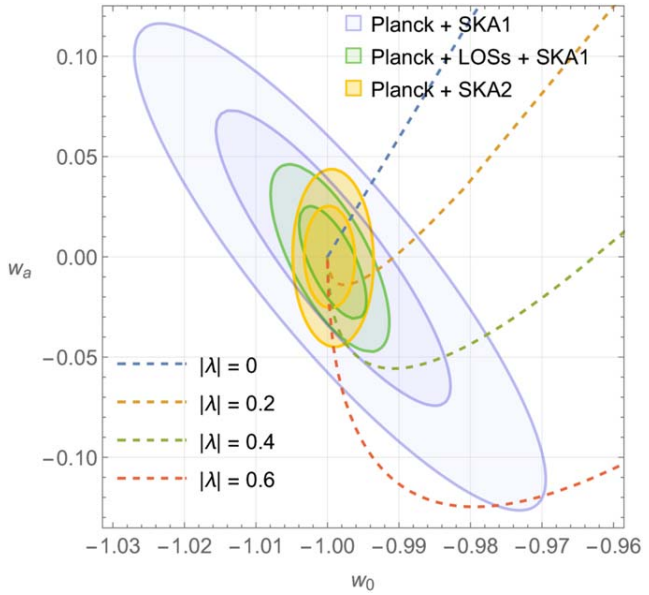


Figure 3. Relation between w_0, w_a and $|\lambda_0| = |V'(0)|/V(0)$. The ellipses represent the expected 67% and 95% CL constraints on w–w_a from various future experiments, while the dashed lines are the isolines $\lambda =$ constant as calculated with Equation (8) evaluated at $N = 0$.

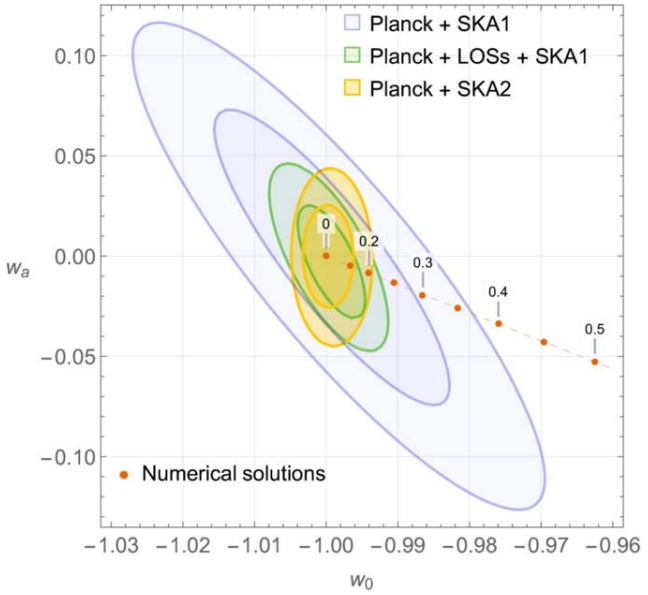


Figure 4. Comparison between the expected 67% and 95% CL constraints on w_0–w_a for various future experiments (blue, green, and yellow ellipses) together with numerical results for $|\lambda|$ for an exponential potential, as given by Equation (11) (orange dots). The parameter λ varies from 0 to 0.5.

evaluate the "exclusion power" of future surveys, we assume that the actual cosmological behavior will be de Sitter–like and investigate how this might contradict the relevant swampland conjecture. For a given potential family, we vary both the initial conditions (leading to trajectories compatible with the known features of the universe) and the values of the parameters entering the model. For each simulation, we evaluate $|\lambda| = |V'|/V$ along the trajectory and keep its most relevant (that is, smallest) value. Then, to remain conservative, we keep the highest of those $|\lambda|$ values within a given confidence level (CL) ellipse in the w_0–w_a plane for different forecasts. This sets the result given in the corresponding table. To summarize intuitively: for each parameter choice and initial conditions, we compute (w_0, w_a) and $-V'/V$, and evaluate how observational constraints on the first can constrain the second, in the most reliable and conservative way. In practice, the effect of the initial conditions—provided that the model accounts for the current observations—can hardly be noticed.

4.2. Model-independent Analysis and Single-exponential Potential

From the differential Equation (6) evaluated at $N = 0$, we can directly relate w_0, w_a, and $\lambda_0 := \lambda(0)$ as measured today. We draw in Figure 3 the isolines $\lambda =$ constant against the future constraints from Plank, Euclid, and SKA, with $\Omega_{\phi,0} \approx 0.7$. As a first observation, one can see that for higher values of λ, the isoline is getting closer to the line $w_0 = -1$ with $w_a \in \mathbb{R}^-$. In the limit $\lambda \to \infty$, one recovers the vertical axis. This seemingly constitutes a drawback for constraining $|V'|/V$ today using a w_0–w_a analysis. However, this conclusion fails to capture the dynamics of the system. In practice, models leading to $w_0 \approx -1$ and $w_a < 0$ suffer from a major conceptual problem. Because the EOS parameter is bounded from below, $w(N) > -1$ (we do not consider here exotic phantom models known to exhibit an unstable vacuum filled with negative mass particles), such a configuration is highly unstable, which raises strong coincidence and fine-tuning problems. It is, however, not possible to define a model-independent constraint on $|\lambda|$ from this straightforward analysis, and one needs to study different potentials to test the swampland conjecture. This is the aim of the next sections. Nonetheless, it is possible to understand from Figure 3 that producing a theoretical prediction for w_0 and w_a for a specific model leads to a direct constraint on the parameter λ_0.

Table 1
Expected Constraints on $|V'|/V$ for an Exponential Potential, as Given by Equation (11)

	Pl. + SKA1	Pl. + LOSs + SKA1	Pl. + SKA2
67% CL	$\|\lambda\| < 0.28$	$\|\lambda\| < 0.17$	$\|\lambda\| < 0.16$
95% CL	$\|\lambda\| < 0.36$	$\|\lambda\| < 0.22$	$\|\lambda\| < 0.20$

Note. Those constraints are also valid for generic scaling freezing models described by Equation (14).

To better understand the method used to test the de Sitter conjecture, we start with the simplest model based on a single-exponential potential given by Equation (11). As mentioned earlier, the initial conditions are set to follow the tracker solution and to get $\Omega_{\phi,0} \approx 0.7$. Numerical simulations with the parameter λ ranging from 0 to 0.5 are displayed in Figure 4. Simulations for higher values of λ were also carried out to ensure that the results follow the same trend. It is not useful to scan the parameter space for arbitrary $\lambda \in \mathbb{R}^+$, as for $\lambda \gtrsim 2$, one cannot reach $\Omega_{\phi,0} \approx 0.7$. It can indeed be shown (Tsujikawa 2013) that for $\lambda^2 > 3$, the fixed point is $w = 0$, and for $\lambda \gtrsim 2$, it is reached before Ω_ϕ approaches 0.7. Interestingly, the numerical results for the current values of w and $-dw/da$, for different values for λ, are aligned on a straight line (at least up to the point for $\lambda = 2$). The intersections between this line and the ellipses of Figure 4 representing the expected 67% and 95% CL of three different future experiments lead to constraints on $|\lambda| = |V'|/V$. The results are summarized in Table 1.

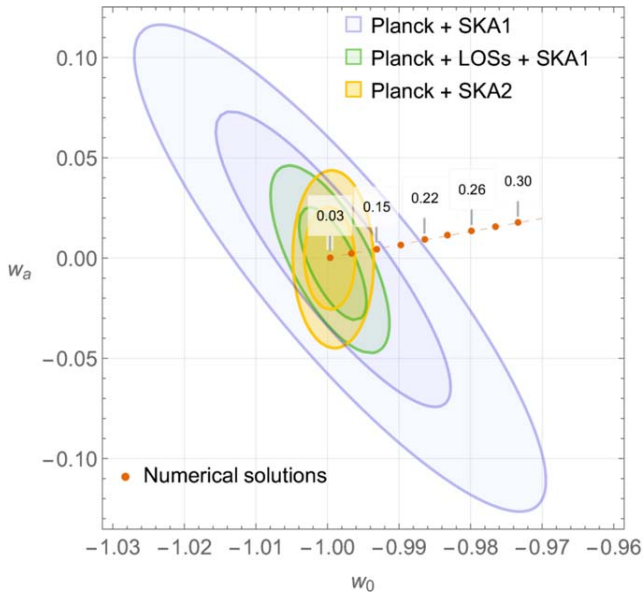


Figure 5. Comparison between the expected 67% and 95% CL constraints on w_0–w_a for various future experiments (blue, green, and yellow ellipses) together with numerical results for $|\lambda|$ for tracking freezing models with potentials given by Equation (9). The parameter α ($\Gamma = VV''/(V')^2 = 1 + 1/\alpha$) ranges from 0.001 to 0.08, and the labels of the points are the values of $|\lambda|$.

Remarkably, the higher upper bound obtained ($|\lambda| < 0.16$ at 67% CL when using all the cosmological information that will be available) is significantly better than the limit currently available ($|\lambda| < 0.6$) and much smaller than the lower bound suggested by the de Sitter conjecture ($|\lambda| > 0.8$).

Nontrivially, as we shall notice for each model, and as we have just mentioned, numerical results are "aligned." One can then calculate the constraint on $|\lambda_0|$ (which is, for the potentials considered so far, the most stringent case) by evaluating the intersection between the line and the ellipses. This is actually not obvious from the shape of the isolines $\lambda_0 =$ constant in Figure 3. One could indeed have expected that if the line of numerical results was exhibiting a strong positive slope, the intersection between the ellipses and the line would not be the relevant constraint on λ_0 but would instead be weaker. However, one can show that this is not the case, as the slope would have to be greater than 6, which is actually the slope of the isoline $\lambda_0 = 0$. Furthermore, one can notice that the lower the slope of the line of numerical results, the weaker the constraint on $|V'|/V$.

4.3. Tracking Freezing Models

From the initial conditions given by Equation (10), for the tracker solution, together with the requirements $\Omega_{\phi,0} \approx 0.7$, numerical solutions can be computed for different values of the free parameter α. The results are displayed in Figure 5. Numerical results for the λ parameter correspond to $\alpha \in [0.001, 0.08]$. Because the EOS parameter $w(N) \approx -2/(2+\alpha)$ is nearly constant and began to unfreeze quite recently, numerical simulations for $\alpha > 0.08$ are unnecessary as they fall outside the 95% CL constraint on (w_0, w_a). One can observe this behavior in Figure 1, where $\alpha = 1$ was chosen. As previously, the numerical solutions—that is, the points in the (w_0, w_a) plane associated with different λ—are aligned. The upper bounds on $|\lambda|$ are summarized in Table 2. In this case also, they are very relevant for the swampland program.

Table 2
Expected Constraints on $|V'|/V$ for Tracking Freezing Models, Based on the Potential Given by Equation (9)

	Pl. + SKA1	Pl. + LOSs + SKA1	Pl. + SKA2
67% CL	$\|\lambda\| < 0.16$	$\|\lambda\| < 0.11$	$\|\lambda\| < 0.11$
95% CL	$\|\lambda\| < 0.21$	$\|\lambda\| < 0.14$	$\|\lambda\| < 0.15$

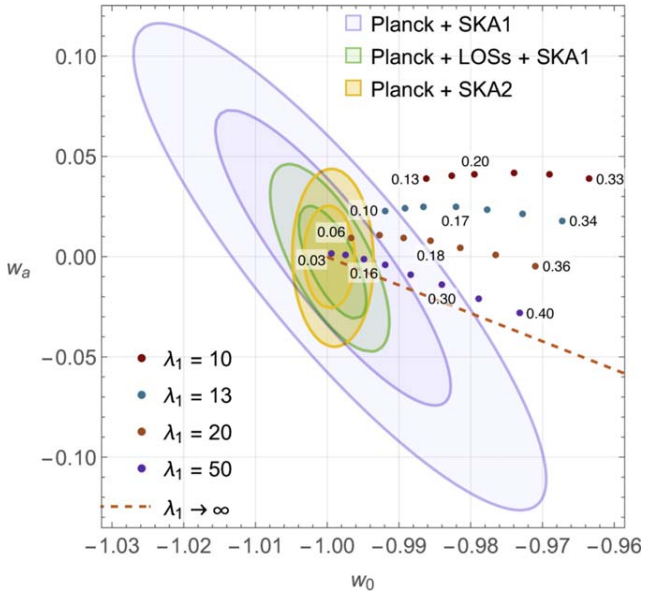


Figure 6. Comparison between the expected 67% and 95% CL constraints on w_0–w_a from various future experiments (blue, green, and yellow ellipses) together with numerical results for $|\lambda|$ for scaling freezing models with potentials given by Equation (14). The parameters (λ_1, λ_2) are, respectively, (10, [0, 0.25]), (13, [0, 0.3]), (20, [0, 0.35]), and (50, [0, 0.4]), and the labels of the points are the values of $|\lambda|$.

4.4. Scaling Freezing Models

In the case of the scaling freezing model, the potential given by Equation (14) has three independent parameters, λ_1, λ_2, and V_1/V_2, making the exhaustive scan a priori subtler. However, the value of the ratio between V_1 and V_2 has nearly no effect on the behavior of the system, hence one can safely set $V_1 = V_2$ and $\lambda_1 > \lambda_2$ without loss of generality. It was shown in Barreiro et al. (2000) and Gupta et al. (2012) that in order to have an asymptotic freezing solution with a transition from w ($N \to -\infty$) ≈ 0 to $w(N \to \infty) \approx -1 + \lambda_2^2/3$, it is necessary that $\lambda_1^2 > 3$ and $\lambda_2^2 < 3$. The problem therefore basically consists of exploring the cosmological dynamics, and the subsequent minimum of $|V'|/V$, along each track and for values of (λ_1, λ_2) satisfying the previous constraints. The results for the parameters (10, [0, 0.25]), (13, [0, 0.3]), (16, [0, 0.3]), (22, [0, 0.35]), and (50, [0, 0.4]) are displayed in Figure 6. Models with $\lambda_1 < 9.4$ are disfavored from nucleosynthesis analysis (Bean et al. 2001), and simulations for 0.25, 0.3, 0.35, $0.4 < \lambda_2 < \sqrt{3}$ (respectively, depending on the value of λ_1) are omitted for clarity as they fall outside the w_0–w_a constraints.

As expected, the higher the parameter λ_1, the closer the results to those obtained with the single-exponential potential

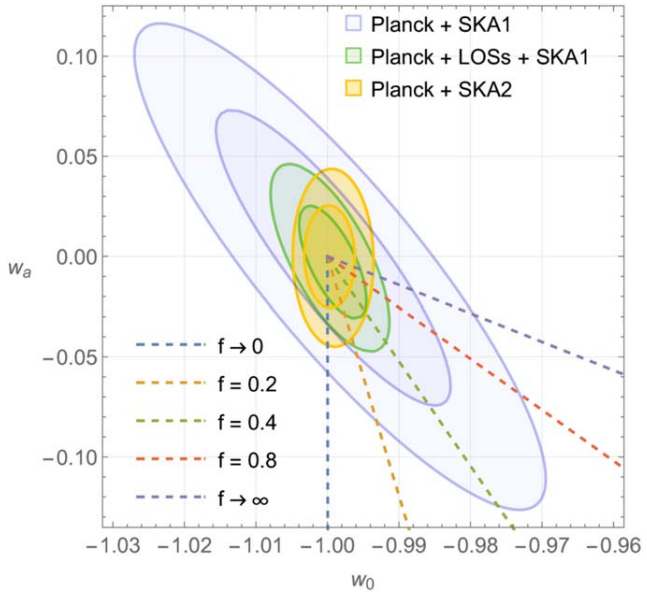


Figure 7. Comparison between the expected 67% and 95% CL constraints on w_0–w_a from various future experiments (blue, green, and yellow ellipses) and the w_0–w_a relation for the thawing model described by the potential given in Equation (16) for different values of the parameter f.

represented by the dashed line in Figure 6. Interestingly, for all finite values of λ_1, the generic trend is such that the constraints obtained are more stringent than for the single exponential—that is, for a constant V' —case. To remain conservative, it is therefore possible to keep the results obtained for a single exponential, and given in Table 1, as generic constraints for scaling freezing models described by the potential given by Equation (14).

4.5. Thawing Models

Unlike previous models, for which numerical solutions on $w(N)$ can be fully trusted, thawing models described by potentials given by Equations (15) and (16) are better understood using the analytical approximation of Equation (17). In this approximation, both potentials are equivalent up to the parameter c_i^2 and differ by a term $-\lambda_i^2/2$, with $\lambda_i = -V'(\phi_i)/V(\phi_i)$. Starting with the potential given by Equation (16), where $c^2 = c_i^2 = 1/f^2$ is constant, one can find an explicit relation between w_0 and w_a for different values of f using the approximation given by Equation (17). The result is shown in Figure 7.

The dashed lines represent the allowed values of w_0 and w_a for different values of the parameter f. In particular, this means that by varying the initial conditions of the system, any point on the line can be reached. As discussed earlier, the slope of these lines defines the constraint on the contemporary value of $\lambda = -V'/V$: the steeper the slope, the weaker the constraint. This is problematic as Figure 7 shows that choosing small enough values of f will lead to a very weak constraint on the de Sitter conjecture. However, in this case, the function $|\lambda|$ is increasing with time. This suggests focusing on the remote past so as to set relevant constraints. It might appear meaningless as the quintessence scalar field was then subdominant and the behavior of the universe was not yet de Sitter–like. If correct, the de Sitter conjecture must however hold at all *points* in the field space, as long as the effective field theory description is

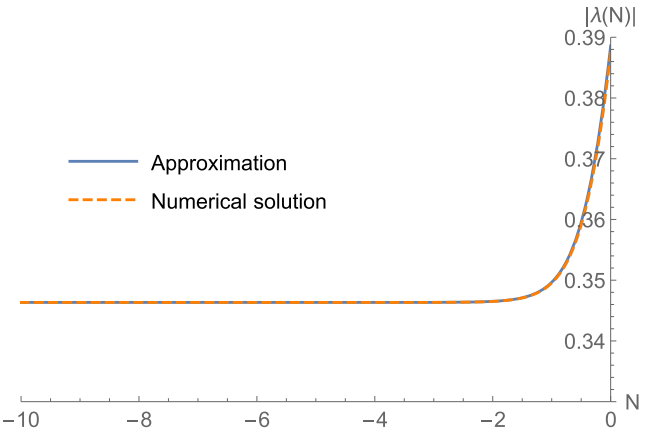


Figure 8. Comparison between the analytical approximation and the numerical solution of $\lambda(N)$ for the thawing model given by Equation (16) with $f = 2$. The value $\lambda(0)$ is found using the expected 95% CL constraint on w_0–w_a from the Planck+SKA1 experiment.

valid. Interesting constraints on the swampland conjecture can therefore be obtained using a contemporary observation of (w_0, w_a) but taking into account the past behavior of the function λ.

Let us call $\lambda_0(f)$ the current value of $-V'/V$ for a given f parameter. A bound can be derived from the intersection between the dashed lines and the ellipses in Figure 7 together with Equation (6). Using the approximations of Equations (17) and (23), as well as the ODE given by Equation (8), one can calculate the behavior of λ with respect to the time N (expressed as e-folds) for any parameter f. In Figure 8, as an example, $\lambda(N)$ is plotted with the approximation mentioned above, for $f = 1$, considering the constraint on $\lambda_0(1)$ from the 95% CL based Planck+SKA1 observations. The numerical solution is superimposed.

An interesting feature that can be observed on this figure is the stability of $-V'/V$ in the past. The parameter $\lambda(N)$ is indeed quasi-constant for $N < -2$ and one can obtain the constraint of the de Sitter parameter for $f = 2$, that is, $|V'|/V < 0.34$. It is also possible to check in the figure that the approximation of $\lambda(N)$ is strongly reliable.

This shows the path to the derivation of a constraint that is actually independent of the parameter f. We first find the current constraint on $-V'/V$, i.e. $\lambda_0(f)$, we then find the behavior of $\lambda(N)$ and fix $N = -10$ to obtain the constraint on the swampland conjecture. We have checked that considering earlier times does not improve the results. The procedure is repeated for different values of f and for the different experimental scenarios. The results are shown in Figure 9 where the limit on $|\lambda(N = -10)|$ is given with respect to f. The numerical values are given in Table 3. In this case also, they are very stringent and meaningful for the swampland program.

4.6. Summary

This establishes that whatever the (reasonable) potential considered, the future generation of experiments should be able —if the actual behavior of the universe is as driven by a cosmological constant—to put string theory under pressure. At least, the sensitivity is such that the original de Sitter conjecture will be tested in the interesting regime where the measured value is in strong conflict with the theoretical limit. We have shown that for all the classes of models considered, the SKA2 observations are expected to lead to $|V'|/V < 0.16$ at 67% CL

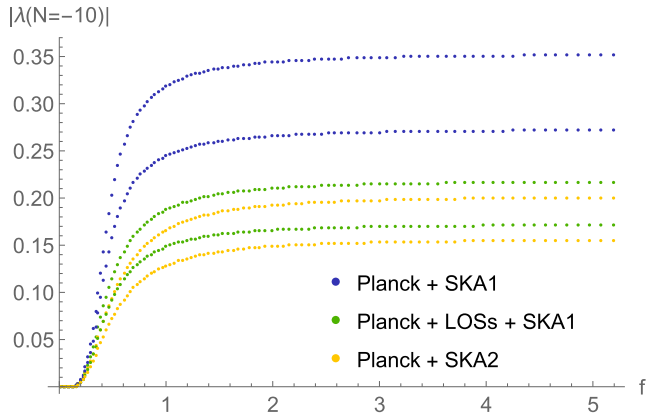


Figure 9. Constraints from various future experiments on $|\lambda(N=-10)|$ as a function of f for thawing models given by Equation (16).

Table 3
Expected Constraints on $|V'|/V$ from Different Sets of Experiments for Thawing Models with the Potential Given by Equation (16)

	Pl. + SKA1	Pl. + LOSs + SKA1	Pl. + SKA2
67% CL	$\|\lambda\| < 0.27$	$\|\lambda\| < 0.17$	$\|\lambda\| < 0.16$
95% CL	$\|\lambda\| < 0.35$	$\|\lambda\| < 0.22$	$\|\lambda\| < 0.20$

Note.
Unlike previous models the limit is obtained in the past. To remain conservative, the limit f$\rightarrow \infty$ was considered, which corresponds to the less stringent case.

and to $|V'|/V < 0.20$ at 95% CL, whereas the de Sitter conjecture requires $|V'|/V \gtrsim 1$.

We have also checked that varying the value of $\Omega_{\phi,0}$ within the observational uncertainties does not change the constraints at the level of accuracy of this work. To summarize, all the considered scenarios for future experiments will contradict the original de Sitter conjecture at a quite high confidence level (unless, of course, the actual dynamics of the universe is revealed not to be driven by a true cosmological constant).

If we consider the refined de Sitter conjecture, results are unchanged for tracking freezing and scaling freezing models. They always fail to satisfy the new condition. However, in the case of thawing models, the system is satisfied and no constraint can be put if both conditions are taken into account: when one condition is violated, the other is satisfied. This is an important issue that should be addressed in the future. However, particle physics arguments are expected to allow, in the future, to constrain the f parameter of the potential, therefore breaking the degeneracy (Marsh 2016).

5. Prospects

Although this work is devoted to the actual estimate of the Vera Rubin, Euclid, and SKA capabilities to improve constraints on the de Sitter conjecture, it is also worth trying to go beyond those experiments. We therefore provide estimates of the upper bound on $|\lambda|$ that could be derived from hypothetical even larger observatories to be possibly constructed in the long run. This also allows one to get an accurate limit on $|\lambda|$ for contours in the w–w_a plane that might differ from the simulations used in this study. We, however, still assume an elliptic approximation for the confidence level isolines and denote, respectively, σ_{w_a} and σ_{w_0} the uncertainties

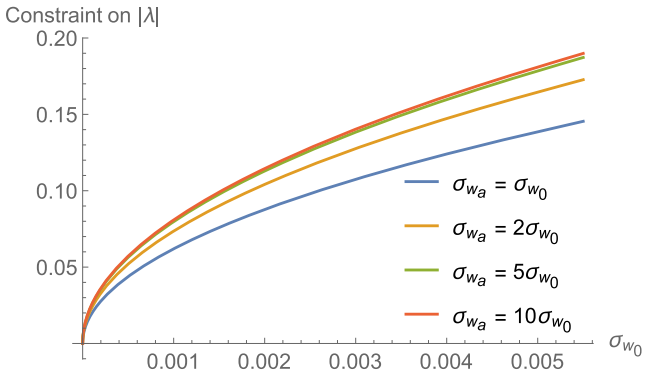


Figure 10. Evolution of the constraint on the de Sitter conjecture with respect to the standard deviation on w_0. The standard deviation on w_a is set such that the shape of the ellipse is unchanged.

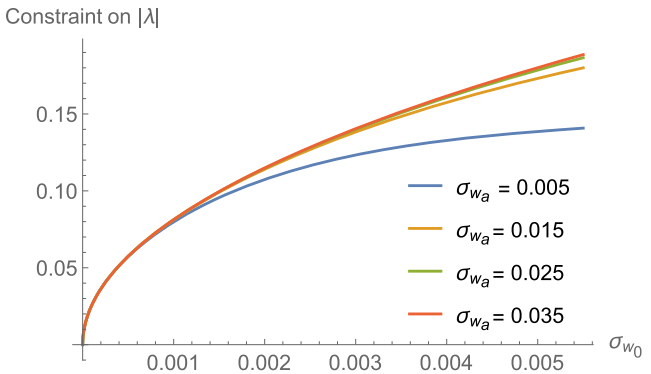


Figure 11. Evolution of the constraint on the de Sitter conjecture with respect to the standard deviation on w_0 for different values of the standard deviation on w_a.

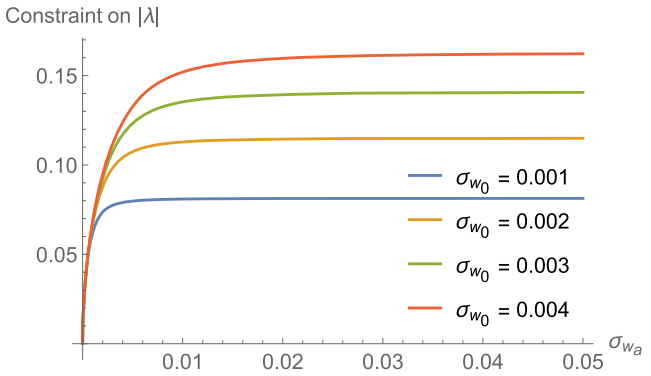


Figure 12. Evolution of the constraint on the de Sitter conjecture with respect to the standard deviation on w_a for different values of the standard deviation on w_0.

on the semimajor and semiminor axes, which are considered aligned with the w_0–w_a axes, similarly to the Planck+SKA2 constraint. Obviously, other hypotheses could be made but this allows the main features to be captured.

In Figure 10, we display the 67% CL limit on $|\lambda|$, keeping the shape of the ellipse, i.e., the ratio between σ_{w_a} and σ_{w_0}, unchanged. For comparison, this ratio ranges between 7.2 and 9.3 for the various expected constraints from the Planck +SKA1, Planck+LOSs+SKA1, and Planck+SKA2 simulations. In Figure 11, we display the limit as a function of σ_{w_0} for different values of σ_{w_a}. The other way round, in Figure 12, we

plot the limit as a function of σ_{w_a} for different values of σ_{w_0}. We always consider the less constraining potential.

This shows that there is still room for improving the limits beyond the next generation of experiments. This also underlines that ameliorating the sensitivity on w_0 or w_a changes the situation differently. Although reducing the error on w_0 has a monotonic and quite regular effect on the improvement on the limit, the σ_{w_a} behavior exhibits a "plateau" as soon as $\sigma_{w_a} > 0.01$.

6. Remarks and Conclusion

6.1. Is the de Sitter Conjecture Reliable?

The de Sitter conjecture is just a conjecture. How reliable is it (see the introduction of Raveri et al. 2019)? It is known that the landscape of string theory does contain Minkowski solutions. The richness of the structure is nearly infinite. Each geometry can support more than 10^{10^5} flux vacua (Taylor & Wang 2015) and the number of compactification geometries is higher than 10^{1000} (Taylor & Wang 2018). However, constructing metastable de Sitter solutions from this huge landscape still appears to be highly problematic, and this is the main underlying motivation for the conjecture. It is indeed known that the de Sitter space is excluded as a solution of fundamental supergravity theories (Maldacena & Nunez 2001). It is also neither a solution of type I/heterotic supergravity (Green et al. 2012; Gautason et al. 2012), nor of heterotic world-sheet conformal field theory (Kutasov et al. 2015). Basically, it seems quite well established that de Sitter solutions cannot be found in regions of parametric control in string theory (Dine & Seiberg 1985a, 1985b). Interesting attempts to evade this conclusion do exist in type IIB string theory (Giddings et al. 2002; Andriot et al. 2020b), together as in type IIA and M-theory. None of them is, however, conclusive either because of the lack of control on the quantum corrections to the effective spacetime action (they ignore the fact that the background is non-static; Sethi 2018) or because of the absence of fully explicit constructions.

Although anti–de Sitter vacua are well understood in string theory, it is a fact that de Sitter space is surprisingly hard to control. Many no-go theorems (see, e.g., Dass 2002; Russo & Townsend 2019; Shukla 2020; Basile & Lanza 2020) were derived and quite a lot of concrete examples support the possibility that string theory just cannot admit any de Sitter vacuum. Still, counterarguments are being built around the Kachru–Kallosh–Linde–Trivedi (KKLT) proposal (Kachru et al. 2003; Kallosh et al. 2019) using a Kähler moduli stabilization or with large volume scenarios (see Cicoli et al. 2008 and references therein). The full picture is still unclear.

The de Sitter swampland program is an active and highly controversial field of research (Andriot 2019). It might very well be that there is enough complexity in the space of string theory vacua and sufficient richness in the unexplored sectors of the model for a landscape of de Sitter solutions to indeed exist. The extraordinary difficulty in exhibiting any convincing de Sitter solution, in spite of the huge number of explored solutions, however, suggests that the de Sitter space lies in the swampland, and that the conjecture holds. Recently, interesting links with trans-Planckian censorship were even built (Bedroya & Vafa 2020). This is not a theorem but, in our opinion, a reasonable guess. The conjecture constitutes, at least, an outstanding way to possibly put string theory under pressure, something that has proven to be extremely difficult in the last decades.

6.2. Other Approaches to Quantum Gravity

In most studies devoted to the swampland (see, e.g., Palti 2019), including this article, the words "string theory" and "quantum gravity" are used as if they were synonymous or, at least, as if quantum gravity was to be understood as a sector of string theory only. This is obviously exaggerated. There are many other roads toward quantum gravity (see Oriti 2009 for an overview): loop quantum gravity (see, e.g., Rovelli 2011), noncommutative geometry (see, e.g., Chamseddine & Connes 1997), group field theory (see, e.g., Baratin & Oriti 2012), causal sets (see, e.g., Bombelli et al. 1987; Sorkin 2009), asymptotic safety (see, e.g., Weinberg 2009; Gubitosi et al. 2019), causal dynamical triangulation (see, e.g., Loll & Quant 2020), etc. Those models are unquestionably speculative—the semiclassical limit is often not known or not clear—but all have preliminary predictions for cosmology (Barrau 2017). Most of them have no problem with de Sitter spaces. The existence of a small and positive cosmological constant was even predicted, before it was observed, by Sorkin within the causal sets framework (see references in Dowker & Zalel 2017). It could also be that gravity does not need to be quantized (Tilloy 2019).

It should therefore be made clear that the de Sitter conjecture is not about any theory of quantum gravity but only about string theory. This is why we believe that it makes sense, as we did here, to evaluate whether this can be useful to potentially falsify string theory in the future. It is however much more hazardous, as sometimes advertised, to use swampland conjectures to rule out some low-energy models. The fact that they lie in the swampland of string theory does not mean they are intrinsically wrong[6]: string theory is far from being established. Although intensively debated, the claimed nonempirical corroboration of string theory is not sufficient to make it a universal paradigm (Dawid 2017; Chall 2018; Cabrera 2018). It seems to us much more fruitful to use the swampland ideas as a first—and very welcome—step toward a possible proof of the effective falsifiability of string theory. This is, however, not even obvious: from the point of view of string theory, it is not clear that one should expect dark energy to be described by a scalar field.

6.3. Conclusion and Future Developments

In this work, we have considered three classes of potentials as benchmarks for quintessence models. We have shown that the SKA network of radio antennas, the Vera Rubin ground-based telescope, and the Euclid satellite could be able to derive a limit on the first de Sitter conjecture, $|V'|/V < 0.16$, that contradicts the most reliable theoretical estimates (Andriot et al. 2020a). This shows that if the swampland conjecture holds, string theory might be on the road of "falsification" in the next decade. This conclusion would require the inclusion of multifield scenarios (Achúcarro & Palma 2019; Bravo et al. 2020; Lin & Solomon 2021).

Our main result is nearly model independent in the sense that the upper bound mentioned above is the less stringent one

[6] Some swampland arguments are very generic and based on solid quantum field theory conclusions about the UV completion. They have to be distinguished from string-inspired guesses.

among the three classes of potential considered here. Scaling tracking, scaling freezing, and thawing models are the main ideas currently on the table for quintessence. We have therefore scanned the panel of intensively discussed models. It is however not impossible that other potentials, in agreement with the cosmological dynamics but less constrained from the viewpoint of the de Sitter conjecture, might be found. Our result is therefore not a theorem but a reasonable conclusion based on consistent potentials. It should however be mentioned that, in principle, the actual potential could be reconstructed by a combined cosmological analysis (Boisseau et al. 2000). This opens the possibility of deriving a precise measurement of $-V'/V$ and V''/V.

We have also investigated possible long run constraints as a function of the shape and size of the ellipse of uncertainty, beyond the main surveys considered in most of this work.

Showing that the real world does not fulfill the de Sitter conjecture would unquestionably not be enough to discard string theory. But, among other indications, it might play a role in a possible paradigm shift. Low-scale supersymmetry as a fully natural solution to the hierarchy problem was not abandoned by part of the community just after one unsuccessful run at the Large Hadron Collider, but after many arguments, the edifice was weakened. Every sign counts.

Obviously, it could also be that string theory is correct, that the de Sitter conjecture holds and that our universe does not conflict with it but follows a not purely de Sitter dynamics. We leave for a future study the investigation of the detection capability (i.e., a measurement $|V'|/V \sim \lambda_D$ with λ_D not too small) of future surveys at the level required by the conjecture. Is it still possible to measure a "nearly but not exactly" de Sitter behavior marginally compatible with the Swampland criteria? The interval of possible values λ_D, not conflicting with any known data, is quite narrow. It would also be important to take into account the Hubble constant tension in this framework (Banerjee et al. 2020).

Finally, on the purely theoretical side, not only should the very validity of the conjecture be better understood but the actual value λ_c would have to be better evaluated. The latest estimates are encouraging but one might also argue that our limit of 0.16 is not that far from the expected value of order 1. Precise numbers from the string side are now needed beyond orders of magnitude.

We thank David Andriot for enlightening comments on the de Sitter conjecture. We also thank the anonymous referee, who helped a lot in improving the article.

Appendix
Intuitive remarks

In this section, we summarize some qualitative arguments, to guide the unfamiliar reader, on the reasons why de Sitter spaces are so problematic in string theory. There are no straightforward and fully intuitive explanations. Rather, there are many "indications" that build up together. The following list (borrowing from Akrami et al. 2019; Raveri et al. 2019; Heisenberg et al. 2021) does not pretend to be exhaustive, and the arguments do not try to be rigorous.

1. As we have explained before, supersymmetry and de Sitter spaces are not easy to reconcile. Exact supersymmetry is incompatible with the de Sitter symmetries. Finding solutions with spontaneously broken supersymmetry is the natural way to go. In principle, it is possible to start with a theory that explicitly breaks supersymmetry but then one generically encounters stability and divergence issues.
2. In the string framework, the AdS/CFT correspondence plays a central role in paths toward quantum gravity as it states that a strongly coupled n-dimensional gauge theory is equivalent to a gravitational theory in an $(n+1)$-dimensional anti–de Sitter spacetime. It, however, cannot be (at least not in a known and controlled way) extended to de Sitter spacetimes.
3. Stability is not ensured. If a de Sitter solution is defined as a positive extremum in the field space, it is not clear that it can correspond to a minimum for all the fields. This question has been addressed statistically from the viewpoint of "random supergravity potentials" and "random matrices" and is not yet fully clear.
4. Some generic arguments suggest that scalar potentials tend to 0 when the string coupling constant (or the volume of internal dimensions) goes to infinity. But if one constructs a de Sitter vacuum—that is, with a positive potential—while the potential vanishes at infinity, it can only be metastable. The relevant question becomes that of the lifetime of this vacuum, but it is expected that the nonperturbative quantum effect will destabilize it.
5. It might be than one of the most "physical" reasons behind the unsuitability of de Sitter spaces in string theory is rooted in the trans-Planckian censorship conjecture (previously mentioned in the text). Some explicit calculations show that it seems strongly related with the de Sitter conjecture.
6. Finally, whatever the detailed framework within string/M-theory, it becomes clear that the question of constructing de Sitter solutions is extremely constrained. Relying on either the KKLT scenario or on more classical approaches, constraints are accumulating, making constructions incredibly difficult—if not totally impossible.

ORCID iDs

Aurlien Barrau https://orcid.org/0000-0003-1670-2966
Cyril Renevey https://orcid.org/0000-0002-0474-8222

References

Abell, P. A., Allison, J., Anderson, S. F., et al. 2009, arXiv:0912.0201
Achúcarro, A., & Palma, G. A. 2019, JCAP, 2019, 041
Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, A&A, 641, A6
Agrawal, P., Obied, G., Steinhardt, P. J., & Vafa, C. 2018, PhLB, 784, 271
Akrami, Y., Kallosh, R., Linde, A., & Vardanyan, V. 2019, ForPh, 67, 1800075
Albrecht, A., & Skordis, C. 2000, PhRvL, 84, 2076
Alonso, D., Awan, H., Biswas, R., et al. 2018, arXiv:1809.01669
Amendola, L., Appleby, S., Bacon, D., et al. 2013, LRR, 16, 6
Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, LRR, 21, 2
Andriot, D. 2018, PhL B, 785, 570
Andriot, D. 2019, ForPh, 67, 1800103
Andriot, D., Cribiori, N., & Erkinger, D. 2020a, JHEP, 2020, 162
Andriot, D., Marconnet, P., & Wrase, T. 2020b, JHEP, 2020, 076
Andriot, D., & Roupec, C. 2019, ForPh, 67, 1800105
Arjona, R., & Nesseris, S. 2021, PhRvD, 103, 063537
Arkani-Hamed, N., Georgi, H., & Schwartz, M. D. 2003, AnPhy, 305, 96
Arkani-Hamed, N., Motl, L., Nicolis, A., & Vafa, C. 2007, JHEP, 2007, 060

Audren, B., Lesgourgues, J., Bird, S., Haehnelt, M. G., & Viel, M. 2013, JCAP, 2013, 026
Baldauf, T., Mirbabayi, M., Simonovi, M., & Zaldarriaga, M. 2016, arXiv:1602.00674
Banerjee, A., Cai, H., Heisenberg, L., Colgáin, E. O., Sheikh-Jabbari, M., & Yang, T. 2020, arXiv:2006.00244
Banks, T., Dine, M., & Gorbatov, E. 2004, JHEP, 2004, 058
Banks, T., & Dixon, L. J. 1988, NuPhB, 307, 93
Baratin, A., & Oriti, D. 2012, PhRvD, 85, 044003
Barrau, A. 2007, CERN Cour., 47, 13
Barrau, A. 2017, CRPhy, 18, 189
Barreiro, T., Copeland, E. J., & Nunes, N. 2000, PhRvD, 61, 127301
Basile, I., & Lanza, S. 2020, JHEP, 2020, 108
Bean, R., Hansen, S. H., & Melchiorri, A. 2001, PhRvD, 64, 103508
Becker, K., Becker, M., & Schwarz, J. H. 2006, String Theory and M-theory: A Modern Introduction (Cambridge: Cambridge Univ. Press)
Bedroya, A., & Vafa, C. 2020, JHEP, 2020, 123
Ben-Dayan, I. 2019, PhRvD, 99, 101301
Bianchi, E., & Rovelli, C. 2010, arXiv:1002.3966
Blake, C., Kazin, E. A., Beutler, F., et al. 2011, MNRAS, 418, 1707
Blumenhagen, R., Lüst, D., & Theisen, S. 2013, Basic Concepts of String Theory, Theoretical and Mathematical Physics (Heidelberg: Springer)
Boisseau, B., Esposito-Farese, G., Polarski, D., & Starobinsky, A. A. 2000, PhRvL, 85, 2236
Bombelli, L., Lee, J., Meyer, D., & Sorkin, R. 1987, PhRvL, 59, 521
Bonaldi, A., Harrison, I., Camera, S., & Brown, M. L. 2016, MNRAS, 463, 3686
Brandenberger, R. H. 2011, IJMPS, 1, 67
Bravo, R., Palma, G. A., & Riquelme, S. 2020, JCAP, 2020, 004
Brax, P. 2018, RPPh, 81, 016902
Brennan, T. D., Carta, F., & Vafa, C. 2017, arXiv:1711.00864
Bull, P. 2016, ApJ, 817, 26
Cabrera, F. 2018, Synthese, doi:10.1007/s11229-018-01987-9
Carr, B. (ed.) 2007, Universe or Multiverse? (Cambridge: Cambridge Univ. Press)
Carroll, S. M. 2018, arXiv:1801.05016
Casadio, R., & Harms, B. 2000, NuPhS, 80, 1207
Chall, C. 2018, SHPMP, 63, 128
Chamseddine, A. H., & Connes, A. 1997, CMaPh, 186, 731
Cheung, C., & Remmen, G. N. 2014, PhRvL, 113, 051601
Chevallier, M., & Polarski, D. 2001, IJMPD, 10, 213
Chiba, T., De Felice, A., & Tsujikawa, S. 2013, PhRvD, 87, 083505
Choi, K., Chway, D., & Shin, C. S. 2018, JHEP, 2018, 142
Chongchitnan, S., & Efstathiou, G. 2007, PhRvD, 76, 043508
Cicoli, M., Conlon, J. P., & Quevedo, F. 2008, JHEP, 2008, 105
Copeland, E. J., Liddle, A. R., & Wands, D. 1998, PhRvD, 57, 4686
Creminelli, P., Nicolis, A., & Trincherini, E. 2010, JCAP, 2010, 021
Danielsson, U. 2001, RPPh, 64, 51
Danielsson, U. H., & Riet, T. Van 2018, IJMPD, 27, 1830007
Dass, N. D. Hari 2002, MPLA, 17, 1001
Dawid, R. 2017, arXiv:1702.01133
Denef, F., Hebecker, A., & Wrase, T. 2018, PhRvD, 98, 086004
Dine, M., & Seiberg, N. 1985a, PhRvL, 55, 366
Dine, M., & Seiberg, N. 1985b, PhLB, 162, 299
Dowker, F., & Zalel, S. 2017, CRPhy, 18, 246
Duary, T., & Banerjee, A. D. N. 2019, EPJC, 79, 888
Durrer, R., & Hasenkamp, J. 2011, PhRvD, 84, 064027
Durrer, R., & Laukenmann, J. 1996, CQGra, 13, 1069
Dutta, S., & Scherrer, R. J. 2008, PhRvD, 78, 123525
Dvali, G., Gomez, C., & Zell, S. 2019, ForPh, 67, 1800094
Dvali, G. R., Gabadadze, G., Kolanovic, M., & Nitti, F. 2002, PhRvD, 65, 024031
Ferreira, P. G., & Joyce, M. 1998, PhRvD, 58, 023503
Freivogel, B. 2011, CQGra, 28, 204007
Frieman, J. A., Hill, C. T., Stebbins, A., & Waga, I. 1995, PhRvL, 75, 2077
Garg, S. K., & Krishnan, C. 2019, JHEP, 2019, 075
Garg, S. K., Krishnan, C., & Zaid Zaz, M. 2019, JHEP, 2019, 029
Garriga, J., Schwartz-Perlov, D., Vilenkin, A., & Winitzki, S. 2006, JCAP, 2006, 017
Gautason, F. F., Junghans, D., & Zagermann, M. 2012, JHEP, 2012, 029
Geng, H. 2020, PhLB, 803, 135327
Giddings, S. B., Kachru, S., & Polchinski, J. 2002, PhRvD, 66, 106006
Gliozzi, F., Scherk, J., & Olive, D. I. 1977, NuPhB, 122, 253
Green, M. B., & Schwarz, J. H. 1984, PhL, 149B, 117
Green, S. R., Martinec, E. J., Quigley, C., & Sethi, S. 2012, CQGra, 29, 075006
Grimm, T. W., Palti, E., & Valenzuela, I. 2018, JHEP, 2018, 143
Gubitosi, G., Ripken, C., & Saueressig, F. 2019, FoPh, 49, 972
Gubser, S. S. 2010, The Little Book of String Theory (Princeton, NJ: Princeton Univ. Press)
Gupta, G., Majumdar, S., & Sen, Mon, A. A. 2012, MNRAS, 420, 1309
Hall, L. J., & Nomura, Y. 2008, PhRvD, 78, 035001
Harlow, D. 2016, JHEP, 2016, 122
Harrison, I., Camera, S., Zuntz, J., & Brown, M. L. 2016, MNRAS, 463, 3674
Heidenreich, B., Reece, M., & Rudelius, T. 2018, EPJC, 78, 337
Heisenberg, L., Bartelmann, M., Brandenberger, R., & Refregier, A. 2018, PhRvD, 98, 123502
Heisenberg, L., Bartelmann, M., Brandenberger, R., & Refregier, A. 2021, PhLB, 812, 135990
Hewett, J. L., Lillie, B., & Rizzo, T. G. 2005, PhRvL, 95, 261603
Hilbert, S., Xu, D., Schneider, P., Springel, V., Vogelsberger, M., & Hernquist, L. 2017, MNRAS, 468, 790
Hollands, S., & Wald, R. M. 2002, GReGr, 34, 2043
Jennings, E., Wechsler, R. H., Skillman, S. W., & Warren, M. S. 2016, MNRAS, 457, 1076
Kachru, S., Kallosh, R., Linde, A. D., & Trivedi, S. P. 2003, PhRvD, 68, 046005
Kallosh, R., Linde, A., McDonough, E., & Scalisi, M. 2019, JHEP, 2019, 134
Kallosh, R., & Linde, A. D. 2007, JCAP, 2007, 017
Kane, G. L. 1998, NuPhS, 62, 144
Klaewer, D., Lüst, D., & Palti, E. 2019, ForPh, 67, 1800102
Klaewer, D., & Palti, E. 2017, JHEP, 2017, 088
Kutasov, D., Maxfield, T., Melnikov, I., & Sethi, S. 2015, PhRvL, 115, 071305
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv:1110.3193
Lidsey, J. E., & Seery, D. 2007, PhRvD, 75, 043505
Lilley, M., & Peter, P. 2015, CRPhy, 16, 1038
Lin, W.-C., & Solomon, R. M. 2021, PhRvD, 103, 063533
Loll, R., & Quant, Class. 2020, CQGra, 37, 013002
Lovelace, C., & Lett, Phys. 1971, PhL, 34B, 500
Maartens, R., Abdalla, F. B., Jarvis, M., & Santos, M. G. 2015, in Proceedings of Advancing Astrophysics with the Square Kilometre Array (AASKA14) (Trieste: SISSA), 16
Maldacena, J. M., & Nunez, C. 2001, IJMPA, 16, 822
Marsh, D. J. E. 2016, PhR, 643, 1
Martin, J. 2008, MPLA, 23, 1252
Mortonson, M. J., Weinberg, D. H., & White, M. 2013, arXiv:1401.0046
Murayama, H., Yamazaki, M., & Yanagida, T. T. 2018, JHEP, 2018, 032
Obied, G., Ooguri, H., Spodyneiko, L., & Vafa, C. 2018, arXiv:1806.08362
Ooguri, H., Palti, E., Shiu, G., & Vafa, C. 2019, PhLB, 788, 180
Ooguri, H., & Vafa, C. 2007, NuPhB, 766, 21
Ooguri, H., & Vafa, C. 2017, AdTMP, 21, 1787
Oriti, D. 2009, Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter (Cambridge: Cambridge Univ. Press)
Palti, E. 2019, ForPh, 67, 1900037
Peebles, P. J. E., & Ratra, B. 1988, ApJL, 325, L17
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Polchinski, J. 2004, IJMPA, A19S1, 145
Polchinski, J. 2006, in The Quantum Structure of Space and Time: Proc. 23rd Solvay Conf. on Physics, 216
Polchinski, J. 2007, String Theory. Volume 1: An Introduction to the Bosonic String (Cambridge: Cambridge Univ. Press)
Popławski, N. J. 2010, PhLB, 694, 181, [Erratum: 2011, PhLB, 701, 672]
Ratra, B., & Peebles, P. J. E. 1988, PhRvD, 37, 3406
Raveri, M., Hu, W., & Sethi, S. 2019, PhRvD, 99, 083518
Rickles, D. 2014, A Brief History of String Theory, The Frontiers Collection (Heidelberg: Springer)
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Rovelli, C. 2011, in 3rd Quantum Gravity and Quantum Geometry School (QGQGS 2011) (Trieste: SISSA), 003
Rovelli, C. 2013, FoPh, 43, 8
Russo, J. G., & Townsend, P. K. 2019, JHEP, 2019, 097
Sahni, V., & Wang, L.-M. 2000, PhRvD, 62, 103517
Santos, M. G., Bull, P., Alonso, D., et al. 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14) (Trieste: SISSA), 019
Scherrer, R. J., & Sen, A. A. 2008, PhRvD, 77, 083515
Schneider, A., & Teyssier, R. 2015, JCAP, 2015, 049
Schwarz, J. H. 1996, in Astronomy, Cosmoparticle Physics: Proc. 2nd Int. Conf. COSMION'96, 562
Scolnic, D., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Sen, S., Sen, A. A., & Sami, M. 2010, PhLB, 686, 1
Senatore, L. 2017, in Proc. Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015), 447

Seo, M.-S. 2019, PhLB, 797, 134904
Sethi, S. 2018, JHEP, 2018, 022
Shukla, P. 2020, PhRvD, 102, 026014
Smolin, L. 2006, The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next (Boston, MA: Houghton Mifflin)
Sorkin, R. D. 2009, JPhCS, 174, 012018
Sprenger, T., Archidiacono, M., Brinckmann, T., Clesse, S., & Lesgourgues, J. 2019, JCAP, 2019, 047
Steinhardt, P. J., Wang, L.-M., & Zlatev, I. 1999, PhRvD, 59, 123504
Stoeger, W. R., Ellis, G. F. R., & Kirchner, U. 2004, arXiv:astro-ph/0407329
Susskind, L. 2003, arXiv:hep-th/0302219
Tansella, V., Bonvin, C., Durrer, R., Ghosh, B., & Sellentin, E. 2018, JCAP, 2018, 019
Taylor, W., & Wang, Y.-N. 2015, JHEP, 2015, 164
Taylor, W., & Wang, Y.-N. 2018, JHEP, 2018, 111
Tilloy, A. 2019, JPhCS, 1275, 012006
Tong, D. 2009, arXiv:0908.0333
Tsujikawa, S. 2013, CQGra, 30, 214003
Vafa, C. 2005, arXiv:hep-th/0509212
Veneziano, G. 2002, JHEP, 2002, 051
Veneziano, G. 2003, in Proc., 21st Texas Symp. on Relativistic Astrophysics (Texas in Tuscany), 1
Vilenkin, A. 2007, JPhA, 40, 6777
Weinberg, S. 2009, in 6th International Workshop on Chiral Dynamics (CD09) (Trieste: SISSA), 001
Wilson-Ewing, E. 2013, JCAP, 2013, 026
Witten, E. 2000, in Proc. 6th Int. Workshop on Chiral Dynamics, PoS(CD09) (Trieste: SISSA), 001
Yahya, S., Bull, P., Santos, M. G., et al. 2015, MNRAS, 450, 2251
Zlatev, I., Wang, L.-M., & Steinhardt, P. J. 1999, PhRvL, 82, 896

2.3 Extension to Horndeski theories

At first, it is interesting to restrict ourselves to theories with minimal coupling between the scalar and tensor sectors such as in Quintessence or more generally in Brans-Dicke theories [91]. But one should eventually consider all possible ways to include a scalar field in order to be exhaustive. In the context of the EFT of dark energy, non-minimal couplings allow for non-linear terms in the perturbations of the time-time component of the metric [92]. To be more precise, in the EFT approach the time-translation symmetry is broken and hence it is possible to add more terms in the Einstein-Hilbert action that are invariant under the remaining spatial diffeomorphism, such as g^{00}. Since we are interested only in small contributions, we proceed with a perturbative expansion in δg^{00} around the FLRW background. At zeroth and linear order, one can recover the action of quintessence and Brans-Dicke models, while at quadratic order one finds non-minimal coupling terms when reconstructing scalar-tensor theories from the EFT of dark energy [92].

Therefore, if we are interested in understanding the whole range of possible time symmetry breaking, we need to study the most general scalar-tensor theory that is invariant under diffeomorphism and Lorentz transformation. It might feel somewhat confusing that we can describe a gravity theory with broken time symmetry using a diffeomorphism invariant scalar-tensor theory. But the invariance under general coordinate transformations can be restored using the Stueckelberg trick [84, 93] at the cost of an extra scalar degree of freedom. Both approaches can be seen as complementary. Beyond diffeomorphism and Lorentz invariance, we limit ourselves to theories with second order equations of motion in order to avoid Ostrogradsky instabilities [94]. The resulting general scalar-tensor theory is called *Horndeski theory* and it is described by the action [95]

$$\begin{aligned} S_H[\phi, g] =& \frac{M_p^2}{2}\int \mathrm{d}^4x\sqrt{-g}\Big\{G_2(\phi,X) - G_3(\phi,X)\Box\phi \\ &+ G_4(\phi,X)R + G_{4X}(\phi,X)[(\Box\phi)^2 - (\nabla_\mu\nabla_\nu\phi)^2] \\ &+ G_5(\phi,X)G_{\mu\nu}\nabla^\mu\nabla^\nu\phi - \frac{G_{5X}(\phi,X)}{6}[(\Box\phi)^3 - 3\Box\phi(\nabla_\mu\nabla_\nu\phi)^2 + 2(\nabla_\mu\nabla_\nu\phi)^3]\Big\} \\ &+ S_m[g]\,, \end{aligned} \tag{2.11}$$

where $G_i(\phi, X)$, with $i = 2, ..., 5$ are arbitrary functions of the scalar field ϕ and the kinetic term $X = -\partial_\mu\phi\partial^\mu\phi/2$ and the notation G_{iX} stands for the derivative with respect to X.

Such a flexible theory allows for a wide range of possible behaviour for dark energy, but a large subset of Horndeski theories has already been tightly restricted not to say excluded by experiments. The recent detection of the binary neutron star merger through gravitational waves and electromagnetic signal set a very solid constraint on the speed of gravitational waves, $1 - c_g/c < 10^{-15}$ [96, 97]. This measurement has the consequence that $G_{4X} = G_5 = 0$, hence strongly reducing the set of possible scalar-tensor theories [98]. Moreover, the numerous pulsars and solar system experiments restrict a lot the dynamics of the scalar field [99, 100]. This means that the remaining modified gravity theories that are still studied seriously exhibit what is called a *screening mechanism*. It is a non-linear feature of a theory that hides the scalar field at short range, strongly reducing its observability in astrophysical experiments. There are different types of screening mechanisms that have been

discovered so far, large field value screening such as the chameleon mechanism and kinetic screening such as the Vainshtein mechanism. In the following subsection, we introduce these mechanisms using the following resources [7, 101–105].

2.3.1 Large field value screening

A screening mechanism can already emerge from a scalar-tensor theory of the Brans-Dicke type, with a suitable scalar potential $V(\phi)$ and coupling $A(\phi)$ to the matter field ψ_m. The general action of the Brans-Dicke theory can be written under the form

$$S_{\rm BD}\big[\phi, g, \psi_m\big] = \int \mathrm{d}^4x\sqrt{-g}\left(\frac{M_p^2}{2}R + X - V(\phi)\right) + S_m\big[A^2(\phi)g_{\mu\nu}, \psi_m\big]\,. \tag{2.12}$$

A scalar-tensor theory under this form is said to be in the *Einstein frame*, which means that the gravitational and scalar sectors are only minimal coupled, but the scalar field is directly coupled to the matter field in the matter action S_m through the *Jordan frame* metric $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$. Note that the matter field is coupled minimally to the Jordan frame metric. It is possible to decouple the scalar field from the matter field using the change of variables $g_{\mu\nu} \to \tilde{g}_{\mu\nu}$ and $\phi \to \varphi = A^{-2}(\phi)$, at the cost of introducing non-minimal couplings between the new scalar and tensor fields. The resulting scalar-tensor theory is said to be in the Jordan frame.

In the limit of non-relativistic matter, i.e. $T \approx -\rho$, the scalar field equation becomes

$$\Box\phi = V_{\mathrm{eff},\phi}(\phi)\,, \quad V_{\rm eff}(\phi) = V(\phi) + \rho\ln A(\phi)\,, \tag{2.13}$$

with $\Box = \partial^\mu\partial_\mu$, taking the Newtonian limit for simplicity, and the subscript $,\phi$ representing the derivative with respect to the scalar field. The dependence of the effective potential $V_{\rm eff}$ on the matter density makes it possible to hide the scalar field in high density regions by choosing suitable potential and coupling to matter. In the *chameleon mechanism*, this is achieved by making the effective mass of the scalar field large in high density regions. To see this, let's take a simple example where we have

$$V(\phi) = \frac{\Lambda^{n+4}}{\phi^n} \quad \text{and} \quad A(\phi) = e^{\xi\phi/M_p}\,, \tag{2.14}$$

where $n > 0$ is an integer, leading to the effective potential

$$V_{\rm eff}(\phi) = \frac{M^{n+4}}{\phi^n} + \xi\phi\rho\,. \tag{2.15}$$

We call the minimum of this effective potential $\phi_{\rm min}(\rho)$ and it obviously depends on the density. Let's now assume that we have a large and massive object of radius R and mass M with density $\rho_{\rm in}$ bathed in a low density region $\rho_{\rm out}$, such that $\rho_{\rm in} \gg \rho_{\rm out}$. The goal is then to calculate the contribution $\delta\phi$ of the massive object on the outside scalar field, which takes the form $\phi(r) = \phi_\infty + \delta\phi(r)$, ϕ_∞ being the scalar field far from the object. It is further fair to assume that far away, the scalar field corresponds to the minimum of the effective potential $\phi_\infty = \phi_{\rm min}(\rho_{\rm out})$, while deep within the massive object, the scalar field is $\phi_0 = \phi_{\rm min}(\rho_{\rm in})$. The potential and coupling to matter of the chameleon theory is chosen

such that the effective mass $m_{\rm eff}(\phi) = V_{{\rm eff},\phi\phi}(\phi)$ of the scalar field inside the object is much greater than the effective mass of the outside scalar field, $m_{\rm in} \gg m_{\rm out}$. We are interested in looking at the strength of the contribution of different volume element $\mathrm{d}V(r) = 4\pi r^2 \mathrm{d}r$. The contribution from the core of the object $\delta\phi_c$ behave as a massive scalar field of mass $m_{\rm in}$ and therefore its effect is exponentially suppressed with the distance to the source, namely

$$\delta\phi_c(\tilde{r}) \sim e^{-m_{\rm in}\tilde{r}}\,, \tag{2.16}$$

$\tilde{r}$ being the distance between the core and the outside region. This means that inner regions of the massive object do not significantly contribute to $\delta\phi$ and only a thin layer of thickness ΔR near the surface does. The contribution from the thin shell also behaves as a massive scalar field, but with mass $m_{\rm out}$ and the exterior solution to the scalar field is then well approximated by

$$\phi(r) \approx -\frac{\xi}{4\pi M_p}\frac{3\Delta R}{R}\frac{M}{r-R}e^{-m_{\rm out}(r-R)} + \phi_\infty\,. \tag{2.17}$$

We see than since only a thin shell contributes to the outside scalar field, $\Delta R/R \ll 1$, the force associated to its presence, $F_\phi \sim \phi'(r)$, can be neglected.

2.3.2 Cubic Galileon theory and Vainshtein mechanism

In the realm of Horndeski theories, the *cubic Galileon theory* stands out by its attractiveness in the modified gravity community. This popularity emerged for several reasons. First, it is the simplest scalar field theory that is invariant under a Galilean transformation, which makes the theory stable under radiative quantum correction [7]. When promoted to a scalar-tensor theory, the cubic Galileon theory is the decoupling limit of DGP models [106] and massive gravity [79]. In the context of massive gravity, the decoupling limit corresponds to taking a small graviton mass $m_g \to 0$ compared to the Planck $M_p \to \infty$, while keeping the energy scale $\Lambda = (M_p m^2)^{1/3}$ constant. Having a small graviton mass relative to the Planck mass is justified experimentally, since one requires $m_g \sim 10^{-33}$ eV to explain the observed accelerated expansion and so far experimental constraints lie at $m_g < 10^{-27}$ eV [107].

An other important feature of the cubic Galileon theory is its ability to screen the scalar field at short scales through the *Vainshtein mechanism* [108] and make it negligible in astrophysical experiments. To see this, let us start with the Lagrangian of the theory, namely

$$\mathcal{L} = 6X + \frac{2}{\Lambda^3}X\Box\phi + M_p^{-1}\phi T^\mu_{\ \mu}\,, \tag{2.18}$$

with $\Box = \partial^\mu\partial_\mu$. We further derive the equation of motion using the Euler-Lagrange equation

$$6\Box\phi + \frac{2}{\Lambda^3}\left((\Box\phi)^2 - (\partial_\mu\partial_\nu\phi)^2\right) = -M_p^{-1}T^\mu_{\ \mu}\,. \tag{2.19}$$

Note that we study the scalar sector in flat space-time for simplicity. Let us now assume spherical symmetry and staticity, such as it is the case in the solar system for example. The

field equation becomes

$$\boldsymbol{\nabla}\cdot\left(6\boldsymbol{\nabla}\phi+\hat{\mathbf{r}}\frac{4}{\Lambda^3}\frac{(\boldsymbol{\nabla}\phi)^2}{r}\right)=\frac{M}{M_p}\delta(\mathbf{x})\,, \tag{2.20}$$

where we set the energy-momentum tensor of a point mass object, $T^{\mu}_{\ \mu}=-M\delta(\mathbf{x})$. This equation can be integrated and taking into account that the scalar field only depends on the radial coordinate, we find

$$6\phi'(r)+\frac{4}{\Lambda^3 r}\phi'^2(r)=\frac{M}{4M_p\pi r^2}\,, \tag{2.21}$$

which can be solved algebraically for $\phi'(r)$ leading to

$$\phi'(r)=\frac{3\Lambda^3 r}{4}\left(-1+\sqrt{1+\frac{1}{9\pi}\left(\frac{r_V}{r}\right)^3}\right)\,, \tag{2.22}$$

with $r_V=(M/M_p)^{1/3}/\Lambda$ the Vainshtein radius. In the limit where we are well within the Vainshtein radius, $r\ll r_V$, the force due to the scalar field behaves as

$$F_\phi(r):=\frac{1}{M_p}\phi'(r)\sim\frac{\Lambda^3 r_V}{4\sqrt{\pi}M_p}\sqrt{\frac{r_V}{r}}\,, \tag{2.23}$$

while the gravitational force F_g is well approximated by Newtonian gravity, leading to

$$\frac{F_\phi}{F_g}\sim 2\sqrt{\pi}M_p\left(\frac{r}{r_V}\right)^{3/2}\,. \tag{2.24}$$

The scalar force is negligible in the limit $r\ll r_V$. For the case of the decoupling limit of massive gravity, the Vainshtein radius of a sun-like object is expected to be $\sim 10^4$ pc, making the planets far within the region where the scalar field can be neglected.

2.4 Testing screened Horndeski models with gravitational waves

The cosmological behaviour of Horndeski theories is well understood, in part thanks to the EFT approach to dark energy. However, in order to use the numerous astrophysical tests of General Relativity in the context of Horndeski theories, we need a post-Newtonian (PN) formalism that includes all possible candidates for dark energy. Until recently, this has proven to be a difficult task due to the non-linear nature of screening mechanisms and only theory dependent approaches were developed [109, 110]. New effective approaches to screening mechanisms emerged in recent years. First using a dual Lagrangian obtained with a Legendre transformation [111] and later using what is called the *scaling method* [112], which describes screening effects in a perturbative way. The latter method was later used to derive the parameterized post-Newtonian (PPN) parameters for the cubic Galileon and chameleon theories [113], hence leading the way to a general PN approach of Horndeski theories with screening mechanisms. Ultimately, one would like to constrain the set of all possible EFTs of dark energy, that are compatible with current cosmological observations,

using solar system and pulsar tests of GR. Therefore, we are interested in the development of a PN formalism for Horndeski theories that are reconstructed from the EFT approach. A PN formalism for all Horndeski theories is not necessary. Moreover, only reconstructed Horndeski theories exhibiting screening mechanisms can reasonably satisfy astrophysical tests. The PN formalism for screened reconstructed Horndeski theories, the most relevant scalar-tensor theories describing dark energy, was developed by Renevey et al. [114] with the use of the scaling method.

Heuristically the scaling method can be understood as follow. Let's assume that the scalar field equation takes the general form

$$\alpha^s F_1\big[\phi, X\big] + \alpha^t F_2\big[\phi, X\big] = \frac{T}{M_p^2}\,, \tag{2.25}$$

where $F_{1,2}$ depend on the scalar field and its derivatives, $s, t \in \mathbb{R}$ and α is a coupling constant that controls the scale at which F_1 or F_2 is dominant. The scaling parameter α could represent for example some distance that separate the screened region from the cosmological scale or an energy density scale above which screening is at play. The scaling method consists in making a perturbative expansion of the scalar field as

$$\phi(x^\mu) \approx \phi_0 \left(1 + \alpha^q \psi(x^\mu)\right)\,, \tag{2.26}$$

where the exponent $q \in \mathbb{R}$ is added because we do not know yet how the scaling parameter is related to the remaining dynamical scalar perturbation ψ. One can show [112] that q can only take a finite number of possible values in order for the expansion to be well-defined and the goal is to choose the unique value for q such that in the regime $\psi \ll \alpha^{-q} \implies \alpha^q\psi \ll 1$, we recover Einstein gravity as desired in the screened region. This regime can also be set by taking $\alpha \to \infty$ or $\alpha \to 0$, depending on the sign of q.

To better understand the use of the scaling method, one can apply it on the cubic Galileon and chameleon theories. Starting with the latter, its action in the Jordan frame can be written as

$$S[\phi, g] = \frac{M_p^2}{2}\int d^4x\sqrt{-g}\left(\phi R + \frac{2\omega}{\phi}X - \alpha(\phi - \phi_0)^n\right)\,, \tag{2.27}$$

where ω and α are coupling parameters, $n > 0$ and ϕ_0 some background value of the scalar field. The scaling method consists in making a perturbative expansion of the scalar field as

$$\phi(x^\mu) \approx \phi_0 \left(1 + \alpha^q \psi(x^\mu)\right)\,, \tag{2.28}$$

which we can insert in the action to get

$$S = \frac{M_p^2}{2}\int d^4x\sqrt{-g}\left(\phi_0 R + \phi_0\alpha^q\psi R + 2\omega\phi_0\alpha^{2q}\tilde{X} - \alpha^{1+nq}(\phi_0\psi)^n\right)\,, \tag{2.29}$$

with $\tilde{X} = -\partial_\mu\psi\partial^\mu\psi/2$. We see that in the limit $\alpha^q\psi \ll 1$ we indeed recover the Einstein-Hilbert action. We are left to find the dynamics of the scalar field and show that we can indeed find q such that the expansion (2.28) is well defined. In order to have a well defined action in expansion of α^q, we need that the last term on the RHS to have an integer power

of α^q, thus $1+nq=mq$, $m\in\mathbb{N}^+$. If $m>1$, the resulting scalar field equation at leading order would be $R=0$, which is in contradiction with the Einstein field equations. This implies that the only choice we have left is $q=(1-n)^{-1}$. Deriving the scalar field equation at leading order α^q using the Euler-Lagrange equation gives

$$n\psi^{n-1}=R^{(0)}\approx\frac{T}{M_p^2}\,, \tag{2.30}$$

the second equality is coming from the leading order field equations for the metric $g^{(0)}_{\mu\nu}$, which are the Einstein field equations. We see that if we want chameleon screening to be effective, i.e. $\psi\sim 0$, in a high density regime we need the exponent in Eq. (2.30) to be negative, $n-1<0$. Hence q is in fact positive and the screened regime is found with the limit $\alpha\to 0$.

Next we can use the scaling method in the case of a derivative screening such as in the cubic Galileon theory. The Jordan frame action is

$$S[\phi,g]=\frac{M_p^2}{2}\int \mathrm{d}^4x\sqrt{-g}\left(\phi R+\frac{2\omega}{\phi}X-\alpha\frac{X}{4\phi^3}\Box\phi\right), \tag{2.31}$$

and plugging in the scalar field expansion, we get

$$S=\frac{M_p^2\phi_0}{2}\int \mathrm{d}^4x\sqrt{-g}\left(R+\alpha^q\psi R+2\omega\alpha^{2q}\tilde{X}-\alpha^{1+3q}\frac{\tilde{X}}{4\phi_0}\Box\psi\right). \tag{2.32}$$

Once again, we recover GR in the limit $\alpha^q\psi$, but we need to show that the scalar field expansion is well defined by finding a unique q. Using a similar argument as in the previous paragraph, we only have one choice for the exponent, $q=-1/2$. The screened regime is therefore described using the limit $\alpha\to\infty$. The scalar field equation at leading order can then be found using the Euler-Lagrange equation. One can of course compare this approach to the previous section 2.3.2. We found that the cubic Galileon field behaves as

$$\phi'(r)\sim\frac{\Lambda^{3/2}}{\sqrt{r}}\,, \tag{2.33}$$

in the screened regime, for a spherically symmetric source. By comparing both actions, we know that $\alpha=\Lambda^{-3}$. Furthermore, one can show by solving the Euler-Lagrange equation with spherical symmetry that the scalar perturbation $\psi(r)$ behaves as

$$\psi(r)\sim\sqrt{r}\,, \tag{2.34}$$

in the same regime. Comparing both approaches we get

$$\phi'(r)\approx\alpha^q\psi'(r)\sim\frac{\Lambda^{-3q}}{\sqrt{r}}\sim\frac{\Lambda^{3/2}}{\sqrt{r}}\,, \tag{2.35}$$

indeed recovering the same behaviour.

Using the scaling method, we constructed a perturbative approach to screening mechanisms, where the scalar field and the metric decouple at leading order and we recover General Relativity in the screened regime. At this level of precision, the PN formalism of

the screened model is equivalent to the PN expansion of GR, but we are not restricted to leading order. One can also expand the metric in powers of α^q, namely $g_{\mu\nu} \approx g^{(0)}_{\mu\nu} + \alpha^q g^{(q)}_{\mu\nu}$, where $g^{(0)}_{\mu\nu}$ solves the Einstein field equations and one can find the field equation for $g^{(q)}_{\mu\nu}$ with the Euler-Lagrange equation with respect to the metric for the term proportional to α^q. It is then possible to build a PN expansion for the metric correction term $g^{(q)}_{\mu\nu}$ and build the PN formalism for screened theories at next-to-leading order [113, 114], which includes the small effect of the scalar field on the metric. A PN approach deriving the gravitational waveform produced by two compact objects was developed in Ref. [115] for GR and in Ref. [116] for Brans-Dicke theories. Hence, one direct application of the PN formalism of screened Horndeski theories is the calculation of the gravitational waveform produced by a compact binary system for theories with screening mechanism. In the following article, we made use of the scaling method and the derived PN formalism for screened theories to calculate the gravitational waves resulting from the coalescence of two compact objects.

Effect of screening mechanisms on black hole binary inspiral waveforms

Cyril Renevey[*]
Laboratoire de Physique Subatomique et Cosmologie, 53 Avenue des Martyrs, 38000 Grenoble, France

Ryan McManus[†]
School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

Charles Dalang[‡] and Lucas Lombriser[§]
Département de Physique Théorique, Université de Genève, 24 quai Ernest Ansermet, 1211 Genève 4, Switzerland

(Received 6 December 2021; accepted 4 March 2022; published 28 April 2022)

Scalar-tensor theories leaving significant modifications of gravity at cosmological scales rely on screening mechanisms to recover general relativity (GR) in high-density regions and pass stringent tests with astrophysical objects. Much focus has been placed on the signatures of such modifications of gravity on the propagation of gravitational waves (GWs) through cosmological distances while typically assuming their emission from fully screened regions with the wave generation strictly abiding by GR. Here, we closely analyze the impact of screening mechanisms on the inspiral GW waveforms from compact sources by employing a scaling method that enables a post-Newtonian (PN) expansion in screened regimes. Particularly, we derive the leading-order corrections to a fully screened emission to first PN order in the near zone, and we also compute the modifications in the unscreened radiation zone to second PN order. For a concrete example, we apply our results to a cubic Galileon model. The resulting GW amplitude from a binary black hole inspiral deviates from its GR counterpart at most by one part in 10^2 for the modifications in the radiation zone and at most one part in 10^{11} due to next-order corrections to the fully screened near zone. We expect such modifications to be undetectable by the current generation of GW detectors, but the deviation is not so small as to remain undetectable in future experiments.

DOI: 10.1103/PhysRevD.105.084059

I. INTRODUCTION

General relativity (GR) and quantum field theory constitute the cornerstones of theoretical physics and allow for precise predictions which have been successfully tested experimentally [1–3]. In the context of cosmology, the observational viability of GR invokes a universe dominated by dark fluids [4]: a dark energy, explaining the late-time accelerated expansion of our Universe, and cold dark matter, describing the cosmic microwave background anisotropies, galaxy clustering and lensing, as well as galactic rotation curves. For a dark energy in the form of a cosmological constant this composition constitutes the concordance cosmology, the Λ-cold-dark-matter (ΛCDM) model. Despite its observational success, a number of smaller and larger tensions have recently arisen that remain unexplained to date [5–13]. Perhaps an even larger enigma is posed by our lack of understanding of the cosmological constant. If attributed to the quantum fluctuations of vacuum energy, theoretical estimates for its magnitude from quantum field theory are off by more than 50 orders of magnitude [14–17]. One of the leading hypotheses for this discrepancy is the existence of new symmetries that are thought to cancel vacuum fluctuations. This, however, leaves the need for an explanation of the small residual cosmological constant driving the observed late-time accelerated cosmic expansion.

In this context, a range of alternative explanations for cosmic acceleration to this residual cosmological constant has (re)emerged over the past two decades, which involve adding new degrees of freedom to the gravitational action [18–23]. Of these alternative theories of gravity, the Horndeski action [19] constitutes the most general class of Lorentz-invariant four-dimensional scalar-tensor theories that lead to second-order equations of motion, which turns out to be sufficient [24] but not necessary [25–27] to avoid Ostrogradski instabilities. A significant parameter space of this class of models has traditionally been explored as an explanation of cosmic acceleration. Often, a

[*]renevey@lpsc.in2p3.fr
[†]ppzrm@exmail.nottingham.ac.uk
[‡]charles.dalang@unige.ch
[§]lucas.lombriser@unige.ch

significant modification of gravity on cosmological scales was invoked as the driver of cosmic acceleration, while so-called *screening mechanisms* would allow a recovery of GR in high-density environments or at short distances such as in galaxies or in the Solar System, where tight constraints on deviations from GR have been established [1,28–33]. However, this concept has become severely challenged by the luminal propagation of gravitational waves (GWs) [34,35], and cosmic acceleration may be limited to the dark energy aspects of a Horndeski field rather than its direct dynamical change of gravity. Nevertheless, dark energy fields may be accompanied by cosmologically significant universal interactions with the matter fields, which require screening on small scales and for which screening effects can provide distinctive observational signatures [33].

Prominent screening mechanisms include the chameleon [36], k-mouflage [37], and Vainshtein [38] mechanisms, which rely on nonlinear terms in the equations of motion to recover GR in their respective screened regimes. Consequently, the nonlinearity renders post-Newtonian approaches difficult to implement due to the linearization of the relevant field equations. However, it was shown that this problem can be circumvented with the employment of the so-called *scaling method* [39–42] that enables an expansion in screened regions. More specifically, the method can be used to identify both the leading corrections from the scalar field to the metric field equations in the screened small-scale regions or conversely of the unscreened theory on large scales.

The first direct detection of GWs in 2015 [43] has opened a new observational window on gravity. To exploit this new wealth of data, a range of analytical and numerical techniques has been developed to study the inspiral-merger ringdown of two heavy compact objects [44–46]. Among them is the post-Newtonian (PN) formalism for the inspiral phase, where a term in the waveform is said to be of PN order n if it is of order $\mathcal{O}((v/c)^{2n})$, abbreviated $\mathcal{O}(\epsilon^n)$, where v is the speed of any of the two objects. In GR, the waveform from an inspiral system has already been determined to 5PN [47,48], with multiple complementary methods of calculation. In the case of Brans-Dicke theories [18], the calculations have been performed up to 2PN by Lang [49] following the method developed by Will and Wiseman [50]. In this study, we will use these two references extensively and refer to them from now on as Lang14 and WW96, respectively. In the context of modified gravity, numerous studies have explored how propagation effects modify the amplitude or polarization of the GW [51–64]. Predominantly, these studies assume that the generation of the waveform can be modeled according to GR, invoking the operation of a screening mechanism at the source as motivation for that.

In this paper, we compute the leading-order corrections to the GW waveform expected for the three aforementioned screening mechanisms [65], including effects on both propagation and the leading correction to the screened GW generation. To this end, we follow the method of WW96 and cast the field equations in their relaxed form, i.e., as a flat-space wave equation sourced by quadratic self-interactions and the stress-energy tensor. In this form, the solutions are retarded Green's integrals, for which the integrands depend on the GW itself, the quantity we are trying to solve for. The problem can be solved perturbatively by first neglecting the self-interactions, providing an approximate solution $h^{\mu\nu}_{(1)}$, where the only source is the energy-momentum tensor, and introducing those in the retarded integrals to obtain a new, more accurate solution $h^{\mu\nu}_{(2)}$. In practice, the integrals are split between the so-called *near* and *radiation* zones, which are depicted in Fig. 1. These are complementary spacelike volumes with their boundary carrying no physical meaning. In GR, the volume separation is introduced to cancel apparently divergent surface integrals between the near and radiation zones. In this work, we use this splitting in a physically motivated sense such that the near zone corresponds to a screened region, where nevertheless small deviations with respect to GR may arise at the next-to-leading order, and where an unscreened scalar-tensor modification applies in the radiation zone. Besides determining the overall modified waveform, this strategy enables us to compute the leading-order corrections on the generation of the waveform in modified gravity theories, which so far has typically been assumed fully screened and abiding exactly by GR.

The paper is organized as follows. In Sec. II, we derive the equations of motion for a subset of Horndeski theories

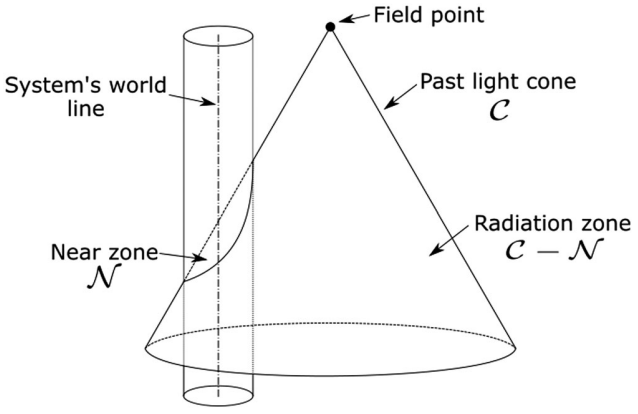


FIG. 1. Adapted from [50]. The past light cone $\mathcal{C}$ of a field point far from the binary system. The near zone, $\mathcal{N}$, is the intersection of the past light cone and the world tube of the system's world line with radius $\mathcal{R}$. The radiation zone, $\mathcal{C}-\mathcal{N}$, is the remainder of the light cone. We assume that the metric field equations take on their screened form within $\mathcal{N}$, whereas they assume their unscreened form in $\mathcal{C}-\mathcal{N}$. When calculating the waveform at the field point, we treat the contribution from the near zone and radiation zone separately.

that describe tensorial GWs propagating at light speed and that have sufficient freedom to allow for one of the three types of screening mechanisms. We express those in their relaxed form, briefly review the scaling method, and explain how it applies to the retarded integrals of Green's function. In Sec. III, we combine the results of Lang14 and WW96, assuming first the fully screened scenario, where the scalar field has no effect in the near zone and GR applies exactly. We compute the leading terms that arise from Brans-Dicke corrections in the radiation zone to 2PN. In Sec. IV, we then compute the leading correction to the fully screened case from the scalar field in the near zone by employing the scaling method. Further, we calculate the gravitational waveform up to 1PN beyond the quadrupole formula. In Sec. V, we compute those corrections for a black hole binary system and apply it to an example of Galileon cosmology. Finally, Sec. VI is dedicated to our conclusions, and Appendix provides further details on our calculation of the gravitational waveform from a binary system in the case of GR.

Throughout the paper, we adopt the signature $(-,+,+,+)$ for the metric. Greek indices run from 0 to 3 and Latin indices from 1 to 3. A comma indicates a partial derivative, $A_{,\mu} \equiv \partial_\mu A$, while a semicolon denotes a covariant derivative associated with the Levi-Civita connection, $A_{\mu;\nu} \equiv \nabla_\nu A_\mu$. Bold symbols $\boldsymbol{x}$ represent Euclidean three-dimensional vectors with Euclidean norm $|\boldsymbol{x}|$. Symmetrization of indices follow the notation $T_{(\mu_1\cdots\mu_n)} \equiv \frac{1}{n!}\sum_{i\in S_n} T_{\sigma_i(\mu_1\cdots\mu_n)}$, where $\{\sigma_i\}_{i\in S_n}$ is the set of all permutations of $\{\mu_1,\ldots,\mu_n\}$. Units are such that $c=\hbar=1$.

II. MODIFIED FIELD EQUATIONS IN THE PRESENCE OF SCREENING

We will consider scalar-tensor theories that can be described by the following action:

$$S[g,\phi] = \frac{1}{16\pi G}\int \sqrt{-g}\Big(\phi R + 2\frac{\omega(\phi)}{\phi}X + \alpha\delta_C V(\phi) + \alpha\delta_K\Gamma_2(\phi)X^2 + \alpha\delta_V\Gamma_3(\phi)X\Box\phi\Big), \tag{1}$$

where R is the Ricci scalar of the metric $g_{\mu\nu}$ with determinant $g \equiv \det g_{\mu\nu}$. $V, \omega, \Gamma_2, \Gamma_3$ are arbitrary functions of the scalar field ϕ, $X = -g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi/2$ is its kinetic term, and $\Box \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu$. The action (1) corresponds to a reduced set of Horndeski theories [19] with luminal propagation of tensor gravitational waves [39], satisfying arrival time constraints [51] from GW170817 [35]. The functions $V(\phi)$, $\Gamma_2(\phi)$, and $\Gamma_3(\phi)$ allow for large field value as well as first and second derivative screening, respectively [22]. The parameters δ_C, δ_K, and δ_V are unity when implementing the chameleon [36], k-mouflage [37], and Vainshtein [38] screening mechanisms, respectively, and vanish otherwise. While Eq. (1) in principle admits a superposition of screening mechanisms, in this work, we assume that exactly one of the screening effects is operating at a time. One can impose this condition by requiring $\delta_C+\delta_K+\delta_V=1$. We emphasize that therefore our results are not applicable to basic Brans-Dicke models that do not screen such as treated in Lang14. The *scaling parameter* α will be utilized to identify the most relevant terms beyond GR in the screened regime. The field equations derived from the action (1) are of the form

$$\phi G_{\mu\nu} = 8\pi G T_{\mu\nu} + \frac{\omega(\phi)}{\phi}\left(\phi_{,\mu}\phi_{,\nu} - \frac{1}{2}g_{\mu\nu}\phi_{,\lambda}\phi^{,\lambda}\right) + (\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\Box\phi) + \alpha A_{\mu\nu}[g,\phi], \tag{2}$$

$$\Box\phi = \frac{1}{3+2\omega(\phi)}\left(8\pi G T - 16\pi G\phi\frac{\partial T}{\partial\phi} - \frac{\mathrm{d}\omega}{\mathrm{d}\phi}\phi_{,\lambda}\phi^{,\lambda} + \alpha A^{\lambda}{}_{\lambda}[g,\phi]\right) + \alpha B[g,\phi], \tag{3}$$

where $G_{\mu\nu}$ is the Einstein tensor, $T_{\mu\nu}$ is the stress-energy (or energy-momentum) tensor of the matter fields and $T \equiv g^{\mu\nu}T_{\mu\nu}$ denotes its trace. The second term on the right-hand side of the scalar field equation of motion appears if we assume that the energy-momentum tensor effectively depends on the scalar field [69]. The additional two functionals $A_{\mu\nu}[g,\phi]$ and $B[g,\phi]$ include the contributions of the functions V, Γ_2, and Γ_3, namely

$$A_{\mu\nu} = \delta_C\frac{1}{2}g_{\mu\nu}V(\phi) + \delta_K\Gamma_2(\phi)\left(\frac{1}{2}g_{\mu\nu}X^2 + X\nabla_\mu\phi\nabla_\nu\phi\right) + \delta_V\frac{\partial\Gamma_3}{\partial\phi}X\nabla_\mu\nabla_\nu\phi + \delta_V\Gamma_3(\phi)\left(\frac{1}{2}g_{\mu\nu}\nabla^\alpha\phi\nabla_\beta\nabla_\alpha\phi\nabla^\beta\phi + \Box\phi\nabla_\mu\phi\nabla_\nu\phi - \nabla^\alpha\phi\nabla_\alpha\nabla_{(\mu}\phi\nabla_{\nu)}\phi\right), \tag{4}$$

$$B = -\delta_C\frac{\partial V}{\partial\phi} - \delta_K\Gamma_2(\phi)(2X\Box\phi - 2\nabla_\alpha\phi\nabla_\beta\phi\nabla^\alpha\nabla^\beta\phi) + 3\delta_K\frac{\partial\Gamma_2}{\partial\phi}X^2 - \delta_V\Gamma_3(\phi)((\Box\phi)^2 - \nabla_\alpha\nabla_\beta\phi\nabla^\alpha\nabla^\beta\phi) + 2\delta_V\left(\frac{\partial\Gamma_3}{\partial\phi}\nabla_\alpha\phi\nabla_\beta\phi\nabla^\alpha\nabla^\beta\phi + \frac{\partial^2\Gamma_3}{\partial\phi^2}X^2\right). \tag{5}$$

The role of screening mechanisms is to suppress the gravitational effect of the scalar field in high-density regions. Therefore, the functionals $A_{\mu\nu}$ and B are expected to have a significant effect in the near zone, so as to recover GR, whereas they are expected to vanish far away from the source.

For the matter source, we consider a collection of point particles, labeled with capital indices, which are described by the following energy-momentum tensor at the field point $x^\gamma = (t,\boldsymbol{x})$:

$$T^{\alpha\beta}(x^\gamma) = \sum_A m_A(\phi)(-g)^{-1/2}\frac{u_A^\alpha u_A^\beta}{u_A^0}\delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t)), \quad (6)$$

where u_A^α is the four-velocity of body A, $\boldsymbol{x}_A(t)$ is its time-dependent spatial position, and m_A its gravitational mass including the gravitational binding energy of the Ath body. An unfortunate consequence of modeling the matter source as point particles is that we neglect tidal effects, but see [70] for recent work on tidal deformability of black holes.

Note that the masses can depend on the scalar field, even in the Jordan frame we have adopted [69]. With the assumption of massive pointlike objects, we neglect nontrivial effects of ϕ on the matter content, which violate the strong equivalence principle. For Brans-Dicke theories, one needs to take into account the dependence of the gravitational constant on the scalar field since it affects the total gravitational energy. As suggested in Ref. [69], when considering the inertial mass of objects in the Jordan frame, the masses of compact objects are replaced with functions of the scalar field where the total mass is given by

$$m \equiv \sum_A m_A(\phi). \quad (7)$$

The inertial mass can then be expanded about the local background value of the scalar field ϕ_0,

$$m_A(\phi) \approx m_A(\phi_0)(1 + s_A(\phi-\phi_0)/\phi_0), \quad (8)$$

where

$$s_A = -\left.\frac{\partial \ln m_A}{\partial \ln G_N}\right|_{\phi_0} = \left.\frac{\partial \ln m_A}{\partial \ln \phi}\right|_{\phi_0} \quad (9)$$

is the sensitivity of compact objects in scalar-tensor theories and, in particular, $s_A = 1/2$ for black holes. However, for theories exhibiting Vainshtein or k-mouflage mechanisms the scalar field equation in the screened region effectively becomes a total derivative. One can then use the generalized Gauss theorem to stipulate that the scalar field cannot depend on the composition of the object, but only on its total mass [71]. Therefore, we only need to take the sensitivity into account in the case of the chameleon mechanism and the inertial mass becomes

$$m_A(\phi) \approx m_A(1 + \delta_C s_A(\phi-\phi_0)/\phi_0), \quad (10)$$

where we renamed $m_A(\phi_0) \to m_A$ for simplicity.

With the equations of motion at hand and an explicit matter source specified, we proceed to compute the waveform at a generic spacetime point far away from the source, in the so-called *radiation zone*. To this end, we make use of the field equations (2) and (3) in their relaxed forms. These are obtained through a series of field redefinition. As the scalar field couples nonminimally to the metric, their derivation is different from that in GR and we follow Ref. [72] to find them. The scalar field is redefined as

$$\varphi(t,\boldsymbol{x}) \equiv \frac{\phi(t,\boldsymbol{x})}{\phi_0}. \quad (11)$$

We assume it to be static for the purpose of this calculation as the timescales over which the background value varies are much larger than the orbital timescales of the system. Note that a change in the scalar field value between the spacetime positions of the source and observer may be detected as an effective modification of the luminosity distance [57,58]. We define the following symmetric rank-two tensor, which describes the deviation from flat space [73]:

$$h^{\mu\nu} = \eta^{\mu\nu} - \varphi\sqrt{-g}g^{\mu\nu}. \quad (12)$$

Note that we have not yet assumed that $h^{\mu\nu}$ is small. Next, we impose the flat-space harmonic gauge on $h^{\mu\nu}$,

$$h^{\mu\nu}{}_{,\nu} = 0. \quad (13)$$

The purpose for the field redefinition (12) is to rewrite the metric field equation (2) as a flat-space wave equation,

$$\Box_\eta h^{\mu\nu} = -16\pi G\tau^{\mu\nu} - 2\alpha(-g)A^{\mu\nu}\frac{\phi}{\phi_0} \equiv S^{\mu\nu}, \quad (14)$$

where $\Box_\eta \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ and

$$\tau^{\mu\nu} \equiv (-g)\varphi T^{\mu\nu} + \frac{1}{16\pi G}(\Lambda^{\mu\nu} + \Lambda_s^{\mu\nu}). \quad (15)$$

One can think of $\Lambda^{\mu\nu}$ and $\Lambda_s^{\mu\nu}$ as the stress-energy tensors held in the gravitational wave and scalar field, respectively. These in turn are given by

$$\Lambda^{\mu\nu} \equiv 16\pi G(-g)t_{LL}^{\mu\nu} + h^{\mu\alpha}{}_{,\beta}h^{\nu\beta}{}_{,\alpha} - h^{\alpha\beta}h^{\mu\nu}{}_{,\alpha\beta}, \quad (16)$$

where

$$\begin{aligned}(-g)t_{LL}^{\mu\nu} \equiv \frac{1}{16\pi G}\Big[&g_{\lambda\alpha}g^{\beta\rho}h^{\mu\lambda}{}_{,\beta}h^{\nu\alpha}{}_{,\rho} + \frac{1}{2}g_{\lambda\alpha}g^{\mu\nu}h^{\lambda\beta}{}_{,\rho}h^{\rho\alpha}{}_{,\beta}\\ &- 2g_{\alpha\beta}g^{\lambda(\mu}h^{\nu)\beta}{}_{,\rho}h^{\rho\alpha}{}_{,\lambda} + \frac{1}{8}(2g^{\mu\lambda}g^{\nu\alpha} - g^{\mu\nu}g^{\lambda\alpha})\\ &\times(2g_{\beta\rho}g_{\sigma\tau} - g_{\rho\sigma}g_{\beta\tau})h^{\beta\tau}{}_{,\lambda}h^{\rho\sigma}{}_{,\alpha}\Big]\end{aligned} \quad (17)$$

is the Landau-Lifshitz tensor and

$$\Lambda_s^{\mu\nu} \equiv \frac{3+2\omega}{\varphi^2}\varphi_{,\alpha}\varphi_{,\beta}\Big((\eta^{\mu\alpha}-h^{\mu\alpha})(\eta^{\nu\beta}-h^{\nu\beta}) - \frac{1}{2}(\eta^{\mu\nu}-h^{\mu\nu})(\eta^{\alpha\beta}-h^{\alpha\beta})\Big). \tag{18}$$

The scalar field satisfies the equation

$$\Box_\eta \varphi = -8\pi G\tau_s + \alpha B \equiv S, \tag{19}$$

where the source of the scalar field equation is given by

$$\tau_s = -\frac{1}{3+2\omega}\sqrt{-g}\varphi\left(T - 2\varphi\frac{\partial T}{\partial\phi}\right) - \frac{1}{8\pi G}h^{\alpha\beta}\varphi_{,\alpha\beta} + \frac{1}{16\pi G}\frac{\mathrm{d}}{\mathrm{d}\varphi}\left[\ln\left(\frac{3+2\omega}{\varphi^2}\right)\right]\varphi_{,\alpha}\varphi_{,\beta}(\eta^{\alpha\beta}-h^{\alpha\beta}). \tag{20}$$

One may worry that second-order derivatives of the metric and scalar fields appear within both $A^{\mu\nu}$ and B in Eqs. (14) and (19). However, the α prefactor in both terms ensures that the system in the screened and unscreened regime is appropriately diagonalized with respect to the second derivatives, with those appearing within $A^{\mu\nu}$ and B acting as source terms for the next order in α corrections to the metric and scalar fields.

Since the field equations take the form of a flat-space wave equation, one can express them as retarded integrals with the appropriate Green's function, namely

$$h^{\mu\nu}(t,\boldsymbol{x}) = -\int_{\mathcal{C}}\frac{S^{\mu\nu}(t-|\boldsymbol{x}-\boldsymbol{x}'|,\boldsymbol{x}')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\mathrm{d}^3\boldsymbol{x}', \tag{21}$$

$$\varphi(t,\boldsymbol{x}) = -\int_{\mathcal{C}}\frac{S(t-|\boldsymbol{x}-\boldsymbol{x}'|,\boldsymbol{x}')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\mathrm{d}^3\boldsymbol{x}', \tag{22}$$

where $\mathcal{C}$ is the past light cone of the field point x^μ. Note that the source terms S depend on both φ and $h^{\mu\nu}$, which make Eqs. (21) and (22) coupled integral equations for φ and $h^{\mu\nu}$.

As a first step, in this work, we specify to cases where the contribution of the scalar wave can be neglected. We leave the computations for the more general case to future work. Depending on the screening mechanism, the effect of an incoming scalar wave on our detectors may or may not be attenuated by screening in the Solar System. For example, in the chameleon screening mechanism, the suppression of scalar waves depends on the ratio between the effective mass of the scalar field and the GW frequency [74]. However, not all mechanisms screen incoming scalar waves. For example, the Vainshtein mechanism does not [75], and scalar waves might play an important role in such scenarios, even if their amplitude at emission has been shown numerically to be suppressed [66].

For an observer located far away from the source, in the radiation zone, the contributions of $A_{\mu\nu}$ and B to the field equations can be neglected as we adopted a Brans-Dicke field at large scales. This reduces the integrands to $S_{\mu\nu} \to -16\pi\tau_{\mu\nu}$ and $S \to -8\pi\tau_s$. The assumption of Brans-Dicke behavior at leading order on cosmological scales can be motivated by the effective field theory (EFT) approach to dark energy [76–85]. Under this formalism, the cosmological behavior of Horndeski theories is described as an expansion in the metric perturbations and the scalar field represents the Goldstone boson of the symmetry breaking of time translation. The leading-order terms of a general Horndeski action at low energies are then equivalent to those of a Brans-Dicke theory with scalar field potential [86–88]. Furthermore, we split the integration domain into two parts: the *near zone* $\mathcal{N}$, which we assume to be screened, and the *radiation zone* $\mathcal{C}-\mathcal{N}$, assumed unscreened. For the tensor wave, this splitting reads

$$h^{\mu\nu}(t,\boldsymbol{x}) = 4\int_{\mathcal{N}}\frac{\tau^{\mu\nu}_{\mathcal{N}}(t-|\boldsymbol{x}-\boldsymbol{x}'|,\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|}\mathrm{d}^3\boldsymbol{x}' + 4\int_{\mathcal{C}-\mathcal{N}}\frac{\tau^{\mu\nu}_{\mathcal{C}-\mathcal{N}}(t-|\boldsymbol{x}-\boldsymbol{x}'|,\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|}\mathrm{d}^3\boldsymbol{x}'. \tag{23}$$

Importantly, the effective stress-energy tensor takes different forms in the two different regions, represented by $\tau^{\mu\nu}_{\mathcal{N}}$ and $\tau^{\mu\nu}_{\mathcal{C}-\mathcal{N}}$ for the near and radiation zones, respectively. The latter is given by the Brans-Dicke effective stress-energy tensor in Eq. (15), and we describe $\tau^{\mu\nu}_{\mathcal{N}}$ in the following.

In the near zone, screening shall operate and we expect a close recovery of GR. However, while the effect of the scalar field is strongly attenuated, it nevertheless does not completely vanish. One can describe this contribution of the scalar field using the *scaling method* developed in Refs. [39–42], in which the metric and the scalar field are expanded as

$$\phi(t,\boldsymbol{x}) = \phi_0(1+\alpha^q\psi(t,\boldsymbol{x}) + \mathcal{O}(\alpha^{2q}\psi^2)), \tag{24}$$

$$g_{\mu\nu}(t,\boldsymbol{x}) = g^{(0)}_{\mu\nu}(t,\boldsymbol{x}) + \alpha^q g^{(q)}_{\mu\nu}(t,\boldsymbol{x}) + \mathcal{O}(\alpha^{2q}\psi^2). \tag{25}$$

Here, ϕ_0 is the local background value of the scalar field sourced by the local environment. It varies sufficiently slowly so as to be treated as a constant. The dynamical contribution of the scalar field is captured by ψ. The scaling parameter α is the relevant quantity describing the operation of a screening mechanism and was already introduced in Eq. (1), where it represents the couplings to different terms in the scalar-tensor action. Finally, the exponent $q \in \mathbb{R}$ is uniquely determined for each screening mechanism [39]. In particular, the exponent is chosen such that we have consistent equations of motion recovering GR in the limit $\alpha^q\psi \to 0$ and that the scalar field is sourced by matter, ensuring a viable perturbative expansion in powers of $\alpha^q\psi$. Describing the screened regime with the scaling method [39–42] was shown to be equivalent, but significantly less cumbersome, to techniques employing a dual Lagrangian

[89] or Lagrange multipliers [90] and was carefully studied in the context of the Vainshtein mechanism [91].

The metric is decomposed into two parts. The first contribution $g^{(0)}_{\mu\nu}$ satisfies the Einstein field equations, where the local background scalar field effectively modifies the coupling to matter,

$$G^{(0)}_{\mu\nu} = \frac{8\pi G}{\phi_0} T_{\mu\nu}. \tag{26}$$

The second term $g^{(q)}_{\mu\nu}$ includes the leading contributions to the metric of the correction to the scalar field $\alpha^q\psi$. Its equation of motion is found by substituting $\phi \to \phi_0(1+\alpha^q\psi)$ in Eq. (2) and keeping only the leading contributions that are proportional to α^q. The equation for ψ is found using the same substitution in Eq. (3), where the functional $B[g,\psi]$ becomes the leading contribution together with the stress-energy tensor. The resulting expressions are

$$\begin{aligned} G^{(q)}_{\mu\nu} &= \psi G^{(0)}_{\mu\nu} + \nabla_\mu\nabla_\nu\psi - g^{(0)}_{\mu\nu}\Box\psi + \delta_C \frac{1}{2} g^{(0)}_{\mu\nu} V_0 \phi_0^{n_V-1}\psi^{n_V} \\ &+ \delta_K \phi_0^3 \Gamma_2(\phi_0)\left(\frac{1}{2} g^{(0)}_{\mu\nu}\tilde{X}^2 + \tilde{X}\nabla_\mu\psi\nabla_\nu\psi\right) \\ &+ \delta_V \phi_0^2 \Gamma_3(\phi_0)\left(\frac{1}{2} g^{(0)}_{\mu\nu}\nabla^\alpha\psi\nabla_\beta\nabla_\alpha\psi\nabla^\beta\psi\right. \\ &\left. + \Box\psi\nabla_\mu\psi\nabla_\nu\psi - \nabla^\alpha\psi\nabla_\alpha\nabla_{(\mu}\psi\nabla_{\nu)}\psi\right), \end{aligned} \tag{27}$$

$$\begin{aligned} \frac{8\pi G}{\phi_0}\left(T - 2\frac{\partial T}{\partial\phi}\right) &= \delta_C n_V \frac{\psi^{n_V-1}}{\phi_0^{n_V-1}} V_0 + 2\delta_K\phi_0^3\Gamma_2(\phi_0) \\ &\times (\tilde{X}\Box\psi + \nabla_\mu\tilde{X}\nabla^\mu\psi) + \delta_V\phi_0^2\Gamma_3(\phi_0) \\ &\times ((\Box\psi)^2 - \nabla_\mu\nabla_\nu\psi\nabla^\mu\nabla^\nu\psi), \end{aligned} \tag{28}$$

where $\tilde{X} = -1/2 g^{(0)}_{\mu\nu}\partial^\mu\psi\partial^\nu\psi$ and the covariant derivatives as well as the contractions are obtained using $g^{(0)}_{\mu\nu}$. Allowing for an embedding of the chameleon, k-mouflage, and Vainshtein screening mechanisms, we have assumed that $V(\phi) = V_0\alpha^{n_V q}\psi^{n_V}$ with $0 < n_V < 1$, $\Gamma_2(\phi) = \Gamma_2(\phi_0) \neq 0$, and $\Gamma_3(\phi) = \Gamma_3(\phi_0) \neq 0$ at leading order in $\alpha^q\psi$ [39–42,92]. For the remainder of this work, we shall furthermore adopt a normalization of the gravitational constant by $G/\phi_0 = 1$. By looking at the right-hand side of Eq. (28), one can see that in the case of k-mouflage and Vainshtein screening, the scalar contribution enters as a total derivative. In these cases, this holds true for any matter distribution, and therefore only the total mass of the objects can affect the scalar field. Under the scaling formalism, one can expand the near-zone effective stress-energy tensor $\tau^{\mu\nu}_{\mathcal{N}}$ as

$$\tau^{\mu\nu}_{\mathcal{N}} = \tau^{\mu\nu}_{\rm GR} + \alpha^q\tau^{\mu\nu}_{(q)} + \mathcal{O}(\alpha^{2q}\psi^2), \tag{29}$$

where $\tau^{\mu\nu}_{\rm GR}$ is the part fully described by GR and $\tau^{\mu\nu}_{(q)}$ will be calculated to first post-Newtonian order (1PN) in Sec. IV A. We also note that in both regions, $\mathcal{N}$ and $\mathcal{C}-\mathcal{N}$, the effective stress-energy tensor is conserved to leading order in α, i.e., $\tau^{\mu\nu}{}_{,\mu} = 0 + \mathcal{O}(\alpha)$.

The gravitational waveform h^{ij} can be decomposed into three parts,

$$h^{ij}(t,\boldsymbol{x}) = h^{ij}_{(0)}(t,\boldsymbol{x}) + h^{ij}_{\rm BD}(t,\boldsymbol{x}) + \alpha^q h^{ij}_{(q)}(t,\boldsymbol{x}) + \mathcal{O}(\alpha^{2q}\psi^2). \tag{30}$$

Hereby, $h^{ij}_{(0)}$ denotes the contribution that is described by GR, corresponding to the result of the retarded integral of $\tau^{\mu\nu}_{\rm GR}$ in the near zone and the GR part of $\tau^{\mu\nu}_{\mathcal{C}-\mathcal{N}}$ in the radiation zone. This term has been computed to 2PN beyond the quadrupole formula in WW96. In Sec. III, we also calculate the contribution from the scalar part of $\tau^{\mu\nu}_{\mathcal{C}-\mathcal{N}}$ to 2PN, namely $h^{ij}_{\rm BD}$. The last term, $\alpha^q h^{ij}_{(q)}$, originates from the screened effect of the near-zone scalar field, and it will be derived to 1PN in Sec. IV. In models where gravitational objects can be self-screened relative to an ambient scalar field, such as is the case for the chameleon mechanism, black holes are, in fact, insensitive to this extra degree of freedom [93,94]. Such behavior motivates the distinction between the fully screened case where the near-zone contribution $\alpha^q h^{ij}_{(q)}$ is neglected and the case where it is not.

III. GRAVITATIONAL WAVEFORM OF THE FULLY SCREENED THEORY

We shall first consider the case where a system is completely screened in the near zone; i.e., it is fully described by GR in this region, and reduces to Brans-Dicke theory in the unscreened radiation zone. We determine the gravitational waveform up to 2PN for an arbitrary source distribution in this scenario. Hereby, we differentiate between the waveform described by GR and the extra terms from the Brans-Dicke scalar field, namely $h^{ij} = h^{ij}_{(0)} + h^{ij}_{\rm BD}$ with $\alpha^q h^{ij}_{(q)} \to 0$. The solution $h^{ij}_{(0)}$ can be found in WW96, and we calculate $h^{ij}_{\rm BD}$. In Sec. III A, we find the instantaneous metric and stress-energy tensor in the near zone. In Sec. III B, we derive the metric and scalar fields in the radiation zone as sourced by the near zone. Finally in Sec. III C, we calculate the waveform contribution from the Brans-Dicke scalar field.

A. Metric and effective stress-energy tensor in the near zone

To compute the gravitational waveform to 2PN beyond the quadrupole approximation, one must determine the gravitational field in the near zone. The contributions made

to the waveform by the metric in the near zone are twofold. First, they directly act as a source for higher-order contributions to the gravitational stress-energy tensor as $\Lambda^{\mu\nu} \in \tau^{\mu\nu}$. See the first term in Eq. (23). Second, they contribute indirectly by sourcing the metric at a generic point in the radiation zone, which in turn is used in the second part of Eq. (23). This can be understood as the light cone being modified by the gravitational field, changing the propagation of the gravitational wave.

For a many-body system, we define its characteristic size $\mathcal{S} = \max\{r_{AB} = |\boldsymbol{x}_A - \boldsymbol{x}_B|, \ \forall \ A, B\}$ as being the largest distance between any pair of point particles. The radius $\mathcal{R}$ of the near zone is defined to match the scale below which the screening mechanism is effective and where we assume that GR applies exactly. We will also refer to it as the screening radius. Since any object in the system lies well within the screening radius, $\mathcal{R} \gg \mathcal{S}$, the retardation of the fields is negligible at short scale and time derivatives can be treated as a higher-order term as in the standard PN approximation. In other words, the metric can be found in terms of instantaneous potentials. We denote the constant spacelike hypersurface at time $u = t - R$ as $\mathcal{M}$, with R the distance between the source and the observer, and it is bounded by a sphere of radius $\mathcal{R}$. We have assumed that the near zone is fully screened and, hence, that the metric field equations in the screened region are equivalent to those of GR (see Sec. II). We skip the calculation of the metric and reproduce here the results of Sec. III of WW96. The gravitational field components are given by

$$h^{00} = 4\left(U + \frac{1}{2}\partial_t^2\mathcal{X} - P + 2U^2\right) + \mathcal{O}(\epsilon^{5/2}), \tag{31a}$$

$$h^{0i} = 4U_i + \mathcal{O}(\epsilon^{5/2}), \tag{31b}$$

$$h^{ij} = 4P_{ij} + \mathcal{O}(\epsilon^{5/2}), \tag{31c}$$

where we recall that $\epsilon = v^2/c^2$ and for the background metric

$$g^{00} = -(1 + 2U + \partial_t^2\mathcal{X} + 2U^2) + \mathcal{O}(\epsilon^{5/2}), \tag{32a}$$

$$g^{0i} = -4U_i + \mathcal{O}(\epsilon^{5/2}), \tag{32b}$$

$$g^{ij} = (1 - 2U - \partial_t^2\mathcal{X})\delta^{ij} + \mathcal{O}(\epsilon^2), \tag{32c}$$

from which the determinant is found to be

$$(-g) = 1 + 4\left(U + \frac{1}{2}\partial_t^2\mathcal{X}\right) - 8(P - U^2) + \mathcal{O}(\epsilon^{5/2}). \tag{33}$$

The instantaneous potentials are defined as

$$U(u,\boldsymbol{x}) \equiv \int_{\mathcal{M}} \frac{\mathrm{d}^3\boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|}(T^{00} + T^{ii})(u,\boldsymbol{x}'), \tag{34a}$$

$$\mathcal{X}(u,\boldsymbol{x}) \equiv \int_{\mathcal{M}} \mathrm{d}^3\boldsymbol{x}'|\boldsymbol{x} - \boldsymbol{x}'|(T^{00} + T^{ii})(u,\boldsymbol{x}'), \tag{34b}$$

$$U_i(u,\boldsymbol{x}) \equiv \int_{\mathcal{M}} \frac{\mathrm{d}^3\boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|}T^{0i}(u,\boldsymbol{x}'), \tag{34c}$$

$$P_{ij}(u,\boldsymbol{x}) \equiv \int_{\mathcal{M}} \frac{\mathrm{d}^3\boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|}\left[T^{ij} + \frac{1}{4\pi}\left(U_{,i}U_{,j} - \frac{1}{2}\delta_{ij}U_{,k}U_{,k}\right)\right] \times (u,\boldsymbol{x}'). \tag{34d}$$

With the metric in the near zone found to $\mathcal{O}(\epsilon^{5/2})$, the stress-energy tensor can also be found to $\mathcal{O}(\epsilon^{5/2})$, which will be required to determine the metric in the radiation zone. The stress-energy tensor is given by

$$T^{00} = \sum_A m_A\left[1 - U + \frac{1}{2}\partial_t^2\mathcal{X} + \frac{1}{2}v_A^2 + \frac{1}{2}U^2 + \frac{3}{2}Uv_A^2 + 4P + \frac{3}{8}v_A^4 - 4U_i v_A^i + \mathcal{O}(\epsilon^{5/2})\right]\delta^3(\boldsymbol{x} - \boldsymbol{x}_A(t)), \tag{35a}$$

$$T^{ij} = \sum_A m_A v_A^i v_A^j\left[1 - U + \frac{1}{2}v_A^2 + \mathcal{O}(\epsilon^2)\right]\delta^3(\boldsymbol{x} - \boldsymbol{x}_A(t)), \tag{35b}$$

$$T^{0j} = \sum_A m_A v_A^j\left[1 - U - \frac{1}{2}\partial_t^2\mathcal{X} + \frac{1}{2}v_A^2 + \mathcal{O}(\epsilon^{5/2})\right] \times \delta^3(\boldsymbol{x} - \boldsymbol{x}_A(t)). \tag{35c}$$

We will only need the source $\tau^{\mu\nu}$ for the calculations performed here to the orders $\mathcal{O}(\rho\epsilon)$ for τ^{00}, $\mathcal{O}(\rho\epsilon)$ for τ^{ij}, and $\mathcal{O}(\rho\epsilon^{3/2})$ for τ^{0i}, where $\rho \sim \mathcal{O}(\epsilon)$. With the background stress-energy tensor determined by Eq. (16), we are left with finding $\Lambda^{\mu\nu}$ to the appropriate orders. The noncompact part of the effective stress-energy tensor in Eq. (16) with the near-zone metric of Eqs. (31) and (32) at the background level is given by

$$\Lambda^{00} = 14\left(U + \frac{1}{2}\partial_t^2\mathcal{X}\right)_{,k}^2 + 16\left[-U\ddot{U} + U_{,k}\dot{U}_k - 2U_k\dot{U}_{,k} + \frac{5}{8}\dot{U}^2 + \frac{1}{2}U_{m,k}(U_{m,k} + 3U_{k,m}) + 2P_{,k}U_{,k} - P_{kl}U_{,kl} - \frac{7}{2}UU_{,k}^2\right] + \mathcal{O}(\rho\epsilon^3), \tag{36a}$$

$$\Lambda^{0i} = 16\left[(U + \partial_t^2\mathcal{X})_{,k}(U_{k,i} - U_{i,k}) + \frac{3}{4}\dot{U}(U + \partial_t^2\mathcal{X})_{,i}\right] + \mathcal{O}(\rho\epsilon^{5/2}), \tag{36b}$$

$$\Lambda^{ij} = 4\Big[(U+\partial_t^2\mathcal{X})_{,i}(U+\partial_t^2\mathcal{X})_{,j} - \frac{1}{2}\delta_{ij}(U+\partial_t^2\mathcal{X})_{,k}(U+\partial_t^2\mathcal{X})_{,k}\Big] + 16[2U_{,(i}\dot{U}_{j)} - U_{k,i}U_{k,j} - U_{i,k}U_{j,k} + 2 - U_{k,(i}U_{j),k} - \delta_{ij}(3/2\dot{U}^2 + U_{,k}\dot{U}_k - U_{m,k}U_{[m,k]}] + \mathcal{O}(\rho\epsilon^{5/2}). \tag{36c}$$

We do not need Λ_s in this region as we have assumed that it vanishes due to screening. In Sec. III B, we will also need τ_s evaluated in the near zone. Neglecting the effect of the near-zone scalar field, i.e., $\varphi \to 1$, the scalar source term substantially simplifies to

$$\tau_s = -\frac{\sqrt{-g}}{(3+2\omega_0)}\left(T - 2\frac{\partial T}{\partial\phi}\right) = \frac{1}{(3+2\omega_0)}\sum_A m_A\left(1 - 2\delta_C s_A - \frac{1}{2}v_A^2 + U\right) \times \delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t)). \tag{37}$$

B. Metric in the radiation zone

At leading order, there is no contribution to the gravitational waveform from the radiation zone. This is due to the leading order GW being sourced by the compact matter component, which lies solely in the near zone. However, as we calculate the gravitational wave to higher orders, the noncompact stress-energy components, Λ^{ij} and Λ_s^{ij} in Eq. (15), contribute to the waveform through the scattering of gravitational waves on the background. As such, we need to compute the metric in the radiation zone as sourced by the near zone, which we will denote by $h^{\mu\nu}_{\mathcal{N}}$ in line with Refs. [49,72]. To find $h^{\mu\nu}_{\mathcal{N}}$, we compute the retarded integral

$$h^{\mu\nu}_{\mathcal{N}}(t',\boldsymbol{x}') = 4\int_{\mathcal{N}'}\frac{\tau^{\mu\nu}(t'-|\boldsymbol{x}'-\boldsymbol{x}''|,\boldsymbol{x}'')}{|\boldsymbol{x}'-\boldsymbol{x}''|}\mathrm{d}^3\boldsymbol{x}''. \tag{38}$$

Since the integration is performed over the screened region, one should expect that the results that we thereby obtain mimic those of GR. Note that the domain $\mathcal{N}'$ of the integral is slightly different from the near zone $\mathcal{N}$ as it is part of the past light cone of another field point, x'^μ instead of x^μ. The main difference resides in the modification of the time component in the near zone, which becomes $x'^0 = t' - |\boldsymbol{x}'-\boldsymbol{x}''|$ and $\boldsymbol{x}'' \in \mathcal{N}'$. We define the distance and the direction of the arbitrary field point x'^μ as R' and $\hat{N}'^i = x'^i/R'$, respectively. The metric at the detector is then found by taking $R' = R$, and the detector lies in the direction $\hat{N}$ from the source.

In the screened region, the source $\tau^{\mu\nu}$ takes the same form as in GR, and hence the metric solution in the radiation zone is the same as in WW96. As we will see, deviations from GR still arise from the radiation zone. We omit the details of such a calculation and restrict the discussion here to an outline of how this computation is performed. The retarded integral (38) is evaluated over the domain $\mathcal{N}'$ and so it can be expressed as a sum of moments. This expansion is valid as long as the distance to the field point, R', is much greater than the characteristic size of the system, $\mathcal{S}$, leading to an expansion in R'^{-1}. As such,

$$h^{\mu\nu}(t,\boldsymbol{x}) = 4\sum_{q=0}^{\infty}\frac{(-1)^q}{q!}\left(\frac{1}{R'}M^{\mu\nu k_1\cdots k_q}\right)_{,k_1\cdots k_q}, \tag{39}$$

where

$$M^{\mu\nu k_1\cdots k_q} = \int_{\mathcal{M}'}\tau^{\mu\nu}x'^{k_1}\cdots x'^{k_q}\mathrm{d}^3x'. \tag{40}$$

We evaluate these moments across a constant time hypersurface $\mathcal{M}'$ at the retarded time $u = t' - R'$. Since the integral is still performed over the near zone, the effective stress-energy tensor takes its GR form and Λ_s does not contribute.

To 2PN beyond the quadrupole formula only the compact part of $\tau^{\mu\nu}$ contributes, where higher-order compact contributions arise through inserting the instantaneous potentials into the determinant $-g$ in Eq. (15). This renders these integrals relatively simple because of the Dirac δ-functions in the stress-energy tensor. Furthermore, many of the higher-order moments can be written in terms of the lower moments together with the momentum currents

$$\mathcal{J}^{iQ} = \epsilon^{iab}\int_{\mathcal{M}}\tau^{0b}x^{aQ}\mathrm{d}^3x, \tag{41}$$

where ϵ^{iab} is the Levi-Civita symbol. Following the derivations in WW96 [Eq. (5.5)] and Lang14 [Eq. (4.33)], we find

$$h^{00}_{\mathcal{N}} = 4\frac{\tilde{m}}{R'} + 7\left(\frac{m}{R'}\right)^2 + 2\left(\frac{M^{ij}}{R'}\right)_{,ij} - \frac{2}{3}\left(\frac{M^{ijk}}{R'}\right)_{,ijk}, \tag{42a}$$

$$h^{0i}_{\mathcal{N}} = -2\left(\frac{\dot{M}^{ij} - \epsilon^{ija}\mathcal{J}^a}{R'}\right)_{,j} + \frac{2}{3}\left(\frac{\dot{M}^{ijk} - 2\epsilon^{ika}\mathcal{J}^{aj}}{R'}\right)_{,jk}, \tag{42b}$$

$$h^{ij}_{\mathcal{N}} = \left(\frac{m}{R'}\right)^2\hat{N}'^{ij} + 2\frac{\ddot{M}^{ij}}{R'} - \frac{2}{3}\left(\frac{\ddot{M}^{ijk} - 4\epsilon^{(i|ka}\dot{\mathcal{J}}^{a|j)}}{R'}\right)_{,k}, \tag{42c}$$

where $\hat{N}'^{ij} = \hat{N}'^i\hat{N}'^j$, all moments are evaluated on the constant $u' - R'$ hypersurface $\mathcal{M}'$, $\tilde{m} = \sum_A m_A + \frac{1}{2}m_A v_A^2 - \frac{1}{2}m_A U_A$, and U_A is the total gravitational potential

at the point $\boldsymbol{x}_A$ neglecting the infinite self-energy contribution of body A.

We are left with finding the scalar field in the radiation zone, $\varphi_{\mathcal{N}}$, to 2PN. The integral (22) evaluated over the near zone can also be expanded into the sum of moments,

$$\varphi_{\mathcal{N}}(t,\boldsymbol{x}) = 2\sum_{q=0}^{\infty}\frac{(-1)^q}{q!}\left(\frac{1}{R}M_s^{k_1\ldots k_q}\right)_{,k_1\ldots k_q} \tag{43}$$

with

$$M_s^{k_1\ldots k_q} = \int_{\mathcal{M}}\tau_s x'^{k_1}\ldots x'^{k_q}\mathrm{d}^3x', \tag{44a}$$

where τ_s is given by Eq. (20). The calculation is otherwise equivalent to that of the metric expansion. The $\mathcal{O}(\rho\epsilon^{3/2})$ moments are found to be

$$M_s = \frac{1}{(3+2\omega)}\sum_A m_A\left(1-2\delta_C s_A-\frac{1}{2}v_A^2+U_A\right), \tag{44b}$$

$$M_s^i = \frac{1}{(3+2\omega)}\sum_A m_A x_A^i\left(1-2\delta_C s_A-\frac{1}{2}v_A^2+U_A\right), \tag{44c}$$

$$M_s^{ij} = \frac{1}{(3+2\omega)}\sum_A m_A x_A^{ij}(1-2\delta_C s_A), \tag{44d}$$

$$M_s^{ijk} = \frac{1}{(3+2\omega)}\sum_A m_A x_A^{ijk}(1-2\delta_C s_A). \tag{44e}$$

We will not need moments larger than the octopole. Note that these differ from the results of Lang14 [Eq. (6.11)] since the scalar part of the stress-energy tensor $\Lambda_s^{\mu\nu}$ is modified due to the screening effect. Because of this modification, when evaluated in the center of mass frame corrected to 2PN, the scalar dipole moment vanishes, unlike in the Brans-Dicke case. We will see in Sec. III C 2 that this has a major effect on the waveform as many terms present in Brans-Dicke theory no longer contribute. The scalar field in the radiation zone is then determined by inserting these moments into Eq. (43),

$$\varphi_{\mathcal{N}}(t,\boldsymbol{x}') = 2\frac{M_s}{R'}-2\left(\frac{M_s^i}{R'}\right)_{,i}+\left(\frac{M_s^{ij}}{R'}\right)_{,ij}-\frac{1}{3}\left(\frac{M_s^{ijk}}{R'}\right)_{,ijk}. \tag{45}$$

C. Finding the waveform

With the metric in the near and radiation zones at hand, we are fully equipped to compute the gravitational waveform. As we are only interested in the metric waves, we focus our attention on the retarded integral of the gravitational field in Eq. (21) to 2PN beyond the quadrupole approximation. Moreover, the calculation applies to a field point where $R \gg \mathcal{R} \gg \mathcal{S}$. As a result, we may discard terms that decay faster than $1/R$ and will refer to this region of spacetime as the far zone to distinguish it from the radiation zone.

In Sec. III C 1, we outline how the near zone contributes to the gravitational wave. No further calculation is needed as the source in the near zone is equivalent to that of GR and we can use the results of WW96. In Sec. III C 2, we discuss the contribution from the radiation zone. As the metric found from Λ^{ij} again takes a form equivalent to GR, we refer to WW96 for the bulk of the calculation. However, unlike in GR, the wave is also sourced by Λ_s^{ij}, which leads to new terms.

1. Waveform contribution from the near zone

The calculation of the gravitational waveform to 2PN beyond the quadrupole at the detector is a continuation of the calculation performed in Sec. III B. The main difference is that now we only need to keep terms in the metric expansion that decay as R^{-1} since terms decreasing faster are negligible if the observer lies sufficiently far from the source. Furthermore, as we are evaluating the waveform and not the whole metric, we must remember that only the spatial part of the metric solution is needed and perform a transverse and traceless projection. This projection moreover cancels terms proportional to either δ_{ij} or the directional vector to the detector $\boldsymbol{N}$. The calculation of the gravitational waveform due to the near zone is analogous to WW96, and so we refer to Ref. [50] for the details, only providing here a brief overview of the derivation.

The moments in the expansion (39) can be expanded in R^{-1}, and since we are calculating the metric in the far zone, only the leading order is required. The expansion can then be expressed in terms of *Epstein-Wagoner* moments [95] by employing the conservation of the effective stress-energy tensor. This simplifies Eq. (39) to

$$h_{\mathcal{N}}^{ij} = \frac{2}{R}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\sum_{m=0}^{\infty}\hat{N}_{k_i}\cdots\hat{N}_{k_m}I_{\mathrm{EW}}^{ijk_1\cdots k_m}, \tag{46}$$

where

$$I_{\mathrm{EW}}^{ij} = \int_{\mathcal{M}}\tau^{00}x^i x^j\mathrm{d}^3\boldsymbol{x}+I_{\mathrm{EW\,(surf)}}^{ij}, \tag{47a}$$

$$I_{\mathrm{EW}}^{ijk} = \int_{\mathcal{M}}(2\tau^{0(i}x^{j)}x^k-\tau^{0k}x^i x^j)\mathrm{d}^3\boldsymbol{x}+I_{\mathrm{EW\,(surf)}}^{ijk}, \tag{47b}$$

$$I_{\mathrm{EW}}^{ijk_1\cdots k_m} = \frac{2}{m!}\frac{\mathrm{d}^{m-2}}{\mathrm{d}t^{m-2}}\int_{\mathcal{M}}\tau^{ij}x^{k_1}\cdots x^{k_m}\mathrm{d}^3\boldsymbol{x}\quad(m\geq 2), \tag{47c}$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} I^{ij}_{\mathrm{EW\,(surf)}} = \oint_{\partial\mathcal{M}} (4\tau^{l(i}x^{j)} - (\tau^{kl}x^i x^j)_{,k})\mathcal{R}^2 \hat{n}^l \mathrm{d}^2\Omega, \quad (48a)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} I^{ijk}_{\mathrm{EW\,(surf)}} = \oint_{\partial\mathcal{M}} (2\tau^{l(i}x^{j)}x^k - \tau^{kl}x^i x^j)\mathcal{R}^2 \hat{n}^l \mathrm{d}^2\Omega \quad (48b)$$

with $\hat{n}$ denoting a radially outward-directed unit vector.

The explicit calculation of these moments is an involved task, making up the bulk of both the WW96 and Lang14 articles. Because of the similarity between the calculation performed in WW96 and here, we will not replicate it. In brief, τ^{ij} is expressed in terms of the instantaneous potentials using the results in Eqs. (35) and (36). The potentials are themselves found using the matter distribution in Eq. (6). Finally, the Epstein-Wagoner moments are inserted in the expansion (46).

The solutions for the Epstein-Wagoner moments are given in Eqs. (6.6) of WW96. However, in their calculation of the gravitational waveform, terms from the near-zone and the radiation-zone contributions that depend on $\mathcal{R}$ cancel out. There is no guarantee for this in our model, and so we highlight the terms proportional to $\mathcal{R}$,

$$I^{ij}_{\mathrm{EW}} \supset -\frac{14}{5} m\mathcal{R}\ddot{M}^{ij} = -\frac{28}{5} m\mu\mathcal{R}(v^{ij} - x^{(i}a^{j)}), \quad (49a)$$

$$I^{ijkl}_{\mathrm{EW}} \supset -\frac{8}{35} m\mathcal{R}\ddot{M}^{ij}\delta^{kl} = -\frac{8}{35} m\mu\mathcal{R}(v^{ij} - x^{(i}a^{j)})\delta^{kl}, \quad (49b)$$

$$I^{ijklmn}_{\mathrm{EW}} \supset -\frac{2}{315} m\mathcal{R}\ddot{M}^{ij}\delta^{(kl}\delta^{mn)} = -\frac{2}{315} m\mu\mathcal{R}(v^{ij} - x^{(i}a^{j)})\delta^{(kl}\delta^{mn)}. \quad (49c)$$

These terms depend on the size of the near zone $\mathcal{R}$, which we have set to the screening radius. In GR these terms were not expected to appear in the final waveform as the value of $\mathcal{R}$ is unphysical. Their cancellations with the boundary terms of the radiation-zone integration is checked in the following.

2. Waveform contribution from the radiation zone

When determining the metric in the radiation zone in Sec. III B, we found that the metric sourced by the near zone remains equivalent to that in GR. Deviations in the metric occur due to the self-interaction in the radiation zone. The modification of the gravitational theory in the radiation zone by Brans-Dicke gravity only begins to affect the metric of this region at $\mathcal{O}(\epsilon^2)$.

We first want to know if the extra terms (49a)–(49c) remain in the gravitational waveform or are canceled by contributions from the radiation zone. The relevant part of Λ^{ij} in the radiation zone up to $\mathcal{O}(\epsilon^3)$ is found to be

$$\Lambda^{ij} \supset -h^{00}_{(1)}\ddot{h}^{ij}_{(2)} + \frac{1}{4}h^{00}_{(1),(i}h^{00}_{(2),j)} + \frac{1}{2}h^{00}_{(1),(i}h^{kk}_{(2),j)} + 2h^{00,(i}_{(1)}\dot{h}^{j)0}_{(2)}, \quad (50)$$

where the subscripts denote the PN orders used. We may use the results in Eq. (5.8) of WW96 for the metric components h^{ij} to 2PN beyond the quadrupole order,

$$h^{ij} \supset \frac{4m}{R}\int_0^\infty \mathrm{d}s\,\partial_t^4 M^{ij}(u-s)\left[\ln\left(\frac{s}{2R+s}\right) + \frac{11}{12}\right]. \quad (51)$$

This is the first tail term arising from the interaction of the gravitational wave with the background Newtonian potential and is of order 1.5PN. Therefore, only one more iteration is required to reach the desired precision of 2PN. The components of the gravitational stress-energy tensor that we need to consider are

$$\Lambda^{ij} \supset -h^{00}_{(1)}\ddot{h}^{ij}_{(1.5)} + \frac{1}{2}h^{00}_{(1),i}h^{00}_{(1.5),j} + \frac{1}{2}h^{00}_{(1),(i}h^{kk}_{(1.5),j)}. \quad (52)$$

As can be seen with Eq. (46), the first term, the source is proportional to δ_{ij} and so is not transverse-traceless. For the second and third terms, the monopole-monopole couplings scale as $R^{-1}{}_{,(i}R^{-1}{}_{,j)}$, and when the derivatives are evaluated, they have an angular dependence of the form $\hat{N}^{ij}$, which is not transverse-traceless. Hence, only the terms from monopole-quadrupole and monopole-current quadrupole are relevant in Λ^{ij}. The metric components due to the near zone again mimic those of GR, leading to the following result:

$$h^{ij} \supset \frac{4m}{3R}\hat{n}^k\int_0^\infty \mathrm{d}s\,\partial_t^5 M^{ijk}(u-s)\left[\ln\left(\frac{s}{2R+s}\right) + \frac{97}{60}\right] - \frac{16m}{3R}\epsilon^{(i|ka}\hat{n}^k\int_0^\infty \mathrm{d}s\,\partial_t^4 J^{a|j)}(u-s)\left[\ln\left(\frac{s}{2R+s}\right) + \frac{7}{6}\right] + \frac{1912}{315}\frac{m}{R}\partial_t^4 M^{ij}(u)\mathcal{R}, \quad (53)$$

which is the remainder of Eq. (5.8) in WW96. We can check that the extra contributions linear in $\mathcal{R}$ from the boundary exactly cancel the contributions from Eq. (49). This means that to 2PN, the radius $\mathcal{R}$ does not enter the waveform as in GR.

The remaining contributions arise from Λ^{ij}_s in Eq. (18). The calculation is similar to that performed in Sec. VI C of Lang14, but with differing prefactors. In Sec. III B, we found that the spatial part of the gravitational field h^{ij} scales as R^{-2} to leading order, leaving no contribution at the quadrupole order to the gravitational waveform in the far zone. For the next order, we expand Λ_s to 0.5PN beyond the quadrupole to find

$$\tau^{ij}_{\rm BD} = \frac{\pi}{16}(3+2\omega)\varphi^{,(i}_{\rm mono}\varphi^{,j)}_{\rm di} + \mathcal{O}(\epsilon^2). \tag{54}$$

We denote by $\varphi_{\rm mono}$, $\varphi_{\rm di}$, $\varphi_{\rm quad}$, and $\varphi_{\rm oct}$, the monopole, dipole, quadrupole, and octopole contributions to φ in Eq. (45), respectively. The subscript BD refers to the corrections from the Brans-Dicke field. Since the dipole term vanishes in the center of mass frame, Λ_s does not contribute at this order. At 1PN, we find

$$\tau^{ij}_{\rm BD} = \frac{\pi}{16}(3+2\omega)(\varphi^{,(i}_{\rm mono}\varphi^{,j)}_{\rm quad} + \varphi^{,(i}_{\rm mono}\varphi^{,j)}_{\rm mono}) + \mathcal{O}(\epsilon^{5/2}), \tag{55}$$

where the cross-term monopoles are evaluated at $\mathcal{O}(\epsilon)$. The expansion in terms of the moments of the scalar field is analogous to that in Eq. (6.29) of Lang14 and leads to the 1.5PN contribution to the waveform

$$P^{1.5}h^{ij}_{\rm BD} = \frac{4m_s}{R}\left(-\frac{1}{12}\partial_t^3 M_s^{ij}\right), \tag{56}$$

where P^k stands for the kth PN order beyond the quadrupole formula and $m_s = \sum m_A(1-2\delta_C s_A)$. This term enters at 1.5PN beyond the quadrupole and is an extra contribution to the waveform arising from the scalar quadrupole. Interestingly, in pure Brans-Dicke theory there is a new tail term that arises from the dipole-dipole interactions, which, however, does not appear in our screened theory as in the near zone we are sourced by the GR solution for the metric.

Finally, to find the contribution to the metric attributed to Λ_s at 2PN beyond the quadrupole, we need to consider only the monopole-octopole terms as all other contributions at this order vanish when taking the transverse-traceless projection,

$$\tau^{ij}_{\rm BD} \supset \frac{\pi}{16}(3+2\omega)\varphi^{,(i}_{\rm mono}\varphi^{,j)}_{\rm oct} + \mathcal{O}(\epsilon^3), \tag{57}$$

where $\varphi_{\rm oct}$ is the octopole term. Calculating the relevant contributions, we are left with the final 2PN term of the waveform,

$$P^2 h^{ij}_{\rm BD} = \frac{4m_s}{R}\left(-\frac{1}{60}\partial_t^4 M_s^{ija}\hat{N}^a\right), \tag{58}$$

which enters at 2PN beyond the quadrupole. Again, we lose the additional tail terms found in Brans-Dicke theory due to the vanishing scalar dipole moment. Therefore, the waveform up to 2PN beyond the quadrupole formula corresponds to that of GR with the additional contributions at 1.5PN and 2PN from $\Lambda_s^{\mu\nu}$ of the form

$$h^{ij}_{\rm BD} = \frac{4m_s}{R}\left(-\frac{1}{12}\partial_t^3 M_s^{ij}\right) + \frac{4m_s}{R}\left(-\frac{1}{60}\partial_t^4 M_s^{ija}\hat{N}^a\right) + \mathcal{O}(P^{2.5}h^{ij}_{\rm BD}). \tag{59}$$

IV. LEADING CORRECTIONS TO THE WAVEFORM FROM THE NEAR-ZONE SCALAR FIELD

We have now established that the effect of a scalar field on the gravitational waveform from a fully screened source enters at 1.5PN beyond the quadrupole formula and is due to its radiation-zone contribution. In the following, we shall compute the leading correction of the near-zone scalar field on the waveform of a screened source. In a Brans-Dicke model with no screening, the near-zone scalar field affects the waveform formula at 1PN [49]. We expect the screened scalar field to also impact the system at 1PN, but this effect should be smaller than in the unscreened model, attenuated by the scaling parameter α^q. Thus, α^q is expected to suppress the new near-zone contributions to the waveform formula. In this work, we shall neglect higher-order contributions $\mathcal{O}(\alpha^{2q}\psi^2)$, and we limit the post-Newtonian calculations to 1PN. In Sec. IV A, we find the scalar contributions to the instantaneous metric and stress-energy tensors in the near zone. Then in Sec. IV B, we show that the radiation-zone self-interactions do not affect the gravitational waveform, and we find the scalar corrections to the Epstein-Wagoner moments.

A. Corrections to the metric and effective stress-energy tensors in the near zone

Similar to the fully screened case in Sec. III, we first need to find the near-zone metric $g_{\mu\nu}$ and gravitational field $h^{\mu\nu}$ up to 1PN. These tensor fields are described by GR with the added corrections, $\alpha^q g^{\mu\nu}_{(q)}$ and $\alpha^q h^{\mu\nu}_{(q)}$, respectively. From there, we may calculate the leading α correction to the stress-energy tensors $\tau^{\mu\nu}$ and τ_s at the required PN order.

Starting with the metric, the 1PN formalism in the screened regime was introduced in Ref. [41] and later extended to Horndeski theories with luminal propagation of gravitational waves in Ref. [42]. One can use these results to write the correction to the metric field up to 1PN,

$$\alpha^q g^{(q)}_{00} = \alpha^q(\psi + \tilde{\Phi}_1 - 3\mathcal{A}_\psi - \mathcal{B}_\psi + 6\tilde{\Phi}_2 + \Phi_{Sc}) + \mathcal{O}(\epsilon^3), \tag{60a}$$

$$\alpha^q g^{(q)}_{0i} = \alpha^q \gamma_{0i} + \mathcal{O}(\epsilon^{5/2}), \tag{60b}$$

$$\alpha^q g^{(q)}_{ij} = (-\alpha^q\psi)\delta_{ij} + \mathcal{O}(\epsilon^2). \tag{60c}$$

The post-Newtonian retarded potentials involving the scalar field are of the form

$$\tilde{\Phi}_1(\boldsymbol{x}) = -\frac{1}{4\pi}\int \frac{\psi(\boldsymbol{x}')v(\boldsymbol{x}')^2}{|\boldsymbol{x}-\boldsymbol{x}'|^3}\mathrm{d}^3\boldsymbol{x}',$$
$$\mathcal{A}_\psi(\boldsymbol{x}) = -\frac{1}{4\pi}\int \frac{\psi(\boldsymbol{x}')[\boldsymbol{v}(\boldsymbol{x}')\cdot(\boldsymbol{x}-\boldsymbol{x}')]^2}{|\boldsymbol{x}-\boldsymbol{x}'|^5}\mathrm{d}^3\boldsymbol{x}', \tag{61}$$

$$\tilde{\Phi}_2(\boldsymbol{x}) = \int \frac{\rho(\boldsymbol{x}')\psi(\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|}\mathrm{d}^3\boldsymbol{x}',$$

$$\mathcal{B}_\psi(\boldsymbol{x}) = -\frac{1}{4\pi}\int \frac{\psi(\boldsymbol{x}')[\boldsymbol{a}(\boldsymbol{x}')\cdot(\boldsymbol{x}-\boldsymbol{x}')]}{|\boldsymbol{x}-\boldsymbol{x}'|^3}\mathrm{d}^3\boldsymbol{x}', \quad (62)$$

$$\boldsymbol{\nabla}^2\gamma_{0i} = -1/2\partial_i\dot{\psi}, \quad (63)$$

$$\boldsymbol{\nabla}^2\Phi_{Sc} = -\delta_C\frac{V_0}{2\phi_0^{-n_V}}\psi^{n_V} + \delta_K\frac{\phi_0^4}{2}\Gamma_2(\phi_0)(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi)^2 - \delta_V\phi_0^3\Gamma_3(\phi_0)\boldsymbol{\nabla}^2\psi(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi), \quad (64)$$

where $\boldsymbol{\nabla}$ is the three-gradient, $\boldsymbol{\nabla}^2$ is the three-Laplacian, and the α-order of the scalar field $\psi \sim \psi^{(q)}$ is kept implicit throughout this work.

Notice that the scalar field ψ appears abundantly in the 1PN correction to the metric, but we have not yet described its behavior in the near zone, which we shall resolve now. We recall the effective scalar field equation in the screened region as exposed in Eq. (28),

$$8\pi\left(T - 2\frac{\partial T}{\partial\phi}\right) = \delta_C n_V\frac{\psi^{n_V-1}}{\phi_0^{n_V-1}}V_0 + 2\delta_K\phi_0^3\Gamma_2(\phi_0)(\tilde{X}\Box\psi + \nabla_\mu\tilde{X}\nabla^\mu\psi) + \delta_V\phi_0^2\Gamma_3(\phi_0)((\Box\psi)^2 - \nabla_\mu\nabla_\nu\psi\nabla^\mu\nabla^\nu\psi), \quad (65)$$

where $\tilde{X} \equiv -g^{\mu\nu}_{(0)}\partial_\mu\psi\partial_\nu\psi/2$ and all contractions or covariant derivatives are made using the GR metric $g^{(0)}_{\mu\nu}$. At first sight, the PN order of the scalar field appears to be inconsistent among the three different terms on the right-hand side of Eq. (65). This prevents us from considering a superposition of screening mechanisms. Since the leading PN order of the trace of the stress-energy tensor and its derivative is $\mathcal{O}(\epsilon)$, the leading PN order of the scalar dynamical part ψ depends on the screening mechanism at play. As discussed in Refs. [42,91], the scaling parameter α actually carries a PN order, so that the full scalar field term $\alpha^q\psi$ has the leading order $\mathcal{O}(\epsilon)$. The PN order of the scalar field $\mathcal{O}(\epsilon^k)$ is determined through Eq. (65), which in turn determines the PN order of $\alpha^q \sim \mathcal{O}(\epsilon^{(1-k)})$. From now on, we refer to the leading PN contribution of the scalar field as ψ_k, such that $\psi = \psi_k + \mathcal{O}(\epsilon^{k+1})$. The field equation describing the behavior of ψ_k then reads

$$8\pi\sum_A m_A(1-2\delta_C s_A)\delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t)) = -\delta_C n_V\phi_0^{n_V-1}\psi_k^{n_V-1}V_0 - \delta_K\phi_0^3\Gamma_2(\phi_0)((\boldsymbol{\nabla}\psi_k\cdot\boldsymbol{\nabla}\psi_k)\boldsymbol{\nabla}^2\psi_k + \boldsymbol{\nabla}(\boldsymbol{\nabla}\psi_k\cdot\boldsymbol{\nabla}\psi_k)\cdot\boldsymbol{\nabla}\psi_k) - \delta_V\phi_0^2\Gamma_3(\phi_0)((\boldsymbol{\nabla}^2\psi_k)^2 - \nabla_i\nabla_j\psi_k\nabla^i\nabla^j\psi_k). \quad (66)$$

Since $n_V < 1$, the field equation for theories exhibiting the chameleon mechanism is ill-defined. This is due to the fact that we considered the matter content to be described by massive pointlike objects. In chameleon screening, compact objects with high energy densities are screened against the ambient scalar field. In the pointlike approximation, we assume vacuum everywhere, while the density of the body follows a Dirac delta distribution, which is incompatible with the screening mechanism at play. Because GWs that are presently detected originate from very compact objects, one can simply assume that the scalar degree of freedom is fully screened by the very high internal density $\psi_k \propto \rho^{1-n_V} \sim 0$, and we recover the case developed in Sec. III.

We now have all the ingredients to proceed with the calculation of the α correction to the gravitational field $h^{(q)\mu\nu}$ as well as the effective stress-energy tensors $\tau^{(q)\mu\nu}$ and $\tau_s^{(q)}$ in the near zone. Starting with the former, we can use Eq. (12) to deduce its correction term. However, at 1PN only the $0i$-components are affected,

$$h^{(q)00} = h^{(q)ij} = 0 + \mathcal{O}(\alpha^{-q}\epsilon^2), \quad (67a)$$

$$h^{(q)0i} = -\gamma_{0i}^{(q)} + \mathcal{O}(\alpha^{-q}\epsilon^{5/2}). \quad (67b)$$

Note that we included α^{-q} in the neglected PN orders since the scale parameter carries a PN order itself. Once we take the full correction term $\alpha^q h^{(q)ij}$, we recover the usual PN orders. The effective stress-energy tensors $\tau^{(q)\mu\nu}$ and τ_s^q are then found by inserting Eqs. (60) and (67) into Eqs. (16) and (20). For the tensor $\Lambda^{\mu\nu}$, no contribution from the scalar field is present at 1PN. Furthermore, corrections to $\Lambda_s^{\mu\nu}$ are of $\mathcal{O}(\alpha^{2q}\psi^2)$ and we neglect them. The compact part of the effective stress-energy tensor in terms of the scalar field and gravitational potentials can be computed at 1PN using Eq. (6), which gives

$$T^{(q)00} = \sum_A m_A\delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t))\frac{5}{2}\psi_k + \mathcal{O}(\alpha^{-q}\rho\epsilon^2), \quad (68a)$$

$$T^{(q)0i} = T^{(q)ij} = 0 + \mathcal{O}(\alpha^{-q}\rho\epsilon^{3/2}). \quad (68b)$$

We notice that only the compact part of the stress-energy tensor is modified at 1PN, which allows for a straightforward integration. The 1PN contribution of the near-zone scalar field to the effective stress-energy tensors can therefore be obtained using Eqs. (15) and (20):

$$\tau^{(q)00} = -\sum_A m_A\delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t))\frac{\psi_k}{2} + \mathcal{O}(\alpha^{-q}\rho\epsilon^2), \quad (69a)$$

$$\tau^{(q)0i} = 0 + \mathcal{O}(\alpha^{-q}\rho\epsilon^{3/2}), \quad (69b)$$

$$\tau^{(q)ij} = 0 + \mathcal{O}(\alpha^{-q}\rho\epsilon^2), \quad (69c)$$

and

$$\tau_s^{(q)} = \frac{3}{2(3+2\omega_0)}\sum_A m_A(1-2\delta_C s_A)\delta^3(\boldsymbol{x}-\boldsymbol{x}_A(t))\psi_k + \mathcal{O}(\alpha^{-q}\rho\epsilon^2). \quad (70)$$

B. Metric and scalar field corrections in the radiation zone and contribution to the waveform

The metric in the radiation zone is found using a sum of moments as defined in Eq. (39). The leading corrections from the near-zone scalar field can be calculated from the stress-energy tensor $\tau^{(q)}_{\mu\nu}$ and, at 1PN, only the 00-component of the gravitational field is affected. The Dirac δ-function in the compact part of the stress-energy tensor ensures that we can easily integrate the moments and find

$$h^{(q)00}_{\mathcal{N}}(t,\boldsymbol{x}') = -\frac{2}{R'}\sum_A m_A \psi_k(\boldsymbol{x}_A(t)) + \mathcal{O}(\alpha^{-q}\rho\epsilon^2). \quad (71)$$

We recall that R' is the distance between an arbitrary field point x'^μ in the radiation zone and the source. When taking the transverse-traceless gauge, this correction term vanishes and therefore does not contribute to the gravitational waveform.

One follows a similar strategy to find the corrected scalar field in the radiation zone. Its moment expansion is defined in Eq. (43) and at 1PN, we find

$$\varphi^{(q)}_{\mathcal{N}} = \frac{3}{2(3+2\omega_0)R'}\sum_A m_A(1-2\delta_C s_A)\psi_k(\boldsymbol{x}_A(t)) + \mathcal{O}(\alpha^{-q}\rho\epsilon^2). \quad (72)$$

The corrected radiation-zone scalar field is important to take into account the scalar wave. However, this correction only affects the gravitational wave beyond 1PN, and we do not need to compute its contribution to the gravitational waveform.

As we are restricting ourselves to the leading corrections at 1PN beyond the quadrupole formula, the radiation-zone contributions can be neglected. Therefore, we focus our attention to the near-zone effects and the Epstein-Wagoner moments. Only the 00-component of the stress-energy tensor is corrected at this order, and this means that the near-zone scalar field does not significantly affect the moments beyond the quadrupole. Furthermore, the contributing part of $\tau^{(q)}_{\mu\nu}$ is compact and easy to integrate over the near zone. The result is simply

$$I^{(q)ij}_{\mathrm{EW}} = -\frac{1}{2}\sum_A x^i_A x^j_A m_A \psi_k(\boldsymbol{x}_A(t)). \quad (73)$$

This term is to be added to the usual GR result of the Epstein-Wagoner moments and additional terms in the waveform emerge from a modification of the equations of motion of the source as will become clear in Sec. V.

V. BLACK HOLE BINARY WAVEFORM

In Secs. III and IV, we have derived the deviation from the GR waveform for the N-body problem. However, the targeted sources of the LIGO-Virgo Collaboration are the late inspirals of binary objects, composed of black holes and/or neutron stars. In the following, we shall restrict our discussion to the case of black hole–black hole binaries on quasicircular orbits. The quasicircular assumption is well motivated by the continuing loss of angular momentum circularizing the orbits [96] for the late inspiral.

In Sec. V A, we treat the case of the fully screened source, which is described by GR, and give the corrections to 2PN beyond the quadrupole approximation to the waveform from a Brans-Dicke scalar field in the radiation zone. In Sec. V B, we find the effect of the near-zone scalar field on the equations of motion of the two-body problem. In Sec. V C, we give the contributions up to 1PN beyond the quadrupole formula of the screened near-zone scalar field to the gravitational waveform. Finally, in Sec. V D, we calculate the order of magnitude of the correction in the fully screened case as well as in the case with an effective screened scalar field for the cubic Galileon model [97].

A. Fully screened behavior at 2PN

Since the scalar field in the near zone is first assumed to be fully screened with no effect on the binary system, we can use the equations of the two-body system in GR to describe the behavior of the source to 2PN (see Appendix or Ref. [31]). We work in the center of mass coordinates that are valid to second PN order. We define the usual two body variables: the total mass $m = m_1 + m_2$, mass difference $\delta m = m_1 - m_2$, reduced mass $\mu = m_1 m_2/m$, symmetric mass ratio $\eta = \mu/m$, relative position $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$, relative distance $r = |\boldsymbol{x}|$, unit vector $\hat{\boldsymbol{n}} = \boldsymbol{x}/r$, relative velocity $\boldsymbol{v} = \dot{\boldsymbol{x}}$, and direction to the observer $\hat{\boldsymbol{N}}$. The GR equations of motion to second PN order are presented in Eq. (A4a).

With these quantities in the center of mass frame and the equations of motion for the two-body system, we are now equipped to find the new contributions to the GW in terms of the Newtonian parameters. We recall that the sensitivity of black holes is $1/2$ and, hence, the scalar field in the chameleon case leaves no contribution to the gravitational waveform. Up to 2PN beyond the quadrupole approximation, the gravitational waveform becomes

$$h^{ij} = h^{ij}_{\mathrm{GR}} + h^{ij}_{\mathrm{BD}} + \mathcal{O}(\alpha^q\psi), \quad (74)$$

$$h^{ij}_{\mathrm{GR}} = \frac{\mu}{2R}(Q^{ij}_{\mathrm{GR}} + P^{0.5}Q^{ij}_{\mathrm{GR}} + P^1 Q^{ij}_{\mathrm{GR}} + P^{1.5}Q^{ij}_{\mathrm{GR}} + P^2 Q^{ij}_{\mathrm{GR}} + \cdots), \quad (75)$$

$$h^{ij}_{\mathrm{BD}} = (1-\delta_C)\frac{\mu}{2R}(P^{1.5}Q^{ij}_{\mathrm{BD}} + P^2 Q^{ij}_{\mathrm{BD}} + \cdots), \quad (76)$$

where

$$P^{1.5}Q^{ij}_{\mathrm{BD}} = \frac{8m^2}{3(3+2\omega_0)r^2}[2\hat{n}^{(i}v^{j)} - \dot{r}\hat{n}^i\hat{n}^j], \quad (77)$$

$$P^2 Q^{ij}_{\rm BD} = \frac{2m\delta m}{15(3+2\omega_0)r^2}\left(-12(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})v^i v^j - 24(\hat{\boldsymbol{N}}\cdot\boldsymbol{v})\hat{n}^{(i}v^{j)} + \frac{16}{r}\dot{r}(\hat{\boldsymbol{N}}\cdot\boldsymbol{v})\hat{n}^{(i}x^{j)}\right.$$
$$+\frac{12m}{r^2}(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})\hat{n}^{(i}x^{j)} + \frac{16\dot{r}}{r}(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})v^{(i}x^{j)} - \frac{16}{r}(\hat{\boldsymbol{N}}\cdot\boldsymbol{v})v^{(i}x^{j)} - \frac{2m}{r^3}(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})x^i x^j$$
$$-\frac{15\dot{r}^2}{r^2}(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})x^i x^j + \frac{3v^2}{r^2}(\hat{\boldsymbol{N}}\cdot\hat{\boldsymbol{n}})x^i x^j + \frac{6\dot{r}}{r^2}(\hat{\boldsymbol{N}}\cdot\boldsymbol{v})x^i x^j + \frac{6m}{r^2}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})\hat{n}^i\hat{n}^j$$
$$+\frac{16\dot{r}}{r}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})\hat{n}^{(i}v^{j)} - \frac{8}{r}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})v^i v^j - \frac{4m}{r^3}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})\hat{n}^{(i}x^{j)} - \frac{30\dot{r}^2}{r^2}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})\hat{n}^{(i}x^{j)}$$
$$\left.+\frac{6v^2}{r^2}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})\hat{n}^{(i}x^{j)} + \frac{12\dot{r}}{r^2}(\hat{\boldsymbol{N}}\cdot\boldsymbol{x})v^{(i}x^{j)}\right). \tag{78}$$

In Eq. (76), the factor $(1-\delta_C)$ ensures that only the k-mouflage and Vainshtein cases are taken into account. We note that these contributions to the gravitational waveform originate from the scattering of the GW with the radiation-zone scalar field, which is itself sourced by the matter source in the near zone. The waveform $h^{ij}_{\rm GR}$ can be found in WW96 (Sec. VI), where the gravitational constant should be modified as $G \to G/\phi_0$.

B. Two-body problem and the effect of the near-zone scalar field

The screened scalar field effectively modifies the metric at Newtonian and 1PN order; hence, it is expected to affect the motion of the two black holes at these orders as well. Since we restrict the calculations of the scalar contribution to 1PN, we do not need higher PN orders of the two-body problem. Following the strategy in WW96 and Lang14, we need to find the relative acceleration $\boldsymbol{a}$ between the two black holes using their equations of motion. Furthermore, using the effective Lagrangian that recovers the equations of motion, we find the positions $\boldsymbol{x}_j$ and velocities $\boldsymbol{v}_j$ of black holes $j=1$ and $j=2$ with respect to the relative position $\boldsymbol{x}$ and velocity $\boldsymbol{v}$ in the center of mass frame.

Starting with the relative acceleration, it can be found using the geodesic equation for both black holes and taking the difference between the two. Mathematically this gives

$$\boldsymbol{a} = \ddot{\boldsymbol{x}}_1 - \ddot{\boldsymbol{x}}_2, \tag{79}$$

$$\ddot{x}^i_{1,2} = -\Gamma^i_{\alpha\beta}\dot{x}^\alpha_{1,2}\dot{x}^\beta_{1,2} + \Gamma^0_{\alpha\beta}\dot{x}^\alpha_{1,2}\dot{x}^\beta_{1,2}\dot{x}^i_{1,2}, \tag{80}$$

where $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols. We recall that the dot notation represents the derivative with respect to coordinate time t, such that the geodesic equation is expressed using coordinate time, rather than proper time, thus taking a different form than the usual geodesic equation. Rewriting the equations of motion at Newtonian and PN order leads to the result

$$\boldsymbol{a} = \boldsymbol{a}^{(0)}_{\rm N} + \alpha^q \boldsymbol{a}^{(q)}_{\rm N} + \boldsymbol{a}^{(0)}_{\rm PN} + \alpha^q \boldsymbol{a}^{(q)}_{\rm PN}, \tag{81}$$

$$\boldsymbol{a}^{(q)}_{\rm N} = -\frac{1}{2}(\psi'_1(r) + \psi'_2(r))\hat{\boldsymbol{n}}, \tag{82}$$

$$\boldsymbol{a}^{(q)}_{\rm PN} = \frac{1}{r^2}(m_2\psi_1 + m_1\psi_2 + 3m_1\psi_1 + 3m_2\psi_2)\hat{\boldsymbol{n}} + \dot{r}(\boldsymbol{\gamma}'_1 - \boldsymbol{\gamma}'_2) + (\Phi_{Sc,1}{}' + \Phi_{Sc,2}{}')\hat{\boldsymbol{n}}$$
$$+\frac{1}{r}(m_2\psi'_1 + m_1\psi'_2 - 3m_1\psi'_1 - 3m_2\psi'_2)\hat{\boldsymbol{n}} + (m_2^2\psi'_1 + m_1^2\psi'_2)\left(\frac{v^2}{2m^2}\hat{\boldsymbol{n}} - \frac{1}{2m^2}\dot{r}\boldsymbol{v}\right)$$
$$-\frac{1+\mu}{2m}(m_2\psi'_1 + m_1\psi'_2)\dot{r}\boldsymbol{v} + \frac{m_2}{2m}(\dot{r}\boldsymbol{\gamma}'_1 - (\boldsymbol{v}\cdot\boldsymbol{\gamma}'_1)\hat{\boldsymbol{n}}) - \frac{m_1}{2m}(\dot{r}\boldsymbol{\gamma}'_2 - (\boldsymbol{v}\cdot\boldsymbol{\gamma}'_2)\hat{\boldsymbol{n}})$$
$$-\frac{1}{m^2}(m_2^2\psi'_1 + m_1^2\psi'_2)\dot{r}\boldsymbol{v} - \frac{1}{2\pi r^5}\left(m_1\psi_1 + m_2\psi_2 - \frac{r}{4}m_1\psi'_1 + \frac{r}{4}m_2\psi'_2\right)\hat{\boldsymbol{n}}$$
$$+\frac{1}{8m^2\pi r^4}[(m_2^2\psi_1 + m_1^2\psi_2)(6\dot{r}\boldsymbol{v} + 9\dot{r}^2\hat{\boldsymbol{n}}) - 3(m_2^2\psi'_1 + m_1^2\psi'_2)r\dot{r}^2\hat{\boldsymbol{n}}$$
$$+16\pi(m_1^2\psi_2 + m_2^2\psi_1)v^2\hat{\boldsymbol{n}} - 4\pi(m_2^2\psi'_1 + m_1^2\psi'_2)rv^2\hat{\boldsymbol{n}}], \tag{83}$$

and $\boldsymbol{a}^{(0)}_{\rm N}$ and $\boldsymbol{a}^{(0)}_{\rm PN}$ are the Newtonian and post-Newtonian accelerations from GR and can be found in Appendix. Primes denote derivatives with respect to r. When calculating the Christoffel symbols, we made use of the Newtonian relations $\boldsymbol{a}_1 = -m_1/r^2\hat{\boldsymbol{n}}$ and $\boldsymbol{a}_2 = m_2/r^2\hat{\boldsymbol{n}}$ as well as

$$\frac{\partial \boldsymbol{v}_1}{\partial x_2^j} = \frac{1}{2r}\boldsymbol{v}_1\hat{n}^j \quad \text{and} \quad \frac{\partial \boldsymbol{v}_2}{\partial x_1^j} = -\frac{1}{2r}\boldsymbol{v}_2\hat{n}^j \tag{84}$$

from $|\boldsymbol{v}_i| \propto 1/\sqrt{r}$.

To find the velocities $\boldsymbol{v}_j$, $j = 1, 2$ with respect to the relative velocity, we can make use of the vanishing total momentum $\boldsymbol{P} = 0$ in the center of mass frame. The latter is derived from the total Lagrangian that is deduced using the equations of motion (80). Since we are only interested in the waveform at 1PN, it is sufficient to find the velocities at Newtonian order. The Lagrangian at this order can be written as

$$\mathcal{L}_{\rm N} \equiv \mathcal{L}_{\rm N}^{(0)} + \alpha^q \mathcal{L}_{\rm N}^{(q)} = \frac{m_1}{2}\boldsymbol{v}_1^2 + \frac{m_2}{2}\boldsymbol{v}_2^2 + \frac{m_1 m_2}{r} - \alpha^q \frac{1}{2}(\psi_1(r) + \psi_2(r)), \tag{85}$$

where $\mathcal{L}_{\rm N}^{(q)}$ can easily be found using Eq. (82). Therefore, since we take the derivative of the Lagrangian with respect to the velocities to find the momentum and $\mathcal{L}_{\rm N}^{(q)}$ only depends on the position of both objects, we recover Newtonian gravity, namely

$$\boldsymbol{P}_{\rm N} \equiv \frac{\partial \mathcal{L}_{\rm N}}{\partial \boldsymbol{v}_1} + \frac{\partial \mathcal{L}_{\rm N}}{\partial \boldsymbol{v}_2} = \boldsymbol{P}_{\rm N}^{(0)}. \tag{86}$$

Hence, the velocities with respect to the relative velocity at the required order simply become

$$\boldsymbol{v}_1 = \frac{m_2}{m}\boldsymbol{v} \quad \text{and} \quad \boldsymbol{v}_2 = -\frac{m_1}{m}\boldsymbol{v}. \tag{87}$$

The individual positions $\boldsymbol{x}_j$ with respect to the relative position are found by integrating the velocities (87) and we simply find

$$\boldsymbol{x}_1 = \frac{m_2}{m}\boldsymbol{x} \quad \text{and} \quad \boldsymbol{x}_2 = -\frac{m_1}{m}\boldsymbol{x}. \tag{88}$$

C. Near-zone corrections at 1PN beyond quadrupole formula

The equations of motion for the binary system are already modified at Newtonian order through $\boldsymbol{a}_{\rm N}^{(q)} \neq 0$ in Eq. (83). Therefore, we expect the quadrupole formula at 0PN to be affected. Nonetheless, the effect of this extra scalar force is suppressed by the small parameter α^q with respect to the GR contribution. Using the correction term (73) to the Epstein-Wagoner quadrupole moment and the results from Sec. V B, we can find the gravitational waveform up to 1PN beyond the quadrupole approximation of a binary black hole inspiral affected by a screened scalar field. The part of the waveform $h_{\rm GR}^{ij}$ described by GR can be found in WW96 (Sec. VI), where the gravitational constant is replaced by $G \to G/\phi_0$ and $h_{\rm BD}^{ij}$ enters only at higher PN order. We calculate the extra scalar contributions $\alpha^q h^{(q)ij}$ for which the result to 1PN is

$$h^{ij} = h_{\rm GR}^{ij} + \alpha^q h^{(q)ij} + \mathcal{O}(\alpha^{2q}\psi^2), \tag{89}$$

$$h^{(q)ij} = \frac{\mu}{2R}[Q^{(q)ij} + P^{0.5}Q^{(q)ij} + P^1 Q^{(q)ij} + \cdots], \tag{90}$$

where

$$Q^{(q)ij} = -r(\psi_1' + \psi_2')\hat{n}^i\hat{n}^j, \tag{91}$$

$$P^{0.5}Q^{(q)ij} = \frac{\delta m r}{2m}[(\psi_1' + \psi_2')((\boldsymbol{v}\cdot\hat{\boldsymbol{N}} - \dot{r}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})\hat{n}^i\hat{n}^j - 6(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})v^{(i}\hat{n}^{j)}) + (\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})r\dot{r}(\psi_1'' + \psi_2'')\hat{n}^i\hat{n}^j], \tag{92}$$

$$\begin{aligned} P^1 Q^{(q)ij} = {} & 2r a_{\rm PN}^{(q)(i}\hat{n}^{j)} + \frac{\hat{n}^i\hat{n}^j}{12}[(16 - 42\eta + (2 - 6\eta)(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})^2)m(\psi_1' + \psi_2') - 6(1 - 3\eta)r^2\dot{r}^2(\psi_1'' + \psi_2'') \\ & + 6(1 - 3\eta)(\psi_1' + \psi_2')(r\dot{r}(\hat{\boldsymbol{N}}\cdot\boldsymbol{v}) - 2rv^2 + 4(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})m) \\ & + \frac{6}{m}((m + r(\dot{r}^2 - v^2))(m_1\psi_1' + m_2\psi_2') - r^2\dot{r}^2(m_1\psi_1'' + m_2\psi_2''))] \\ & + r\hat{n}^{(i}v^{j)}[(3\eta - 1)(\psi_1' + \psi_2')(2\dot{r} + (\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})(4\hat{\boldsymbol{N}}\cdot\boldsymbol{v} - \dot{r}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})) \\ & - \frac{2\dot{r}}{m}(m_1\psi_1' + m_2\psi_2') + (3\eta - 1)(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})r\dot{r}(\psi_1'' + \psi_2'')] \\ & - \frac{v^i v^j}{2m}[2(m_1\psi_1 + m_2\psi_2) + 4(1 - 3\eta)mr(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{N}})^2(\psi_1' + \psi_2')]. \end{aligned} \tag{93}$$

The corrections at 0PN and 0.5PN are only due to the modification of the equations of motion of the binary system at Newtonian order, while the scalar contributions at 1PN originate from the extra scalar force at 0PN and 1PN as well as from the modification of the Epstein-Wagoner quadrupole moment.

D. Example: Galileon cosmology

With the leading corrections to GR from a scalar field in the fully screened case and in the case where near-zone effects are considered at hand, we are ready to give an order of magnitude estimate of those. As mentioned earlier, the near-zone scalar field affects the waveform at the quadrupole order, but the modification is suppressed with α^q. Some of the most frequently studied models exhibiting the Vainshtein mechanism are the Galileons [97]. We focus on the cubic Galileon model described by the action

$$S[\phi, g] = \frac{M_p^2}{2}\int \mathrm{d}^4x\left(\phi R + 2\frac{\omega}{\phi}X - \frac{\alpha}{\phi^2}X\Box\phi\right), \quad (94)$$

where α and ω are real parameters. To have an impact on late-time cosmology and contribute to cosmic acceleration (cf. [34]), a possibility is to have $\alpha \sim H_0^{-2}$, when[1] $\omega \sim -10$ [98]. This is the same order of magnitude as in Ref. [71], where the cubic Galileon associated with the Dvali-Gabadadze-Porrati (DGP) model [99,100] is considered. We therefore use these orders of magnitude for our comparison. In the following, we adopt units such that $G/\phi_0 = c = \hbar = 1$.

We start by calculating the near-zone contribution of the scalar field. Following the scaling method [39], we find that the unique value for q in the case of the Vainshtein mechanism is $q = -1/2$ and, hence, the damping parameter is of the order $\alpha^q \sim H_0$. The field equation for the PN scalar field ψ is found using Eq. (66) and is written as

$$4\pi\sum_{A=1}^{2} m_A\delta^3(\boldsymbol{x}-\boldsymbol{x}_A) = ((\boldsymbol{\nabla}^2\psi)^2 - \nabla_i\nabla_j\psi\nabla^i\nabla^j\psi). \quad (95)$$

To solve this equation for the binary case, we follow the strategy presented in Ref. [101] and we define $\psi(\boldsymbol{x}) = \psi_1(\boldsymbol{x}) + \psi_2(\boldsymbol{x}) + \psi_\Delta(\boldsymbol{x})$, where $\psi_{1,2}$ is the scalar field generated by the body 1,2, respectively, and ψ_Δ includes the nonlinearities. $\psi'_{1,2}(r)$ are obtained from Eq. (95) assuming spherical symmetry, and we solve them for each body separately. Then, we find $\psi'_\Delta(r)$ by substituting the solutions $\psi'_{1,2}(r)$ into Eq. (95), and we assume cylindrical symmetry around the axis perpendicular to the orbital plane for the whole system. We require the solution to vanish asymptotically, setting an integration constant to zero, and get

$$\psi'(r) = \sqrt{\frac{m_1+m_2}{2r}}, \quad (96)$$

for which the PN order is $\mathcal{O}(\epsilon^{1/2})$, which is what we expect from Eq. (95). The contribution of this extra degree of freedom on the gravitational waveform is given by Eq. (91) as

$$\begin{aligned}\alpha^q Q^{(q)ij} &= -\alpha^q r(\psi'(\boldsymbol{x}_1) + \psi'(\boldsymbol{x}_2))\hat{n}^i\hat{n}^j \\ &= -\alpha^q m\sqrt{\frac{2r}{\mu}}\hat{n}^i\hat{n}^j, \quad \forall\ r < \mathcal{R}. \end{aligned} \quad (97)$$

We recall that α^q also carries a PN order, in this case $\mathcal{O}(\epsilon^{1/2})$, so that the overall PN order of the correction to the quadrupole formula is the same as that of GR.

The leading contribution from the radiation zone in the fully screened case can be found directly from the waveform formula (77). We consider quasicircular orbits, for which the relative velocity $\boldsymbol{v}$ is perpendicular to the relative position $\boldsymbol{x}$ at Newtonian order and $\dot{r} \simeq 0$. Furthermore, using Newtonian approximation, the norm of the relative velocity is $v \approx \sqrt{\mu/r}$. The correction to the gravitational waveform is then given by

$$P^{1.5}Q^{ij}_{\mathrm{BD}} = \frac{16}{3(3+2\omega)}\frac{m^2}{r^2}\frac{\mu}{r}\hat{n}^{\langle i}\hat{v}^{j\rangle}, \quad (98)$$

where $\hat{\boldsymbol{v}} \equiv \boldsymbol{v}/\|\boldsymbol{v}\|$ is the unit velocity vector. This correction is comparable to the 1.5PN order terms of GR. In comparison, the GR quadrupole of the gravitational waveform may be approximated as

$$Q^{ij}_{\mathrm{GR}} = 2\left(v^{\langle i}v^{j\rangle} - \frac{m}{r}\hat{n}^i\hat{n}^j\right) \simeq -2\frac{m}{r}\hat{n}^i\hat{n}^j. \quad (99)$$

The relative amplitude of each contribution, Eqs. (97) and (98), may be compared to Eq. (99) by neglecting factors of order unity. We obtain

$$|\alpha^q Q^{(q)ij}| = \frac{\alpha^q}{\sqrt{\mu}}r^{3/2}|Q^{ij}_{\mathrm{GR}}|, \quad (100)$$

$$|P^{1.5}Q^{ij}_{\mathrm{BD}}| = \frac{1}{3+2\omega}\frac{m\mu}{r^2}|Q^{ij}_{\mathrm{GR}}|. \quad (101)$$

By comparing $\alpha^q Q^{(q)ij}$ and Q^{ij}_{GR}, one can deduce that for a binary system with higher masses and higher orbital period (i.e., higher r), the effect of the screened scalar field is maximized compared to the GR quadrupole formula. The scalar contribution is also strengthened when the difference between the two masses is high, minimizing μ in the denominator of Eq. (100). In contrast, the effect of the Brans-Dicke scalar field in $P^{1.5}Q^{ij}_{\mathrm{BD}}$ is amplified compared to that of GR for smaller and heavier systems, but attenuated for asymmetric masses.

Hence, we consider two different systems for the comparison: the binary black holes that produced the first detected GW [43] and the inspiral of a stellar black hole with $m_1 = 10m_\odot$ around Sagittarius A* with $m_2 = 4\times10^6 m_\odot$, as may be observed in the close future by LISA [102,103]. The maximal amplitudes for GW150914 are

[1]The cubic interaction gives an additional contribution to the kinetic term, avoiding ghost instabilities [98].

$$\alpha^q Q^{(q)ij} \sim 4 \times 10^{-20} Q^{ij}_{\rm GR}, \tag{102}$$

$$P^{1.5} Q^{ij}_{\rm BD} \sim 3 \times 10^{-3} Q^{ij}_{\rm GR} \tag{103}$$

for the frequency range 35–200 Hz. In particular, Brans-Dicke effects are stronger at higher frequencies, while near-zone contributions are amplified at lower ones. In the case of a stellar black hole around Sagittarius A*, the observable frequency range of LISA is expected to be between 0.1 mHz and 1 Hz, which leads to

$$\alpha^q Q^{(q)ij} \sim 5 \times 10^{-12} Q^{ij}_{\rm GR}, \tag{104}$$

$$P^{1.5} Q^{ij}_{\rm BD} \sim 5 \times 10^{-9} Q^{ij}_{\rm GR}. \tag{105}$$

In the second case, the radiation-zone contribution is weakened due to the mass asymmetry, but in the case of a supermassive black hole merger with objects of comparable masses, the Brans-Dicke effects would be of the order $(10^{-3} - 10^{-2}) Q^{ij}_{\rm GR}$. The long-range effect of the scalar field has therefore a stronger impact on the gravitational waveform with the 1.5PN term lying about 1 order of magnitude below LISA sensitivity [104]. The near-zone correction on the waveform of the screened scalar field can be considered negligible in all cases of interest.

VI. CONCLUSIONS

In this work, we have investigated the impact of screening mechanisms on the GW waveform from compact sources. Our results are explicitly applicable to the three most prominent screening mechanisms: the Vainshtein, k-mouflage, and chameleon effects. For our calculations we employed the modified field equations in their relaxed form and solved them with retarded integrals over the manifold. We first considered the case where the system is fully screened, i.e., described by GR in the near zone and Brans-Dicke theory in the radiation zone. The radius distinguishing these zones is identified as the screening radius of the underlying gravity theory. It was found that the boundary terms exactly cancel between the near and radiation zones and the screening radius does not enter the waveform. The deviations from the GR waveform that this model induces appear at 1.5PN beyond the quadrupole formula.

Further, we found the leading corrections in the near zone to the fully screened results induced by the scalar field. In this case, we derived the leading corrections in the stress-energy tensor modifying the Epstein-Wagoner moments at 1PN. The near-zone metric was also found to be modified at Newtonian and post-Newtonian orders, leading to 0PN, 0.5PN, and 1PN corrections beyond the quadrupole formula from the scalar force. Although these contributions have low PN order, they are suppressed by the nature of screening, which consequently makes these deviations subdominant to the Brans-Dicke corrections found at 1.5PN from the unscreened radiation zone. Importantly, in Chameleon models the waveform is unaffected for a black hole source by the scalar field, but the effect is present for other compact objects, such as neutron stars.

Finally, we applied our results to a Galileon cosmology toy model and estimated the amplitude of the deviation from GR for a black hole binary. We found that in the most optimistic scenarios the waveform amplitude, and hence luminosity distance, deviates from its GR counterpart by one part in 10^{11} from the modification of the near-zone metric by the screened scalar field and one part in 10^2 from the radiation-zone Brans-Dicke field. While the near-zone corrections are not realistically measurable in the foreseeable future, radiation-zone contributions might be observable with future detectors. In particular, LISA and the Einstein telescope are forecasted to have an order 10^{-2} error on the luminosity distance [104,105]. Hence, the effect of a screened Horndeski theory on the waveform may be detectable at the 1-σ level in the near future. The predicted radiation-zone contributions depend on the Brans-Dicke parameter, and we expect future GW measurements to set new constraints on this parameter.

While the effect of screening on the waveform is small, it may still play an important role in multiband gravitational wave detection [106]. The LISA mission will detect binaries several years before merger [102,103,107]. Consequently, as the binary evolves, it exits the LISA sensitivity band only to enter ground-based detectors sensitivity bands a few years to decades later. The cumulative effect of the modified waveform energy emission would accumulate over this time causing a dephasing of the screened signal and one predicted by GR. Further analysis on the energy loss due to the modified GW and the suppressed scalar wave would allow one to predict a different phase and time of merger in comparison to GR, a potential test of any given screening mechanism despite the difference in the signal itself being too small to detect directly in ground-based detectors [108]. In addition, recent relativistic numerical analysis has shown that kinetic screening might not play a role in the strong field regime leaving the scalar contributions as strong as in Brans-Dicke theories [109,110].

In the analysis presented, we made several simplifying assumptions. First, we omitted the contribution of the scalar waves on the detectors, which may play a significant role. Even though such waves are screened in the emission and detection regions weakening their effects, future work should be devoted to their study. We made the standard assumption that the orbit was circularized for late inspiral calculations. However, even if the system were elliptical, one would expect that the modification to the waveform caused by dropping this assumption remains small since the system still lies within the screened region. In this work, we restricted our final calculations to the case of black hole

binaries, and one may wonder how the results would change for neutron star binaries or mixed systems. For the k-mouflage and Vainshtein mechanisms, one can argue that the field equation governing the behavior of the near-zone scalar field satisfies Gauss's law, thus the composition of the objects does not affect their motion, only their mass does. This is not the case for the chameleon mechanism, where the sensitivity to the scalar field varies with the compactness. For theories with such screening, one would need to include near-zone effects of the scalar field in the case of neutron stars, but we expect these contributions to be strongly negligible compared to the Brans-Dicke corrections from the radiation zone. Finally, we have not considered the contributions originating from the intrinsic spin of each object. In GR, these contributions modify the Epstein-Wagoner moments and enter in the waveform formula at 1PN beyond quadrupole order [50]. Therefore, we expect that, in the fully screened scenario treated in Sec. III, only GR contributions affect the waveform (see Appendix F of WW96). When calculating the leading corrections from the near-zone scalar field in Sec. IV, one could expect new spin contributions involving the scalar field entering at 1PN, but these terms would also be strongly suppressed by the screening mechanism.

In conclusion, this work supports the use of GW emission processes abiding by GR in the study of screened modified gravity theories. Nonetheless, we raised the possibility to probe the effect of a screened scalar field on GWs with the sensitivity of our future detectors. Even though we have only investigated cubic Galileon interactions as a specific case, this study offers an efficient method to check for a wide range of Horndeski theories whether scalar field effects can be neglected in the GW emissions of the respective models.

ACKNOWLEDGMENTS

The authors thank Jonathan Gair and Jorge Peñarrubia for useful discussions at an early stage of this work in 2016/17. C. R. was supported by the University of Grenoble-Alpes. R. M. acknowledges support from the STFC Consolidated Grant No. ST/T000732/1. C. D. and L. L. were supported by a Swiss National Science Foundation (SNSF) Professorship Grant (Nos. 170547 and 202671). L. L. also acknowledges the support by a SNSF Advanced Postdoc.Mobility Fellowship (No. 161058) during an early stage of this work. Please contact the authors for access to research materials.

APPENDIX: TWO-BODY PROBLEM IN GENERAL RELATIVITY

For completeness we shall provide here the details of the calculation of the gravitational waveform from a binary system with masses $m_{1,2}$ and positions $\boldsymbol{x}_{1,2}$ conducted in Sec. V A. In performing this calculation we change coordinates to place the center of mass of the system at the origin. This is a common trick in classical physics where the mass of each body is only its rest mass. However, we consider PN corrections to the system, and so the respective mass of an object is also changed by its potential and kinetic energy. Using the corrected mass changes the center of mass, and so its definition changes accordingly.

The center of mass is found to be

$$\boldsymbol{X} = m^{-1}(m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2) + \boldsymbol{f}^{(1)} + \mathcal{O}(\epsilon^2), \tag{A1}$$

$$\boldsymbol{f}^{(1)} = -\frac{1}{2}\eta\frac{\delta m}{m}\left(v^2 - \frac{m}{r}\right)(\boldsymbol{x}_1 - \boldsymbol{x}_2), \tag{A2}$$

where $m = m_1 + m_2$, $\mu = m_1 m_2/m$, $\eta = \mu/m$, $\delta m = m_1 - m_2$, and $r = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$. One may worry that as we are calculating the metric to $\mathcal{O}(\epsilon^4)$, we may need the center of mass up to $\mathcal{O}(\epsilon^2)$, but the only place the correction enters is in the leading-order quadrupole term, where it cancels out exactly.

We wish to describe the system in terms of the relative separation $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ and its derivatives. In the center of mass coordinate system, $\boldsymbol{X} = 0$, we find

$$\boldsymbol{x}_1 = (m_2/m)\boldsymbol{x} - \boldsymbol{f}^{(1)} + \mathcal{O}(\epsilon^2), \tag{A3a}$$

$$\boldsymbol{x}_2 = -(m_1/m)\boldsymbol{x} - \boldsymbol{f}^{(1)} + \mathcal{O}(\epsilon^2). \tag{A3b}$$

It will be useful to define the system in terms of Newtonian objects in the final expression. So we define the Newtonian angular momentum, $\boldsymbol{L} = \mu\boldsymbol{x} \times \boldsymbol{v}$, and unit normal to the orbital plane $\hat{\boldsymbol{\lambda}} = \boldsymbol{L} \times \hat{\boldsymbol{n}}$.

When computing the time derivatives of the center of mass coordinates, we will make use of the equations of motion for the black hole binary. As the near zone is described by GR, the equations of motion can be taken readily from sources such as Ref. [96]. The accelerations due to the gravitational forces to first PN order are found to be

$$\boldsymbol{a} = \boldsymbol{a}_{\mathrm{N}} + \boldsymbol{a}_{\mathrm{PN}} + \mathcal{O}(\epsilon^2), \tag{A4a}$$

$$\boldsymbol{a}_{\mathrm{N}} = -\frac{m}{r^2}\hat{\boldsymbol{n}}, \tag{A4b}$$

$$\boldsymbol{a}_{\mathrm{PN}} = -\frac{m}{r^2}(A_{\mathrm{PN}}\hat{\boldsymbol{n}} + B_{\mathrm{PN}}\dot{r}\boldsymbol{v}), \tag{A4c}$$

where for convenience we define

$$A_{\mathrm{PN}} = -2(2+\eta)\frac{m}{r} + (1+3\eta)v^2 - \frac{3}{2}\eta\dot{r}^2, \tag{A5a}$$

$$B_{\mathrm{PN}} = -2(2-\eta). \tag{A5b}$$

Derivatives of the equations of motion are also needed. In order to calculate $\dot{a}^j$ and $\ddot{a}^j$, we make use of the identities

$$\dot{\hat{\boldsymbol{n}}} = \frac{\boldsymbol{v}}{r} - \frac{\dot{r}\boldsymbol{n}}{r}, \tag{A6a}$$

$$\ddot{r} = \ddot{r}_{\rm N} + \ddot{r}_{\rm PN}, \tag{A6b}$$

$$\ddot{r}_{\rm N} = -\frac{m}{r^2} + \frac{v^2}{r} - \frac{\dot{r}^2}{r}, \tag{A6c}$$

$$\ddot{r}_{\rm PN} = -\frac{m}{r^2}\left(\alpha\frac{m}{r} + \beta v^2 + (\gamma+\sigma)\dot{r}^2\right), \tag{A6d}$$

which give $\dot{a}^j$ to the required order as

$$\dot{a}^i = (\dot{a}_{\rm N})^i_{\rm N} + (\dot{a}_{\rm PN})^i_{\rm N}, \tag{A7a}$$

$$(\dot{a}_{\rm N})^i_{\rm N} = \frac{2m\dot{r}}{r^3}\hat{n}^i - \frac{m}{r^3}v^i, \tag{A7b}$$

$$\begin{aligned}(\dot{a}_{\rm PN})^i_{\rm N} = {}& -18\frac{m^2\dot{r}}{r^4}\hat{n}^i + 4\eta\frac{m^2}{r^4}v^i + (3+12\eta)\frac{m\dot{r}v^2}{r^3}\hat{n}^i \\ & - \frac{15}{2}\eta\frac{m\dot{r}^3}{r^3}\hat{n}^i - \left(12 - \frac{15}{2}\eta\right)\frac{m\dot{r}^2}{r^3}v^i \\ & + (3-5\eta)\frac{mv^2}{r^3}v^i,\end{aligned} \tag{A7c}$$

and $\ddot{a}^j$ to the required order as

$$\ddot{a}^i = (\ddot{a}^i_{\rm N})_{\rm N} + (\ddot{a}^i_{\rm N})_{\rm PN} + (\ddot{a}^i_{\rm PN})_{\rm N}, \tag{A8a}$$

$$(\ddot{a}^i_{\rm N})_{\rm N} = -2\frac{m^2}{r^5}\hat{n}^i - 15\frac{m\dot{r}^2}{r^4}\hat{n}^i + 3\frac{mv^2}{r^4}\hat{n}^i + 6\frac{m\dot{r}}{r^4}v^i, \tag{A8b}$$

$$\begin{aligned}(\ddot{a}^i_{\rm N})_{\rm PN} = {}& (8+4\eta)\frac{m^3}{r^6}\hat{n}^i - (2+6\eta)\frac{m^2v^2}{r^5}\hat{n}^i \\ & + (12-3\eta)\frac{m^2\dot{r}^2}{r^5}\hat{n}^i - (4-2\eta)\frac{m^2\dot{r}}{r^5}v^i,\end{aligned} \tag{A8c}$$

$$\begin{aligned}(\ddot{a}^i_{\rm PN})_{\rm N} = {}& -36\eta\frac{\dot{r}m^2}{r^5}v^i - (30-42\eta)\frac{m\dot{r}v^2}{r^4}v^i \\ & - (60-45\eta)\frac{m\dot{r}^3}{r^4}v^i + (126-57\eta)\frac{m^2\dot{r}^2}{r^5}\hat{n}^i \\ & - (15+165/2\eta)\frac{mv^2\dot{r}^2}{r^4}\hat{n}^i + 102/2\eta\frac{m\dot{r}^4}{r^4}\hat{n}^i \\ & + (22+2\eta)\frac{m^3}{r^6}\hat{n}^i + (8-10\eta)\frac{m^2v^2}{r^5}\hat{n}^i \\ & + (3+15\eta)\frac{mv^4}{r^4}\hat{n}^i.\end{aligned} \tag{A8d}$$

[1] C. M. Will, The confrontation between general relativity and experiment, Living Rev. Relativity **17**, 4 (2014).

[2] S. Chatrchyan *et al.* (CMS Collaboration), Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. B **716**, 30 (2012).

[3] G. Aad *et al.* (ATLAS Collaboration), Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716**, 1 (2012).

[4] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020).

[5] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, Cosmic distances calibrated to 1% precision with Gaia EDR3 parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM, Astrophys. J. Lett. **908**, L6 (2021).

[6] K. C. Wong *et al.*, H0LiCOW—XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3σ tension between early- and late-Universe probes, Mon. Not. R. Astron. Soc. **498**, 1420 (2020).

[7] E. Di Valentino, A. Melchiorri, and J. Silk, Planck evidence for a closed Universe and a possible crisis for cosmology, Nat. Astron. **4**, 196 (2020).

[8] W. Handley, Curvature tension: Evidence for a closed universe, Phys. Rev. D **103**, L041301 (2021).

[9] M. Asgari *et al.* (KiDS Collaboration), KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics, Astron. Astrophys. **645**, A104 (2021).

[10] L. Lombriser, Consistency of the local Hubble constant with the cosmic microwave background, Phys. Lett. B **803**, 135303 (2020).

[11] B. Boseand and L. Lombriser, Easing cosmic tensions with an open and hotter universe, Phys. Rev. D **103**, L081304 (2021).

[12] E. Di Valentino, O. Mena, S. Pan, L. Visinelli, W. Yang, A. Melchiorri, D. F. Mota, A. G. Riess, and J. Silk, In the realm of the Hubble tension—a review of solutions, Classical Quantum Gravity **38**, 153001 (2021).

[13] L. Perivolaropoulosand and F. Skara, Challenges for ΛCDM: An update, arXiv:2105.05208.

[14] S. Weinberg, The cosmological constant problem, Rev. Mod. Phys. **61**, 1 (1989).

[15] J. Martin, Everything you always wanted to know about the cosmological constant problem (but were afraid to ask), C. R. Phys. **13**, 566 (2012).
[16] N. Kaloperand and A. Padilla, Sequestering the Standard Model Vacuum Energy, Phys. Rev. Lett. **112**, 091304 (2014).
[17] L. Lombriser, On the cosmological constant problem, Phys. Lett. B **797**, 134804 (2019).
[18] C. Bransand and R. Dicke, Mach's principle and a relativistic theory of gravitation, Phys. Rev. **124**, 925 (1961).
[19] G. W. Horndeski, Second-order scalar-tensor field equations in a four-dimensional space, Int. J. Theor. Phys. **10**, 363 (1974).
[20] C. de Rham, G. Gabadadze, and A. J. Tolley, Resummation of Massive Gravity, Phys. Rev. Lett. **106**, 231101 (2011).
[21] L. Heisenberg, Generalization of the Proca action, J. Cosmol. Astropart. Phys. 05 (2014) 015.
[22] A. Joyce, L. Lombriser, and F. Schmidt, Dark energy versus modified gravity, Annu. Rev. Nucl. Part. Sci. **66**, 95 (2016).
[23] L. Heisenberg, A systematic approach to generalisations of general relativity and their cosmological implications, Phys. Rep. **796**, 1 (2019).
[24] M. Ostrogradsky, Mem. Ac. St. Petersbourg **VI 4**, 385 (1850).
[25] C. de Rham, G. Gabadadze, and A. J. Tolley, Helicity decomposition of ghost-free massive gravity, J. High Energy Phys. 11 (2011) 093.
[26] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, Healthy Theories Beyond Horndeski, Phys. Rev. Lett. **114**, 211101 (2015).
[27] S. Jana, C. Dalang, and L. Lombriser, Horndeski theories and beyond from higher dimensions, Classical Quantum Gravity **38**, 025003 (2021).
[28] I. H. Stairs, Testing general relativity with pulsar timing, Living Rev. Relativity **6**, 5 (2003).
[29] R. Manchester, Pulsars and gravity, Int. J. Mod. Phys. D **24**, 1530018 (2015).
[30] A. M. Archibald, N. V. Gusinskaia, J. W. Hessels, A. T. Deller, D. L. Kaplan, D. R. Lorimer, R. S. Lynch, S. M. Ransom, and I. H. Stairs, Universality of free fall from the orbital motion of a pulsar in a stellar triple system, Nature (London) **559**, 73 (2018).
[31] C. M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, Cambridge, England, 2018).
[32] C. Renevey, Review on tests of general relativity and modified gravity using pulsar timing, arXiv:1905.13720.
[33] T. Baker, A. Barreira, H. Desmond, P. Ferreira, B. Jain, K. Koyama, B. Li, L. Lombriser, A. Nicola, J. Sakstein, and F. Schmidt, Novel probes project: Tests of gravity on astrophysical scales, Rev. Mod. Phys. **93**, 015003 (2021).
[34] L. Lombriserand and N. A. Lima, Challenges to self-acceleration in modified gravity from gravitational waves and large-scale structure, Phys. Lett. B **765**, 382 (2017).
[35] B. Abbott *et al.* (LIGO Scientific, Virgo, Fermi-GBM, INTEGRAL Collaborations), Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A, Astrophys. J. **848**, L13 (2017).
[36] J. Khouryand and A. Weltman, Chameleon cosmology, Phys. Rev. D **69**, 044026 (2004).
[37] E. Babichev, C. Deffayet, and R. Ziour, k-Mouflage gravity, Int. J. Mod. Phys. D **18**, 2147 (2009).
[38] A. Vainshtein, To the problem of nonvanishing gravitation mass, Phys. Lett. B **39**, 393 (1972).
[39] R. McManus, L. Lombriser, and J. Peñarrubia, Finding Horndeski theories with Einstein gravity limits, J. Cosmol. Astropart. Phys. 11 (2016) 006.
[40] L. Lombriser, A parametrisation of modified gravity on nonlinear cosmological scales, J. Cosmol. Astropart. Phys. 11 (2016) 039.
[41] R. McManus, L. Lombriser, and J. Peñarrubia, Parameterised post-Newtonian expansion in screened regions, J. Cosmol. Astropart. Phys. 12 (2017) 031.
[42] C. Renevey, J. Kennedy, and L. Lombriser, Parameterised post-Newtonian formalism for the effective field theory of dark energy via screened reconstructed Horndeski theories, J. Cosmol. Astropart. Phys. 12 (2020) 032.
[43] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116**, 061102 (2016).
[44] K. S. Thorne, The theory of gravitational radiation—an introductory review, in *Gravitational Radiation* (North-Holland Publishing Co., Amsterdam, 1983), pp. 1–57.
[45] M. Maggiore, *Gravitational Waves. Vol. 1. Theory and Experiments* (Oxford University Press, New York, 2007), p. 574.
[46] M. Maggiore, *Gravitational Waves. Vol. 2: Astrophysics and Cosmology* (Oxford University Press, New York, 2018).
[47] S. Foffa, P. Mastrolia, R. Sturani, C. Sturm, and W. J. Torres Bobadilla, Static Two-Body Potential at Fifth Post-Newtonian Order, Phys. Rev. Lett. **122**, 241605 (2019).
[48] J. Blümlein, A. Maier, and P. Marquard, Five-loop static contribution to the gravitational interaction potential of two point masses, Phys. Lett. B **800**, 135100 (2020).
[49] R. N. Lang, Compact binary systems in scalar-tensor gravity. II. Tensor gravitational waves to second post-Newtonian order, Phys. Rev. D **89**, 084014 (2014).
[50] C. M. Willand and A. G. Wiseman, Gravitational radiation from compact binary systems: Gravitational wave forms and energy loss to second postNewtonian order, Phys. Rev. D **54**, 4813 (1996).
[51] L. Lombriserand and A. Taylor, Breaking a dark degeneracy with gravitational waves, J. Cosmol. Astropart. Phys. 03 (2016) 031.
[52] A. Nishizawa, Generalized framework for testing gravity with gravitational-wave propagation. I. Formulation, Phys. Rev. D **97**, 104037 (2018).
[53] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, Gravitational-wave luminosity distance in modified gravity theories, Phys. Rev. D **97**, 104066 (2018).
[54] L. Amendola, I. Sawicki, M. Kunz, and I. D. Saltas, Direct detection of gravitational waves can measure the time variation of the Planck mass, J. Cosmol. Astropart. Phys. 08 (2018) 030.

[55] M. Lagos, M. Fishbach, P. Landry, and D. E. Holz, Standard sirens with a running Planck mass, Phys. Rev. D **99**, 083504 (2019).
[56] E. Belgacem *et al.* (LISA Cosmology Working Group), Testing modified gravity at cosmological distances with LISA standard sirens, J. Cosmol. Astropart. Phys. 07 (2019) 024.
[57] C. Dalangand and L. Lombriser, Limitations on Standard Sirens tests of gravity from screening, J. Cosmol. Astropart. Phys. 10 (2019) 013.
[58] C. Dalang, P. Fleury, and L. Lombriser, Horndeski gravity and standard sirens, Phys. Rev. D **102**, 044036 (2020).
[59] A. Garoffolo, G. Tasinato, C. Carbone, D. Bertacca, and S. Matarrese, Gravitational waves and geometrical optics in scalar-tensor theories, J. Cosmol. Astropart. Phys. 11 (2020) 040.
[60] S. Mastrogiovanni, D. Steer, and M. Barsuglia, Probing modified gravity theories and cosmology using gravitational-waves and associated electromagnetic counterparts, Phys. Rev. D **102**, 044009 (2020).
[61] T. Bakerand and I. Harrison, Constraining scalar-tensor modified gravity with gravitational waves and large scale structure surveys, J. Cosmol. Astropart. Phys. 01 (2021) 068.
[62] C. Dalang, P. Fleury, and L. Lombriser, Scalar and tensor gravitational waves, Phys. Rev. D **103**, 064075 (2021).
[63] G. Fanizza, G. Franchini, M. Gasperini, and L. Tedesco, Comparing the luminosity distance for gravitational waves and electromagnetic signals in a simple model of quadratic gravity, Gen. Relativ. Gravit. **52**, 111 (2020).
[64] S. Mukherjee, B. D. Wandelt, and J. Silk, Testing the general theory of relativity using gravitational wave propagation from dark standard sirens, Mon. Not. R. Astron. Soc. **502**, 1136 (2021).
[65] Note that numerical attempts to describe GW radiation for theories exhibiting the Vainshtein mechanism, have been conducted in Refs. [66–68], but its inspiral waveform and that of other screening mechanisms have not been described previously.
[66] F. Dar, C. De Rham, J. T. Deskins, J. T. Giblin, and A. J. Tolley, Scalar gravitational radiation from binaries: Vainshtein mechanism in time-dependent systems, Classical Quantum Gravity **36**, 025008 (2019).
[67] C. de Rham, A. J. Tolley, and D. H. Wesley, Vainshtein mechanism in binary pulsars, Phys. Rev. D **87**, 044025 (2013).
[68] C. de Rham, A. Matas, and A. J. Tolley, Galileon radiation from binary systems, Phys. Rev. D **87**, 064024 (2013).
[69] D. M. Eardley, Observable effects of a scalar gravitational field in a binary pulsar, Astrophys. J. **196**, L59 (1975).
[70] A. Le Tiecand and M. Casals, Spinning Black Holes Fall in Love, Phys. Rev. Lett. **126**, 131102 (2021).
[71] N. Chowand and J. Khoury, Galileon cosmology, Phys. Rev. D **80**, 024037 (2009).
[72] S. Mirshekariand and C. M. Will, Compact binary systems in scalar-tensor gravity: Equations of motion to 2.5 post-Newtonian order, Phys. Rev. D **87**, 084070 (2013).
[73] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973).
[74] T. Katsuragawa, T. Nakamura, T. Ikeda, and S. Capozziello, Gravitational waves in $F(R)$ gravity: Scalar waves and the chameleon mechanism, Phys. Rev. D **99**, 124050 (2019).
[75] A. Joyce, B. Jain, J. Khoury, and M. Trodden, Beyond the cosmological standard model, Phys. Rep. **568**, 1 (2015).
[76] P. Creminelli, G. D'Amico, J. Norena, and F. Vernizzi, The effective theory of quintessence: The $w < -1$ side unveiled, J. Cosmol. Astropart. Phys. 02 (2009) 018.
[77] M. Park, K. M. Zurek, and S. Watson, A unified approach to cosmic acceleration, Phys. Rev. D **81**, 124008 (2010).
[78] J. K. Bloomfieldand and E. E. Flanagan, A class of effective field theory models of cosmic acceleration, J. Cosmol. Astropart. Phys. 10 (2012) 039.
[79] G. Gubitosi, F. Piazza, and F. Vernizzi, The effective field theory of dark energy, J. Cosmol. Astropart. Phys. 02 (2013) 032.
[80] J. K. Bloomfield, E. E. Flanagan, M. Park, and S. Watson, Dark energy or modified gravity? An effective field theory approach, J. Cosmol. Astropart. Phys. 08 (2013) 010.
[81] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, Essential building blocks of dark energy, J. Cosmol. Astropart. Phys. 08 (2013) 025.
[82] J. Gleyzes, D. Langlois, and F. Vernizzi, A unifying description of dark energy, Int. J. Mod. Phys. D **23**, 1443010 (2014).
[83] E. Belliniand and I. Sawicki, Maximal freedom at minimum cost: Linear large-scale structure in general modifications of gravity, J. Cosmol. Astropart. Phys. 07 (2014) 050.
[84] L. Lombriserand and A. Taylor, Semi-dynamical perturbations of unified dark energy, J. Cosmol. Astropart. Phys. 11 (2015) 040.
[85] N. Fruscianteand and L. Perenon, Effective field theory of dark energy: A review, Phys. Rep. **857**, 1 (2020).
[86] J. Kennedy, L. Lombriser, and A. Taylor, Reconstructing Horndeski models from the effective field theory of dark energy, Phys. Rev. D **96**, 084051 (2017).
[87] J. Kennedy, L. Lombriser, and A. Taylor, Reconstructing Horndeski theories from phenomenological modified gravity and dark energy models on cosmological scales, Phys. Rev. D **98**, 044051 (2018).
[88] J. Kennedy, L. Lombriser, and A. Taylor, Screening and degenerate kinetic self-acceleration from the nonlinear freedom of reconstructed Horndeski theories, Phys. Rev. D **100**, 044034 (2019).
[89] G. Gabadadze, K. Hinterbichler, and D. Pirtskhalava, Classical duals of derivatively self-coupled theories, Phys. Rev. D **85**, 125007 (2012).
[90] A. Padillaand and P. M. Saffin, Classical duals, legendre transforms and the Vainshtein mechanism, J. High Energy Phys. 07 (2012) 122.
[91] N. Bolis, C. Skordis, D. B. Thomas, and T. Zośnik, Parametrized post-Newtonian-Vainshteinian formalism for the Galileon field, Phys. Rev. D **99**, 084009 (2019).
[92] L. Lombriser, Constraining chameleon models with cosmology, Ann. Phys. (Berlin) **526**, 259 (2014).
[93] X. Zhang, T. Liu, and W. Zhao, Gravitational radiation from compact binary systems in screened modified gravity, Phys. Rev. D **95**, 104027 (2017).

[94] T. Liu, X. Zhang, W. Zhao, K. Lin, C. Zhang, S. Zhang, X. Zhao, T. Zhu, and A. Wang, Waveforms of compact binary inspiral gravitational radiation in screened modified gravity, Phys. Rev. D **98**, 083023 (2018).
[95] R. Epsteinand and R. V. Wagoner, Post-newtonian generation of gravitational waves, Astrophys. J. **197**, 717 (1975).
[96] L. Blanchet, Gravitational radiation from post-Newtonian sources and inspiralling compact binaries, Living Rev. Relativity **17**, 2 (2014).
[97] A. Nicolis, R. Rattazzi, and E. Trincherini, The Galileon as a local modification of gravity, Phys. Rev. D **79**, 064036 (2009).
[98] F. P. Silvaand and K. Koyama, Self-accelerating universe in Galileon cosmology, Phys. Rev. D **80**, 121301 (2009).
[99] G. R. Dvali, G. Gabadadze, and M. Porrati, 4-D gravity on a brane in 5-D Minkowski space, Phys. Lett. B **485**, 208 (2000).
[100] L. Lombriser, W. Hu, W. Fang, and U. Seljak, Cosmological constraints on DGP braneworld gravity with brane tension, Phys. Rev. D **80**, 063536 (2009).
[101] T. Hiramatsu, W. Hu, K. Koyama, and F. Schmidt, Equivalence principle violation in vainshtein screened two-body systems, Phys. Rev. D **87**, 063525 (2013).
[102] E. Barausse *et al.*, Prospects for fundamental physics with LISA, Gen. Relativ. Gravit. **52**, 81 (2020).
[103] P. Amaro-Seoane *et al.* (LISA Collaboration), Laser interferometer space antenna, arXiv:1702.00786.
[104] E. Belgacem, G. Calcagni, M. Crisostomi, C. Dalang, Y. Dirian, J. M. Ezquiaga, M. Fasiello, S. Foffa, A. Ganz, J. Garcí-Bellido *et al.*, Testing modified gravity at cosmological distances with LISA standard sirens, J. Cosmol. Astropart. Phys. 07 (2019) 024.
[105] M. Maggiore *et al.*, Science case for the Einstein telescope, J. Cosmol. Astropart. Phys. 03 (2020) 050.
[106] E. Barausse, N. Yunes, and K. Chamberlain, Theory-Agnostic Constraints on Black-Hole Dipole Radiation with Multiband Gravitational-Wave Astrophysics, Phys. Rev. Lett. **116**, 241104 (2016).
[107] D. Gerosa, S. Ma, K. W. K. Wong, E. Berti, R. O'Shaughnessy, Y. Chen, and K. Belczynski, Multiband gravitational-wave event rates and stellar physics, Phys. Rev. D **99**, 103004 (2019).
[108] G. Gnocchi, A. Maselli, T. Abdelsalhin, N. Giacobbo, and M. Mapelli, Bounding alternative theories of gravity with multiband GW observations, Phys. Rev. D **100**, 064024 (2019).
[109] M. Bezares, L. ter Haar, M. Crisostomi, E. Barausse, and C. Palenzuela, Kinetic screening in non-linear stellar oscillations and gravitational collapse, Phys. Rev. D **104**, 044022 (2021).
[110] M. Bezares, R. Aguilera-Miret, L. ter Haar, M. Crisostomi, C. Palenzuela, and E. Barausse, No Evidence of Kinetic Screening in Merging Binary Neutron Stars, Phys. Rev. Lett. **128**, 091103 (2022).

Chapter 3

Phenomenology of bouncing scenarios

3.1 The inflationary paradigm

The big bang model of the universe is strongly supported by experimental evidences and well established in today's cosmological paradigm. However, it fails to answer some questions about the way the universe is arranged, which is connected in a way or another to an initial condition problem. If the big bang model is the final theory to explain the early universe, then the initial conditions of the universe must be fine tuned to account for the observed cosmological signatures. One of such peculiarities is the *flatness problem*. The spatial curvature is yet unobserved by CMB and BAO experiments and a strong constraint is already set on the curvature parameter today $\left|\Omega_{k,0}\right| < 0.003$ [24], which means that $|\Omega_k| < 10^{-16}$ at the time of nucleosynthesis. It seems that the matter content of the universe perfectly matches the critical value necessary for a flat space universe, which appears as a coincidence in the big bang model. Moreover, the temperature of the CMB is almost exactly isotropic, which means that different regions of the universe, that have never been in causal contact, happen to have the same temperature. This *horizon problem* is again a peculiarity that is unexplained by the big bang model.

Interestingly, both problems are alleviated if the universe had undergone a period of accelerated expansion at its beginning [117,118]. An accelerated expansion is mathematically defined as $\ddot{a} > 0$, with the dot standing for the derivative with respect to time. This is achieved if the equation of state parameter of the energy content is small enough, i.e. $w < -1/3$ from the Raychaudhury equation. This means that the density parameter of the universe would evolve as $\Omega \sim a^q$, with $q = -3(1+w) > -2$, hence it grows faster than the curvature parameter $\Omega_k \sim a^{-2}$, which becomes negligible compared to the matter content for a long enough period of accelerated expansion. Furthermore, one can make far away regions of the universe in causal contact with such an accelerated expansion. Two regions in space-time are in causal contact if they lie in each others particle horizon. The comoving particle horizon was defined as

$$\mathcal{R}_P = \int_{t_{bb}}^{t_0} \frac{\mathrm{d}t'}{a(t')}\,, \tag{3.1}$$

which can be recast under the form

$$\mathcal{R}_P = \int_{\ln 0}^{\ln a_0} (aH)^{-1} \mathrm{d} \ln a \,. \tag{3.2}$$

$\mathcal{R}_H = (aH)^{-1}$ is in fact the comoving Hubble horizon and evolves as $\mathcal{R}_H \sim a^{(1+3w)/2}$. Hence, for $w < -1/3$ the comoving Hubble horizon increases when going backward in time, making the comoving particle horizon bigger than it would have been for a universe dominated by radiation for example. In order to resolve both problems with an accelerated expansion, the scale factor must grow by at least a factor e^{60} during that period. If a_i is the scale factor at the onset of the inflationary period and a_{rh} the scale factor at the time of *reheating*, the end of the period, then $a_{rh} = e^N a_i$, where the number of *e-folds* N should be greater than 60. This comes from the fact that, if what we call *inflation* is the correct model to describe the very early universe, then the largest scale that are seen in the CMB temperature anisotropies were created 60 e-folds before the reheating.

While a period of inflation can solve some of the main issues with the big bang model, we are left to find a physical mechanism that could explain an accelerated expansion. In the context of dark energy, we already saw that a scalar field can play such a role and popular models of inflation consist of a universe filled with a scalar field that is commonly called *inflaton* [119]. As a first model to understand the dynamics of inflation we can use scalar field ϕ with a general potential $V(\phi)$. The energy-momentum tensor of such a field is

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi + g_{\mu\nu} \left(X + V(\phi) \right) \,, \tag{3.3}$$

which can then be used in the Einstein field equations. We recall that $X = -\partial_\mu \phi \partial^\mu \phi / 2$ is the kinetic term of the scalar field. The resulting field equations for the homogeneous, isotropic and flat space case gives

$$H^2 = \frac{8\pi}{3} \left(\frac{1}{2} \dot{\phi}^2 + V(\phi) \right) \,, \tag{3.4}$$

$$\dot{H} = -\frac{1}{16\pi} \dot{\phi}^2 \,, \tag{3.5}$$

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi}(\phi) = 0 \,. \tag{3.6}$$

Out of these three equations, only two are independent. The scalar field actually behaves as a perfect fluid with varying equation of state and its fluid properties are

$$\rho_\phi = \frac{1}{2} \dot{\phi}^2 + V(\phi) \quad \text{and} \quad p_\phi = \frac{1}{2} \dot{\phi}^2 - V(\phi) \,. \tag{3.7}$$

We saw that in order to have an accelerated expansion with $w < -1/3$, we need

$$\epsilon := -\frac{\dot{H}}{H^2} = \frac{\dot{\phi}^2}{16\pi H^2} < 1 \,, \tag{3.8}$$

and since we want the inflationary period to last long enough, we want ϵ to remain small

during that period. Hence, we require that

$$\bar{\eta} := \frac{\dot{\epsilon}}{H\epsilon} < 1\,. \tag{3.9}$$

Therefore, we need to choose a scalar potential that satisfies these conditions during a sustained period.

In what we call the *slow roll approximation*, we go one step further and require that $\epsilon, \bar{\eta} \ll 1$. These more restrictive conditions make it simpler to have a long period of inflation. One can then show using the field equations that

$$\epsilon \approx \epsilon_V := \frac{1}{2}\left(\frac{V_{,\phi}}{V}\right)^2 \quad \text{and} \quad \bar{\eta} \approx \eta_V := \frac{\left|V_{,\phi\phi}\right|}{V}\,, \tag{3.10}$$

where ϵ_V and η_V are called the *slow roll parameters*. From the definition of ϵ (3.8), we see that during the slow roll phase the universe behaves as a de-Sitter space-time with $w \approx -1$, hence with an exponential growth of the scale factor. To measure how long inflation lasts, we can derive the number of e-folds by which the scale factor grows

$$N := \int \mathrm{d}\ln a \approx \int_{\phi_i}^{\phi_{rh}} \frac{\mathrm{d}\phi}{\sqrt{2\epsilon_V}}\,. \tag{3.11}$$

While the slow roll approximation and the use of the scalar field in constructing a long lasting accelerated expansion seem somewhat arbitrary, it is so far well motivated by the temperature anisotropies of the CMB as we will see in the next section 3.2.

3.2 The origin of structures

The existence of inhomogeneities in the distribution of matter today and the temperature anisotropies proved us that the universe does not fully satisfy the cosmological principle. Moreover, these small fluctuations are essential in modern cosmology, from the BAO to the study of large scale structures. Beyond its ability to solve the flatness and horizon problems, inflation offers a great explanation for the origin of the fluctuations as quantum perturbations at the onset of the inflationary period. In this section, we introduce the concept of cosmological perturbation theory in the context of inflation, discuss the experimental consequences of these perturbations on the CMB and explain how all structures in the universe might have emerged from quantum fluctuations.

3.2.1 Cosmological perturbation theory for inflation

Temperature anisotropies in the CMB are small compared to the background radiation, supporting the use of perturbation theory in the context of the early universe. During inflation the universe is assumed to be filled with a scalar field and the action of the system is

$$S\big[\phi, g\big] = \frac{1}{8\pi}\int \mathrm{d}^4x\sqrt{-g}\left(\frac{1}{2}R + X - V(\phi)\right)\,. \tag{3.12}$$

The goal is now to write this action at linear order in the perturbation of the metric and scalar field, namely

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \quad \text{and} \quad \phi = \bar{\phi} + \delta\phi \,, \tag{3.13}$$

where $\bar{g}_{\mu\nu}$ is the background FLRW metric and $\bar{\phi}$ is the background scalar field responsible for the accelerated expansion. As a symmetric rank two tensor, the metric perturbations $\delta g_{\mu\nu}$ can be decomposed into ten degrees of freedom, four scalars, four vectors and two tensors. Together with the scalar field perturbation $\delta\phi$, we have a priori eleven degrees of freedom. The interesting consequence of this decomposition of the perturbations is that the perturbations of different spin are decoupled and evolve independently at leading order. Thanks to the gauge invariance under general coordinate transformation, which is a 4-dimensional gauge group, we can get rid of $2 \cdot 4 = 8$ degrees of freedom. The elimination of twice the amount of degrees of freedom as compared to the dimensional size of the gauge group can be understood as follow. Out of the ten Einstein field equations describing the metric perturbations in their canonical form, four are non-dynamical, meaning that they do not contain second derivatives. This implies that the momentum associated to the variable is zero leading to four initial constraint equations. Moreover, the field equations for these non-dynamical variables add four other constraints, reducing the total number of degrees of freedom by eight [7]. Eventually, we are left with three degrees of freedom. One can show that for single scalar field inflation, the vector degrees of freedom don't propagate [15], we are therefore left with one scalar and two tensor degrees of freedom. Fortunately, the tensor perturbations are already gauge invariant, but one needs to make sure that the scalar perturbation that is used is invariant as well. An interesting but non-unique choice can be [15]

$$\delta\phi = 0 \,, \quad \delta g_{ij} = a^2 \left((1 - 2\mathcal{R})\delta_{ij} + h_{ij}\right) \,, \quad \partial_i h_{ij} = h^i_i = 0 \,, \tag{3.14}$$

where $\mathcal{R}$ is the scalar perturbation and h_{ij} the tensor ones.

Since both types of perturbations are decoupled at leading order, we can work on them independently. Starting with the scalar perturbation, its action at second order is

$$S_{(2)}\big[\mathcal{R}\big] = \frac{1}{16\pi} \int \mathrm{d}^4x \; a^3 \frac{\dot{\phi}^2}{H^2} \left(\dot{\mathcal{R}}^2 - a^{-2}\partial_i\mathcal{R}\partial^i\mathcal{R}\right) \,, \tag{3.15}$$

where we renamed $\bar{\phi} \to \phi$ for simplicity. This action is obtained after expanding the action (3.12) in powers of $\mathcal{R}$ and $\partial_\mu\mathcal{R}$. It is common practice to rewrite the scalar perturbation using the Mukhanov-Sasaki variable

$$v_S = z\mathcal{R} \,, \quad \text{with} \quad z = \frac{a\dot{\phi}}{H} \,, \tag{3.16}$$

and using the conformal time η as the new time coordinate, giving us

$$S_{(2)}\big[v_S\big] = \frac{1}{16\pi} \int \mathrm{d}\eta \mathrm{d}^3x \left(v_S'^2 + \partial_i v_S \partial^i v_S + \frac{z''}{z} v^2\right) \,, \tag{3.17}$$

where the prime notation stands for the derivative with respect to the conformal time. Since

we assume flat space, any functions defined on $\mathbb{R}^3$ can be written as a Fourier expansion of the form

$$v_S(\eta, \mathbf{x}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3k \; v_S(\eta, \mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \,. \tag{3.18}$$

The equation of motion for the scalar modes can be found using the Euler-Lagrange equation of the action for v_S (3.17) and we get

$$v_S''(\eta, k) + \left(k^2 - \frac{z''}{z}\right) v_S(\eta, k) = 0 \,. \tag{3.19}$$

Note that the scalar modes only depend on the norm $k := |\mathbf{k}|$ of the wave vector, due to the isotropic assumption. Eq. (3.19) is also called the *Mukhanov-Sasaki* equation for scalar modes.

Let's now move on to the tensor perturbations. We follow a similar path as in the scalar case but in a different order, i.e. find the action at second order for the perturbations h_{ij}, rewrite the action in terms of the Fourier modes

$$h_{ij} = \frac{1}{(2\pi)^3} \int \mathrm{d}^3k \sum_{s=+,\times} \varepsilon_{ij}^s h^s(\eta, \mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \,, \tag{3.20}$$

define the Mukhanov-Sasaki variable $v_T^s(\eta, k) = a h^s(\eta, k)/\sqrt{32\pi}$ and find the equation of motion for the modes using the Euler-Lagrange equation, namely

$$v_T^{s\prime\prime}(\eta, k) + \left(k^2 - \frac{a''}{a}\right) v_T^s(\eta, k) = 0 \,. \tag{3.21}$$

We found the evolution equation for both types of perturbations in Fourier space. We are left to describe what are the physical consequences of these perturbations and where do they come from.

But first, we want to understand how these fluctuations evolve in a quasi-de Sitter spacetime, which is the behaviour during slow roll inflation. If the scale factor grows exponentially fast, one can show that

$$\frac{z''}{z} \approx \frac{a''}{a} \approx \frac{2}{\eta^2} \,. \tag{3.22}$$

Therefore, both types of fluctuations evolve as

$$v''(\eta, k) + \left(k^2 - \frac{2}{\eta^2}\right) v(\eta, k) = 0 \,. \tag{3.23}$$

Moreover, one can normalize η such that $\eta(t_{rh}) = 0$, the conformal time is zero at reheating and $\eta(t \to t_{bb}) \to -\infty$. This implies that the evolution equation (3.23) can be separated into two different regimes. At very early times when $|\eta| \gg k^{-1}$, the perturbations oscillate as a harmonic oscillator and this will be useful in the quantum treatments of the fluctuations. In the second regime, when the mode k becomes large relative to the Hubble horizon, i.e. $|\eta| \ll k^{-1}$, the fluctuations can behave in two different ways, either $v(a) \sim a$ or $v(a) \sim a^{-2}$.

We will see in Sec. 3.2.2 that in the case of perturbations that are created by quantum fluctuations of the Bunch-Davies vacuum, the perturbations behave as the former case and they are said to be *adiabatic*. This means that the quantities

$$\mathcal{R}(\eta,k) \sim \frac{v_S(\eta,k)}{z(\eta)} \quad \text{and} \quad h^s(\eta,k) \sim \frac{v_T^s(\eta,k)}{a(\eta)}\,, \tag{3.24}$$

are frozen at super-Hubble scales until inflation comes to an end at reheating. These are therefore interesting properties to study and constrain using CMB measurements.

3.2.2 Quantum origin of the perturbations

In the previous subsection, we found the equations (3.19) and (3.21) that describe the evolution on a flat FLRW space-time of the scalar and tensor perturbations, respectively, during the inflationary period, when the universe was assumed to be dominated by the inflaton field. The theory of single field inflation predicts that the universe started with a quasi-de Sitter phase, with the scale factor growing exponentially fast. In particular, we showed that if we go far enough in the past $\eta \to -\infty$, we can find a time when all the relevant modes k were far inside the Hubble horizon and the fluctuations behaved as a harmonic oscillator. This is fortunate because we know very well how to quantize a Harmonic oscillator and we can therefore describe the fluctuations quantum mechanically.

Let's proceed to the canonical quantization of the general perturbation $v_k(\eta) := v(\eta,k)$, whose evolution equation is Eq. (3.23), and which can represent both scalar and tensor perturbations. The fluctuation v_k and its canonical momentum $\pi_k := \partial\mathcal{L}/\partial v_k'$ are promoted to operators, which satisfy the commutation relation

$$[\hat{v}_k, \hat{\pi}_{k'}] = i\delta(k+k')\,. \tag{3.25}$$

Furthermore, the fluctuations can be expanded into a creation and annihilation operator, namely

$$\hat{v}_k(\eta) = v_k(\eta)\hat{a}_k + v_k^*(\eta)\hat{a}_k^\dagger\,, \tag{3.26}$$

with v_k^* the conjugate of v_k, the latter satisfying the classical equation of motion (3.23). The operators $\hat{a}_k$ and $\hat{a}_k^\dagger$ must satisfy the commutation relation

$$W(v_k)\left[\hat{a}_k, \hat{a}_k^\dagger\right] = 1\,, \tag{3.27}$$

where

$$W(v_k) = i\left(v_k^* v_k' - v_k^{*\prime} v_k\right)\,, \tag{3.28}$$

is the *Wronskian* of v_k. Through a renormalisation of $\hat{v}_k$, we can choose $W(v_k) = 1$, without loss of generality. This gives us the first initial conditions for the function $v_k(\eta)$. With the use of the creation operator, provided that we define a vacuum state $|0\rangle$, we can construct the Hilbert space of quantum states. The last ingredient to fully describe the quantum behaviour of the fluctuations is therefore the vacuum state. In the case of the harmonic oscillator the vacuum state is uniquely defined as the minimum energy eigenstate of the Hamiltonian

operator. The requirement that the vacuum state is an eigenstate of the Hamiltonian fixes the second initial condition for $v_k(\eta)$. We can see this from

$$\hat{H}\,|0\rangle = \frac{1}{2}\left(v_k'^2 + k^2 v_k^2\right)^* \hat{a}_k^\dagger \hat{a}_k^\dagger\,|0\rangle + \left(\left|v_k'^2\right| + k^2|v_k|^2\right)|0\rangle \propto |0\rangle\ , \tag{3.29}$$

which implies that

$$v_k' = -ikv_k\,. \tag{3.30}$$

With these two initial conditions, the quantum fluctuation operator $\hat{v}_k$ that defines a unique vacuum when $|\eta| \gg k^{-1}$ is fully determined and its mode functions are

$$v_k(\eta \to -\infty) = \frac{1}{\sqrt{2k}} e^{-ik\eta}\,. \tag{3.31}$$

This choice of mode functions is called the *Bunch-Davies* or *Minkowski vacuum*. We have found a great candidate for the initial condition of our fluctuations. We can now evolve the mode functions using Mukhanov-Sasaki equation in quasi-de Sitter space-time (3.23) and the only solution satisfying the initial condition (3.31) is

$$v_k(\eta) = \frac{e^{-ik\eta}}{\sqrt{2k}}\left(1 - \frac{i}{k\eta}\right)\,. \tag{3.32}$$

When the fluctuations of mode k exit the Hubble horizon, i.e. $|\eta|k \ll 1$, we indeed have that the mode functions behave as

$$v_k(\eta) \propto \eta^{-1} \implies \frac{v_k(a)}{a} \sim 1\,. \tag{3.33}$$

The curvature and tensor perturbations are frozen on super-Hubble scales.

We have found the quantum operator $\hat{v}_k(\eta)$ describing the evolution of the perturbations after unambiguously defining the vacuum state $|0\rangle$. We have made the assumption that the space-time evolution was quasi-de Sitter so that the Mukhanov-Sasaki equations for both types of perturbations are the same, but the derivation of the quantum fluctuations is analogous when using the different equations of motion for scalar and tensor modes. We stress out here that this is only valid if in the far past, there was a time when the fluctuations could be described by a time-independent harmonic oscillator. While this holds in flat space single field inflation, this is not always guaranteed in bouncing models for example, in which there often is a non quasi-de Sitter phase before the inflationary period. One needs to be careful to initiate quantum fluctuations at a time when the visible modes are far within the Hubble horizon in order to guarantee the uniqueness of the prediction. If this is not possible, one needs to keep in mind that the vacuum state is not unique and the predictions are ambiguous [15, 120].

3.2.3 Primordial power spectra from inflation

It is now time to calculate the cosmological consequences of the quantum fluctuations. The measurable quantities are the expectation values of operators and in particular the

probabilistic moments of the fluctuation operator

$$\hat{v}(\eta, \mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}} \hat{v}_k(\eta) e^{i\mathbf{k}\cdot\mathbf{x}} \,. \tag{3.34}$$

The first order moment $\langle\hat{v}\rangle$, or in other word the mean of $\hat{v}$, is zero. This is expected since we assumed small perturbations around the background. The variance can be calculated as follows,

$$\langle 0|\, \hat{v}^\dagger(\eta, \mathbf{x})\hat{v}(\eta, \mathbf{x})\, |0\rangle = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^{3/2}} f_k(\eta) f^*_{k'}(\eta)\, \langle 0|\, \hat{a}_{\mathbf{k}} \hat{a}^\dagger_{\mathbf{k}'}\, |0\rangle\, e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \tag{3.35}$$

$$= \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^{3/2}} f_k(\eta) f^*_{k'}(\eta) \delta_{\mathbf{k}'\mathbf{k}} \tag{3.36}$$

$$= \int \mathrm{d}\ln k\ \frac{k^3}{2\pi^2} |f_k|^2 \tag{3.37}$$

$$:= \int \mathrm{d}\ln k\ \mathcal{P}_v(k, \eta) \,, \tag{3.38}$$

where we defined the *dimensionless power spectrum* $\mathcal{P}_v(k, \eta)$. We saw that the curvature perturbation $\mathcal{R}$ is frozen at super-Hubble scales, which makes it an interesting and easy variable to work with. The same is true for the tensor perturbations h^s. We therefore give the prediction of inflation for the primordial power spectra at reheating of the curvature fluctuation and the gravitational waves. The former is found to be

$$\mathcal{P}_{\mathcal{R},rh}(k) = \frac{k^3}{2\pi^2} \left| \frac{v_k(\eta)}{z_S(\eta)} \right|^2 \Bigg|_{\eta=\eta_\star} = \frac{H^4_\star}{4\pi^2 \dot{\phi}^2_\star} \,. \tag{3.39}$$

Thanks to the freezing effect at horizon crossing $\eta = \eta_\star = k^{-1}$, we can evaluate the power spectrum only up to that time and not all the way to reheating. Similarly for the primordial power spectrum for the tensor modes, we get

$$\mathcal{P}_{T,rh}(k) = 2\mathcal{P}_{h^s,rh}(k) = \frac{H^2_\star}{4\pi^3} \,. \tag{3.40}$$

One could now calculate the next order moment, but in the case of single field inflation, all higher order moments cancel. This is the behaviour of a Gaussian statistics, it is fully defined by its two first moments. At present, there is no evidence of non-Gaussianities from cosmological surveys [24].

Using what we know about the slow-roll behaviour of inflation, we can make a prediction on the shape of the primordial power spectra. Starting with the scalar perturbations, we

have

$$\mathcal{P}_{\mathcal{R},rh}(k) = \frac{H_\star^2}{64\pi^3\epsilon_\star} \tag{3.41}$$

$$\implies \frac{\mathrm{d}\mathcal{P}_{\mathcal{R},rh}}{\mathrm{d}k} = \left(\frac{2\dot{H}H}{64\pi^3\epsilon} - \frac{H^2\dot{\epsilon}}{64\pi^3\epsilon^2}\right)_{t=t_\star} \frac{\mathrm{d}t_\star}{\mathrm{d}k} \tag{3.42}$$

$$= \left(-2H_\star\epsilon_\star\mathcal{P}_{\mathcal{R},rh} - H_\star\bar{\eta}_\star\mathcal{P}_{\mathcal{R},rh}\right) a_\star k^{-2} \tag{3.43}$$

$$= (-2\epsilon_\star - \bar{\eta}_\star)k^{-1}\mathcal{P}_{\mathcal{R},rh}(k)\,, \tag{3.44}$$

where we used $\mathrm{d}t = a\mathrm{d}\eta$ and $\eta_\star = k^{-1} \implies a_\star H_\star = k$. This means that the primordial power spectrum for the scalar perturbations behaves as

$$\mathcal{P}_{\mathcal{R},rh}(k) = A_s\left(\frac{k}{k_*}\right)^{n_s-1}\,, \tag{3.45}$$

where A_s is the amplitude of the power spectrum at the *pivot scale* $k = k_*$, with $k_* = 0.05$ Mpc^{-1}, and $n_s - 1 = -2\epsilon_\star - \bar{\eta}_\star$, n_s being the *scalar spectral index*. Slow-roll inflation predicts a small but non-zero value for $n_s - 1$ and current observations from the Planck satellite confirms that prediction with $n_s = 0.9665 \pm 0.0038$ [24]. The amplitude of the scalar primordial power spectrum has been measured to be $\ln\left(10^{10}A_s\right) = 3.047 \pm 0.014$ [24].

A similar reasoning can be made for the primordial power spectrum of the tensor modes. It follows a power law of the form

$$\mathcal{P}_{T,rh}(k) = A_t\left(\frac{k}{k_*}\right)^{n_t}\,, \tag{3.46}$$

where $n_t = -2\epsilon$ and A_t is its amplitude at the pivot scale. The tensor perturbations have so far not been observed, the spectral index is therefore still unknown, but should be small according to slow-roll inflation. A constraint on its amplitude can be made and is expressed by the tensor-to-scalar ratio $r_{0.002}$ at the scale $k = 0.002$ Mpc^{-1}, namely

$$r_{0.002} = \frac{\mathcal{P}_{T,rh}(k=0.002)}{\mathcal{P}_{\mathcal{R},rh}(k=0.002)} < 0.06\,. \tag{3.47}$$

The prediction and confirmation of the small scalar spectral index as well as the absence of non-Gaussianities have been great successes of the theory of inflation and put it in the front scene of early universe cosmology. The observation of the tensor perturbations as predicted by the theory would easily place it in the concordance model.

3.3 Bounce from positive curvature and inflation

In our introduction of the theory of inflation, we have so far assumed that the universe is spatially flat. After all, one of the reasons to hypothesize an inflationary period is the extreme flatness of space. On the other hand, this means that we do not expect the curvature of space to be small at the onset of inflation. Assuming no priors on the initial conditions of the universe, the latter is equally likely to have a positive or a negative curvature, but an almost flat spatial section is very unlikely. Once again, inflation has been invented so that

no finely tuned priors are needed to explain cosmological observations. Today's constraint on the curvature parameter, $\Omega_K = 0.0007 \pm 0.0019$, gives us almost no clue about the sign of the spatial curvature. A potential conflict between Planck's measurement and other cosmological surveys about the curvature was risen in Ref. [121] and the increasing number of cosmological tensions [122–124] might shift tomorrow's cosmological paradigm. However, the consensus concerning the value of Ω_K is still favouring a flat universe [125].

In this context, it makes sense to look for other cosmological signatures of a possible positive curvature and the early inflationary period, where the effect of the spatial curvature is at its highest, is an interesting place to start. Ignoring for now the perturbations, the Friedmann equations of an inflaton field bathed in a closed universe read

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho_\phi + 3p_\phi) \tag{3.48}$$

$$H^2 = \frac{8\pi}{3}\rho_\phi - \frac{K}{a^2}\,. \tag{3.49}$$

During inflation, the density of the scalar field is roughly constant since the equation of state parameter is $w \approx -1$. Therefore, when going backward in time, the curvature term grows faster than the quasi-constant density term. In particular, this explains why inflation is efficient at reducing the contribution of curvature to the energy content of the universe. But this also implies that at very early times, the Hubble parameter might have been zero, or in other words, our expanding universe, if closed, started with a big bounce instead of a big bang [126]. Beyond solving the big bang singularity problem, the *curvature bounce* would happen at the energy scale of inflation, which is expected to be several orders of magnitude higher than the Planck scale, and therefore the need of a quantum theory of gravity would be alleviated. It is remarkable that one can resolve the big bang singularity only with the use of non-exotic physics. It is fair to say that inflation is part of today's cosmological paradigm and we emphasize that from our understanding of physics, assuming the initial conditions of the universe are randomly set, the existence of a positive curvature is reasonably likely.

Depending on today's value of the curvature parameter $\Omega_{K,0}$ and the energy scale of inflation E_{infl}, one can predict the number of e-folds as [126]

$$N \approx \frac{1}{2}\ln\left(\frac{\Omega_{R,0}}{\Omega_{K,0}}\left((1+z_{eq})\frac{T_{infl}}{T_{eq}}\right)^2\right)\,, \tag{3.50}$$

where the subscript "eq" stands for the time of matter-radiation equilibrium and $\Omega_{R,0}$ is the radiation parameter evaluated today. This prediction is derived using the assumptions that the energy density of the inflaton is constant and $\rho_\phi(t_b) = K/a^2(t_b)$ at the onset of inflation, which is at the time of the bounce t_b. Using the constraint on the curvature density parameter and $E_{infl} \sim 10^{15}$ GeV, we find that the inflationary period must have lasted at least $N > 68$, if the universe is positively curved.

Thanks to the background dynamics, one can solve the big bang singularity and make further predictions on the duration of inflation. However, in order to see potential signatures of the big bounce we need to move towards cosmological perturbation theory. When calculating the evolution of the perturbations in a closed universe as opposed to the spatially flat

case, several differences emerge [127–130]. While the number of degrees of freedom stays the same, the action and the equations of motion in position space are different. Moreover, we cannot perform a Fourier expansion of the perturbations, but we must use the hyperspherical harmonics $\mathcal{Q}_{nlm}(\mathbf{x})$ to decompose the solutions to the equations. Such expansions for both modes are

$$v(\eta, \mathbf{x}) = \sum_{n=2}^{\infty} \sum_{l=0}^{n-1} \sum_{m=-l}^{l} v_{nlm} \mathcal{Q}_{nlm}(\mathbf{x}) \tag{3.51}$$

$$h_{ij}(\eta, \mathbf{x}) = \sum_{n=2}^{\infty} \sum_{l=0}^{n-1} \sum_{m=-l}^{l} \sum_{s=1}^{2} h_{ij}^{nlm,s} \mathcal{Q}_{ij}^{nlm,s}(\mathbf{x}) \tag{3.52}$$

and are now discrete. Overall, this means that the Mukhanov-Sasaki equations take the form

$$v_S''(\eta, n) + A_n(\eta) v_S'(\eta, n) + B_n(\eta) v_S(\eta, n) = 0 \tag{3.53}$$

$$v_T^{s\prime\prime}(\eta, n) + \left((n^2 - 1)K - \frac{a''}{a} \right) v_T^s(\eta, n) = 0\,, \tag{3.54}$$

where the functions A_n and B_n are defined later in this section. Interestingly, the scalar and tensor mode equations do not depend on the wave number l or m as well as the polarisation s for the tensor case, which simplifies the calculations. While the form of the equation for the tensor modes is basically unchanged, the scalar mode equation has become much more complex to work with. In both cases, the fact that the Hubble horizon does not only increase when going backward in time can be a problem when choosing the vacuum state. In the pre-bounce universe, the Hubble horizon decreases when going backward in time. In the case of the scalar perturbations, both A_n and B_n need to satisfy the conditions of the Minkowski vacuum in order to have an unambiguous definition of the vacuum state. The derivation of the tensor primordial power spectrum is made in the first paper below. The analysis of the behaviour of the bounce at the background level is also explored in the same paper, where we show in particular that a period of deflation in the pre-bounce universe is necessary. The calculation of the scalar primordial power spectrum as well as the power spectrum of the temperature anisotropies in the CMB is exposed in the second paper below. In the latter paper, we also look at the behaviour of the pre-bounce universe and its effect on the scalar perturbations depending on the length of the pre-bounce deflation.

Curvature bounce in general relativity: background and primordial spectrum

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Curvature bounce in general relativity: background and primordial spectrum

Cyril Renevey,[a,b] **Aurélien Barrau,**[a] **Killian Martineau**[a] **and Selim Touati**[a]

[a]Laboratoire de Physique Subatomique et de Cosmologie,
Université Grenoble-Alpes, CNRS/IN2P3,
53, avenue des Martyrs, 38026 Grenoble cedex, France
[b]Institute for Theoretical Physics, ETH Zürich,
Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland

E-mail: renevey@lpsc.in2p3.fr, barrau@in2p3.fr, martineau@lpsc.in2p3.fr, touati@lpsc.in2p3.fr

Received October 27, 2020
Accepted November 17, 2020
Published January 13, 2021

Abstract. Recent data suggest that the Universe could be positively curved. Combined with an inflationary stage, this might lead to a curvature bounce instead of the Big Bang. The background evolution is presented, as a function of the parameters controlling the cosmic evolution. The primordial tensor spectrum is also calculated and possible observational footprints of the model are underlined. Several potentials are considered and general remarks are made about "naturalness" in this context.

Keywords: alternatives to inflation, CMBR theory, physics of the early universe

ArXiv ePrint: 2010.13542

JCAP01(2021)018

 https://doi.org/10.1088/1475-7516/2021/01/018

Contents

1 Introduction

It is often argued that, unless radically new physics is used, the classical origin of the Universe is a singularity. This is not rigorously true. Following arguments given in [1–3], we show that a curvature bounce might have occurred in the past. This does not require any exotic physics. The key ingredients of this scenario are the existence of an inflationary period — which is obviously part of the standard cosmological model — and a positive curvature for the spatial sections of the Universe. Although still actively debated and not fully consensual, a positive curvature seems to be favored by cosmological microwave background (CMB) measurements. If not truly established, this hypothesis is anyway compatible with current data and this constitutes the core of this work. The idea of a "curvature bounce" is not new but deserves a fresh look, taking into account the latest analyses.

This obviously raises some "fine-tuning" issues. Although meaningful, those considerations are extremely difficult to formalize rigorously, due to the ambiguity associated with any arbitrary chosen measure. In this article, we do not address in details the naturalness of the model, instead we try to evolve backward in time the current state of the Universe and investigate the resulting history.

Many different bouncing models have been considered in cosmology (see [4–6] for recent reviews). In particular, the bounce can be associated with an inflationary phase (see, e.g., [7–15]). This leads to a wide phenomenology depending on the detailed physical processes involved in the scenarii. This work focuses on the specific case were *no* new physics in involved, which possible thanks to the curvature.

The consequences of a positive curvature are profound for the history of the Universe and might change the way we understand its "origin". We first explain what are the motivations for a positive curvature. We then study in details the background behavior. The primordial tensor power spectrum is finally calculated. Throughout all the article we use Planck units.

2 On the curvature of space

The question of the spatial curvature of the Universe is an old one. It has been debated for decades. Although a flat space is usually considered as one of the key prediction of inflation, the actual situation is not that simple. Obviously, whatever the "initial" curvature, inflation

JCAP01(2021)018

will make it negligible at the reheating. However, as the Universe expands, the curvature contribution to the Friedmann equation will decrease more slowly ($\propto a^{-2}$) than the matter ($\propto a^{-3}$) and radiation ($\propto a^{-4}$) ones. Curvature will then eventually dominate in the future, unless — as it is the case in our Universe — the cosmological constant ($\propto a^{0}$) overwhelms everything before this happens.

Otherwise stated: the end of inflation might be the moment in cosmic history when the curvature contribution to the dynamics of the Universe is the smallest one.

As recalled above, whatever its non-vanishing value, the relative contribution of the curvature increases as time goes on after inflation. But, importantly, it also increases when going backward in time during inflation: while the scalar field density remains roughly constant in the quasi-de Sitter stage, the curvature contribution increases (as it still scales as $\propto a^{-2}$) when the scale factor decreases. This might trigger a bounce instead of the singularity.

This question becomes especially important in the current context. First, the Planck data suggests that, using the CMB alone, the Universe might be positively curved [16]. The statistical significance of this result is weak. However, in addition, recent analyses presented in [17–19] considerably strengthen this possibility. In particular, it is shown in [19] that the enhanced lensing amplitude in primordial power spectra, when compared to the prediction of the standard ΛCDM model, can be explained by this effect. This would also remove the tension within the Planck data about the values of cosmological parameters measured at different angular scales. The study concludes that Planck data favor a closed Universe with a probability of nearly 99.99%.

Counter-arguments where given in [20]. The actual conclusion heavily depends on the priors used for the analysis. At this stage, it is fair to say that a positive curvature is not firmly established but is worth being seriously considered. First, because some studies — in particular using the CMB alone — point in this direction and, second, because theoretical arguments — in particular grounded in quantum gravity — favor this hypothesis.

3 Background behavior of a closed Universe

We assume a homogeneous and isotropic background with closed spatial sections. The topology is $\mathbb{R}\times\mathbb{S}^3$, where $\mathbb{S}^3$ represents a hypersphere. The spatial curvature parameter $K>0$ is related to the physical radius $r(t)$ of the 3-sphere by $r^2(t)=a^2(t)/K$, where $a(t)$ is the dimensionless scale factor. Under these assumptions, the FLRW metric can be written as

$$\mathrm{d}s^2=-\mathrm{d}t^2+\frac{a^2(t)}{K}\left(\mathrm{d}\chi^2+\sin^2(\chi)\mathrm{d}\Omega^2\right). \tag{3.1}$$

The matter content of the Universe is represented by a perfect fluid of density ρ and pressure p, such that the Einstein field equations lead to the Friedmann and Raychaudhury equations

$$H^2=\frac{8\pi}{3}\rho-\frac{K}{a^2}, \tag{3.2}$$

$$\dot{H}=-4\pi(\rho+p)+\frac{K}{a^2}, \tag{3.3}$$

where $H=\dot{a}/a$ is the usual Hubble parameter. During the inflationary period, the matter content of the Universe can be described by a scalar field. In this work, we assume the inflaton to be a massive scalar field, that is to be described by the potential $V=m^2\phi^2/2$. In principle, it would make sense to consider other potentials — a few remarks will be made

JCAP01(2021)018

on this point in the last section — but the massive case makes the comparison with other models easier and captures most of the relevant phenomenology. The Klein-Gordon equation in the expanding (or contracting) Universe reads

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial V(\phi)}{\partial \phi} = 0. \tag{3.4}$$

The density and pressure of the inflaton field can be written as

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) \quad \text{and} \quad p = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{3.5}$$

As well known, eqs. (3.2), (3.3) and (3.4) are not independent and only two of them are necessary to solve the system.

Let us first state that, when going backward in time, if the curvature density dominates the dynamics, while the Universe was in a de Sitter phase (or nearly so), a bounce inevitably takes place, as shown by the trivial analytical solution of the equations of motion:

$$a = \sqrt{\frac{3K}{\Lambda_i}} \cosh\left(\sqrt{\frac{\Lambda_i}{3}} t\right), \tag{3.6}$$

where $\Lambda_i = 8\pi\rho_{vac}$, with ρ_{vac} the vacuum-like energy density of the field. Among bouncing models [4–6], this one is specific in the sense that it does not require any exotic physics.

Following [3], it can easily be shown that the number of inflationary e-folds between the bounce and the reheating is here given by:

$$N \approx \frac{1}{2} \ln\left(\frac{\rho_{R,0}}{\rho_{K,0}}\left[(1+z_{eq})\frac{T_{RH}}{T_{eq}}\right]^2\right), \tag{3.7}$$

where $\rho_{R,0}$ and $\rho_{K,0}$ are respectively the current densities of radiation and curvature, z_{eq} is the redshift at the equilibrium time, T_{RH} is the reheating temperature and T_{eq} is the equilibrium temperature. The sudden reheating approximation is obviously crude [21] but sufficient for this study. This formula does *not* use any dynamical feature of inflation: it just counts the amount of contraction (when thinking backward in time) "needed" for the curvature density to equal the scalar field density (with the opposite sign in the Friedmann equation), which ensures the vanishing of the Hubble parameter at some point. When normalizing the curvature density at the value suggested in [19], this leads to $N \approx 65$ for $T_{RH} \approx 10^{16}\,\mathrm{GeV}$. The inflation energy scale cannot be much above, otherwise this would conflict with current data, in particular with the tensor-to-scalar ratio upper limit. On the other hand, the number of e-folds cannot be much below this value so as to solve the usual cosmological paradoxes.[1] This basically means that those parameters are somehow fixed in this model, which is to be contrasted with the usual cosmological framework where no upper bound exists on the number of inflationary e-folds and where the lower bound ont the reheating temperature is extremely weak.

As a consequence, from the purely numerical viewpoint, the value of the curvature can therefore be tuned so that the number of e-folds fits eq. (3.7), at the considered density:

[1] In principle, one could relax this constraint by just requiring the number of inflationary e-folds to be equal to the number of post-inflationary e-folds but, in this specific framework, this also conflicts with data if it is much below 65, as shown in [3].

JCAP01(2021)018

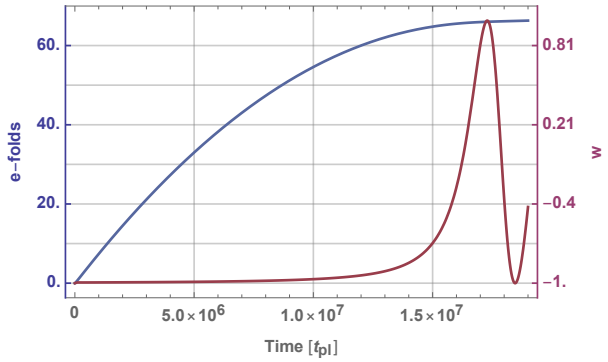


Figure 1. *In blue:* scale factor (variation measured in e-folds) during the inflationary period; *In purple:* $w = p/\rho$ during the same period. The density at t_0 has been chosen such that the number of e-folds during inflation is $N \approx 65$.

this will automatically account for the (possibly) observed curvature. We have checked that it is indeed possible to find an initial value for the field — associated with a given energy density — so that the duration of inflation predicted by the model coincides with the intrinsic inflationary dynamics.

In order to have a strong control on the behavior of the inflationary period, we set the initial conditions (IC) of the background close to the onset of inflation, when $w_0 := p_0/\rho_0 \approx -1$ at t_0.

Interestingly, the amount of inflation predicted in this framework is both high enough to be compatible with the observational lower bound and small enough so that non-trivial effects associated with the bounce could be seen in the CMB. In most bouncing models, the duration of inflation is so high [15, 22, 23] that subtle footprints of the bounce are deeply super-Hubble and therefore non-observable. In this model, one might expect non-trivial features at large, but sub-Hubble, physical scales, that is in the low-k part of the primordial spectra.

In the case of a massive scalar field, the number of e-folds can easily be analytically calculated and is given by $N = 2\pi\phi^2 - 1/2$ in flat space. As the effect of the curvature is only relevant very close to the bounce, this result can also be safely used in a closed universe at the level of accuracy required here. At t_0, the onset of inflation, N is directly connected to the density of the scalar field, using eq. (3.5). In order to have $N \approx 65$, as required for consistency, we choose $\rho_0 = 7.3 \times 10^{-12}$. Finally, we normalize the scale factor so that $a(t_0) := a_0 = 1$. Figure 1 shows the usual inflationary behavior, with $K = 0$. Expanding the timeline to earlier times, $t < t_0$, would exhibit a singularity, where the scale factor tends to zero. At the time of reheating t_{RH}, the parameter w starts to oscillate, as expected.

A positively curved Universe filled with an inflaton field can avoid the Big Bang singularity if the curvature density $\rho_K := 3K/(8\pi a^2)$ is strong enough to compensate for the scalar field density at some point. This is basically due to the form of the Friedmann equation (3.2), where the curvature density appears with the opposite sign (for a positive curvature) than the field density ρ. Furthermore, during inflation one has $w \approx -1$, which means that matter density is evolving as $\rho \propto a^0$ whereas $\rho_K \propto a^{-2}$. Hence, there should be have been a time in the past when $H^2 = \frac{8\pi}{3}(\rho - \rho_K)$ vanishes, which represents a bounce. However, close to the singularity, the parameter w is rising when going backward in time and if its value reaches $w = -1/3$ before both densities cancel each other, the Big Bang scenario is inevitable. This

JCAP01(2021)018

is analytically obvious as, if $w \geq -1/3$, the inflaton density varies faster than the curvature one. If the curvature density is high enough at t_0, there will be a time in the past $t_B < t_0$, where $\rho_K(t_B) = \rho$.

The strategy of this study should now be clear: we basically fix a positive curvature for the spatial sections of the Universe and "impose" (as suggested by countless arguments) the existence of an inflationary stage. When evolving the equations of motion toward the past, a bounce takes place. Of course, if one thinks from the state of the Universe in the contracting branch, it is well known that the bounce will *not* be dynamically favored. The strong potential energy domination required for the (contracting) quasi-de Sitter stage to emerge, which is itself required for the bounce, is possible only for a small fraction of the phase space. We believe that two different questions should be disentangled. One is to figure out what the past of the Universe was. Taking into account what we know, the answer provided here makes sense ans strongly suggest the existence of a bounce. Another one is to try to understand why things happened in this way. We do not address here this second question which is extremely complicated and often ill-defined, in particular because of problems appearing when trying to deal with a reliable measure in cosmology. This issue is discussed in [3]. We anyway stress here that, whatever its probability, the bouncing trajectory is indeed possible and even likely, if we think backward in time from our current knowledge. The instability with which it is associated when thinking forward in time [24] is a very relevant question, but distinct from the viewpoint adopted here.

To ivestigate the likelihood of a bouncing trajectory when going backward in time, we consider the interval of K values leading to this scenario. A first upper limit for the curvature comes from the consistency of the Friedmann equation. Indeed, since the left-hand side of eq. (3.2) is squared, the right-hand side must be positive. Taking this condition at the time t_0 leads to the requirement

$$K \leq \frac{8\pi}{3}\rho_0 a_0^2 \approx 6.1 \times 10^{-11}. \tag{3.8}$$

The lower bound to the curvature parameter is set by requiring the existence of a bounce instead of the usual singularity. As the non-linearity of the equations of motion makes the analytical study of w with respect to a tedious, we have turned to numerical computations to determine the bound. The numerical study leads to the lower limit

$$1.9 \times 10^{-11} \lesssim K. \tag{3.9}$$

Importantly, the values given by eqs. (3.8) and (3.9) are close to one another. This interval is only valid when we take $w(t_0) = -1$. If one sets $w > -1$, it is possible to generate a longer period of inflation, before the initial conditions, as we will see later on.

A range of possible values for K leading to a bouncing scenario has therefore been found, when considering the initial conditions stated above. As expected, different values of K lead to different bouncing solutions. In particular, the duration of the pre-bounce deflation is impacted by the strength of the curvature at the bounce. This is an important feature. In figure 2, we have drawn the deflation period for different values of K. The weaker the curvature at the initial time $t_0 = 0$, the shorter the deflation period. This raises an interesting point. One of the appealing features of this bouncing scenario is to avoid the Big Bang singularity. However, if the deflation period was too brief, one will inevitably face another singularity *in the past*. The cosmological constant will indeed never dominate over the curvature density when going backward in time before the bounce. Another bounce (at

JCAP01(2021)018

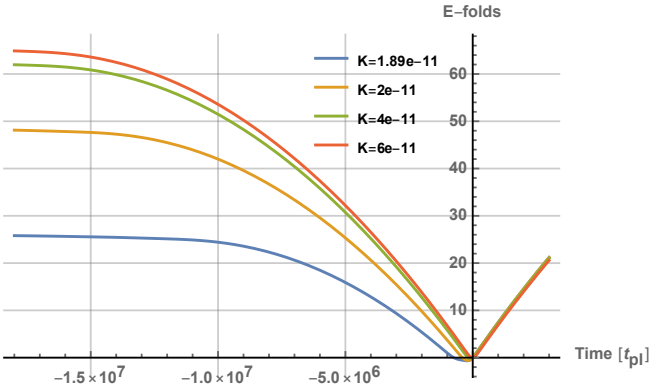


Figure 2. Deflation behavior (e-folds) for different values of the curvature K.

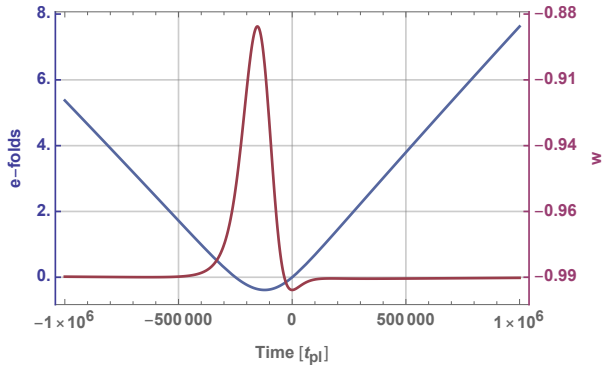


Figure 3. Bouncing scenario with $K = 3 \times 10^{-11}$. *In blue:* scale factor (variation measured in e-folds) close to the bounce; *In purple:* $w = p/\rho$.

low density) will take place, leading to a contraction — still thinking backward in time — and to a singularity when going further to the past. This might be evaded by selecting very specific conditions, but the argument cannot be meaningfully repeated for each contracting phase. A long enough deflation is therefore mandatory so that the cosmological constant protects the Universe from a re-contraction in the remote past. Figure 2 shows that this is possible. There also exist scenarios where the period of deflation is longer than the period of inflation. One can indeed notice on figure 3 that the behavior of w is not symmetric around the bounce, which causes the difference in duration between deflation and inflation. If one sets w slightly over -1 at the initial time t_0, it is possible to flip the behavior of w and to obtain a longer period of deflation. This effect can be seen in figures 4 and 5.

It is obviously important to check that the allowed interval for K, corresponding to a bouncing scenario, is consistent with experimental data. Having a concrete answer to this question is however difficult. Since the experimental curvature is measured in the contemporary Universe, one needs to know the number of e-folds from the beginning of inflation to the present time to set the initial value which, itself, influences the dynamics. Some specific initial conditions were suggested in [25, 26], where the effect of curvature during inflation was also studied (but ignoring the presence of a bounce). In particular, the time $t_\star$ at which the modes $k = 0.05\,\mathrm{Mpc}^{-1}$ exited the horizon is used to set the initial conditions. If one normalizes the scale factor at the time $t_\star$ and uses $\Omega_{k,0} = -0.044$ [19], the resulting curva-

JCAP01(2021)018

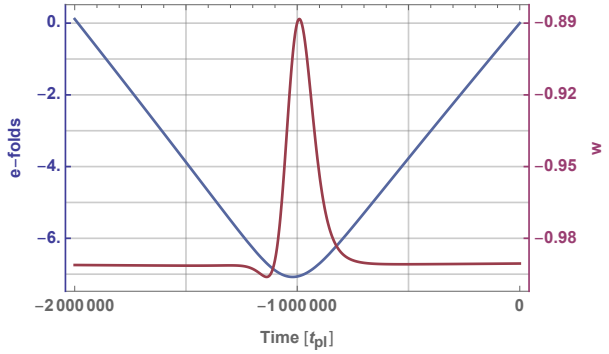


Figure 4. Bouncing scenario with the experimentally valid curvature and inflation parameters. *In blue:* scale factor (variation measured in e-folds) close to the bounce; *In purple:* $w = p/\rho$.

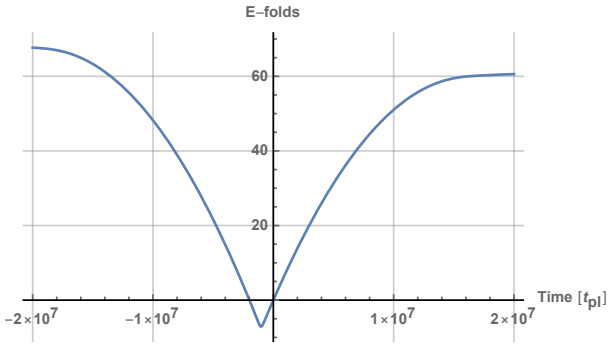


Figure 5. Deflation and inflation with experimentally valid curvature and inflation parameters at time $t_\star = 0$.

ture is $\bar{K} \approx 5 \times 10^{-17}$. The other parameters, such as the density and w, are also different from what we used. Even though the density is slightly less than what was used earlier, leading to a number of e-folds $N \approx 60$, the experimental curvature $\bar{K}$ is way lower than the range we previously calculated. However the initial conditions of [25, 26] are not set at the onset of inflation, but at a later time when $w > -1$. Due to the difference in evolution between ρ and ρ_K, there is no inconsistency. We thus rely on numerical computations to check the plausibility of the bouncing scenario. Simulating a long period of inflation when going backward in time is unstable, even in flat space. Very fine-tuned initial conditions can lead to a long period of inflation in the past, but in most cases, inflation stops abruptly in a singularity. Varying the values of $\phi_\star$ and $\phi'_\star$ given in [25, 26] within a 1% interval does indeed lead to bouncing scenarios with a long enough inflation period, as shown in figures 4 and 5. However, as explained before and as pointed out in [3], the interval considered here for K is anyway compatible with data by construction of the model. It indeed leads to the amount of inflation precisely needed for the current curvature to be given by the value estimated in [19].

4 Primordial power spectrum for tensor modes

The calculation of perturbations in a closed Universe has been considered in many studies. Interesting results were recently derived, e.g. in [27–29], to cite only a few. We shall here focus

JCAP01(2021)018

specifically on the approach developed in [25, 26]. One considers a family of metrics $g_{\mu\nu}(\varepsilon)$:

$$g_{\mu\nu}(\varepsilon) = a^2(\eta)\left(\bar{g}_{\mu\nu} + \varepsilon h_{\mu\nu}\right) , \tag{4.1}$$

where $a^2\bar{g}_{\mu\nu}$ is the background metric, $h_{\mu\nu}$ is a linear perturbation, ε is a small parameter, and η is the conformal time. The relevant spatial components read:

$$\bar{\boldsymbol{\nabla}}^2 h_{ij}^{TT} - 2\frac{a'}{a}h_{ij}^{'TT} - \frac{2}{r_o^2}h_{ij}^{TT} = 0\,, \tag{4.2}$$

where the prime denotes here a derivative with respect to η, $r_o^2 = 1/K$, and $i, j = 1, 2, 3$. Instead of expanding the equation of motion in Fourier space, as usually done, one expands the perturbations on the tensor hyperspherical harmonics $\mathcal{Q}_{ij}^{nlm,s}(\chi, \theta, \phi)$ [30]. They are eigenfunctions of the Laplacian operator D^2:

$$D^2\mathcal{Q}_{ij}^{nlm,s}(\chi,\theta,\phi) = -\frac{(n^2-3)}{r_o^2}\mathcal{Q}_{ij}^{nlm,s}(\chi,\theta,\phi)\,. \tag{4.3}$$

Importantly, the factor $(n^2 - 3)/r_o^2$ plays a role equivalent to the one of the wavenumber squared k^2 in flat space [31]. Rescaling the coefficients of the development of h (we skip here the integer labels for simplicity) by $\mu = a(\eta)h$ and promoting the resulting field to be an operator, one is led for the coefficients of the creation and annihilation operators to:

$$e_n'' + \left(\frac{n^2-1}{r_o^2} - \frac{a''}{a}\right)e_n = 0\,, \tag{4.4}$$

with the normalization

$$ee^{\star\prime} - e'e^{\star} = i\ . \tag{4.5}$$

The main effects of curvature are therefore twofolds. First, the curvature obviously changes the background evolution which leaves a footprint on the perturbations. Second, it discretizes the effective wavenumber.

Setting initial conditions for the perturbations is a key question when evaluating power spectra. The problem is known for being very difficult in bouncing models. In particular the case of scalar perturbations is problematic because the term $z''(\eta)/z(\eta)$ entering the harmonic oscillator equation does *not* generically tends to zero in the remote past (see, e.g., [32]). This means that it is hard to disentangle effects specifically due to the bounce from effects associated with the lack of a unique privileged vacuum. In addition, the equation of motion is usually very complicated and the relation between z and a is non-trivial and depends on the potential. For those reasons, we focus here on tensor modes.

In a bouncing scenario, there is *stricto sensu* no possibility to define an "initial time". In the specific context of loop quantum cosmology, the point of view that the bounce time should be chosen to set initial conditions was advocated, e.g., in [33], whereas the opposite vision was proposed, e.g., in [34].

In the set-up considered here, there is necessarily a deflation stage before the bounce. This is obviously necessary for the bounce to take place and this is anyway necessary for the curvature to dominate: without deflation, any amount of (pre-bounce) matter or radiation density would grow faster than curvature. This deflation stage makes it impossible to set initial conditions in the past as $a''(\eta)/a(\eta)$ does not anymore decrease when going backward

JCAP01(2021)018

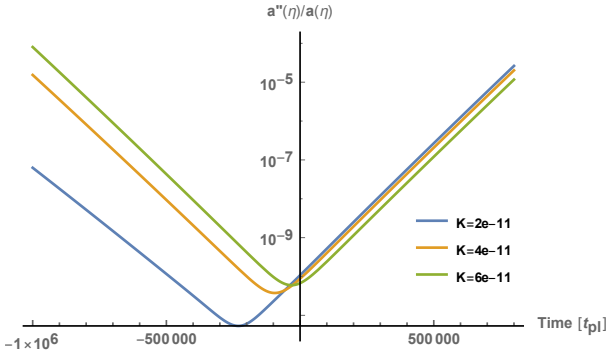


Figure 6. Potential of the tensor modes for different values of the curvature.

in time before the bounce and no Bunch-Davis-like vacuum is approached. However, the $a''(\eta)/a(\eta)$ term is very small close to the bounce and this selects, at least at the heuristic level, a preferred time to set initial conditions. The behavior of $a''(\eta)/a(\eta)$ is shown for different values of K in figure 6.

Considering eq. (4.4), one can recognize a harmonic oscillator with a time-dependent frequency

$$\omega_n^2(\eta) = \frac{n^2-1}{r_o^2} - \frac{a''(\eta)}{a(\eta)} := k_{\rm eff}(n) - \frac{a''(\eta)}{a(\eta)}.$$

In our case, since $r_o^2 \sim 10^{10}$ and $a''(t_B)/a(t_B) \sim 10^{-10}$, one cannot use a rigorous Minkowski vacuum $k_{\rm eff} \gg a''/a$ at the initial time t_B for small discret wavenumbers n. As in [26], we choose an instantaneous vacuum state that minimizes the Hamiltonian at the initial time. In order for waves to propagate with a positive frequency and using the normalization of eq. (4.5), we can set the basic mode functions at t_B as

$$e_n(t_B) = \frac{1}{\sqrt{2\omega_n(t_B)}}, \tag{4.6}$$

$$e_n'(t_B) = -i\sqrt{\frac{\omega_n(t_B)}{2}}. \tag{4.7}$$

In order to compute the primordial power spectrum for tensor modes $\mathcal{P}_T(n)$, we will proceed in the usual way. Starting from the initial time t_B, we simulate the mode functions $e_n(t)$ up to t_e, where $k_{\rm eff} = a(t_e)H(t_e)$. The primordial power spectrum can then be calculated [35]:

$$\mathcal{P}_T(n) = \frac{32k_{\rm eff}^3}{\pi}\left|\frac{e_n}{a}\right|^2\Big|_{t=t_e}. \tag{4.8}$$

The results of this analysis for different values of K are shown in figure 7. The power spectra for different values of K are nearly parallel for $n > 20$. As the initial time when we set the initial conditions varies with respect to K, the amplitude of the power spectra differs slightly. More importantly, one can notice that the curvature significantly affects the large-scale modes. They are damped due to the effect of the closure of space, which is in concordance with the result of [26], where the bounce itself is however not considered.

It is interesting to try to figure out whether the non-trivial effects appearing in the low-$k_{\rm eff}$ part of the power spectrum are observable in the cosmological microwave background (CMB). To this aim, one needs to relate the comoving wavenumber $k_c(t_r) \equiv k_{\rm eff}(t_r)$

JCAP01(2021)018

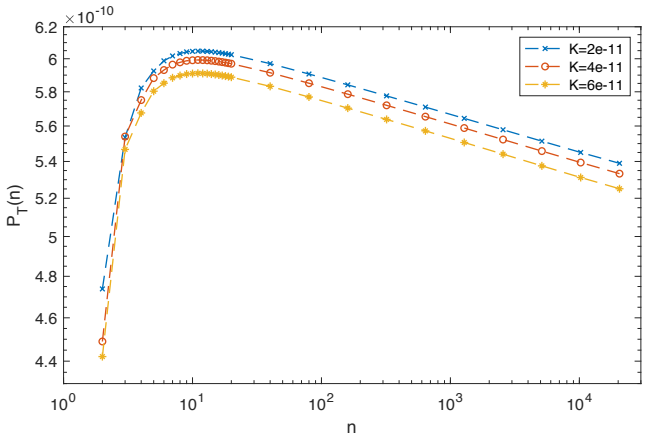


Figure 7. Primordial power spectrum for the tensor modes for different values of the curvature. The discrete wavenumber n has been calculated with $k^2 = (n^2 - 1)\, 2 \times 10^{-11}$.

with its physical counterpart $k_p(t_r)$ at the time of recombination t_r. Using $N \approx 65$ and $T_{RH} \approx 10^{16}\,\mathrm{GeV}$ one can get the orders of magnitude. The physical wavenumber observed in the CMB are of the order $\bar{k}_p \approx 0.05\,\mathrm{Mpc}^{-1} \approx 3 \times 10^{-59}$ and the scale factor at recombination can be estimated using $a(t_r) \approx e^N \times T_{RH}/T_r$, where $T_r \approx 0.3\,\mathrm{eV}$ [16]. Straightforwardly, it follows that $\bar{k}_c \approx 2 \times 10^{-5}$ and with a curvature constant of the order $K \sim 10^{-11}$, we obtain the discrete comoving wavenumber of the CMB, $\bar{n} \approx 4$. This is just in the range where the dampening effect is observable. This result is however very sensitive to the number of e-folds. As the tensor modes have not yet been observed, we leave the accurate simulation for future studies. This work however shows that some features of the curvature bounce might be observable thanks to the limited number of inflationary e-folds inherent to this model.

5 Curvature bounce with different potentials

The massive scalar field is a good toy model to study the inflationary period, but it is disfavored by recent data, relatively to other more realistic potentials, such as the Starobinsky one [36, 37]. It is therefore meaningful to study the cosmological background behavior in the presence of spatial curvature, as well as the likelihood of a bouncing scenario, with such a potential. In addition, a flat potential is expected to make the bounce more "likely" when starting from the contracting phase. Another flat potential especially designed to this aim was also considered in [38].

The Starobinsky potential is of the form

$$V(\phi) = \frac{3M^2}{32\pi}\left(1 - e^{-\sqrt{16\pi/3}\phi}\right)^2, \tag{5.1}$$

where M is a mass scale parameter. A realistic value of the parameter M can be determined following the reasoning of [39], that is using the slow-roll parameters as well as the spectral index n_s and the amplitude A_s. The potential (5.1) was also studied in [26], where no bouncing scenario was encountered. However, following the same strategy as we did before, i.e. varying the initial conditions for the scalar field at the initial time $t_\star$ in a small interval, one can generate a sufficiently long period of inflation prior to the initial time to have a curvature bounce. This shows, as expected, a high sensitivity to initial conditions.

JCAP01(2021)018

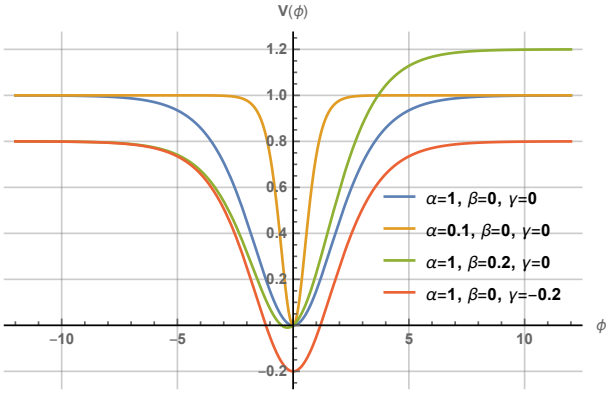


Figure 8. Scalar potential (5.2) for different values of the parameters α, β and γ, where we set $V_0 = 1$.

It is extremely difficult to have a quantitative estimate for the "probability of the bounce" for a given potential. It is well known (see, e.g., the discussion in [23]) that depending on the chosen measure and selected surface for the initial conditions one can be led to completely opposite conclusions. This is a lively debate in cosmology [40]. Nevertheless, it is possible to get a vague qualitative idea of the likelihood of the bounce scenario for the Starobinsky potential compared to the massive scalar field case. We set the initial conditions at time t_0, the onset of inflation, where $w = -1$ and the initial density ρ_0 is chosen such that we have $N \approx 65$ e-folds. In this new case, this means $\rho_0 \approx 1.6 \times 10^{-13}$. As before, one can find a range of possible spatial curvature K such that a bounce takes place. The upper bound is found using (3.8), leading to $K < 13 \times 10^{-13}$, and using numerical computations we find the lower bound $K > 1.1 \times 10^{-13}$. Relatively to the density, the range of curvature values is indeed larger in the case of a Starobinsky potential than for the massive scalar field. It, however, remains quite narrow. The behavior of the background is similar to the previous case.

A potential of the form

$$V(\phi) = V_0 \left(\tanh^2 \left[\frac{\phi}{\sqrt{6\alpha}} \right] + \beta \tanh \left[\frac{\phi}{\sqrt{6\alpha}} \right] + \gamma \right), \tag{5.2}$$

where $V_0 > 0$, $\alpha > 0$, $-1 < \beta < 1$ and $-1 < \gamma \leq 0$, was proposed in [38] to make the bounce "natural". The effect of varying the parameters is illustrated in figure 8. As one can see, α controls the steepness of the drop at $\phi = 0$, β changes the symmetry of the potential and γ sets the amplitude. We have indeed checked that some values of the parameters, e.g. $\gamma = \beta = 0$ and $V_0 \approx K$, do indeed systematically ensure a bounce. This is an important way to circumvent the apparent fine-tuning of the model when thinking forward in time. In principle, if this bouncing scenario were established, this might even allow one to select some families of potentials compatible with a reasonable probability of occurence of the bounce.

6 Conclusion

In this article, we have shown that under the experimentally valid assumption of a positively curved space (and of an inflationary stage in the past), a bounce might have naturally taken place. This basically fixes the number of e-folds — which can be anything above 65 in standard cosmology — and the reheating temperature. This removes the Big Bang singularity without any exotic physics. The view adopted here is the one of a historian who

JCAP01(2021)018

tries to figure out what was the past knowing what we know. We have not addressed the question of determining *why* the Universe might have followed this specific trajectory, which is a different question.

It should be noticed that the model is consistent in the sense that the density never approaches the Planck density. It is therefore legitimate to neglect quantum gravity effects. The physical size of the Universe also remains much larger than the Planck length, even at the bounce time. Both from the density and from the size points of view, the quantum geometry regime is never reached.

We have studied in details the way in which the background behavior depends on the contingent parameters and on the inflaton potential. We have also calculated the primordial tensor power spectrum and shown that footprints of the curvature bounce might be observable at large scales. We have checked that some potentials make the bounce apparently more "natural".

Many open questions remain to be addressed in the future. The issue of fine-tuning and the backreaction effect of fluctuations should obviously be considered seriously. Both because of the measurement problem and because of possible anthropic considerations those points are clearly difficult to answer. The primordial scalar power spectrum should also be calculated. The main trend — that is a suppression of power in the infrared part — is expected to be the same than for tensor modes. This should make the agreement with data even better (although not significantly, due to the cosmic variance) than in the usual model. In the spirit of [41, 42], the question of non-gaussianities should be addressed. Our predictions should also be confronted to the ones obtained in [43] for a symmetric bounce.

References

[1] J. Martin and P. Peter, *On the causality argument in bouncing cosmologies*, *Phys. Rev. Lett.* **92** (2004) 061301 [astro-ph/0312488] [INSPIRE].

[2] J.-P. Uzan, U. Kirchner and G.F.R. Ellis, *WMAP data and the curvature of space*, *Mon. Not. Roy. Astron. Soc.* **344** (2003) L65 [astro-ph/0302597] [INSPIRE].

[3] A. Barrau, *A pure general relativistic non-singular bouncing origin for the Universe*, *Eur. Phys. J. C* **80** (2020) 579 [arXiv:2005.04693] [INSPIRE].

[4] D. Battefeld and P. Peter, *A Critical Review of Classical Bouncing Cosmologies*, *Phys. Rept.* **571** (2015) 1 [arXiv:1406.2790] [INSPIRE].

[5] M. Lilley and P. Peter, *Bouncing alternatives to inflation*, *Comptes Rendus Physique* **16** (2015) 1038 [arXiv:1503.06578] [INSPIRE].

[6] R. Brandenberger and P. Peter, *Bouncing Cosmologies: Progress and Problems*, *Found. Phys.* **47** (2017) 797 [arXiv:1603.05834] [INSPIRE].

[7] Y.-S. Piao, B. Feng and X.-m. Zhang, *Suppressing CMB quadrupole with a bounce from contracting phase to inflation*, *Phys. Rev. D* **69** (2004) 103520 [hep-th/0310206] [INSPIRE].

[8] Y.-F. Cai, T. Qiu, J.-Q. Xia, H. Li and X. Zhang, *A Model Of Inflationary Cosmology Without Singularity*, *Phys. Rev. D* **79** (2009) 021303 [arXiv:0808.0819] [INSPIRE].

[9] Z.-G. Liu, Z.-K. Guo and Y.-S. Piao, *Obtaining the CMB anomalies with a bounce from the contracting phase to inflation*, *Phys. Rev. D* **88** (2013) 063539 [arXiv:1304.6527] [INSPIRE].

[10] Y. Wan, T. Qiu, F.P. Huang, Y.-F. Cai, H. Li and X. Zhang, *Bounce Inflation Cosmology with Standard Model Higgs Boson*, *JCAP* **12** (2015) 019 [arXiv:1509.08772] [INSPIRE].

JCAP01(2021)018

[11] S. Ni, H. Li, T. Qiu, W. Zheng and X. Zhang, *Probing signatures of bounce inflation with current observations*, *Eur. Phys. J. C* **78** (2018) 608 [arXiv:1707.05570] [INSPIRE].

[12] S.F. Bramberger and J.-L. Lehners, *Nonsingular bounces catalyzed by dark energy*, *Phys. Rev. D* **99** (2019) 123523 [arXiv:1901.10198] [INSPIRE].

[13] A. Anabalón, S.F. Bramberger and J.-L. Lehners, *Kerr-NUT-de Sitter as an Inhomogeneous Non-Singular Bouncing Cosmology*, *JHEP* **09** (2019) 096 [arXiv:1904.07285] [INSPIRE].

[14] A. Ashtekar and D. Sloan, *Loop quantum cosmology and slow roll inflation*, *Phys. Lett. B* **694** (2011) 108 [arXiv:0912.4093] [INSPIRE].

[15] K. Martineau, A. Barrau and S. Schander, *Detailed investigation of the duration of inflation in loop quantum cosmology for a Bianchi-I universe with different inflaton potentials and initial conditions*, *Phys. Rev. D* **95** (2017) 083507 [arXiv:1701.02703] [INSPIRE].

[16] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [arXiv:1807.06209] [INSPIRE].

[17] C.-G. Park and B. Ratra, *Using the tilted flat-ΛCDM and the untilted non-flat ΛCDM inflation models to measure cosmological parameters from a compilation of observational data*, *Astrophys. J.* **882** (2019) 158 [arXiv:1801.00213] [INSPIRE].

[18] W. Handley, *Curvature tension: evidence for a closed universe*, arXiv:1908.09139 [INSPIRE].

[19] E. Di Valentino, A. Melchiorri and J. Silk, *Planck evidence for a closed Universe and a possible crisis for cosmology*, *Nature Astron.* **4** (2019) 196 [arXiv:1911.02087] [INSPIRE].

[20] G. Efstathiou and S. Gratton, *The evidence for a spatially flat Universe*, *Mon. Not. Roy. Astron. Soc.* **496** (2020) L91 [arXiv:2002.06892] [INSPIRE].

[21] M.A.G. Garcia, K. Kaneta, Y. Mambrini and K.A. Olive, *Reheating and Post-inflationary Production of Dark Matter*, *Phys. Rev. D* **101** (2020) 123507 [arXiv:2004.08404] [INSPIRE].

[22] L. Linsefors and A. Barrau, *Duration of inflation and conditions at the bounce as a prediction of effective isotropic loop quantum cosmology*, *Phys. Rev. D* **87** (2013) 123509 [arXiv:1301.1264] [INSPIRE].

[23] B. Bolliet, A. Barrau, K. Martineau and F. Moulin, *Some Clarifications on the Duration of Inflation in Loop Quantum Cosmology*, *Class. Quant. Grav.* **34** (2017) 145003 [arXiv:1701.02282] [INSPIRE].

[24] D.N. Page, *A Fractal Set of Perpetually Bouncing Universes?*, *Class. Quant. Grav.* **1** (1984) 417 [INSPIRE].

[25] B. Bonga, B. Gupt and N. Yokomizo, *Inflation in the closed FLRW model and the CMB*, *JCAP* **10** (2016) 031 [arXiv:1605.07556] [INSPIRE].

[26] B. Bonga, B. Gupt and N. Yokomizo, *Tensor perturbations during inflation in a spatially closed Universe*, *JCAP* **05** (2017) 021 [arXiv:1612.07281] [INSPIRE].

[27] J. Haro, *Bouncing cosmologies in geometries with positively curved spatial sections*, *Phys. Lett. B* **760** (2016) 605 [arXiv:1511.05048] [INSPIRE].

[28] A.A. Asgari, A.H. Abbassi and J. Khodagholizadeh, *Curvature and topology dependency of the cosmological spectra*, *Results Phys.* **12** (2019) 2225 [arXiv:1801.01912] [INSPIRE].

[29] S. Akama and T. Kobayashi, *General theory of cosmological perturbations in open and closed universes from the Horndeski action*, *Phys. Rev. D* **99** (2019) 043522 [arXiv:1810.01863] [INSPIRE].

[30] U.H. Gerlach and U.K. Sengupta, *Homogeneous Collapsing Star: Tensor and Vector Harmonics for Matter and Field Asymmetries*, *Phys. Rev. D* **18** (1978) 1773 [INSPIRE].

JCAP01(2021)018

[31] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [astro-ph/9911177] [INSPIRE].

[32] A. Barrau, P. Jamet, K. Martineau and F. Moulin, *Scalar spectra of primordial perturbations in loop quantum cosmology*, *Phys. Rev. D* **98** (2018) 086003 [arXiv:1807.06047] [INSPIRE].

[33] I. Agullo, A. Ashtekar and W. Nelson, *The pre-inflationary dynamics of loop quantum cosmology: Confronting quantum gravity with observations*, *Class. Quant. Grav.* **30** (2013) 085014 [arXiv:1302.0254] [INSPIRE].

[34] S. Schander, A. Barrau, B. Bolliet, L. Linsefors, J. Mielczarek and J. Grain, *Primordial scalar power spectrum from the Euclidean Big Bounce*, *Phys. Rev. D* **93** (2016) 023531 [arXiv:1508.06786] [INSPIRE].

[35] K. Martineau, *Quelques aspects de cosmologie et de physique des trous noirs en gravitation quantique à boucles*, Ph.D. Thesis, Laboratoire de Physique Subatomique et de Cosmologie, France, (2019).

[36] PLANCK collaboration, *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10 [arXiv:1807.06211] [INSPIRE].

[37] A.A. Starobinsky, S. Tsujikawa and J. Yokoyama, *Cosmological perturbations from multifield inflation in generalized Einstein theories*, *Nucl. Phys. B* **610** (2001) 383 [astro-ph/0107555] [INSPIRE].

[38] H. Matsui, F. Takahashi and T. Terada, *Non-singular bouncing cosmology with positive spatial curvature and flat scalar potential*, *Phys. Lett. B* **795** (2019) 152 [arXiv:1904.12312] [INSPIRE].

[39] B. Bonga and B. Gupt, *Phenomenological investigation of a quantum gravity extension of inflation with the Starobinsky potential*, *Phys. Rev. D* **93** (2016) 063513 [arXiv:1510.04896] [INSPIRE].

[40] J.S. Schiffrin and R.M. Wald, *Measure and Probability in Cosmology*, *Phys. Rev. D* **86** (2012) 023521 [arXiv:1202.1818] [INSPIRE].

[41] X. Gao, M. Lilley and P. Peter, *Production of non-Gaussianities through a positive spatial curvature bouncing phase*, *JCAP* **07** (2014) 010 [arXiv:1403.7958] [INSPIRE].

[42] X. Gao, M. Lilley and P. Peter, *Non-Gaussianity excess problem in classical bouncing cosmologies*, *Phys. Rev. D* **91** (2015) 023516 [arXiv:1406.4119] [INSPIRE].

[43] M. Lilley, L. Lorenz and S. Clesse, *Observational signatures of a non-singular bouncing cosmology*, *JCAP* **06** (2011) 004 [arXiv:1104.3494] [INSPIRE].

JCAP01(2021)018

Eur. Phys. J. C (2022) 82:775
https://doi.org/10.1140/epjc/s10052-022-10745-8

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Detailed analysis of the curvature bounce: background dynamics and imprints in the CMB

Cyril Renevey, Aurélien Barrau[a], Killian Martineau

Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53, avenue des Martyrs, 38026 Grenoble cedex, France

Received: 18 June 2022 / Accepted: 22 August 2022

Abstract If the spatial sections of the Universe are positively curved, extrapolating the inflationary stage backward in time inevitably leads to a classical bounce. This simple scenario, non-singular and free of exotic physics, deserves to be investigated in details. The background dynamics exhibits interesting features and is shown to be mostly insensitive to initial conditions as long as observational consequences are considered. The primordial scalar power spectrum is explicitly computed, for different inflaton potentials, and the subsequent CMB temperature anisotropies are calculated. The results are compatible with current measurements. Some deviations with respect to the standard paradigm can however appear at large scales and we carefully disentangle what is associated with the vacuum choice with what is more fundamentally due to the bounce itself.

1 Introduction

The question of the shape of the spatial sections of the Universe is an old one. It is still unanswered but it has attracted a lot of attention recently, based on the refined observations of the cosmic microwave background (CMB). The latest data released by the Planck collaboration [1] might favor a positively curved universe, described by a curvature density $\Omega_K = -0.044$, with a high confidence level [1–4]. Although quite convincing when using CMB data alone, this result is in tension with other measurements, such as the baryonic acoustic oscillations. It is also dependent upon specific statistical priors that can be questioned [5]. A balanced discussion can be found in [6–8]. Nonetheless, even ignoring the claims of [4], the possibility that the universe is positively curved is worth being phenomenologically considered as a possible situation. Especially when taking into account, as we shall see, that this might cure naturally the initial singularity problem. Not to mention that some – speculative but reasonable – arguments from quantum gravity also favor this possibility (a finite space is a natural infrared regulator).

It is believed in the standard cosmological paradigm that at its earliest stage the universe grew exponentially fast in a quasi-de Sitter phase. While not fully confirmed, the theory of inflation is supported by strong observational evidences [9], the most obvious being the slightly red-tilted spectrum of primordial fluctuations. Strong constraints on non-Gaussianities in the CMB [10] favor a single field scenario.

In the inflationary model, the temperature anisotropies observed in the CMB are explained by quantum fluctuations of the inflaton scalar field. These initial fluctuations can be evolved using the theory of gauge-invariant linear perturbations [11] and the primordial power spectrum can then be unambiguously calculated. The study of scalar and tensor perturbations was carried out in the case of a closed universe in [12,13]. It was concluded that a smaller-than-usual power is to be expected at very large scales.

Although straightforward to show, it is often forgotten that a de Sitter space-time with closed spatial sections evades the big bang singularity and instead, naturally leads to a bounce. The possible implementations of this scenario were studied in [14–16]. The singular origin of our universe is naturally regularized thanks to the positive curvature combined with inflation. A further numerical investigation was carried out in [17]. Quite impressively, the usual claims about the need of a quantum (or modified) theory of gravity to escape the unavoidable singularity are simply contradicted using only the hypotheses of the standard cosmological scenario (if space is positively curved). This happens without any exotic physics. In [17], we have studied the primordial tensor power spectrum with quantum fluctuations originating prior to the bounce. No noticeable difference was found when compared

[a] e-mail: barrau@in2p3.fr (corresponding author)

Published online: 01 September 2022

to the no-bounce version of [13]. However, the equation of motion for the scalar perturbations being very different from the one of the tensor modes, we must investigate the scalar sector to predict reliably the possible imprint of the model on the CMB temperature anisotropies. This is the main goal of this work which builds on our previous papers [16,17]. The key-point consists in disentangling the effects due to the curvature (as in [12]) from those due to the bounce itself.

This paper is organized as follows. First, we describe the behaviour of the background dynamics during the period of inflation, around the bounce, and during the pre-bounce contracting universe. We analyze in details the dependence of the background dynamics upon initial conditions. Two different inflationary models are considered : a massive scalar field – which is disfavoured by data but useful for comparisons – and the Starobinsky potential [18–20]. Then, we introduce the theory of linear perturbations in curved space and discuss the meaning of the Bunch–Davies vacuum before and after the bounce. Finally, we move on to the calculation of the primordial power spectra of scalar perturbations, considering different cases for the initial vacuum. We compare our results with previous works, mainly with [12], and with the available CMB data.

Planck units are used throughout the paper except otherwise stated.

2 Background dynamics

To describe the background behaviour, homogeneity and isotropy, together with the closure of the spatial sections, are assumed. The topology of space-time is therefore $\mathbb{R}\times\mathbb{S}^3$, where $\mathbb{S}^3$ represents a hypersphere. The spatial curvature parameter $K>0$ is related to the physical radius $r(t)$ of the 3-sphere by $r^2(t)=a^2(t)/K$, where $a(t)$ is the dimensionless scale factor. Under these hypotheses, the FLRW metric can be written as

$$\mathrm{d}s^2=-\mathrm{d}t^2+\frac{a^2(t)}{K}\left(\mathrm{d}\chi^2+\sin^2(\chi)\mathrm{d}\Omega^2\right). \tag{1}$$

The matter content of the universe is represented by a perfect fluid of density ρ and pressure p, such that the Einstein field equations lead to the Friedmann and Raychaudhury equations:

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

$$\dot{H}=-4\pi(\rho+p)+\frac{K}{a^2}, \tag{3}$$

where $H=\dot{a}/a$ is the Hubble parameter and the dot notation stands for the derivative with respect to the cosmic time. From Eq. (2), it can immediately be seen that during an inflationary period, described by a quasi-constant energy density ρ, the curvature term on the right hand side (RHS) increases when the scale factor decreases. This behaviour inevitably leads to a bounce, at the time when $H=0$, instead of a singularity. A detailed analysis of the basic ingredients of the curvature bounce is given in [16,17]. Once again, it is worth emphasizing that the curvature of the universe is not yet firmly measured and could actually be positive, negative or null. A recent study [4] claimed that the CMB measurements do favor a closed universe with $\Omega_K:=-K/(3\bar{H}^2)=-0.044$, $\bar{H}=54.4$ km/s/Mpc being the current Hubble parameter. While the validity of this result is strongly debated [5], we assume here that it is correct. Even if it is not, a positive curvature remains anyway a possible situation – somehow favored by theoretical arguments and not excluded by data – deserving attention.

During the inflationary period, the matter content is described by a scalar field ϕ. In this work, we consider two commonly used potentials for the inflaton: the quadratic and Starobinsky potentials, respectively given by

$$V(\phi)=\frac{1}{2}m^2\phi^2;\quad V(\phi)=\frac{3M^2}{32\pi}\left(1-e^{-\sqrt{16\pi/3}\phi}\right)^2, \tag{4}$$

where m and M are free parameters that can be constrained by CMB measurements, as we shall see below. The Klein-Gordon equation in the homogeneous FLRW metric given Eq.(1) reads

$$\ddot{\phi}+3H\dot{\phi}+\frac{\partial V(\phi)}{\partial\phi}=0. \tag{5}$$

The density and pressure of the inflaton field can be written as

$$\rho=\frac{1}{2}\dot{\phi}^2+V(\phi)\quad\text{and}\quad p=\frac{1}{2}\dot{\phi}^2-V(\phi). \tag{6}$$

As well known, Eqs. (2), (3) and (5) are not independent.

To perform the simulation, the initial conditions (IC) for the background are set at the bounce, which is chosen to be the origin of the time coordinate, $t=0$. This is only a matter of convenience and this does not mean that the system "looses" memory of the pre-bounce phase. At the bounce, the derivative of the scale factor vanishes, $\dot{a}(0)=0$, and we choose, without any loss of generality, to set $a(0)=1$, fully fixing the IC for the homogeneous gravitational sector. In order to set the IC for the matter sector (and to fix the free parameter), we use the measurements of the amplitude of the scalar power spectrum A_s and its running n_s. Those parameter shed light, in particular, on the cosmological behavior at the time t_*, when the pivot scale $k_*=0.05\,\mathrm{Mpc}^{-1}$ exited the horizon. At that time, the universe can be considered as flat [12] and we further assume that the slow-roll approximation is satisfied [19], i.e.

$$\epsilon_* := -\frac{\dot{H}_*}{H_*^2} \simeq \frac{1}{16\pi}\left(\frac{V_{,\phi}(\phi_*)}{V(\phi_*)}\right)^2 \ll 1, \tag{7}$$

$$\eta_* := \left.\frac{\mathrm{d}\ln\epsilon}{\mathrm{d}N}\right|_{t_*} \simeq 4\,\epsilon_* - \frac{1}{4\pi}\frac{V_{,\phi\phi}(\phi_*)}{V(\phi_*)} \ll 1, \tag{8}$$

where N is the number of e-folds and the subscript "$,\phi$" means a derivative with respect to the scalar field. Using the standard definitions of A_s and n_s, together with the field equations at the time t_*, namely

$$H_* \simeq \sqrt{\pi A_s \epsilon}, \quad n_{s*} \simeq 1 - 2\epsilon_* - \eta_*, \tag{9}$$

$$3H_*\dot{\phi}_* + V_{,\phi}(\phi_*) \simeq 0\,, \tag{10}$$

$$H_*^2 \simeq \frac{8\pi}{3}\left(\frac{\dot{\phi}_*^2}{2} + V(\phi_*)\right), \tag{11}$$

one can solve for ϕ_*, $\dot{\phi}_*$, and for the free parameter included in the definition of $V(\phi)$, either m or M. It is then possible to choose $\phi(t=0)$ and $\dot{\phi}(t=0)$ at the time of the bounce so as to satisfy conditions (9), (10) and (11) at the time t_*.

Actually, because inflation is a strong attractor, some freedom is left in choosing the IC at the time of the bounce. To better understand the physical meaning of the IC, we discuss them using the energy density of the scalar field ρ and the equation of state $w := p/\rho$, instead of ϕ and $\dot{\phi}$. The density ρ allows one to immediately read the energy scale of the considered phase, while w quantifies how far away from a pure de Sitter space-time the universe is. In practice, whatever the chosen value at the bounce $w(t=0) \in [-1, -1/3]$, inflation does occur. Should the initial value of w be higher than $-1/3$, the bounce would simply not happen. On the other hand, the energy density at the bounce $\rho(0)$ is constrained by the physical conditions at the time t_*, which in turn constrain the number of e-folds N of inflation for a given model. As shown in details in [17], when $w > -1/3$ the energy density of matter strongly dominates at the origin of our universe and the big bang singularity cannot be avoided. To be more precise, since $\mathrm{d}w/\mathrm{d}t \neq 0$ at $t=0$ and since the requirement $w < -1/3$ needs to be satisfied right before the bounce, the actual practical upper limit for $w(0)$ is slightly smaller than $-1/3$. Setting the IC for ρ and w is in fact equivalent to setting the IC for ϕ and $|\dot{\phi}|$. We therefore are also free to choose the sign of the derivative of the scalar field. The background IC are given by $\{w(0), \mathrm{sgn}(\dot{\phi}(0))\}$, $\rho(0)$ being fixed by the conditions at t_*.

Figure 1 illustrates the consequences of different choices of IC, by displaying the evolution of the scale factor during the pre-bounce phase and during the inflationary period (with the Starobinsky potential), for different choices of IC. First, one can indeed see that the choice of w and $\mathrm{sgn}(\dot{\phi})$ at the bounce has no influence on the inflationary period, as long as $\rho(0)$ is chosen accordingly. Remarkably, while the post-bounce behaviour is unaffected by the choice of IC, the pre-bounce universe, more precisely the duration of deflation, significantly is. For $\dot{\phi}(0) > 0$ the deflationary period is shortened, while for $\dot{\phi}(0) < 0$ it is lengthened. The choice of w affects the asymmetry between deflation and inflation. Since deflation (just before the bounce) and inflation (just after it) are both required by the model, the scalar field lies high on its potential at the time of the bounce. The choice $\phi(0) > 0$ can therefore be made without loss of generality. Furthermore, the derivative of the scalar field being non-vanishing at $t=0$, ϕ is not at its maximum value, but is instead slightly offset. This is the origin of the asymmetry between the length of deflation and the length of inflation. The quadratic potential leads to similar behaviours.

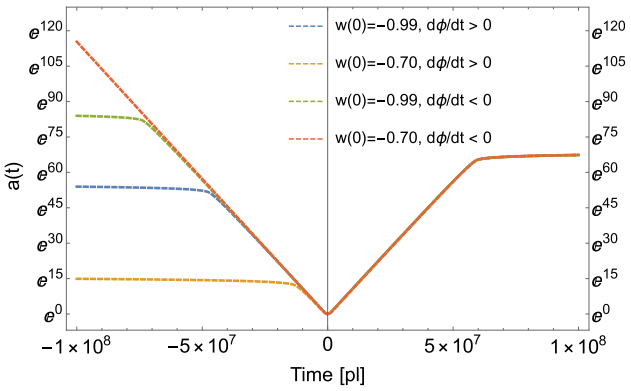


Fig. 1 Evolution of the scale factor $a(t)$ before, during and after the bounce, for different initial conditions

The critical aspect of this otherwise very appealing scenario – built only on the standard model of cosmology together with a positive spatial curvature – is obviously the deflation phase. It raises two important questions that are both physical and philosophical.

The first question is about the probability for this deflation phase to take place and last long enough. Deflation is unstable. If one thinks from the causal viewpoint of the pre-bounce phase, driving the universe to a de-Sitter contraction (the effective cosmological constant remaining positive) requires a large amount of fine-tuning. Some specific choices of potentials may overcome this difficulty [21] but, clearly, the deflationary stage generally remains a repulsor. Is this a problem? Not necessarily. The a priori probability for most phenomena around us is extraordinary small. The viewpoint adopted here is the one of an archaeologist trying to figure out what has happened taking into account the laws we know and the facts we see. Not the one of a historian trying to determine the deep causes of a strange political event. There is nothing odd with discovering phenomena whose probability of occurrence was a priori incredibly small [16]. The point is not to find an explanation for the path followed by the universe but to determine what this path might have been considering the knowledge we have.

The second question is related to the favored duration of deflation. Only a few e-folds are required for the model to work. This however raises an interesting point. If deflation was brief, it means that, when going backward in time prior to the bounce, the curvature term happened to dominate before the cosmological constant had a chance to overcome the dynamics. This implies that – still thinking backward in time – the universe was contracting again. The whole question of the singularity resolution therefore reappears. A new curvature bounce is obviously possible but still needs fine-tuning. The situation repeats itself as long as the deflation stages last only a few e-folds. On the other hand, if deflation was long enough (typically as long as inflation but the accurate value obviously depends on the unknown content of the contracting branch), the cosmological constant protects the pre-bounce phase from any re-contraction (once again in a time decreasing view). Only one bounce takes place. This might indicate that a single long period of deflation is somehow theoretically preferred. But this highly depends on arbitrary probability priors.

It is worth recalling here that, as we have shown in a previous article [16], the number of inflationary e-folds and the reheating temperature are fixed in this model. Consistency reasons impose $N \sim 70$ and $T_{RH} \sim T_{GUT}$. It is quite impressive that the number of e-folds precisely coincides with the minimum value required to account for observations and that the bounce energy scale (which is the same than the inflation scale) is substantially sub-Plankian, implying a well controlled classical behavior.

3 Linear perturbations

In order to compare the predictions of the model with the temperature anisotropies measured by Planck [1], one needs the theory of linear perturbations. In the following, we very briefly summarize the procedure for positively curved spaces [11,22]. On top of the homogeneous and isotropic background, small perturbations are added according to

$$g_{\mu\nu} \simeq \mathring{g}_{\mu\nu} + \delta g_{\mu\nu}; \quad \phi \simeq \mathring{\phi} + \delta\phi, \tag{12}$$

$\mathring{g}_{\mu\nu}$ and $\mathring{\phi}$ being the background quantities. As a symmetric rank two tensor, the metric perturbations can be decomposed into four scalar, four vector and two tensor degrees of freedom. Some of these degrees of freedom are gauge artifacts ans gauge-invariant quantities have to be constructed. In the case of scalar perturbations of the metric, this leads to the well-known Bardeen variables [23], usually denoted Φ and Ψ. The tensor degrees of freedom are naturally gauge invariant. The matter fluctuations described by $\delta\phi$ have to be connected to the metric fluctuations through the Einstein field equations. Using the perturbed version of the equations, one can show that, for a perfect fluid having no anisotropic stress at linear order, the Bardeen variables are equal: $\Phi = \Psi$. The perturbed equations of motion also give the evolution equations for the two remaining scalar perturbations, linking Φ to $\delta\phi$. The quantization of the remaining scalar degree of freedom is performed in the canonical form using the Mukhanov–Sasaki (MS) variable [24,25], which can be written as

$$v(t, \mathbf{x}) = a\left(\delta\phi + \frac{\mathring{\phi}'}{\mathcal{H}}\Phi\right), \tag{13}$$

where a prime represents the derivative with respect to the conformal time η (such that $d\eta = dt/a$) and $\mathcal{H} = a'/a$. The canonical scalar perturbation is directly related to the curvature perturbation $\mathcal{R}$, through

$$v = \frac{a\,\mathring{\phi}'}{\mathcal{H}}\mathcal{R}. \tag{14}$$

So far, we have worked in position space, where the fluctuations are functions of time and spatial coordinates $v(\eta, \mathbf{x})$. As opposed to the spatially flat case, where one describes functions of spatial coordinates using Fourier expansions, when the positive curvature is taken into account, the discrete basis of hyperspherical harmonics $Q_{nlm}(\mathbf{x})$ has to be used, since the manifold is closed. The detailed procedure is explained in [12]. In short, the hyperspherical harmonics can be separated into a radial part $f_{nl}(\chi)$ and a spherical harmonic $Y_{lm}(\theta, \varphi)$ part according to $Q_{nlm}(\mathbf{x}) = f_{nl}(\chi) \cdot Y_{lm}(\theta, \varphi)$. Using this basis, any function defined on the 3-sphere, including the MS variable v, can be expanded as

$$v(\eta, \mathbf{x}) = \sum_{n=2}^{\infty}\sum_{l=0}^{n-1}\sum_{m=-l}^{l} v_{nlm}(\eta)\,Q_{nlm}(\mathbf{x}). \tag{15}$$

The perturbed field equations describing the behaviour of the MS variable can now be written in momentum space, leading to an equation of motion for each mode v_{nlm}, namely [12,13]

$$v''_{nlm} + A_n(\eta)v'_{nlm} + B_n(\eta)v_{nlm} = 0\,, \tag{16}$$

where

$$\begin{aligned} A_n(\eta) = &\Big(32\pi a^3\dot{a}\dot{\phi}V_\phi(\phi) + 48\pi a^2\dot{a}^2\dot{\phi}^2 \\ &-8\pi a^2\dot{\phi}^2\left(8\pi a^2\left(\dot{\phi}^2 - 2V\right) + 2K\right)\Big) \\ &\times\left(2\dot{a}\left(2(n^2-4)\dot{a}^2 + 8\pi a^2\dot{\phi}^2\right)\right)^{-1}, \end{aligned} \tag{17}$$

and

$$
\begin{aligned}
B_n(\eta) = & \frac{8\pi}{\dot{a}^2\left(2\left(n^2-4\right)\dot{a}^2+8\pi a^2\dot{\phi}^2\right)} \\
& \times\left[\frac{\dot{a}^4(n^2-4)\left((n^2-1)K+a^2V_{\phi\phi}\right)}{4\pi}\right. \\
& +\left(4n^2-7\right)a^3\dot{a}^3\dot{\phi}V_\phi-\pi\frac{n^2-1}{n^2-4}a^4\dot{\phi}^4 \\
& \times\left[8\pi a^2\left(\dot{\phi}^2+2V\right)-6K\right]+(n^2-1)a^2\dot{a}^2 \\
& \times\left(-6\pi\frac{n^2-5}{n^2-4}a^2\dot{\phi}^4+4\pi a^2\dot{\phi}^2V\right. \\
& \left.+\frac{3K}{2}\dot{\phi}^2+\frac{9}{2}\dot{a}^2\dot{\phi}^2\right) \\
& +a^3\dot{a}\left[a\dot{a}\dot{\phi}^2V_{\phi\phi}+2a\dot{a}V_\phi^2\right. \\
& \left.\left.+4\pi a^2\dot{\phi}V_\phi\left(\dot{\phi}^2+2V\right)-K\dot{\phi}V_\phi\right]\right] \\
& -H\left(H+A_n(\eta)/a\right)-\frac{\ddot{a}}{a}, \qquad (18)
\end{aligned}
$$

Analogies with the flat case can easily be drawn. Usually, scalar functions are expanded using the basis $Q_{klm}(\mathbf{x}) = j_l(kr)Y_{lm}(\theta,\varphi)$, where $j_l(kr)$ are the Bessel functions and $k \in \mathbb{R}$ spans a continuous spectrum. In fact, in the limit where n is large and χ is small (that is when positively curved spatial sections are nearly flat), the radial part of the hyperspherical harmonics $f_{nl}(\chi)$ behaves as $j_l(kr)$. Moreover, when comparing the equations of motion for the scalar modes, one can identify

$$(n^2-1)K \to k^2. \qquad (19)$$

Hence, $\sqrt{(n^2-1)K}$ can be understood as the curved space counterpart of the flat space wave number k. From now on, we define the "flat case limit wave number" to be $k_n = \sqrt{(n^2-1)K}$.

The scalar perturbations are to be understood as due to quantum fluctuations taking place close to the bounce. To implement this, we follow the standard quantization techniques of cosmological perturbations, setting the initial quantum state for $v_n(\eta)$ in the so-called Bunch–Davies vacuum. It is unique and equivalent to the minimum energy state if and only if the quantum fluctuations behave as in Minkowski space-time, i.e. $A_n \to 0$ and $B_n \to k_n^2$, at the initial time t_i [26]. Those requirements are not trivial in the considered scenario, as we shall see later on. When they are fulfilled, the IC for the perturbations can be set as

$$v(t_i) = \frac{1}{k_n} \quad \& \quad v'(t_i) = -i\,\frac{k_n}{\sqrt{2}}. \qquad (20)$$

To check the validity of the Bunch–Davies vacuum approximation, the functions $A_n(t)$ for a massive inflaton

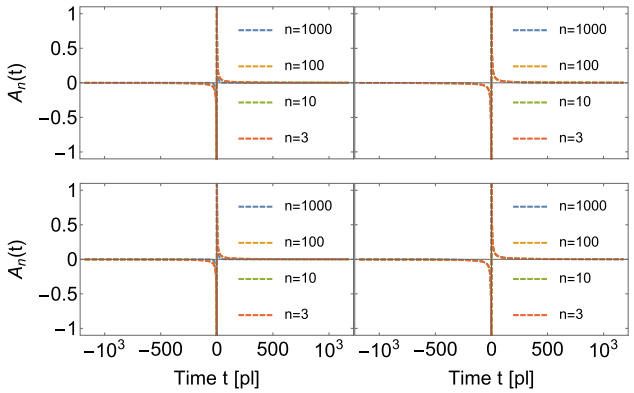


Fig. 2 Behaviour of the functions $A_n(t)$ around the bounce, $t = 0$, for different wave numbers n and initial conditions, in the case of a quadratic potential. *top left:* evolution from the IC $\{-0.99, +1\}$, *top right:* $\{-0.467, +1\}$, *bottom left:* $\{-0.99, -1\}$, *bottom right:* $\{-0.41, -1\}$

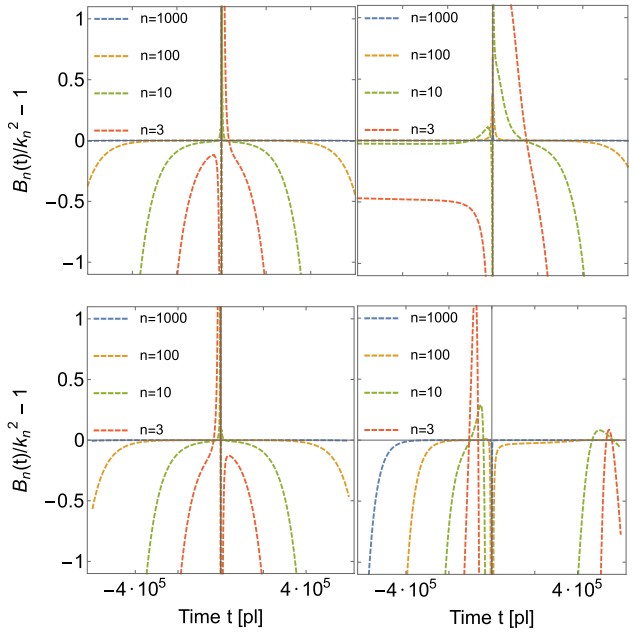


Fig. 3 Behaviour of the functions $B_n(t)$ around the bounce $t = 0$, for different wave numbers n and initial conditions, in the case of a quadratic potential. *top left:* $\{-0.99, +1\}$, *top right:* $\{-0.467, +1\}$, *bottom left:* $\{-0.99, -1\}$, *bottom right:* $\{-0.41, -1\}$

are drawn in Fig. 2, for different values of $\{w(0), \mathrm{sgn}(\dot{\phi}(0))\}$ and n. The "deviations", $B_n(t)/k_n^2 - 1$ are shown in Fig. 3. The case of a Starobinsky potential is considered in Fig. 4. We do not show the functions $A_n(t)$ in the latter case as they are similar to those of the quadratic potential. The values $w(0) = -0.467$ and $w(0) = -0.5$, respectively for the quadratic and starobinsky potentials, with $\dot{\phi}(0) > 0$, are chosen to minimize the number of e-folds of deflation. It is as low as only 1 to 2 e-folds in both cases (we recall that the bounce cannot occur without deflation). For a negative scalar derivative $\dot{\phi}(0) < 0$, the values $w(0) = -0.41$ and $w(0) = -0.5$ are chosen, leading to a very high number

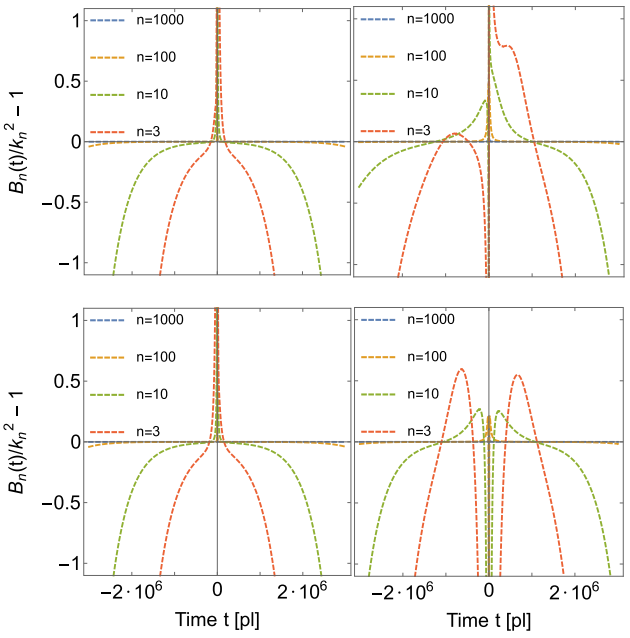


Fig. 4 Behaviour of the functions $B_n(t)$ around the bounce $t = 0$, for different wave numbers n and initial conditions, in the case of a Starobinsky potential. *top left:* evolution from the IC $\{-0.99, +1\}$, *top right:* $\{-0.5, +1\}$, *bottom left:* $\{-0.99, -1\}$, *bottom right:* $\{-0.5, -1\}$

of deflationnary e-folds: 185 and 760, respectively, for the quadratic and the Starobinsky potentials. Clearly, in all cases, the functions $A_n(t)$ and $B_n(t)$ are non-trivial and even divergent at the bounce. This is, partly, why when the fluctuations are initiated before the bounce, they might leave a significant imprint on the scalar perturbations and consequently on the scalar power spectrum. Furthermore, the functions $A_n(t)$ become very small close the bounce, therefore satisfying the first requirement for the Bunch–Davies vacuum. It should also be noticed that they do not significantly vary with respect to the wave number n and to the shape of the inflaton potential. The second requirement, $B_n(t_i) \approx k_n^2$, is harder to fulfill for small values of n, as it can be seen in Figs 3 and 4 – even more so for higher values of $w(0)$. In particular, in the case of the quadratic potential with IC $\{-0.467, +1\}$, the Minkowski vacuum can hardly be met with a precision better than 50%. Although this is *not* problematic for the model itself, it makes the clear analysis of the power spectra less obvious. For the vast majority of IC, it however remains possible to define a time t_i before or after the bounce, such that B_n is close enough to the desired value k_n^2. For larger wave numbers, e.g. $n \geq 10$, the Bunch–Davies vacuum is always very well defined for an appropriate choice of IC.

It should be stressed that, in this model, the Bunch–Davies vacuum can only be chosen for a limited amount of time, either before or after the bounce. We assume that it makes sense to select the minimum energy state when the instantaneous Minkowski vacuum is met. Although reasonable and usual, this is not a straightforward assumption.

Beyond the specific model considered in this study, the question of the origin of cosmological perturbations in bouncing models is a difficult one. For example, in the case of loop quantum gravity, some authors advocate the idea that IC should be implemented at the bounce [27], whereas others insist on putting them in the contracting branch [28]. This raises questions both on the practical side and on the conceptual side. Practically, the issue is the one of the most convenient way to proceed. Conceptually, the question is the one of causality. If time flows in the same direction before and after the bounce,[1] and if the word "initial" is meant literally, the IC should be defined before the bounce. Of course, the detailed underlying process, leading to the inflaton domination (anti-reheating) is still to be described. It is our opinion that, even at the heuristic level, it is meaningful to initiate the perturbations before the bounce, at the time when the Bunch–Davis vacuum is approached as closely as possible.

4 Primordial power spectrum and temperature anisotropies

In previous sections, we have discussed the background behaviour leading to a curvature bounce instead of the big-bang singularity and explained how the cosmological evolution depends upon the IC. We have also introduced the theory of cosmological perturbations in a closed universe and showed that quantum fluctuations can be initiated either before or after the bounce, using in all cases the Bunch–Davies vacuum. The primordial scalar power spectrum and the subsequent effects on the temperature anisotropies measured in the CMB can now be calculated for different hypotheses that will be described later.

To compute the primordial power spectrum, we first evolve the perturbations $v_n(t)$ from the initial time t_i to the reheating using the equation of motion (16). At the reheating, the curvature term is completely negligible when compared to the other fluids and one can use the standard definition of the power spectrum of the scalar perturbations, namely

$$\mathcal{P}_S(n) = \frac{k_n^3}{2\pi^2}\left|\frac{v_n}{z}\right|^2\Bigg|_{t=t_{rh}}, \tag{21}$$

t_{rh} being the reheating time.

[1] This hypothesis is not obvious. If, as suggested by Penrose [29], irreversibility is driven by processes that takes place while approaching (nearly) singular points, it could be that time – in this particular sense (as there are actually many different times at play [30]) – flows in two opposite directions from the bounce.

To calculate the spectrum of temperature anisotropies, we use the Boltzmann code **CAMB**, which is well suited to take into account a positive curvature and a discrete primordial power spectrum [31].

4.1 Post-bounce origin of structures

We start with the assumption that the anisotropies observed in the CMB originates from quantum fluctuations generated after the bounce. This situation is nearly analogous to the one considered in the work of [12] (where no bounce occurs), with the exception of the numerical value of curvature parameter ($\Omega_K = -0.005$ in their case). The specific time at which the IC for the background and for the perturbations are set is also different. One of the main results of [12] was the prediction of a decrease of the primordial power spectrum amplitude at low values of n and, consequently, a decrease of the temperature anisotropies at low multipolar numbers ℓ. In [12], the background IC are chosen at the time t_*, when the modes with $k_* = 0.005\ \mathrm{Mpc}^{-1}$ exited the horizon. The scale factor and the scalar field are then evolved both in the future, $t > t_*$, and in the past, $t < t_*$. Although fully legitimate, this approach might miss the "standard" behaviour as the slow-roll solution is not an attractor when going backward in time (which is equivalent to going forward in time in a contracting universe). This leads to instabilities (the same than those responsible for the low likelihood of a long lasting deflation in the contracting branch) and to the big bang singularity. The fluctuations are set in the Bunch–Davies vacuum before the time t_*, when all the modes are sub-Hubble. However, at that time, the requirements of the Bunch–Davies vacuum are in fact not fully satisfied for low wave numbers n and it might very well be that this is the reason for the calculated damping of the primordial power spectrum at low multipolar numbers.

To support this last argument, we have calculated the primordial scalar power spectrum for both potentials and for different values of the background IC. In each case, we select the best post-bounce time $t_i > 0$, where the requirements for a Bunch–Davies vacuum are satisfied, to initiate the quantum fluctuations. The results for the quadratic and Starobinsky potentials are exposed in Figs. 5 and 6, respectively. It can be seen that, if one chooses the most appropriate time to set the IC for the perturbations, the primordial power spectra for the scalar modes in a positively curved space are almost identical to those of a flat universe, described by as a power law $\mathcal{P}_S(k) = A_s(k/k_*)^{n_s-1}$. Even for the IC $\{-0.99, -1\}$, where the Minkowski vacuum conditions are not exactly met, Fig. 3 shows that one closely recovers the flat space power spectrum. To further support that the lack of power of the primordial power spectrum at low n found in [12] might be due to an ill-defined Bunch–Davies vacuum, one can calculate the scalar power spectrum with a sub-optimal initial time t_i.

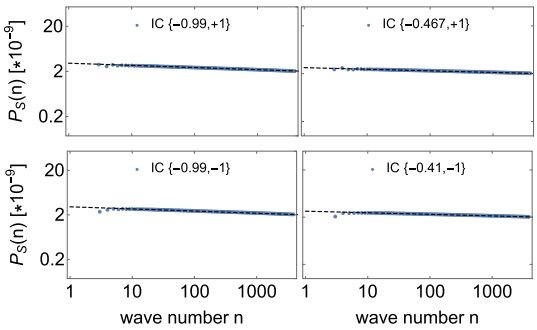


Fig. 5 Primordial scalar power spectrum for different set of background initial conditions, for a quadratic potential, and with fluctuations initiated after the bounce (blue dots). The standard power law is represented as a dashed black line. The IC for the background are: *top left* $\{-0.99, +1\}$, *top right* $\{-0.467, +1\}$, *bottom left* $\{-0.99, -1\}$, *bottom right* $\{-0.41, -1\}$

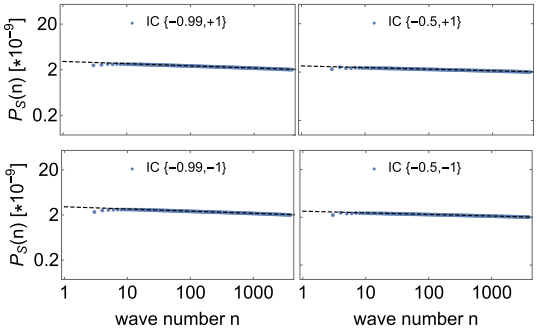


Fig. 6 Primordial scalar power spectrum for different set of background initial conditions, for a Starobinsky potential, and with fluctuations initiated after the bounce (blue dots). The standard power law is represented as a dashed black line. The IC for the background are: *top left* $\{-0.99, +1\}$, *top right* $\{-0.5, +1\}$, *bottom left* $\{-0.99, -1\}$, *bottom right* $\{-0.5, -1\}$

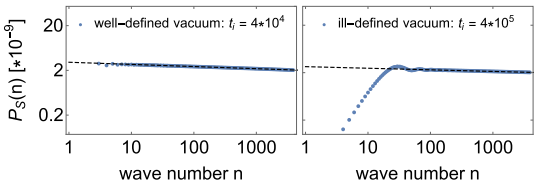


Fig. 7 Comparison between the primordial power spectra of scalar modes from a well-defined (left) and ill-defined (right) Bunch–Davies vacuum. The calculation is performed for the quadratic potential with IC $\{-0.99, +1\}$

The comparison between the power spectra initiated from a well-defined and an ill-defined Bunch–Davies vacuum for the quadratic potential with background IC $\{-0.99, +1\}$ is show on Fig. 7. A significant deficit of power for the ill-defined vacuum, similar to the one of [12], is indeed observed. We do not show the temperature anisotropies for the primordial power spectra shown in Figs. 5 and 6, since they are the same as in the flat space case.

4.2 Pre-bounce origin of quantum fluctuations

In most cases, the Minkowski vacuum requirements are met *before* the curvature bounce and it is reasonable to investigate the situation where the quantum fluctuations are initiated in their minimum energy state at that time. If this assumption is correct, one has to evolve perturbations through the bounce – the later leaving a characteristic imprint. To compute the primordial power spectrum, we proceed as before, i.e. choosing the best initial time $t_i < 0$ to set the IC for the perturbations – that is corresponding to a good quality vacuum. They are then propagated up to the end of inflation and the power spectrum is calculated. This procedure is repeated for both potentials and for different background IC. The divergence of the functions A_n and B_n is an additional numerical difficulty. Fortunately, the divergence being well localized and well behaved, one can always find a time interval around $t = 0$, where both functions can be accurately approximated by $A_n(t) \approx a_n/t$ and $B_n(t) - k^2 \approx b_n/t$, where a_n and b_n can be well determined. With these approximations, the equation for the perturbations (16) can be solved analytically using Bessel functions. Outside this specific time interval around the bounce, the perturbations are evolved with numerical techniques.

The Bunch–Davies vacuum cannot be properly defined in the case of the quadratic potential with background IC $\{-0.467, +1\}$, hence we expect a decline in the primordial scalar power spectrum at low wave numbers. For the other power spectra, we expect the shape to be only modified by the bounce. The primordial power spectrum for scalar fluctuations initiated before the curvature bounce, in the case of the quadratic and Starobinsky potentials, are shown in Figs. 8 and 9 respectively. For both potentials, when the duration of the deflation is similar to the duration of inflation, represented by $w(0) = -0.99$, it can be seen that the primordial power spectrum is not significantly affected. As expected,

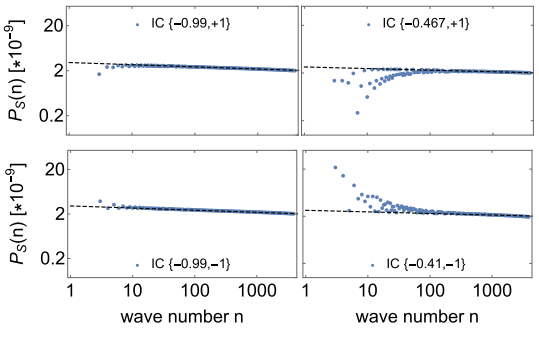


Fig. 8 Primordial scalar power spectrum for different sets of background initial conditions, in the case of the quadratic potential and with fluctuations initiated before the bounce (blue dots). The dashed black line represents the standard power law. The IC for the background are: *top left* $\{-0.99, +1\}$, *top right* $\{-0.467, +1\}$, *bottom left* $\{-0.99, -1\}$, *bottom right* $\{-0.41, -1\}$

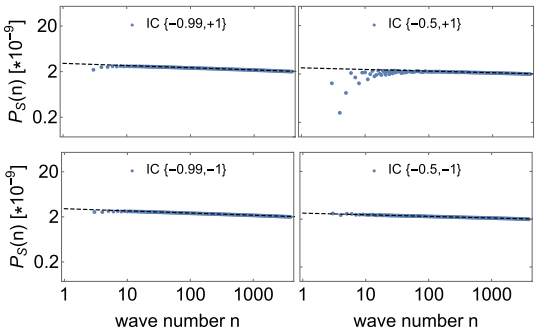


Fig. 9 Primordial scalar power spectrum for different sets of background initial conditions, in the case of the quadratic potential and with fluctuations initiated before the bounce (blue dots). The dashed black line represents the standard power law. The IC for the background are: *top left* $\{-0.99, +1\}$, *top right* $\{-0.5, +1\}$, *bottom left* $\{-0.99, -1\}$, *bottom right* $\{-0.5, -1\}$

for a massive inflaton generating a very short period of deflation, corresponding to IC $\{-0.467, +1\}$, we indeed observe a decline in the power spectrum amplitude at low wave numbers. However, we also observe this behaviour in the case of the Starobinsky potential, for which the vacuum is well-defined. It can then be safely conjectured that the short period of deflation is the origin of the lack of power at low n. When the period of deflation is longer than the period of inflation, the result significantly differ, depending on the potential. For the massive scalar field, the primordial power spectrum at low n is amplified when compared to the flat space case, while for the Starobinsky scalar field, it is not drastically affected.

Three of the eight primordial scalar power spectra deviate from the standard power law $\mathcal{P}_S(k) = A_s(k/k_*)^{n_s-1}$, hence we focus on those for the calculation of the temperature power spectra C_L^{TT}. The deviation mainly appears at low wave number n and the C_ℓ^{TT} are therefore affected at low values of ℓ, where the cosmic variance makes it difficult to have statistically significant measurements. Nonetheless, strong deviations could lead to inconsistencies. Although not statistically significant, it might also be argue that explaining the slight lack of power observed in the CMB at large scales [32] would be welcome. To calculate the C_ℓ^{TT}, or more precisely the $D_\ell^{\mathrm{TT}} := \ell(\ell+1)/(2\pi)C_\ell^{\mathrm{TT}}$, for the three different non-trivial cases, namely quadratic potential with IC $\{-0.467, +1\}$ and $\{-0.41, -1\}$, and the Starobinsky potential with IC $\{-0.5, +1\}$, we proceed as follow. We use the Boltzmann code **CAMB**, modifying the cosmological parameters for the spatial curvature and for the Hubble constant according to [4]. With these appropriate values and the standard power law for the primordial power spectrum, one recovers almost exactly the prediction made with the usual parameters $\Omega_K = 0$ and $H_0 = 67.3$ km/s/Mpc, as found in [1] (this is why, in [4], it can be argued that a positively curved universe is compatible with CMB measurements). We then compute two different power spectra. The first one, named

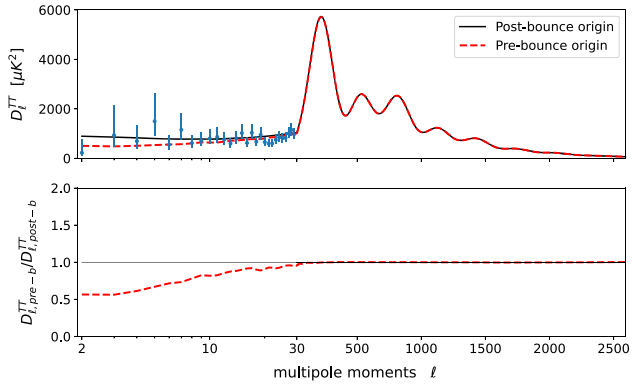


Fig. 10 Temperature angular power spectra D_ℓ^{TT} for a pre-bounce and for a post-bounce origin of structures, with background IC $\{-0.467, +1\}$, and a quadratic potential. *Top:* $D^{\mathrm{TT}}_{L,\text{post-b}}$ in black, $D^{\mathrm{TT}}_{\ell,\text{pre-b}}$ in dashed red, Planck data in blue. *Bottom:* Ratio of the spectra in dashed red

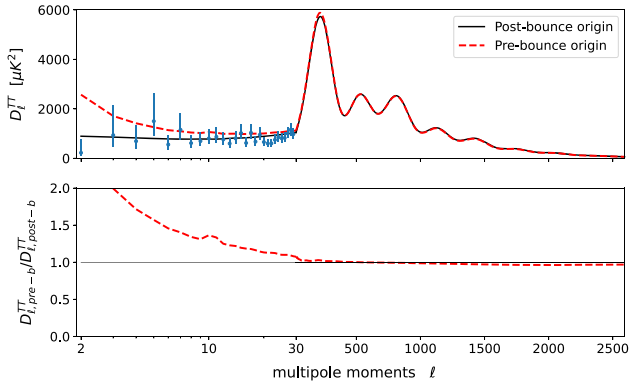


Fig. 11 Temperature angular power spectra D_ℓ^{TT} for a pre-bounce and for a post-bounce origin of structures, with background IC $\{-0.41, -1\}$, and quadratic potential. *Top:* $D^{\mathrm{TT}}_{\ell,\text{post-b}}$ in black, $D^{\mathrm{TT}}_{\ell,\text{pre-b}}$ in dashed red, Planck data in blue. *Bottom:* Ratio of the spectra in dashed red

$D^{\mathrm{TT}}_{\ell,\text{post-b}}$, is calculated using the standard power law for the primordial spectrum, which will serve as a comparison. As the name suggests, it is equivalent to the result calculated for the primordial power spectra derived with post-bounce initial conditions. The second temperature power spectrum $D^{\mathrm{TT}}_{\ell,\text{pre-b}}$ is computed with the modified primordial power spectra from the pre-bounce initial perturbations.

The temperature power spectra for the quadratic potential with a short period of deflation, a long period of deflation, and for the Starobinsky potential with a short period of deflation are shown in Figs. 10, 11 and 12, respectively. While the important increase of power at low values of ℓ, as can be observed in Fig. 11 for a long period of deflation, is disfavored by Planck measurements, the slight decrease at low ℓ for the short deflation fits well the data. It should be underlined that we present here the extreme cases only (either an identical duration for deflation and inflation, or a very long/brief period of deflation). One can of course choose background IC in between these extremes and obtain a weaker deviation. Overall, Planck data do slightly favor a shorter period of deflation than inflation. When compared to the post-bounce initial conditions, the pre-bounce initial conditions with a short period of deflation decrease the χ^2 per degree of freedom by, respectively, 2.25 and 1.40 for a quadratic potential and a Starobinsky potential (performing the analysis with points before the first acoustic peak only). Those numbers are to be considered as indications of the trend and not treated as statistically significant results.

Overall, the curvature bounce hypothesis agrees with current data. Depending on the IC, it can fit slightly better or slightly worst the CMB observations, when compared to the standard model. For the vast majority of IC, it is hardly distinguishable from the usual scenario, which makes it reliable and consistent but hard to falsify. Some non-CMB ideas for experimental probes were however suggested in [16].

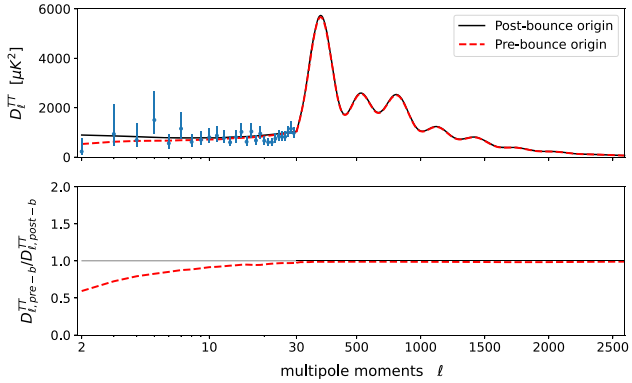


Fig. 12 Temperature angular power spectra D_ℓ^{TT} for a pre-bounce and for a post-bounce origin of structures, with background IC $\{-0.5, +1\}$, and Starobinsky potential. *Top:* $D^{\mathrm{TT}}_{\ell,\text{post-b}}$ in black, $D^{\mathrm{TT}}_{\ell,\text{pre-b}}$ in dashed red, Planck data in blue. *Bottom:* Ratio of the spectra in dashed red

5 Conclusion

In the first part of this article, we have discussed the background behaviour in the pre-bounce universe, at the bounce and during the inflationary period. In particular, we have shown that widely different pre-bounce dynamics can lead to the same inflationary stage, due to the strong attractor status of the slow-roll inflation trajectory for the background equations of motion. To be more precise, the duration of the period of deflation can be chosen freely, without affecting the physics in the expanding universe at the background level. This raises interesting theoretical questions. From the viewpoint of stability for a single bounce, a brief deflation period is favored. But from the viewpoint of the avoidance of multiple bounces, a long deflation is favoured. It should, once again, be emphasized that this work does not focus on the "naturalness" of the considered cosmic evolution. The other way round, it tries to determine the correct solution – what-

ever its a priori probability – taking into account what we know about the Universe.

After having introduced the theory of linear perturbations in curved space and shown that the quantum fluctuations can be initiated in the Bunch–Davies vacuum either before or after the bounce, we have calculated the primordial power spectrum of scalar perturbations for both the quadratic and Starobinsky potentials, for different values of the duration of the deflation stage. We have reached the conclusion that when the initial conditions for the fluctuations are set after the bounce, the standard power law spectrum, $\mathcal{P}_S(k) = A_s(k/k_*)^{n_s-1}$, derived from the flat space case, is recovered nearly exactly. This result is in tension with [12] where a power decrease at low wave numbers was pointed out. We believe that the difference is (maybe partially) due to the fact that the requirements for the Bunch–Davies vacuum are not satisfied at the time when fluctuations are initiated in [12] .

Later on, we have explicitly computed the primordial power spectrum for perturbations that are initially set before the bounce. We have found that, in the case of the quadratic potential, the results significantly differ from the standard power law, when the deflation period is either very long (typically around 185 e-folds) or very short, (around 1 or 2 e-folds). The primordial power spectrum decreases at low wave numbers n for a very short deflation, which is partly explained by an ill-defined Bunch–Davies vacuum. On the other hand, the power is increased at low n for a very long deflation. In the case of the Starobinsky potential, the power spectrum is attenuated at low n for a short deflation, but it is equivalent to the standard power spectrum for a long deflation, as opposed to the massive case. These modified primordial power spectra only affect the temperature power spectrum C_ℓ^{TT} at low values of ℓ, where it is difficult to reach a definitive conclusions due to the cosmic variance. Nonetheless, Planck data slightly favor a short duration of deflation, due to the power decrease at low wave numbers.

It is rather remarkable that, while the pre-bounce dynamics has no effects on the background behaviour of the expanding universe, the duration of the period of deflation can affect the temperature anisotropies observed in the CMB. To reach this conclusion, one needs to assume that fluctuations originate from the Bunch–Davies vacuum before the bounce. While it is a plausible hypothesis, this is obviously not the only possible choice. In addition, in this work, we have extensively used the values of the parameters given in [4] but other possibilities should be considered in future studies.

Finally, it is worth recalling that this entire scenario is also appealing because the density of the Universe never approaches the Planck density. The bounce takes place for $\rho \ll \rho_{Pl}$ and the use of standard quantum field theory techniques is therefore much safer than in quantum gravity scenarios.

Some of the points made in this work are quite generic to bouncing models (see [33] for a review). In particular, the issues that have to be dealt with are basically always the same as far as initial conditions are concerned. The details are however highly depending on the equation of state around the bounce ($w = -1$ in our case), which is different from the one typically encountered in the ekpyrotic bounce ($w \geq 1$) [34], matter bounce [35] ($w = 1/3$) or typical loop quantum cosmology bounce [36] (w varies fast close to bounce). The techniques can therefore be easily generalized to other bouncing paradigms but not the results.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: This is a purely theoretical and conceptual work, there are no relevant data associated.]

Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

1. N. Aghanim et al. (Planck), arXiv:1807.06209 (2018)
2. C.-G. Park, B. Ratra, Astrophys. J. **882**, 158 (2019). arXiv:1801.00213
3. W. Handley (2019). arXiv:1908.09139
4. E. Di Valentino, A. Melchiorri, J. Silk, Nat. Astron. **4**, 196 (2019). arXiv:1911.02087
5. G. Efstathiou, S. Gratton (2020). arXiv:2002.06892
6. S. Vagnozzi, E. Di Valentino, S. Gariazzo, A. Melchiorri, O. Mena, J. Silk, Phys. Dark Universe **33**, 100851 (2021). arXiv:2010.02230
7. S. Vagnozzi, A. Loeb, M. Moresco, Astrophys. J. **908**, 84 (2021). arXiv:2011.11645
8. S. Dhawan, J. Alsing, S. Vagnozzi, Mon. Not. R. Astron. Soc. **506**, L1 (2021). arXiv:2104.02485
9. Y. Akrami et al. (Planck), Astron. Astrophys. **641**, A10 (2020). arXiv:1807.06211
10. Y. Akrami et al. (Planck), Astron. Astrophys. **641**, A9 (2020). arXiv:1905.05697
11. V.F. Mukhanov, H.A. Feldman, R.H. Brandenberger, Phys. Rep. **215**, 203 (1992)
12. B. Bonga, B. Gupt, N. Yokomizo, JCAP **1610**, 031 (2016). arXiv:1605.07556
13. B. Bonga, B. Gupt, N. Yokomizo, JCAP **1705**, 021 (2017). arXiv:1612.07281
14. J. Martin, P. Peter, Phys. Rev. Lett. **92**, 061301 (2004). arXiv:astro-ph/0312488
15. J.-P. Uzan, U. Kirchner, G.F.R. Ellis, Mon. Not. R. Astron. Soc. **344**, L65 (2003). arXiv:astro-ph/0302597

16. A. Barrau, Eur. Phys. J. C **80**, 579 (2020). arXiv:2005.04693
17. C. Renevey, A. Barrau, K. Martineau, S. Touati, JCAP **01**, 018 (2021). arXiv:2010.13542
18. A.A. Starobinsky, Phys. Lett. B **91**, 99 (1980)
19. J. Martin, C. Ringeval, V. Vennin, Phys. Dark Universe **5–6**, 75 (2014). arXiv:1303.3787
20. D. Chowdhury, J. Martin, C. Ringeval, V. Vennin, Phys. Rev. D **100**, 083537 (2019). arXiv:1902.03951
21. H. Matsui, F. Takahashi, T. Terada, Phys. Lett. B **795**, 152 (2019). arXiv:1904.12312
22. A.A. Asgari, A.H. Abbassi, J. Khodagholizadeh, Eur. Phys. J. C **74**, 2917 (2014). arXiv:1405.4466
23. J.M. Bardeen, Phys. Rev. D **22**, 1882 (1980)
24. V.F. Mukhanov, JETP Lett. **41**, 493 (1985)
25. M. Sasaki, Prog. Theor. Phys. **76**, 1036 (1986)
26. S. Schander, A. Barrau, B. Bolliet, L. Linsefors, J. Mielczarek, J. Grain, Phys. Rev. **D93**, 023531 (2016). arXiv:1508.06786
27. I. Agullo, A. Ashtekar, W. Nelson, Phys. Rev. Lett. **109**, 251301 (2012). arXiv:1209.1609
28. B. Bolliet, J. Grain, C. Stahl, L. Linsefors, A. Barrau, Phys. Rev. D **91**, 084035 (2015). arXiv:1502.02431
29. R. Penrose, Singularities and time asymmetry, pp. 581–638 (1980)
30. C. Rovelli (2021). arXiv:2105.00540
31. A. Lewis, A. Challinor, A. Lasenby, Astrophys. J. **538**, 473 (2000). arXiv:astro-ph/9911177
32. E.F. Bunn, A. Bourdon, Phys. Rev. D **78**, 123509 (2008). arXiv:0808.0341
33. R. Brandenberger, P. Peter, Found. Phys. **47**, 797 (2017). arXiv:1603.05834
34. J. Khoury, B.A. Ovrut, P.J. Steinhardt, N. Turok, Phys. Rev. D **64**, 123522 (2001). arXiv:hep-th/0103239
35. R.H. Brandenberger (2012). arXiv:1206.4196
36. A. Ashtekar, P. Singh, Class. Quantum Gravity **28**, 213001 (2011). arXiv:1108.0893

3.4 A bounce in loop quantum cosmology

Quantizing our theory of gravity has been a century long problem in theoretical physics. Due to the non-renormalizability of General Relativity, common QFT techniques are unavailable and the absence of experimental evidence makes it difficult to move towards a quantum description of gravity. Nonetheless, different approaches to quantize General Relativity have emerged. On one side, string theory uses the perturbative techniques of QFT, but in the context of a UV-complete theory. On another side, Loop quantum gravity (LQG) applies the well-studied and robust technique of canonical quantization. While observing quantum gravity in particle physics experiments remains out of reach, one could hope to observe imprints of its effect at very early times. As discussed earlier, if the universe is closed and started with a period of inflation, there might be no need for a quantum theory of gravity. However, the spatial curvature is so far unobserved and our universe is equally likely to be infinitely large. In this case, another solution to the big bang singularity problem must be found and it might lie in the quantum description of General Relativity. In the attempt to understand the very early universe, the techniques of LQG where used on the restricted theory of General Relativity together with the homogeneity and isotropy assumptions, leading to the theory of loop quantum cosmology (LQC). However, several ambiguities, common in quantum mechanics, might be plaguing the predictions of LQC and a consistency check for more general quantum corrections from LQC to classical cosmology should be made. Refs. [7, 131, 132] were used extensively in this section.

3.4.1 Loop quantum gravity in a nutshell

In order to quantize gravity "à la Dirac", we need to find the Hamiltonian of the theory and express it in terms of the canonical variables. In the context of General Relativity, this is not an easy task, because space and time are dissociated in Hamiltonian mechanics, as opposed to relativistic theories. Nonetheless such a separation can be done if we proceed as follows. First, we observe that in the Einstein-Hilbert action

$$S_{EH}[g] = \frac{1}{16\pi}\int \mathrm{d}^4x\sqrt{-g}R\,, \tag{3.55}$$

g_{00} and g_{0i} appear without a time derivative. This means that they can be considered as Lagrange multipliers and they are non-dynamical, their associated field equations behave as constraints. Moreover, if one writes the metric as

$$\mathrm{d}s^2 = \left(-N^2 + N_iN^j\right)\mathrm{d}t^2 + 2N_i\mathrm{d}t\mathrm{d}x^i + q_{ij}\mathrm{d}x^i\mathrm{d}x^j\,, \tag{3.56}$$

the action takes the interesting form [30]

$$S_{ADM}[q_{ij}, N, N_i] = \frac{1}{16\pi}\int \mathrm{d}t\int \mathrm{d}x^3\left(\pi_{ij}\dot{q}^{ij} - N\mathcal{C} - N_i\mathcal{C}^i\right)\,, \tag{3.57}$$

where $\dot{q}_{ij} = \mathcal{L}_t q_{ij}$ is the Lie derivative of the induced metric q_{ij} of the hyper-surfaces Σ_t with normal vector $n_\mu = (N, 0, 0, 0)$,

$$\pi_{ij} = \frac{\delta L}{\delta \dot{q}^{ij}} = \sqrt{q}\left(K_{ij} - K^l{}_l q_{ij}\right) \tag{3.58}$$

is the conjugate momentum of q_{ij}, with $K_{ij} = \mathcal{L}_n q_{ij}/2$ the extrinsic curvature of Σ_t,

$$\mathcal{C} = \left(q_{ik}q_{jl} + q_{il}q_{jk} - q_{ij}q_{kl}\right)\pi_{ij}\pi_{kl}\mathcal{R} \tag{3.59}$$

is the *scalar constraint* with $\mathcal{R}$ the Ricci scalar of Σ_t and

$$C_i = -2\sqrt{q}\,\nabla_l\left(\frac{\pi_i^l}{\sqrt{q}}\right) \tag{3.60}$$

is the *vector constraint*. With the action (3.57) written in terms of the variables q_{ij}, N and N_i, several properties become apparent. N and N_i are directly connected to the components g_{00} and g_{0i} of the metric and they are trivially behaving as Lagrangian multipliers. Their associated field equations straightforwardly give the constraint equations $\mathcal{C} = \mathcal{C}^i = 0$. Since the remaining dynamical variable q_{ij} must be symmetric, it contains a priori six degrees of freedom. But together with the constraints, we are indeed down to two propagating degrees of freedom as expected. The form of the action (3.57) is canonical and so are the variables q_{ij} and π^{ij}. Hence one could proceed with the canonical quantization procedure, which would lead to the Wheeler-DeWitt equation [133] for the wave function $\Psi(q_{ij})$ of the three-metric q_{ij}. One major issue with this approach is that the inner product on the state space is ill-defined and it makes it difficult to derive predictions [132]. Moreover in the context of quantum cosmology, where we reduce the phase-space from the three-metric to the scale factor alone before quantization, different initial conditions lead to different meaningful cosmological predictions [134–136], therefore spoiling the utility of the theory.

In order to evade the problems of the Wheeler-DeWitt approach, one can find new canonical variables with which we then proceed to the Dirac quantization. This is possible if we introduce the concept of *triads* and *co-triads*, denoted e_a^i and e_i^a, respectively. They are three-vectors defined such that we can decompose the dynamical part of the metric as

$$q_{ij} = e_i^a e_j^b \delta_{ab}\,. \tag{3.61}$$

In particular, the triads connect what we call an *internal space* represented by the indices $a, b, ...$, to the hyper-surfaces Σ_t with indices $i, j, ...$. We can also observe that the metric is invariant under any rotation of the internal space, i.e. the physics is gauge invariant under the algebra $su(2)$. The theory expressed in terms of the triads and co-triads is therefore very similar to a Yang-Mills theory. The consequence of this gauge symmetry is that we need to add three constraints to the action when describing the physics using the co-triads.

In order to work with objects that have internal indices, e.g. the vector V^a, we need to define a covariant derivative D_i as

$$D_i V^a = \partial_i V^a + \omega_i{}^a{}_b V^b\,, \tag{3.62}$$

where $\omega_i{}^a{}_b$ is called the *spin connection*. If the object has mixed indices, e.g. E_j^a, then the

covariant derivative reads

$$D_i E^a_j = \partial_i E^a_i + \omega_i{}^a{}_b E^b_j - \Gamma^k_{ij} E^a_k \,. \tag{3.63}$$

We would like that the covariant derivative of the metric cancels, therefore we have the useful condition $D_i e^a_j = 0$.

The action is the integral of a densitized expression, such as the volume form $\sqrt{q}\mathrm{d}x^3$. This means that the densitized triad

$$E^a_i = \sqrt{q}\; e^a_i \,, \tag{3.64}$$

is better suited as a candidate canonical variable. In fact, together with the *Sen-Ashtekar-Barbero (SAB) connection*

$$A^a_i = \omega_{ibc}\epsilon^{bca} + \beta K_{ij}\frac{E^{ja}}{\sqrt{q}} \,, \tag{3.65}$$

where $\gamma \in \mathbb{C}$ is called the *Barbero-Immirzi parameter* and ϵ_{abc} is the Levi-Civita tensor, they form a canonical pair. Therefore, their Poisson brackets are

$$\left\{A^a_i(\mathbf{x}), A^j_b(\mathbf{y})\right\} = 0 \tag{3.66}$$

$$\left\{E^a_i(\mathbf{x}), E^j_b(\mathbf{y})\right\} = 0 \tag{3.67}$$

$$\left\{A^a_i(\mathbf{x}), E^j_b(\mathbf{y})\right\} = 8\pi\beta\delta^j_i\delta^b_a\delta^3(\mathbf{x}-\mathbf{y}) \,, \tag{3.68}$$

and the action in terms of these new variables takes the form

$$S_{SAB}[A^a_i, \lambda^a, N, N_i] = \frac{1}{16\pi}\int \mathrm{d}t \int \mathrm{d}x^3 \left(E^i_a \dot{A}^a_i - N\mathcal{C} - N_i\mathcal{C}^i - \lambda^a\mathcal{G}_a\right) \,, \tag{3.69}$$

with the Gauss, vector and scalar constraints, respectively

$$\mathcal{G}_a = D_i E^i_a \tag{3.70}$$

$$\mathcal{C}_i = F^a_{kl}E^l_a - (1+\beta^2)K^a_i\mathcal{G}_a \tag{3.71}$$

$$\mathcal{C} = \frac{E^i_a E^j_b}{\sqrt{\det E}}\left(\epsilon_c{}^{ab}F^c_{ij} - 2(1+\beta^2)K^a_{[i}K^b_{j]}\right) \,, \tag{3.72}$$

and the curvature tensor of the SAB connection being

$$F^a_{ij} = \partial_i A^a_j - \partial_j A^a_i + \epsilon^a_{bc}A^b_i A^c_j \,. \tag{3.73}$$

The Hamiltonian density is then easily found using $\mathcal{H} = E^i_a\dot{A}^a_i - L_{SAB}$.

We have now all the ingredients to proceed with the Dirac quantization, in which we promote the canonical variables to operators on the Hilbert space of states, or similarly on the space of wave functions, the Poisson brackets become commutation relations and we promote the Hamiltonian to an operator as well. In order to define the space of the wave functions, one can use the very useful *Giles' theorem* [137]. To understand the theorem, we need the concept of *holonomy*. The holonomy $h_\gamma[A]$ of the SAB connection A^a_i on a closed

curve γ is defined as

$$h_\gamma[A] = \mathcal{P}\left[e^{\oint_\gamma A_i^a \tau_a \mathrm{d}x^i}\right], \tag{3.74}$$

the path ordered integral of the exponential of the connection associated to the $su(2)$ generators τ_a. We recall that the theory in terms of the triad is symmetric under $SO(3)$ and, similarly to Yang-Mills theories, the connection $\mathbf{A}_i = A_i^a \tau_a$ is used to make it a gauge symmetry. Hence the full connection $\mathbf{A}_i$ is an element of the vector space $su(2)$ and its components with respect to the basis $\{\tau_a\}_a$ are what we refer to as the connection A_i^a. In the context of Yang-Mills theories, Giles has shown that any function of the connection, such as a wave function $\Psi(A)$, can be decomposed into of sum over all possible loops γ of the traces of the holonomies, namely

$$\Psi(A) = \sum_\gamma \Psi(\gamma) W_\gamma[A], \tag{3.75}$$

with $W_\gamma[A] = \mathrm{Tr}\left(h_\gamma[A]\right)$. In the same way, one can work either in momentum or in position space in quantum mechanics, we can work in the connection or in the "loop" space with $\Psi(\gamma)$. It is this *loop representation* that allows one to resolve the problem of ill-defined inner product encountered in the Wheeler-DeWitt approach. We do not extend the detail of the calculation further, as it is not necessary for the rest of the text, but we must discuss the main consequence of the LQG approach to quantum gravity. Once the kinematical Hilbert space with its inner product has been defined and we promoted the canonical variables as well as the Hamiltonian to operators, we can find the area and volume operators. The minimum value in the spectrum of both operators for physical states is non-zero, meaning that there exists a smallest area and a smallest volume in LQG. The minimum area that one can have is $A_{min} = 4\sqrt{3}\pi\beta l_{pl}^2$, with l_{pl} the Planck length, and it will become important later on.

3.4.2 Loop quantum cosmology with a general holonomy correction

Loop quantum gravity is an attempt to quantize the full theory of General Relativity and we saw that the concept of holonomy is an essential part of the success of the theory. One of the main era when we would expect quantum gravity to be relevant is the early universe. Therefore, one could try to apply LQG techniques on the restricted theory where the cosmological principle of homogeneity and isotropy is considered. When restricting ourselves to this *minisuperspace*, where the metric in flat space is of the form

$$\mathrm{d}s^2 = -N^2(t)\mathrm{d}x_0^2 + a^2(t)\delta_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{3.76}$$

the triads and co-triads simplify to

$$e_i^a = a(t)\delta_i^a \quad \text{and} \quad e_a^i = a^{-1}(t)\delta_a^i, \tag{3.77}$$

respectively. We see that the cosmic time t commonly used in cosmology is connected to the 00-component of the metric as $\mathrm{d}t = N\mathrm{d}x_0$. We now quickly summarize the main steps and cosmological consequences of LQG reduced to the minisuperspace of FLRW cosmology, or otherwise named loop quantum cosmology (LQC). The introductory sections of the following

two papers contain additional details concerning the derivation of the background dynamics of LQC. The densitized triads and the SAB connection can also be rewritten in terms of the scale factor and it is common practice in LQC to work with the new canonical variables $c(t)$ and $p(t)$ defined as follow,

$$E^i_a = a^2(t)\delta^i_a := \frac{p(t)}{V_0^{2/3}}\delta^i_a \quad \text{and} \quad A^a_i = \beta\dot{a}(t)\delta^i_a := \frac{c(t)}{V_0^{1/3}}\delta^a_i \,. \tag{3.78}$$

The fiducial volume V_0 is the volume on which we integrate the Hamiltonian density and is introduced in order to avoid a divergent Hamiltonian. Once its density is integrated on the fiducial volume, the total Hamiltonian including the gravity and scalar sectors is

$$H = N\left(-\frac{1}{16\pi\beta^2}p^{1/2}V_0^{2/3}\delta^i_a\delta^j_b\epsilon_c{}^{ab}F^c{}_{ij} + p^{3/2}\rho\right) . \tag{3.79}$$

Beware that $p(t)$ stands for one of the canonical variables and not the pressure of the scalar fluid. If one uses the definition (3.73) for the curvature of the SAB connection, we find the Hamiltonian

$$H = N\left(-\frac{3}{8\pi\beta^2}\sqrt{p}c^2 + p^{3/2}\rho\right) , \tag{3.80}$$

and the Hamiltonian equations give us the usual Friedmann and Raychaudhury equations. However, in a similar way one can define the Riemann curvature with the parallel transport of a vector around an infinitesimal loop, one can use the trace of the holonomy of a loop to find the curvature F^a_{ij}. Let's take a squared loop denoted $\Box$ with the sides of length μ. Then the curvature tensor can be written as

$$F^c{}_{ij} = -2\lim_{\mu\to 0}\frac{1}{\beta^2\mu^2}\mathrm{Tr}\left(\tau^c(h_{\Box ab}[\mathbf{A}] - \mathbb{1}_d)\right)e^a_i e^b_j \,, \tag{3.81}$$

where d is the dimension of the representation used for the generators τ^a. In particular, if one chooses the fundamental representation where the generators are the Pauli matrices, we get the following result

$$F^c{}_{ij} = -2\lim_{\mu\to 0}\frac{\sin^2(c\mu)}{\beta^2\mu^2}\epsilon^c_{ab}e^a_i e^b_j \,. \tag{3.82}$$

Now, we learned for LQG that the smallest area possible is non-zero, meaning that taking the limit $\mu \to 0$ does not make sense in this context. Instead one should take

$$\mu \to \bar{\mu} = \sqrt{4\sqrt{3}\pi\beta/p}\; l_p := \sqrt{\frac{\Delta}{p}}\; l_p \,, \tag{3.83}$$

which is the square root of the minimum physical area. The effect of such a change on the Hamiltonian (3.80) is the modification

$$c^2 \to \frac{\sin^2(c\bar{\mu})}{\bar{\mu}^2} \,, \tag{3.84}$$

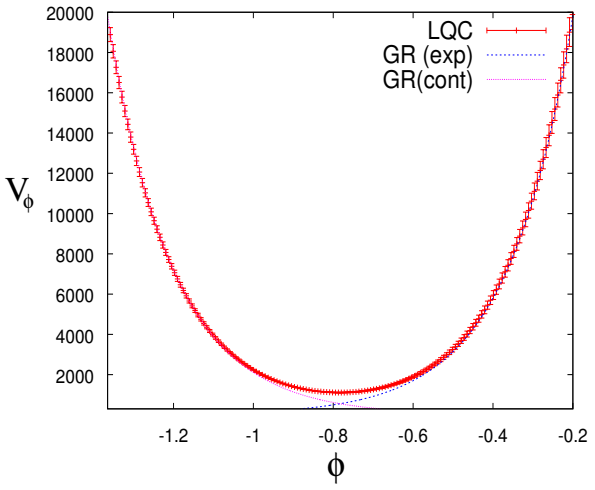


Figure 3.1: Comparison between the quantum evolution as derived in LQC and the classical evolution of the volume observable. The scalar field ϕ is used as the time variable. Unlike classical trajectories leading to singularities, the LQC trajectory is non-singular and leads to a bounce in the Planck regime. The dispersion across the bounce for the quantum approach are included. The state remains peaked and the dispersion approaches a constant at large volumes. Retrieved from [131].

which is called the *holonomy correction*. Using the Hamiltonian equations together with the Hamiltonian constraint $H = 0$ gives us the modified Friedmann of LQC, namely

$$H^2 = \frac{8\pi}{3}\rho\left(1 - \frac{8\pi\beta^2\Delta}{3}\rho\right) := \frac{8\pi}{3}\rho\left(1 - \frac{\rho}{\rho_c}\right) . \tag{3.85}$$

We see that in the limit where $\rho \ll \rho_c$ we recover the standard Friedmann equation. The *critical density* ρ_c is expected to be of order of the Planck density, hence this inequality is satisfied for most of the history of the universe. However, when $\rho \approx \rho_c$ the Hubble parameter cancels, which implies that we have a bounce instead of the big bang singularity.

One could worry that the above derivation of the Friedmann equation employs no quantization techniques. We described here the effective approach to LQC, but a quantum mechanical description can be made. The quantization of the minisuperspace follows the usual canonical method, but on the canonical variables $b = c\bar{\mu}$ and $V = a^3$ instead, the latter representing the volume. First, the canonical variables are promoted to operators, as well as the Hamiltonian $\hat{H}$ and the wave function $\Psi(V, \psi)$ is defined. The evolution equation of the wave function is found with the quantum constraint equation $\hat{H}\Psi = 0$. Numerical calculations of the evolution of the wave function, initiated with a sharply peaked initial condition on the classical solution, were made and the results for the volume observable are exposed in Fig. 3.1. We see that the dispersion of the volume observable stays small, validating the effective approach described earlier, and we indeed have a bounce instead of a big bang.

Another argument in favor of the effective approach to LQC arises when including inhomogeneities. As opposed to the standard cosmological perturbation theory introduced in the previous chapter, we would like to describe quantum fluctuation on a quantum background.

Heuristically, this means that the total wave function would be

$$\Psi(a,\phi,v_S,v_T) = \Psi_{back}(a,\phi)\otimes\Psi_{pert}(a,\phi,v_S,v_T)\,. \tag{3.86}$$

In the absence of back-reactions of the perturbations on the background, an assumption that is commonly used and checked to be consistent [138], the wave function of the background $\Psi_{back}(a,\phi)$ behaves as discussed in the previous paragraph. We are left to find the evolution equation for the wave function of the perturbations. In the case of the tensor modes, one can show that the wave function follows the behaviour

$$\Psi_{back}\otimes i\hbar\frac{\partial\Psi_{pert}}{\partial\phi} = H^T_{pert}\left(\Psi_{back}\otimes\Psi_{pert}\right)\,, \tag{3.87}$$

where H^T_{pert} is the Hamiltonian for the tensor perturbations. By taking the inner product with respect to Ψ_{back}, one can show that the tensor quantum fluctuations evolve as

$$\hat{v}_T'' + \left(k^2 - \frac{\tilde{a}''}{\tilde{a}}\right)\hat{v}_T = 0\,, \tag{3.88}$$

where $\tilde{a}(\eta)$ is an effective scale factor that follows the LQC Friedmann equations. This means that, from the point of view of the perturbations, the quantum background behaves as a classical effective background described by LQC, supporting the effective approach derived earlier. A similar argument can be made for the scalar perturbations.

One issue with the above approach, named *dressed metric*, is that the constraint algebra is distorted at the level of the perturbations, meaning that the Poisson brackets of the constraints do not give a constraint again. An other approach, called *deformed algebra*, to perturbation theory on an LQC background was developed with the requirements that the evolution of the perturbations must be chosen such that the constraint algebra remains closed. For the holonomy correction derived above (3.84), the equations for the perturbations that keep the algebra closed are

$$v_T'' + \left(\Omega^2k^2 - \frac{\tilde{a}''}{\tilde{a}}\right)v_T = 0 \quad\text{and}\quad v_S'' + \left(\Omega^2k^2 - \frac{\tilde{z}''}{\tilde{z}}\right)v_S = 0\,, \tag{3.89}$$

where

$$\Omega = \cos(2\bar{\mu}c) = 1 - 2\frac{\rho}{\rho_c}\,, \tag{3.90}$$

the scale factor from LQC $\tilde{a}$ and $\tilde{z} = \tilde{a}\dot{\tilde{\phi}}/\tilde{H}$. While this approach resolves the distortion of the constraint algebra and recovers the Euclidean period as in Hawking's approach to quantum cosmology [139], it suffers from observational issues. First, one can show that during the phase when Ω is negative, meaning $\rho_c/2 < \rho < \rho_c$, the effective metric "felt" by the perturbations is in fact Euclidean [140]. This makes it complicated to interpret what is the time evolution of the perturbations during that early epoch, but is conceptually interesting as already discussed by Hawking. However, this Euclidean period also prevents the primordial power spectra from being scale invariant, which is in contradiction with observations. However, the problem of time and the problem of inconsistent power spectra might be connected and a better understanding of the effect of the bounce on the perturbations

in the deformed algebra approach is necessary to draw conclusions.

In order to derive the standard behaviour of LQC, several assumptions were made, starting with, for example, the choice of the fundamental representation of $su(2)$ for the calculation of the holonomy. This means in particular that the derived holonomy correction (3.84) is far from unique and a more general expression

$$c^2 \to g^2(c,p)\,, \tag{3.91}$$

should be studied. However, several properties of the standard holonomy corrections remain. First, we should still recover General Relativity and FLRW cosmology in the limit $\rho \ll \rho_c$, otherwise the theory is excluded by observations. It has also been previously discussed that the general holonomy correction must be periodic if derived from LQG arguments [141,142]. This last condition is in fact sufficient to show that the existence of the bounce in LQC does not depend on the holonomy correction. Nonetheless, other predictions of LQC might differ depending on the correction and the robustness of these predictions must be tested. This is the goal of the two following papers. in the first paper, the dependence of the background dynamics on the holonomy correction is studied as well as the scalar primordial power spectrum as predicted by the dressed metric method. In the second paper, the tensor and scalar primordial power spectra in the deformed algebra approach are tested for different holonomy corrections.

Cosmological implications of generalized holonomy corrections

Cyril Renevey, Killian Martineau, and Aurélien Barrau
Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3 53, avenue des Martyrs, 38026 Grenoble cedex, France

(Received 4 October 2021; accepted 2 March 2022; published 18 March 2022)

Most of the phenomenology of loop quantum gravity in the cosmological sector is based on the so-called holonomy correction to the Hamiltonian constraint. It straightforwardly modifies the Friedmann equations. In this work, we investigate the influence of corrections generalizing the one usually used in loop quantum cosmology. We find that a long enough inflation phase can be generated by purely quantum geometrical effects but we also underline the limitations of this scenario. In addition, we study the effects of generalized holonomy corrections on an inflationary phase generated by a massive scalar field. At the level of perturbations, we investigate in detail the consequences on the primordial scalar power spectrum. The results are actually quite general and can be used beyond the "loop" framework.

DOI: 10.1103/PhysRevD.105.063521

I. INTRODUCTION

Loop quantum gravity (LQG) is a nonperturbative framework [1] providing a tentative quantization of general relativity (GR). It has been expressed both in the canonical form [2] and in a covariant way [3]. As for all speculative theories the challenge is twofold. On the one hand, one has to check internal consistency. This is far from being a trivial requirement, especially in quantum gravity. From gauge issues to infrared corrections, quite a lot of questions remain—at least partially—open (see, e.g., [4]). On the other hand, it is mandatory to face the outstanding question of phenomenological consequences [5]. Making links with observations is the key missing ingredient for all attempts in quantum gravity, including string theory [6].

In this article, we address the question of the robustness of some predictions of LQG in the cosmological sector. Many different aspects have already been investigated, taking into account in particular (see, e.g., [7–15])

(i) the way initial conditions are set,
(ii) the validity of the minisuperspace approximation,
(iii) the backreaction effects,
(iv) the deformation of the algebra of constraints,
(v) the inclusion of shear and curvature,
(vi) the quantization of operators associated with negative powers of the volume operator,
(vii) the inclusion of effects inferred from quantum reduced loop gravity or group field theory,
(viii) numerical results beyond the semiclassical approximation, etc.

Here, we tackle a different and somehow underestimated question: the consequences of a generalized holonomy correction. The point is *not* to invent what would be a superexotic theory, with new free parameters, to boost the phenomenological richness. Just the other way round, the aim is to investigate how reliable are the predictions made so far, taking into account implicit assumptions that went mostly unnoticed and may play an important role.

The issue of quantization ambiguities in this framework was pointed out in [16]. Those associated with the quantization of the connection-based holonomy variable might deeply influence the dynamics and constitute the subject of this article. New theoretical arguments are being given in [17], while the present work focuses on potential observable effects. The question is especially important and meaningful when considered from a renormalization point of view.

In the following, the basics of loop quantum cosmology (LQC) are first briefly reminded. We then go into the details of generalized holonomy corrections. In the next section, we show that a long period of inflation can be generated using only a modified holonomy correction, without any massive scalar field. We also highlight the limits of such a model. The consequences of generalized holonomy corrections on both the background inflationary dynamics generated by a massive inflaton field and the scalar primordial power spectrum are finally exposed.

II. FLRW LOOP QUANTUM COSMOLOGY

In order to set the notations and remind the basics to the unfamiliar reader, we summarize the main ideas behind LQC. This also allows the article to be self-contained. In the fully constrained Ashtekar-Barbero formulation of GR, the canonical variables are

$$A_a^i \equiv \Gamma_a^i + \gamma K_a^i \quad \text{and} \quad E_i^a \equiv \frac{1}{2}\varepsilon^{abc}\varepsilon_{ijk}e_b^j e_c^k, \qquad (1)$$

where Γ_a^i is the $su(2)$ spin connection, γ the Barbero-Immirzi parameter, K_a^i the extrinsic curvature, ε^{abc} the totally anti-symmetric tensor, and e_a^i the triads. In this work, we use a, b, c as spacetime indices and i, j, k as internal $su(2)$ algebra indices. Both sets run from 1 to 3. These canonical variables satisfy the relations

$$\{A_a^i(x), E_j^b(y)\} = \kappa\gamma\delta_j^i\delta_a^b\delta^3(x-y), \tag{2}$$

with $\kappa = 8\pi G$.

As GR is a fully constrained theory, its Hamiltonian is written as a sum of constraints, respectively called scalar, vector and Gauss constraints. When homogeneity is assumed, the scalar contribution C_g is the only one to remain. The Hamiltonian can then be written using the lapse function N as

$$C_g^N = \int_\Sigma \mathrm{d}x^3 N C_g \tag{3}$$

$$= \frac{1}{2\kappa}\int_\Sigma \mathrm{d}x^3 \frac{N}{\sqrt{q}} E_i^a E_j^b (\varepsilon_k^{ij} F_{ab}^k - 2(1+\gamma^2) K_{[a}^i K_{b]}^j), \tag{4}$$

with Σ a compact hypersurface. In a homogeneous, isotropic, and flat space, the metric reduces to the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\delta_{ab}\mathrm{d}x^a\mathrm{d}x^b, \tag{5}$$

where the cosmic time is related to the 0-coordinate by $\mathrm{d}t = N\mathrm{d}x^0$. To avoid divergent integrals and an ill-defined symplectic geometry, we perform the integration on an arbitrary cubic fiducial cell of comoving volume V_0. As the homogeneity assumption also implies that spatial derivatives vanish, the spin connection disappears and the canonical variables become simply

$$A_a^i(t) = \gamma\dot{a}(t)\delta_a^i \equiv \frac{c(t)}{V_0^{1/3}}\delta_a^i, \tag{6}$$

where the dot represents a derivative with respect to the cosmic time t, and

$$E_i^a(t) = a^2(t)\delta_i^a \equiv \frac{p(t)}{V_0^{2/3}}\delta_i^a, \tag{7}$$

with the relation

$$\{c, p\} = \frac{\kappa\gamma}{3}. \tag{8}$$

Finally, the scalar constraint in this setting reduces to the simple form

$$C_g^N = -\frac{3}{\kappa\gamma^2} N\sqrt{p}c^2. \tag{9}$$

The lapse function N represents a gauge freedom.

In addition to the gravitational sector, we introduce a scalar field ϕ with an arbitrary potential $V(\phi)$ to investigate an early inflationary period. Using the canonical variables for the scalar field, namely ϕ and $\pi_\phi = p^{3/2}\dot{\phi}$, such that $\{\phi, \pi_\phi\} = 1$, the total Hamiltonian describing the coupled system can be written as

$$C^N = C_g^N + C_m^N = N\left(-\frac{3}{\kappa\gamma^2}\sqrt{p}c^2 + p^{3/2}\rho\right), \tag{10}$$

where $\rho = \pi_\phi^2/(2p^3) + V(\phi)$. The first Friedmann equation can easily be recovered using the evolution equation—that is $\dot{p} = \{p, C^N\}$—with the choice $N = 1$ and the Hamiltonian constraint. In its usual form, it is written as

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\kappa}{3}\rho. \tag{11}$$

Up to now, we have simply recovered the usual GR result within a specific framework. An effective Hamiltonian including corrections from LQG is yet to be constructed. The first step toward canonical quantization with well-defined operators in the quantum theory is to rewrite the Hamiltonian constraint using the holonomy of the connection. In other words, instead of deriving the curvature operator F_{ab}^i directly from the connection c, we use the holonomy $h_{\square_{ij}}$ of the connection on a fiducial square curve $\square_{ij}$ of length $\mu V_0^{1/3}$, with edges in the directions i, j. The holonomy measuring the extent to which the parallel transport of a vector around closed loops fails to preserve the transported vector, its form depends on the chosen $SU(2)$ representation for the parallel transport along the curves. This is known as the spin ambiguity. In standard LQC, the holonomy is calculated using the fundamental 2D representation of $SU(2)$, but the holonomy correction has been calculated for arbitrary representations in [18]. In this section, we describe the procedure for the fundamental spin $1/2$ representation and we recall the procedure to follow for general corrections.

In the harmonic gauge, where $N = p^{3/2}$, the Hamiltonian constraint in terms of the curvature operator reads

$$C_g^h = -\frac{1}{2\kappa\gamma^2} p^2 V_0^{2/3} \bar{e}_i^a \bar{e}_j^b F_{ab}^k, \tag{12}$$

where the superscript h stands for the harmonic gauge and $\bar{e}_i^a$ are the cotriads such that $q_{ab} = a^2(t)\bar{e}_a^i\bar{e}_b^j\delta_{ij}$. The holonomy corrected curvature operator in the fundamental representation is

$$F^k_{ab} = \lim_{\mu\to 0} \frac{-2}{\mu^2 V_0^{2/3}} \mathrm{Tr}\{h_{\square_{ij}} \tau^k\} \bar{e}^i_a \bar{e}^j_b, \tag{13}$$

where τ^k are the generators of the $su(2)$ algebra and, in this case, are represented by the Pauli matrices. By taking the limit $\mu \to 0$, the usual definition of the curvature operator and of the Hamiltonian of standard GR are recovered. However, LQG teaches us that the lowest nonzero eigenvalue of the quantum area operator is $\lambda^2 = 4\sqrt{3}\pi\gamma l_{pl}^2$, hence taking the limit down to zero is in principle not allowed. One therefore chooses $\mu \to \bar{\mu} = \lambda/\sqrt{p}$ as an estimator of the smallest possible length for the edge of the square curve. The holonomy can then be calculated along a closed curve defined by

$$h_{\square_{ij}} = h_i \circ h_j \circ h_{-i} \circ h_{-j}, \tag{14}$$

$$\text{where } h_{\epsilon i} := \exp\{\epsilon \mu c \tau_i\}, \tag{15}$$

with $\epsilon = \pm 1$. Putting everything together, one obtains the Hamiltonian constraint of LQC coupled to a scalar field, that is

$$C^h = -\frac{3}{\kappa\gamma^2\bar{\mu}^2} p^2 \sin^2(\bar{\mu}c) + p^3\rho. \tag{16}$$

This modified Hamiltonian can be recovered from the Hamiltonian of GR using the substitution

$$c^2 \to \frac{\sin^2(\bar{\mu}c)}{\bar{\mu}^2}, \tag{17}$$

usually called "the holonomy correction." Finally, it is possible to derive the modified Friedmann equation of LQC using $\dot{p} = \{p, C^h\}$, together with the Hamiltonian constraint $C^h = 0$. This leads to:

$$H^2 = \frac{\kappa}{3}\rho\left(1 - \frac{\rho}{\rho_c}\right), \tag{18}$$

with $\rho_c = 3/(\kappa\gamma^2\lambda^2)$. The remarkable feature of this new equation is the resolution of the big bang singularity. When $\rho \to \rho_c$ the Hubble parameter vanishes, as obvious from Eq. (18), and a bounce occurs instead of the GR singularity.[1] When choosing the usual value for the Barbero-Immirzi parameter $\gamma = 0.2375$, the critical density is of the order of $\rho_c \approx 0.41\rho_{\mathrm{Pl}}$.

The general case for the curvature operator, calculated using an arbitrary d-dimensional irreducible representation of $SU(2)$, was studied in [21,22]. A new closed formula for the Hamiltonian of flat FLRW models regularized with arbitrary spins was found in [18] and happens to be polynomial in the basic variables, which corresponds to well-defined operators in the quantum theory (taking also into account the inverse-volume corrections). The key-point lies in the fact that the computation in a representation of spin j of the trace of an $SU(2)$ group element does not require the explicit knowledge of all its matrix elements and can be reduced to an expression involving only the trace in the fundamental representation and the class angle. The curvature operator can then be written as

$$F^k_{ab} = \frac{-3}{d(d^2-1)} \frac{1}{\bar{\mu}^2 V_0^{2/3}} \frac{\sin^2(\bar{\mu}c)}{\sin\theta} \frac{\partial}{\partial\theta}\left(\frac{\sin(d\cdot\theta)}{\sin\theta}\right) \varepsilon^k_{ij} \bar{e}^i_a \bar{e}^j_b, \tag{19}$$

with

$$\theta = \arccos\left(\cos(\bar{\mu}c) + \frac{1}{2}\sin^2(\bar{\mu}c)\right). \tag{20}$$

This was derived with a technique quite similar to the one described previously for the holonomy regularization. Another technique to find the curvature operator, called connection regularization, can also be effectively considered [18]. In this approach, a new definition of the curvature operator, only valid in homogeneous space, is used and the result for F^k_{ab} is slightly different. In the literature, higher order holonomy corrections were also investigated in details [22–24]. They arise when higher order terms in powers of $\bar{\mu}$, usually neglected, are taken into account in the expression for the regularized curvature. There could exist a link between these higher order holonomy corrections and the contribution of higher spin representations. However, it was shown in [18] that these effects have actually very different physical consequences. At any order in holonomy corrections, a physical Hilbert space can be rigorously constructed and a complete family of Dirac observable can be identified [22].

It is therefore mandatory to understand the cosmological implications of more general holonomy corrections.

III. COSMOLOGY WITH ARBITRARY HOLONOMY CORRECTIONS

Instead of focusing on specific cases within the LQC framework with either different spin representations or higher order terms, we remain as general as possible for the expression for the holonomy correction. This can be studied by the substitution

$$c^2 \to g^2(c, p), \tag{21}$$

where $g(c, p)$ is an arbitrary function such that, in the low energy limit, standard GR is recovered, that is $g(c, p) \to c$.

[1] We however want to emphasize that contrary to what is often believed, a past singularity is *not* unavoidable in GR, even without exotic matter contents [19,20].

It should be emphasized that this is not only a way of taking into account the lessons from specific situations in LQC, but that this also make sense from a fully generic quantization ambiguity/renormalization viewpoint. Furthermore, in the totally constrained Hamiltonian formalism of GR, all the constraints are first class. We do not relax this requirement so that the evolution operator keeps the subspace of physical states invariant. As shown in [25], keeping only first class constraints and following the usual Dirac prescriptions adds an extra condition on the function $g(c,p)$:

$$g(c,p)=\frac{1}{\bar{\mu}}f(b), \tag{22}$$

where $f(b)$ is an arbitrary function of $b=\bar{\mu}c$, which must behave as $f(b)\approx b$ at low energies (*i.e.,* when $b\ll 1$). As opposed to [25], we used the parameter λ to respect the units of length of $g(c,p)$ and be consistent with LQC. Fortunately, this is fully compatible with the correction given in Eq. (19), ensuring that any holonomy modification coming from an arbitrary spin-representation will keep the algebra of constraints consistent. It is quite remarkable that the "anomaly freedom" requirement (see [26–28] for general considerations) allows one to sharpen the general expression, in a way precisely compatible with known corrections expected in the loop framework. The simple—and mandatory—fact that the evolution vector is asked to be parallel to the submanifold of constraints severely reduces the *a priori* freedom.

The modified equations of motion for the canonical variables c and p are calculated using Hamilton's equations. It is more natural to write them in terms of p and b. Together with the Hamiltonian constraint, they take the form

$$\dot{b}=-\frac{\lambda\kappa\gamma}{2}\rho(1+w), \tag{23}$$

$$\dot{p}=\frac{2}{\gamma\lambda}pf(b)f'(b), \tag{24}$$

$$\frac{\rho}{\rho_c}=f^2(b), \tag{25}$$

where $f'(b)$ should be understood as $\mathrm{d}f(b)/\mathrm{d}b$, $w=P/\rho$ and $P=\pi_\phi^2/(2p^3)-V(\phi)$. To derive Eq. (23) we used the continuity equation

$$\dot{\rho}=-3\frac{\dot{p}}{2p}\rho(1+w). \tag{26}$$

A general modified Friedmann equation can be found using $H=\dot{p}/(2p)$ together with the constraint (25) and can be written as

$$H^2=\frac{\kappa}{3}\rho(f'(b))^2. \tag{27}$$

Since, by construction, $f(b)\to b$ when $b\to 0$, we indeed recover the usual Friedmann equation $H^2=\kappa\rho/3$ in this limit. Let us now investigate the behavior of the model starting in a regime where GR is valid and going backward in cosmic time t, toward a classical singularity associated with $\rho\to\infty$ in GR. Using the null energy condition, $w\geq -1$ together with Eqs. (23) and (24) one can easily show that $\mathrm{d}b/\mathrm{d}t<0\,\forall\, t$. Hence, if we start with $b>0$ in the GR regime, b is always positive and increasing when going backward in time. This is expected as, in the GR limit, the proportionality relation $\rho\propto b$ is satisfied and the density increases in the past direction. Furthermore, since one has $f(b)\approx b$ in the GR regime, the function f is monotonic and strictly increasing with b around $b=0$. In the case where there exists a local maximum $b_{\mathrm{bounce}}>0$, implying $f'(b_{\mathrm{bounce}})=0$, one can see with Eqs. (25) and (27) that the density has to reach a critical value ρ_b, where the Hubble parameter vanishes. It is therefore meaningful to conclude that the big bang singularity is resolved by a bouncing scenario of geometrical origin if and only if the holonomy correction reaches a local maximum. If the function f is strictly monotonic, two different scenarios have to be considered. Either $\lim_{b\to\infty}f(b)=\infty$, in which case the singularity is not resolved, or $\lim_{b\to\infty}f(b)<\infty$ and the situation is similar to eternal inflation where ρ tends to a final constant value behaving as a positive cosmological constant.

IV. FIRST REMARKS ON INFLATION FROM THE HOLONOMY CORRECTION

A natural question arising in this framework is to wonder whether it is possible to describe a long-lasting phase of inflation using only a modification of the holonomy correction without the need for a fluid satisfying the equation of state $w<-1/3$. In the usual LQC framework, the quantum geometrical super-inflation occurring after the bounce cannot account for more than a few e-folds and most of the known inflationary features are due to a hypothetical massive scalar field filling the Universe. Furthermore, one should also investigate if the inflation associated with generalized holonomies could explain the quasi-scale invariance of the power spectrum observed in the cosmological microwave background (CMB). The answer turns out to be positive. However, important drawbacks inherent to the construction will be mentioned in this section. We assume here that the content of the Universe is a massless scalar field, that is $w=1$ at all scales. We also restrict ourselves to holonomy corrections such that there exists a bounce so as to keep the huge benefit of the singularity resolution.

First of all, the correction $f(b)$ has to be chosen so as to ensure an exponential growth of the scale factor a. This is

achieved by $\dot{H} \simeq 0$: the Hubble parameter is nearly constant during inflation. From the modified Friedmann equation (27) and the Hamiltonian constraint (25), one can easily conclude that the holonomy correction should take the following form:

$$f(b) \sim \sqrt{b}. \tag{28}$$

We assume that $f(\pi) = 0$ and call b_{bi} the value of b at the beginning of inflation. We set $b = b_{rh}$ at the reheating, when classical cosmology is recovered. It is therefore necessary to find a function $f(b)$ such that: (i) $f(b) \sim b$ around $b = 0$, to recover GR, (ii) $f(b) \sim \sqrt{b}$ for $b_{rh} < b < b_{bi}$ during the inflationary period, and (iii) $f(b)$ reaches a maximum between b_{bi} and $b = \pi$ to induce a bounce. In order to make explicit that such a function can be constructed, we give an arbitrary example:

$$f^2(b) = (1 + C_1 b)^{1-\alpha} \sin^2(b) \frac{1}{C_1 + 1} \sum_{n=0}^{C_1} \cos^{2n}(b) \tag{29}$$

$$= \frac{(1 + C_1 b)^{1-\alpha}}{1 + C_1} (1 - \cos^{2(C_1+1)}(b)), \tag{30}$$

where C_1 can be chosen in accordance with the parameters b_{bi} and b_{rh} so as to obtain the desired inflationary duration. The parameter α can be chosen to make the power spectrum slightly red. To illustrate the behavior of this holonomy correction, Eq. (30) is plotted on Fig. 1 for different values of C_1, assuming $\alpha = 0$. The higher the value of C_1, the lower the value of b at the end of inflation (denoted b_{rh}) and the closer b_{bi} to π. Let us take a closer look at Eq. (30) in the case $C_1 \gg 1$. If $b \gg 1/\sqrt{C_1}$, one indeed recovers $f^2(b) \sim b$, because the cosine term is small compared to 1 and $C_1 b \gg 1$. On the other hand, when $b \ll 1/C_1$ one has $(1 + C_1 b) \to 1$ while $(1 - \cos^{2C_1}(b)) \sim C_1 b^2$, leading to $f^2(b) \sim b^2$. However, in the intermediate case, $1/C_1 \ll b \ll 1/\sqrt{C_1}$, $(1 + C_1 b) \sim C_1 b$ while $(1 - \cos^{2C_1}(b))$ still behaves as $C_1 b^2$, leading to $f^2(b) \sim b^3$. Hence, the transition at reheating does not straightforwardly go from $f^2(b) \sim b$ to $f^2(b) \sim b^2$, but goes through transition phase, behaving as $f^2(b) \sim b^3$. This is not problematic for our purpose, but it is worth being pointed out. Equation (25) shows that in order to have N e-folds of inflation in a matter dominated universe, one needs $f^2(b_{rh}) e^{3N} = f^2(b_{bi}) \sim 1$. Therefore, if one chooses $C_1 \sim e^{6N}$, the inflationary period ends when $b_{rh} \sim e^{-3N} \Rightarrow f^2(b_{rh}) \sim e^{-3N}$, leading to the desired N e-folds. Then, the transition phase takes place while $e^{-6N} < b < e^{-3N}$—that can call reheating—and finally we recover classical cosmology for $b < e^{-6N}$. Obviously, a quite strong fine-tuning is needed but solutions matching all the requirements can be found. We shall discuss this issue later.

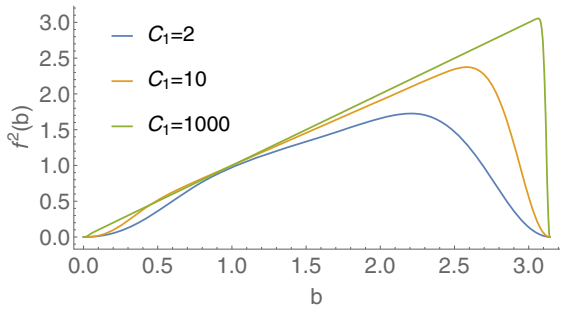


FIG. 1. Holonomy corrections as described by Eq. (30) for different values of the parameter C_1.

Another potential problem to consider is related with the evolution of the energy density. In usual models of inflation, where the exponentially accelerating expansion is produced by a negative pressure fluid, the density of the latter stays roughly constant around (slightly above) the density of reheating $\rho \sim \rho_{rh}$. However, the density of radiation, dust or a massless scalar field evolve as a^{-4}, a^{-3}, and a^{-6}, respectively, thus increase exponentially during inflation as well (when thinking backward in time). If inflation is of quantum geometrical origin and not caused by the standard mechanism of a field slowly rolling on its potential, there is no reason to assume that the usual "contents" are not present and possibly dominant. This means that if one requires a reheating around the GUT scale $T_{\mathrm{GUT}} \sim 10^{15}$ GeV and imposes roughly $N = 65$ e-folds of exponential expansion, the energy at the beginning of inflation T_{bi} would be vastly trans-Planckian. On the other hand, one can impose the maximum energy at the Planck scale $T_{bi} \sim T_{pl}$ and simply ask for more e-folds before the reheating than after, to solve the usual cosmological problems. However this translates into $T_{rh} \leq 1$ TeV. Although unusual, this value is not strictly ruled out by observations.

Finally, it is worth emphasizing that the scalar power spectrum, as observed in the CMB, cannot be easily recovered for other contents than a massive scalar field. To see this, let us consider a massless scalar field and a holonomy correction of the form (30) with a sufficiently long inflationary phase. In this example, we choose $T_{bi} \sim T_{pl}$ and roughly $N = 35$ e-folds. The gauge invariant scalar perturbation can, in this context, be described by the Mukhanov-Sasaki (MS) variables $v_k(\eta)$ and $z(\eta) = \dot{\phi} a/H$, where η is the conformal time, satisfying the equation

$$v_k''(\eta) + \left(k^2 - \frac{z''(\eta)}{z(\eta)}\right) v_k(\eta) = 0. \tag{31}$$

It should be noticed that this classical equation is also the one fulfilled by perturbations in the dressed metric/hybrid quantization approach to LQC [29,30]. The scalar power spectrum is obtained by

$$\mathcal{P}_S(k) = \frac{k^3}{2\pi^2}\left|\frac{v_k}{z}\right|^2\Big|_{k=aH} \equiv \frac{k^3}{2\pi^2}|\mathcal{R}_k|^2|_{k=aH}, \quad (32)$$

where, in the standard inflation theory, the curvature perturbation $\mathcal{R}_k$ is constant for super-Hubble modes $k \leq aH$. The scalar power spectrum therefore keeps a low amplitude matching the observed value. In the case of the massless scalar field with an inflationary stage due to the holonomy correction, one can show using Eq. (31) that $\mathcal{R}_k$ is not constant for super-Hubble modes anymore. During the inflationary period, the Hubble parameter H is constant, $a \propto e^{Ht}$ and $\dot{\phi} = 2\sqrt{\rho}$, meaning that the second MS variable behaves as

$$z(a) \sim a^{-1/2(1+3w)}. \quad (33)$$

In the considered example $w = 1$ but we keep the equation of state arbitrary so that the conclusion remains general. Rewriting the Mukhanov-Sasaki equation (31) with respect to a and taking the limit of large values of a, as one might expect for super-Hubble modes, we get

$$\frac{\mathrm{d}^2 v_k}{\mathrm{d}a^2} + \frac{1}{a}\frac{\mathrm{d}v_k}{\mathrm{d}a} - \frac{2}{a^2}v_k = 0, \quad (34)$$

with solutions $v(a) \sim a$ and $v(a) \sim 1/a^2$. In standard inflation with a massive scalar field, where $w \approx -1$, choosing the Bunch-Davies vacuum comes down to selecting the behavior $v(a) \sim a$ in quasi–de Sitter space, hence one recovers $\mathcal{R}_k = v_k/z \sim a^0$. In the case of a massless field ($w = 1$), such a choice of vacuum leads to the behavior $\mathcal{R}_k \sim a^3$ and therefore the scalar perturbations with super-Hubble modes exponentially increase during inflation. To illustrate this, the scalar power spectrum computed from Eq. (31) for an inflation induced by a holonomy correction with massless scalar field is shown in Fig. 2. As one can see, the spectrum is scale invariant in the UV and keeps its usual shape but the amplitude is meaningless. In principle, there might exist another vacuum selecting $v(a) \sim 1/a^2$ and one would recover $\mathcal{R}_k \sim a^0$. However, in the case of other matter contents, such as radiation ($w = 1/3$) or dust ($w = 0$), the scalar perturbations $\mathcal{R}_k$ cannot be frozen for $k \geq aH$ in geometric inflation. This is an important point, often forgotten, which should be taken into account in phenomenological studies.

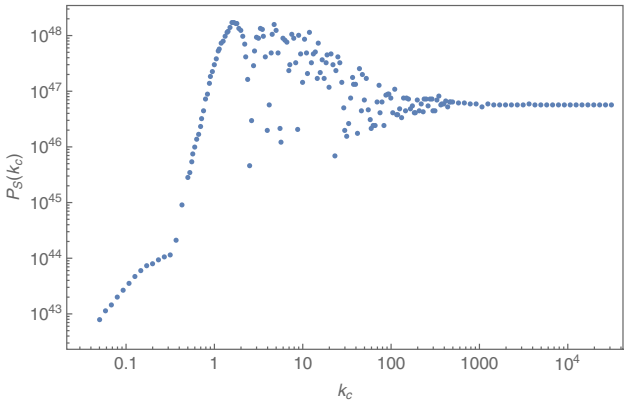


FIG. 2. Scalar power spectrum obtained with a massless scalar field from a stage of "geometrical" inflation entirely produced by the modified holonomy correction.

It should however be emphasized that perturbations were, more than a decade ago, reunderstood in the framework of the effective theory of inflation [31]. This "new" approach somehow disentangles the question of perturbations themselves from the question of the process which generates them. From this point of view, this work is anyway interesting as the perturbations can be treated even without any use of a scalar field. All that is needed is a background and an effective "clock".

V. FOCUS ON THE DURATION OF INFLATION

Detailed studies of the duration of inflation as predicted by LQC have already been made in [14,32,33] for initial conditions set in the remote past. In the case of a massive scalar field playing the role of the inflaton, it was shown that the number of e-folds, as predicted by LQC, is around $N = 140$. Interestingly, the number of e-folds varies only slightly with respect to most contingent parameters. This can easily be seen by studying trajectories in the potential energy $x(t) = \sqrt{V(\phi(t))/\rho_c}$ versus kinetic energy $y(t) = \sqrt{\dot{\phi}/(2\rho_c)}$ plane. The number $N = 140$ corresponds to a quadratic potential $V(\phi) = m^2/2\phi^2$ with a mass parameter $m \sim 1.2 \times 10^{-6}$. For other potentials, the number of e-folds can be different, but remains of this order of magnitude as long as one deals with confining potentials. Although anisotropies can slightly decrease this value, the duration of inflation is a robust property of the background in LQC (a different proposal was however suggested in [34] but relies on conditions set at the bounce, a path that we do not follow here). The fact that this number is way smaller than naïve expectations is a key feature of bouncing scenarios. One can therefore naturally wonder if this prediction changes for different shapes of the holonomy correction. The set of initial prebounce conditions (x_0, y_0) is described using a phase parameter $\delta \in [0, 2\pi[$ such that

$$x_0 = \sqrt{\frac{\rho_0}{\rho_c}}\cos\delta \quad \text{and} \quad y_0 = \sqrt{\frac{\rho_0}{\rho_c}}\sin\delta, \quad (35)$$

where ρ_0 is the initial energy density and defines how far away in the past initial conditions are set. It has been checked that the chosen distribution for δ is conserved over time. For completeness, we reproduce the probability distribution function (PDF) using the settings of [32] in Fig. 3 with a uniform distribution of the phase parameter δ.

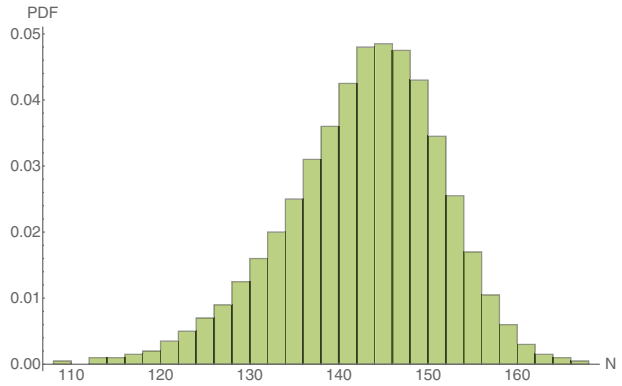


FIG. 3. Probability distribution function of the number of e-folds of inflation for a quadratic potential with $m = 1.2 \times 10^{-6}$ in LQC.

In the following, the mean value of the number of e-folds for different holonomy corrections is calculated using such a PDF.

In this section, we investigate the effect of the form of the holonomy correction on the duration of inflation. For this purpose, we follow the strategy described in [14,33]: we assume a uniformly distributed set of initial conditions for δ and the initial time $t = 0$ sufficiently far in the contracting branch so as to remain away from the quantum gravitational regime. In order to stay conservative about the holonomy correction shape, we study different general properties such as an asymmetry, an increase of the maximum of f^2, and a flattening of the correction on the length of inflation. Overall, we want to keep the asymptotic behavior of LQC, i.e., $f(0+\delta b) = f(\pi+\delta b) = \delta b + \mathcal{O}(\delta b^2)$, hence we choose a correction of the form

$$f^2(b) = \sin^2(b)(1 + A_1 b^{n_1}(b-\pi)^{n_2}), \tag{36}$$

where $n_i \geq 1$, $i = 1, 2$ and $A_1 \geq 0$, or

$$f^2(b) = \sin^2(b)\frac{1}{C_1+1}\sum_{n=0}^{C_1}\cos^{2n}(b), \tag{37}$$

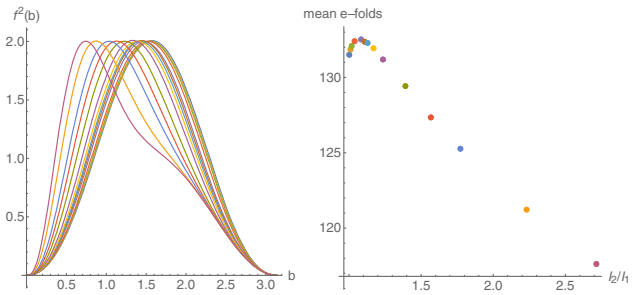


FIG. 4. Left: asymmetric holonomy corrections. Right: mean value of the PDF of the number of inflationary e-folds.

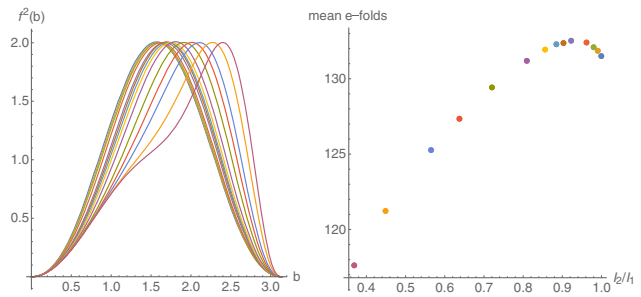


FIG. 5. Left: asymmetric holonomy corrections. Right: mean value of the PDF of the number of inflationary e-folds.

where $C_1 \geq 1$. Such parametrizations ensure the desired behavior at the fix points $b = 0, \pi$.

The results of the numerical simulations for different left and right asymmetries are shown in Figs. 4 and 5. The left panels represent the chosen holonomy corrections whereas the right panels display the mean value of the PDF of the number of e-folds when scanning the full range of initial phases. It should be pointed out that the PDF is narrow enough ($\sigma \sim 10$ e-folds) so that its first moment gives the relevant information. The strength of the asymmetry is measured by the ratio I_2/I_1 between the integral of $f^2(b)$ on $[0, b_{\max}]$ and the integral on $[b_{\max}, \pi]$, where $b_{\max}$ is the value that maximizes $f^2(b)$. The general trend is a decrease of the number of e-folds (although this is not true for very small deformations).

The results of the simulations for different amplitudes and plateaulike functions are displayed in Figs. 6 and 7, respectively. In general, increasing the energy density at the bounce decreases the number of e-folds and flattening the correction (or reducing the energy of the bounce) increases the length of inflation. Since the bounce energy density grows up with $\max f^2$, as shown by Eq. (25), this result seems counter-intuitive at first sight. It is true that the inflation duration generally increases with the bounce energy scale, but this is not the only important ingredient relevant for the number of inflationary e-folds. Another key parameter is the ratio between potential and kinetic energies at the bounce $x(t_b)/y(t_b)$. For kinetic energy dominated bounces, more time will indeed be required before the potential energy finally dominates and the slow-roll

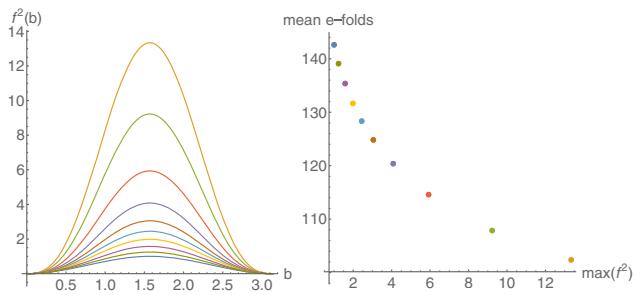


FIG. 6. Left: amplitude-varying holonomy corrections. Right: mean value of the PDF of the number of inflationary e-folds.

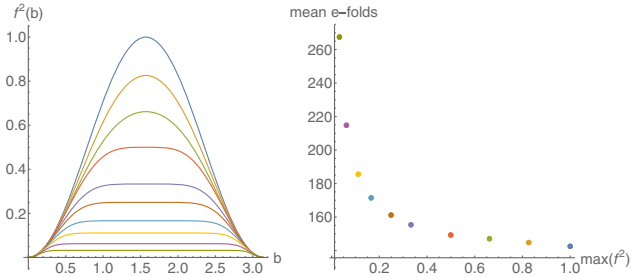


FIG. 7. Left: plateau-like holonomy corrections. Right: mean value of the PDF of the number of inflationary e-folds.

conditions are met. This leads to a lower inflation energy scale than for potential energy dominated bounces. Thus, bounces dominated by kinetic energy will actually lead to shorter inflation phases. This is the reason why the PDF of the number of inflationary e-folds in LQC is peaked around low values of N, close to the experimental lower bound. Setting initial conditions in the remote past of the contracting branch with a flat PDF on the initial phase of the scalar field leads to trajectories with kinetic energy dominated bounces, with ratios $x(t_b)/y(t_b)$ typically of order 10^{-6}. In the present study, the higher $\max f^2$, the more kinetic energy dominated the bounce is. This phenomenon is shown in Fig. 8. First, this indeed confirms in the top right and middle left plots, representing respectively $\max f^2$ with respect to $x(t_b)/y(t_b)$ and the number of e-folds with respect to $x(t_b)/y(t_b)$, that $\max f^2$ has a significant effect on this ratio, which in turn affects the number of e-folds as expected. Using the middle and bottom right plots, one can see that it takes longer for a kinetic energy dominated bounce to start the inflationary period, leading to a lower energy density at inflation, thus lowering the number of e-folds.

There are several lessons to be learnt from this analysis. There is an obvious loss of predictivity associated with a possible generalization of the holonomy correction. This can hardly come as a surprise. Whether this should be considered as good or bad news depends on the point of view. This obviously makes the theory less easily falsifiable. But this also opens up interesting possibilities. It shows that desired cosmological behaviors can be obtained by purely geometrical effects, relaxing the need for exotic matter contents. This also means that, in principle, a good knowledge of the cosmological dynamics can severely constrain the shape of the holonomy correction. In particular, the minimum number of e-folds $N > 60$ required to solve the horizon problem sets a constraint on the maximum of $f^2(b)$ as N decreases when the maximum density at the bounce increases. Most arbitrary shapes are problematic. It should however be underlined that as long as one deals with moderate corrections to the usual $\sin^2(\bar{\mu}c)/\bar{\mu}^2$, the effects on the background dynamics remain quite small and the core of the known LQC predictions remains valid.

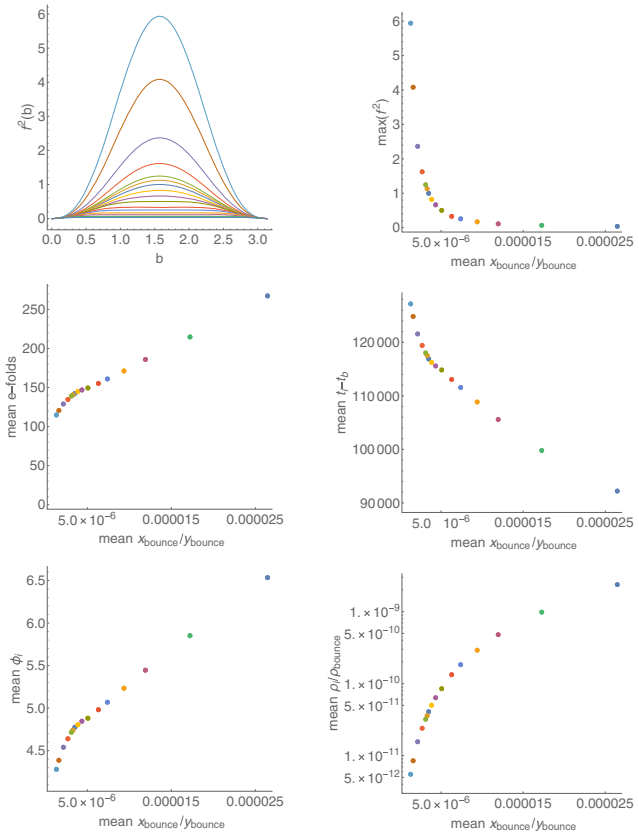


FIG. 8. Top left: amplitude-varying holonomy corrections. Top right: maximum of $f^2(b)$ with respect to the ratio of potential to kinetic energy of the scalar field at the bounce, $x(t_b)/y(t_b)$. Middle left: mean value of the PDF of the number of inflationary e-folds with respect to $x(t_b)/y(t_b)$. Middle right: mean value of the elapsed time between the bounce and inflation with respect to $x(t_b)/y(t_b)$. Bottom left: mean value of the scalar field at the beginning of inflation with respect to $x(t_b)/y(t_b)$. Bottom right: mean value of the ratio of the density at the beginning of inflation over the one at the bounce with respect to $x(t_b)/y(t_b)$.

The ambiguities considered in this work do *not*, by themselves, reveal a theoretical failure. So as to regularize quantum operators associated with nonlinear functionals of the fundamental fields, one relies on the diffeomorphism invariant prescription of "point-splitting" [16]. It happens to be that the regulator can be removed without encountering UV divergences. One is then left with a well defined quantum Hamiltonian constraint, at the price of having many different quantum theories. This is obviously reminiscent of the usual problem of renormalization of quantum fields: the correct theory must be fixed by the renormalization conditions.

VI. EFFECTS ON THE PRIMORDIAL POWER SPECTRA

We now analyze the effects of a modified holonomy correction on the scalar power spectrum when considering

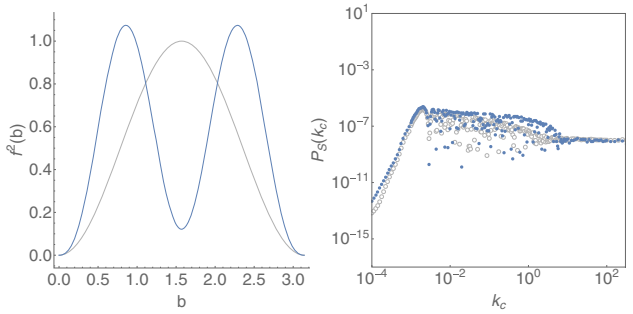


FIG. 9. Left: holonomy corrections for standard LQC (in gray) and for a double bounce (in blue). Right: comparison between the associated scalar power spectra (standard LQC in gray and with modified holonomy correction in blue).

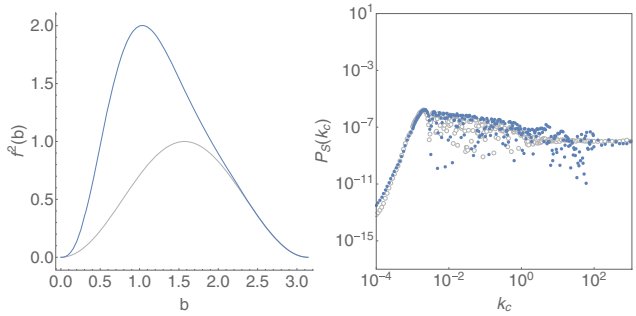


FIG. 11. Left: holonomy corrections for standard LQC (in gray) and for a left asymmetric bounce (in blue). Right: comparison between the associated scalar power spectra (standard LQC in gray and with modified holonomy correction in blue).

an inflationary period generated by a massive scalar field of mass $m = 1.2 \times 10^{-6} m_{\mathrm{Pl}}$. We keep the form of the Mukhanov-Sasaki equation for the scalar perturbations unchanged. This allows to specifically investigate the effects of the modification of the background dynamics, due to the new holonomy correction, on the shape of the spectra, independently of possible modifications of the perturbations propagation equation (which is anyway beyond the scope of this study).

Figures 9 and 10 display comparisons of the usual LQC spectrum (in gray) with spectra obtained using new background dynamics due to modified holonomy corrections (in blue). Figure 9 represents a double bounce while a single bounce at higher energy is shown in Fig. 10. The new features in the spectra corresponding to the modified dynamics are hardly distinguishable from the usual spectrum. The spectra remain (almost) scale invariant and consistent with CMB data in the so-called ultraviolet region, that is the region corresponding to the highest presented wave numbers (and even higher values of k_c), in which the observable scales are located for almost all the parameter space. The modifications to $f^2(b)$ only impact the details of the oscillations in the intermediate regime. This was expected as those oscillations are mainly associated with the bounce and their details depend on the detailed dynamics at this time. The ultraviolet regime, i.e., the one of interest for comparison with data, is mostly independent of the bounce dynamics as the wave numbers of the corresponding modes are much greater than the potential z''/z during the bounce phase. In other words, those modes do not even feel the presence of the bounce(s) It is however worth emphasizing that, as highlighted in [35], the position of the oscillations is dictated by the value of the potential z''/z evaluated at the bounce. Thus, sharper bounces would lead to oscillations in the spectra located further away in the ultraviolet regime. It is also worthwhile to underline that modifications of $f^2(b)$ around $b \approx 0$ or π would modify the low energy behavior and have a deeper impact on the spectra.

We have previously seen that an asymmetry (in b) of the holonomy correction can have a significant influence on the evolution of the inflaton in the high energy regime. It is therefore important to estimate the effects of asymmetries in $f^2(b)$ on the spectra. Associated results are shown in Figs. 11 and 12. The main noticeable effect is an extension of the oscillatory regime toward higher wave numbers.

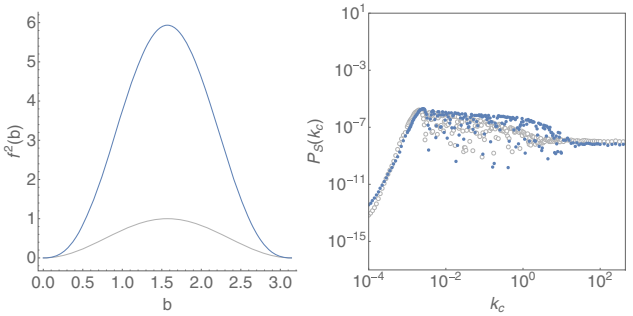


FIG. 10. Left: holonomy corrections for standard LQC (in gray) and for a higher energy bounce (in blue). Right: comparison between the associated scalar power spectra (standard LQC in gray and with modified holonomy correction in blue).

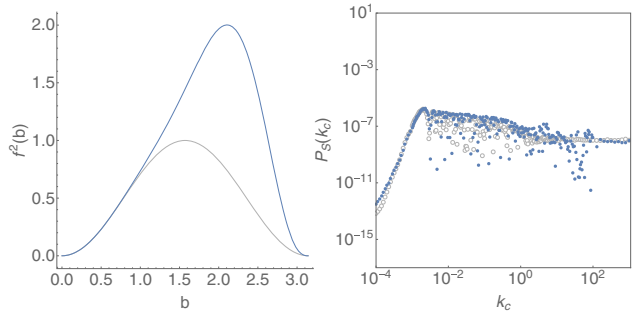


FIG. 12. Left: holonomy corrections for standard LQC (in gray) and for a right asymmetric bounce (in blue). Right: comparison between the associated scalar power spectra (standard LQC in gray and with modified holonomy correction in blue).

This extension is however (for all the functions tested) not sufficient to be game changing for the noticeable features of primordial power spectra in loop inspired cosmologies. All usual conclusions still hold. This effect only makes a possible detection/invalidation of the bounce in CMB data slightly more probable as it decreases a little the level of fine tuning required to bring the window of comoving wave numbers associated with CMB observations toward the oscillatory regime.

It should be underlined that the situation might be totally different in the case of the deformed algebra (DA) approach to perturbations, where the equation of propagation of perturbations is usually given by [36,37]:

$$v_k''(\eta) + \left(\Omega(\eta)k^2 - \frac{z''(\eta)}{z(\eta)}\right) v_k(\eta) = 0, \tag{38}$$

with $\Omega = 1\text{–}2\rho/\rho_c$. The change of sign of Ω at high energies can be interpreted as a switch from a Lorentzian to an Euclidean geometry [38]. The apparition of this Ω function in the MS equation comes from its presence in the anomaly-free algebra of constraints, more precisely in the Poisson bracket between scalar constraints [28]:

$$\{S^{\text{a.f}}[M], S^{\text{a.f}}[N]\} \approx \Omega D^{\text{a.f}}[q^{ab}(M\partial_b N - N\partial_b M)], \tag{39}$$

where S and D correspond respectively to the scalar and diffeomorphism constraints, the label a.f stands for "anomaly-free," and the $\approx$ symbol means that this relation is satisfied on the hypersurface of constraints. But the Ω function appears in this algebra as a consequence of the modification of the kinetic term in the scalar constraint (by the holonomy correction, that is Eq. (17) in the usual case). If we assume a $1+1$ dimensional toy model then $\Omega = (1/2)d^2f^2(b)/db^2$, which, in the case $f^2(b) = \sin^2(b)$, and using the Hamiltonian constraint, gives back the previous expression $\Omega = \cos(2b) = 1\text{–}2\rho/\rho_c$. It is thus clear that modifying the form of $f(b)$ would unavoidably modify the expression of Ω and therefore the Mukhanov-Sasaki equation. As a consequence, the shape of the primordial power spectra could be modified in a much more important way in the deformed algebra scheme than in the approach presented in this manuscript. This point definitely requires deeper investigations that are not the purpose of this paper.

We have focused on the scalar spectrum which is the more interesting one from the viewpoint of observations (and the more intricate to calculate). The general trends however obviously remain true for the tensor spectrum as the new features are due to the modified background dynamics.

VII. CONCLUSION

Quantization ambiguities are unavoidable. The only serious requirement in this framework is to recover general relativity (or the Wheeler-DeWitt equation) in the low energy limit of the theory. This leaves an infinite dimensional set of ambiguities in the choice of the function f. Investigating few of its phenomenological properties was the goal of this study.

We have shown that a long enough inflationary stage can be generated by an appropriate generalized holonomy correction without the need for matter violating the energy conditions. This however requires a serious amount of fine-tuning in the choice of the parameters entering the definition of the function. Nevertheless this concept is ill-defined when no natural measure is available. More profoundly, it is worth emphasizing that the fine-tuning is a real issue only when it is required to produce a situation which is *a priori* singular or specific (in a Bayesian sense). Otherwise, it is just a mere version of the nonproblematic tautology "if the laws had been different, the World would be different." In addition, we have shown that even though a nearly scale invariant primordial power spectrum can be generated, there are some (often forgotten) associated issues to face, making the spectrum hardly compatible with data.

The influence of the shape and amplitude of the holonomy correction on the background dynamics has also been investigated in details. The effects are quite weak and do not change drastically the usual conclusions of LQC as long as the used function remains close enough to the standard case. Interestingly some new effects tend to decrease the number of e-folds. This is a good news for phenomenology. With 140 inflationary e-folds, all the subtle quantum gravitational effects are deeply super-Hubble today. Smaller values bring back the hope that some footprints of the bounce might be seen in the CMB, especially when taking into account nonlinear effects where different modes fail to fully decouple one from the other.

ACKNOWLEDGMENTS

We would like to thank Alejandro Perez who was at the origin of this work and who provided many important inputs and explanations. We also thank Salvatore Ribisi and Lautaro Amedei for valuable discussions.

[1] A. Ashtekar and E. Bianchi, Rep. Prog. Phys. **84**, 042001 (2021).
[2] C. Rovelli, *Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2004).
[3] C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2014), ISBN: 1107069629, 9781107069626, 9781316147290.
[4] C. Kiefer, in *Time and Matter 2007* (University of Nova Gorica Press, 2007).
[5] A. Barrau, T. Cailleteau, J. Grain, and J. Mielczarek, Classical Quant. Grav. **31**, 053001 (2014).
[6] F. Quevedo, arXiv:1612.01569.
[7] A. Ashtekar, A. Corichi, and P. Singh, Phys. Rev. D **77**, 024046 (2008).
[8] A. Ashtekar and P. Singh, Classical Quant. Grav. **28**, 213001 (2011).
[9] I. Agullo, A. Ashtekar, and W. Nelson, Phys. Rev. Lett. **109**, 251301 (2012).
[10] P. Diener, B. Gupt, and P. Singh, Classical Quant. Grav. **31**, 105015 (2014).
[11] A. Ashtekar and A. Barrau, Classical Quant. Grav. **32**, 234001 (2015).
[12] B. Bolliet, J. Grain, C. Stahl, L. Linsefors, and A. Barrau, Phys. Rev. D **91**, 084035 (2015).
[13] E. Alesci and F. Cianfrani, Phys. Rev. D **92**, 084065 (2015).
[14] K. Martineau, A. Barrau, and S. Schander, Phys. Rev. D **95**, 083507 (2017).
[15] S. Gielen and D. Oriti, Phys. Rev. D **98**, 106019 (2018).
[16] A. Perez, Phys. Rev. D **73**, 044007 (2006).
[17] L. Amadei, A. Perez, and S. Ribisi, arXiv:2203.07044.
[18] J. Ben Achour, S. Brahma, and M. Geiller, Phys. Rev. D **95**, 086015 (2017).
[19] A. Barrau, Eur. Phys. J. C **80**, 579 (2020).
[20] C. Renevey, A. Barrau, K. Martineau, and S. Touati, J. Cosmol. Astropart. Phys. 01 (2021) 018.
[21] K. Vandersloot, Phys. Rev. D **71**, 103506 (2005).
[22] D.-W. Chiou and L.-F. Li, Phys. Rev. D **80**, 043512 (2009).
[23] J. Mielczarek and M. Szydlowski, Phys. Rev. D **77**, 124008 (2008).
[24] O. Hrycyna, J. Mielczarek, and M. Szydlowski, Gen. Relativ. Gravit. **41**, 1025 (2009).
[25] Y. Han and M. Liu, Classical Quant. Grav. **35**, 105017 (2018).
[26] J.-P. Wu, M. Bojowald, and Y. Ma, Phys. Rev. D **98**, 106009 (2018).
[27] T. Cailleteau, A. Barrau, J. Grain, and F. Vidotto, Phys. Rev. D **86**, 087301 (2012).
[28] A. Barrau, M. Bojowald, G. Calcagni, J. Grain, and M. Kagan, J. Cosmol. Astropart. Phys. 05 (2015) 051.
[29] I. Agullo, A. Ashtekar, and W. Nelson, Classical Quant. Grav. **30**, 085014 (2013).
[30] M. Fernandez-Mendez, G. A. Mena Marugan, and J. Olmedo, Phys. Rev. D **86**, 024003 (2012).
[31] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, J. High Energy Phys. 03 (2008) 014.
[32] L. Linsefors and A. Barrau, Classical Quant. Grav. **32**, 035010 (2015).
[33] B. Bolliet, A. Barrau, K. Martineau, and F. Moulin, Classical Quant. Grav. **34**, 145003 (2017).
[34] A. Ashtekar and D. Sloan, Phys. Lett. B **694**, 108 (2010).
[35] K. Martineau and A. Barrau, Universe **4**, 149 (2018).
[36] A. Barrau, P. Jamet, K. Martineau, and F. Moulin, Phys. Rev. D **98**, 086003 (2018).
[37] S. Schander, A. Barrau, B. Bolliet, L. Linsefors, J. Mielczarek, and J. Grain, Phys. Rev. D **93**, 023531 (2016).
[38] M. Bojowald and J. Mielczarek, J. Cosmol. Astropart. Phys. 08 (2015) 052.

Impact of generalized holonomy corrections on the cosmological primordial power spectra

Maxime De Sousa, Killian Martineau, Cyril Renevey, and Aurélien Barrau
Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3 53, avenue des Martyrs, 38026 Grenoble cedex, France

(Received 2 December 2022; accepted 25 May 2023; published 14 June 2023)

The propagation of perturbations is studied with generalized holonomy corrections in a fully consistent way, ensuring that the deformed algebra of constraints remains closed. The primordial cosmological power spectra are calculated. It is shown that, although the detailed form of the correction does unavoidably impact the observables, the main known results of loop quantum cosmology are robust in this respect.

DOI: 10.1103/PhysRevD.107.126008

I. INTRODUCTION

Loop quantum gravity (LQG) is a background-independent quantization of general relativity (GR) [1]. It can be expressed in the canonical form [2] or in the covariant way [3]. The theory has been successfully applied to both black holes and the early universe. Many consistency checks—mostly encouraging—have been carried out, although some important questions remain open [4]. An excellent philosophical introduction is given in [5].

The cosmological sector of the theory has received particular attention and numerous complementary aspects were investigated in details (see, e.g., [6–15], and references therein). The main conclusions are the following. The existence of a bounce replacing the usual big bang is a robust result. It has been shown analytically in simplified situations and proven to survive when a cosmological constant is added, when spatial curvature is taken into account, and when quite general potentials for the inflaton field are considered. In addition, the semiclassical states were demonstrated to remain sharply peaked, allowing the safe use of effective equations. Importantly, the duration of inflation is statistically predictable in this framework. Generic features for the primordial power spectra were also derived.

This work deals with a specific—and somehow underestimated—point: the consequences of a generalized holonomy correction. The outstanding issue of quantization ambiguities in LQG was mentioned in [16]. New arguments were recently given in [17]. In particular, the question was addressed from the interesting point of view of renormalization. The quantization ambiguity of the connection-based holonomy variable might influence the associated cosmological predictions. This has been studied in [18]. The main effects are quite weak on the background dynamics and do not change substantially the usual conclusions of loop quantum cosmology (LQC). Interestingly, most new effects tend to decrease the number of e-folds. This makes the situation more phenomenologically promising. Perturbations were also considered in this work. However, the usual Mukhanov-Sasaki equations for gauge-invariant perturbations were used, which are not fully consistent with the underlying deformed algebra. The effects of the holonomy modifications were accounted for at the level of the background and at the level of the potential, but not in the core of the propagation equation. This article fills this gap and shows the calculation of fully reliable primordial spectra (in the deformed algebra approach). The main conclusion is that the known results of LQC are robust.

In the first section, we review the basics of LQC so that this article is self-contained for nonspecialists. Then, the deformed algebra and the propagation equations for perturbations are defined. Finally, the results are shown for different parametrizations of the holonomy correction.

Throughout all the article, we use Planck units.

II. BASICS OF LOOP QUANTUM COSMOLOGY

Loop quantum cosmology is an attempt to perform a symmetry reduction of LQG, mimicking the quantization used in the full theory [19,20]. This section explains the basic ideas for the unfamiliar reader.

The canonical formulation of LQG is based on the Ashtekar connection,

$$A_a^i \coloneqq \Gamma_a^i + \gamma K_a^i, \tag{1}$$

where γ is the Barbero-Immirzi parameter and the extrinsic curvature coefficients are given by $K_a^i = K_{ab} e_j^b \eta^{ij}$ for triads defined such that $q^{ab} = e_i^a e_j^b \delta^{ij}$ at each $x \in \Sigma_t$. The spin connection Γ_a^i reads

$$\Gamma_a^i = -\frac{1}{2}\epsilon^{ij}{}_k e_j^b(\partial_{[a}e_{b]}^k + \delta^{kl}\delta_{ms}e_l^c e_a^m \partial_b e_c^s), \tag{2}$$

e_i^a being the inverse triads such that $e_i^a e_b^j = \delta_i^j$. In order to complete the set of canonical variables, one defines the densitized triads $E_i^a \coloneqq |\det e|^{-1} e_i^a$ that are conjugate to the Ashtekar connection,

$$\{A_a^i(x), E_j^b(y)\} = \kappa\gamma\delta_a^b\delta_j^i\delta^3(x-y), \tag{3}$$

with, $\kappa = 8\pi G$.

The dynamical equations appear as constraints. Namely, the Gauss constraint,

$$G[\Lambda] = (\kappa\gamma)^{-1}\int_{\Sigma_t} d^3x\Lambda^i(\partial_a E_i^a + \epsilon_{ik}^{\ell}A_a^k E_\ell^a), \tag{4}$$

the diffeomorphism constraint,

$$D[N^a] = (\kappa\gamma)^{-1}\int_{\Sigma_t} d^3x N^a F^i_{ab}E_i^b, \tag{5}$$

where $F^i_{ab} = 2\partial_{[a}A^i_{b]} + \epsilon^i_{jk}A^j_a A^k_b$ is defined as the curvature of the Ashtekar connection, and the scalar constraint,

$$C[N] = (2\kappa\gamma)^{-1}\int_{\Sigma_t} d^3x N|\det E|^{-1/2}(\epsilon_{ijk}F^i_{ab}E^a_j E^b_k - 2(1+\gamma^2)K^i_{[a}K^j_{b]}E^a_i E^b_j). \tag{6}$$

For an isotropic, homogeneous, and flat universe, the Friedmann-Lemaître-Robertson-Walker metric can be written as

$$ds^2 = -N^2(dx^0)^2 + a^2(t)dx^a dx^b\delta_{ab}. \tag{7}$$

The Ashtekar variables are rewritten as

$$\begin{aligned} A_a^i(x) &= \gamma\dot{a}(t)\delta_a^i \equiv c(t)\delta_a^i, \\ E_i^a(x) &= a^2(t)\delta_i^a \equiv p(t)\delta_i^a, \end{aligned} \tag{8}$$

where a dot denotes a derivative with respect to the cosmic time $dt = Ndx^0$. Only the scalar constraint contributes to the dynamics of this system. Taking into account the symmetries, it can be written as

$$C[N] = -\frac{3NV_0}{\kappa\gamma^2}p^{1/2}c^2, \tag{9}$$

where V_0 is a fiducial volume element.

The matter sector is assumed to be a scalar field ϕ with an arbitrary potential $V(\phi)$. The full Hamiltonian is

$$H_t[N] = NV_0\left(-\frac{3}{\kappa\gamma^2}p^{1/2}c^2 + p^{3/2}\rho\right), \tag{10}$$

which, after setting $H_t[N] = 0$, leads to the usual Friedmann equation.

The holonomy around the closed fiducial square $\square_{ij}$ can be written as

$$h_{\square_{ij}} = h_{l_i}h_{l_j}h_{l_i}^{-1}h_{l_j}^{-1}, \tag{11}$$

with

$$h_{l_i} = \exp\{|l|k\tau^i\}, \tag{12}$$

where τ^i are base matrices of the fundamental $SU(2)$ representation, which is arbitrary at this point. Hence, the holonomy-corrected curvature is

$$F^k_{ab} = -2\lim_{l\to 0}\mathrm{tr}\left[\frac{h_{\square_{ij}}-1}{l^2}\tau^k\frac{e^i_a e^j_b}{\gamma^2}\right], \tag{13}$$

which is equivalent to

$$F^k_{ab} = \lim_{l\to 0}\frac{\sin^2(|l|k)}{|l|^2}\epsilon^k_{ij}\frac{e^i_a e^j_b}{\gamma^2}. \tag{14}$$

The presence of a minimal area in the theory, given by the smallest nonzero eigenvalue of the area operator in LQG, leads to the introduction of the $\bar{\mu} = l_{\mathrm{Pl}}(4\sqrt{3}\pi\gamma)^{1/2}p^{-1/2}$ parameter which allows to introduces the so-called holonomy correction substitution,

$$c^2 \to \bar{\mu}^{-2}\sin^2(\bar{\mu}c). \tag{15}$$

Gathering everything, the holonomy-corrected Hamiltonian constraint becomes

$$H_t[N] = NV_0\left(-\frac{3}{\kappa\gamma^2\bar{\mu}^2}p^{1/2}\sin^2(\bar{\mu}c) + p^{3/2}\rho\right), \tag{16}$$

which leads to the LQC-modified Friedmann equation,

$$H^2 = \frac{\kappa}{3}\rho\left(1-\frac{\rho}{\rho_c}\right), \tag{17}$$

where $\rho_c = \sqrt{3}/(4\pi\kappa\gamma^3)$ is usually assumed to be close to the Planck density. This is the usual bounce solution.

Many arguments (see, e.g., [21,22]) were given for considering seriously LQC with arbitrary spin representations or higher-order terms. In this work, we will remain as general as possible. To this aim, we will focus on a so-called *polymerization* defined by the substitution

$$c^2 \longrightarrow g^2(c,p). \tag{18}$$

Details about the construction of suitable semiclassical states and associated Dirac observables are given in [6], where the only restriction on the (periodic) $g(c,p)$ function is the low-curvature limit, in which GR should be recovered, i.e., $g(c,p) \longrightarrow c$.

In order to set notations let us recall some known results for the polymerized background dynamics in LQC. As shown in [23], the background dynamics is described by the set of equations

$$\begin{aligned}
\dot{c} &= -\frac{3N}{2\sqrt{p}} g^2(c,p) + \frac{Nk}{\sqrt{p}} \mathcal{G}^{(1)}(c,p) - N\frac{\kappa}{2}\sqrt{p}P, \\
\dot{p} &= 2N\sqrt{p}\mathcal{G}^{(1)}(c,p), \\
\dot{\phi} &= N\pi p^{-3/2}, \\
\dot{\pi} &= -Np^{3/2}\partial_\phi V(\phi),
\end{aligned} \tag{19}$$

where, as defined above, dots correspond to derivatives with respect to the cosmic time t, $\{\phi,\pi\}$ are the canonical variables for a given minimally coupled scalar field with potential $V(\phi)$ and pressure P. We also used the notation $\mathcal{G}^{(1)}(c,p) \coloneqq \partial_c g^2(c,p)/2$.

The background Hamiltonian constraint can be rewritten as

$$3\sqrt{p}g^2(c,p) = \kappa\rho, \tag{20}$$

where $\rho = \frac{\pi^2}{2p^3} + V(\phi)$. We make the usual gauge choice $N = 1$, which allows us to rewrite the above set of equations as a generalized Friedmann equation, together with the usual Klein-Gordon equation for the inflation field:

$$\begin{aligned}
H^2 &= \frac{\kappa}{3}\rho(\partial_c g(c,p))^2, \\
\ddot{\phi} &+ 3H\dot{\phi} + \partial_\phi V(\phi) = 0,
\end{aligned} \tag{21}$$

where $H \coloneqq \frac{1}{2}\dot{p}p^{-1}$ is the Hubble parameter.

For the background dynamics given above, four initial conditions are needed: the scale factor a, the Hubble parameter H, the scalar field ϕ, and its time derivative $\partial_t\phi$, have to be determined at some specific time. The Ashtekar school has usually advocated the (reasonable) idea that the bounce time should be chosen, whereas the Grenoble school prefers the prebounce classical universe. The dynamics at the bounce being dominated by quantum effects, we adopt this second choice (which is anyway meaningful if the bounce state is to be understood as the result of causal evolution from the contracting branch). In addition, in the prebounce phase, one can define a clear measure for probabilities [14,24,25] relying on "safe" equations.

We impose $a(t_i) = 1$. Far in the prebounce phase, the universe is mostly classical and the Hubble parameter is then given by the usual Friedmann equation, i.e., $H(t_i) = -\sqrt{\kappa\rho(t_i)/3}$. To discuss initial conditions for the matter sector, we introduce two parameters x and y defined by

$$\begin{aligned}
x &= (V(\phi)/\rho_c)^{1/2}, \\
y &= (\dot{\phi}^2/2\rho_c)^{1/2},
\end{aligned} \tag{22}$$

satisfying the relation

$$x^2 + y^2 = \frac{\rho}{\rho_c}. \tag{23}$$

In this study, we assume a quadratic potential (that is a simple mass term) for the field. Even though this potential is not favored by observational data [26], it allows easy comparisons with other studies. Our results do not, in any case, significantly depend on the shape of the potential. We have explicitly checked this with the Starobinsky potential. In the remote contracting universe, the dynamics of x and y is described by a harmonic oscillator,

$$\begin{aligned}
x(t) &\approx \left(\frac{\rho(t_i)}{\rho_c}\right)^{1/2} \sin(mt+\delta), \\
y(t) &\approx \left(\frac{\rho(t_i)}{\rho_c}\right)^{1/2} \cos(mt+\delta),
\end{aligned} \tag{24}$$

where the phase parameter δ has been studied in [27] and is not of particular importance in this work. The initial density is

$$\rho(t_i = 0) \approx \rho_c \left(\frac{\Gamma}{\alpha}\right)^2 [1 - (4\alpha)^{-1}\sin(2\delta)], \tag{25}$$

with Γ the ratio of the classical timescale over the quantum one and α a free parameter (set, as usually, to $\alpha = 17\pi + 1$ to ensure that the scalar field oscillates enough during the contracting phase for our approximations to be valid).

III. DEFORMED ALGEBRA AND PERTURBATION EQUATIONS

A. The deformed algebra approach

The treatment of perturbations is less consensual than the one of the background. On the one hand, the so-called *dressed metric* approach (see [28,29] for an introduction) was developed to account for quantum effects as deeply as possible. It is basically equivalent to the hybrid quantization one [30] and the propagation equation is the usual one. On the other hand, the *deformed algebra* framework (see [31] for an introduction) was suggested to put emphasis on covariance. It is the main focus of this study as it constitutes the natural path to investigate the specific effects of the generalized holonomy correction in a self-contained way.

Basically, the deformed algebra approach is a conservative one which relies on consistency. In the canonical formulation of GR, the smeared constraints form a first-class algebra. This closure property—that is, the fact that each Poisson bracket between constraints is proportional to another constraint—ensures that the evolution vectors always remain tangent to the submanifold of constraints. In other words, this makes the constraints compatible with themselves. When holonomy corrections are implemented, the resulting quantum gravity effective constraints do, however, not close anymore for perturbations (the closure is automatically ensured for the background). In the seminal work [32], an elegant and consistent way out was found. The interested reader can find details, e.g., in [33,34]. Important consequences were derived on the allowed shapes of the correction in [35]. To cancel the so-called anomalies, that is $\mathcal{A}_{IJ}$ terms appearing in Poisson brackets between (smeared) corrected constraints C_I^Q,

$$\{C_I^Q, C_J^Q\} = f^K{}_{IJ}(A_b^j, E_i^a)C_K^Q + \mathcal{A}_{IJ}, \tag{26}$$

one adds counterterms physically encoding the deformation of the algebra. A pictorial representation is given in [33]. Those terms are required to vanish in the classical limit and are uniquely determined by the full system of equations (including matter). Quite amazingly, similar conclusions were reached in [23], with a more general holonomy substitution. The fact that an anomaly-free algebra can still be constructed, always requiring the $\bar{\mu}$ scheme, is a strong hint in favor of the consistency of this path.

The idea, when considering linear perturbations, is to perturb constraints (and so the Hamiltonian) up to the quadratic order and to add counterterms (vanishing in the classical limit) to prevent anomalies. The Poisson brackets between all constraints are explicitly calculated. The calculations are quite involved but the final result is surprisingly elegant and simple:

$$\{G[\Lambda], G[\Lambda']\} = 0, \tag{27}$$

$$\{D_{\rm tot}[N^a], G[\Lambda]\} = 0, \tag{28}$$

$$\{H_{\rm tot}[N], G[\Lambda]\} = 0, \tag{29}$$

$$\{D_{\rm tot}[N_1^a], D_{\rm tot}[N_2^a]\} = 0, \tag{30}$$

and

$$\{H_{\rm tot}[N], D_{\rm tot}[N^a]\} = -H_{\rm tot}[\delta N^a \partial_a \delta N], \tag{31}$$

together with

$$\{H_{\rm tot}[N_1], H_{\rm tot}[N_2]\} = \left(\frac{1}{2}\frac{\partial^2 g^2(c,p)}{\partial c^2}\right) \tag{32}$$

$$\times D_{\rm tot}\left[\frac{N}{p}\partial^a(\delta N_2 - \delta N_1)\right]. \tag{33}$$

The factor $\frac{1}{2}\frac{\partial^2 g^2(c,p)}{\partial c^2}$ tends to 1 in the classical limit. When this factor becomes negative, the signature of spacetime changes to Euclidean, in agreement with what happens in the $\bar{\mu}$ scheme near the bounce. This has far-reaching consequences, from a specific phenomenology [9,11,12,36–46] to unforeseen links with the Hartle-Hawking proposal [47,48].

A contradiction with data was noticed in [49] due to the power increase in the UV part of the spectrum, associated with the Euclidean phase. It is very important to underline—as this point is often misunderstood—that this result does *not* mean, in any way, that the deformed algebra approach to LQC is discarded. Just the other way around, it shows that this framework is suited at making potentially testable predictions. It could very well be that the deformed algebra captures the main features of loop gravity and that LQC in itself is falsified. It could also be that the observational window does not fall in the altered part of the comoving spectrum (if inflation is brief). It could finally be that the way perturbations are propagated in the "timeless phase" is incorrect, which has nothing to do with the deformed algebra framework itself [50]. It might even be that initial conditions are not properly set [41].

The main point that has to be underlined at this stage is that as long as the function $g^2(c,p)$ changes concavity, a signature change in unavoidable in this (conservative) approach, as mentioned in [23]. As $g^2(c,p) \sim c^2$ near the origin (to recover GR) and as the function is periodic, the change of concavity automatically happens. This is a strong conclusion. However, contrarily to what is written in [23], this does not necessarily happen near the maximum of the function. Otherwise stated, for generalized holonomy corrections, the signature change is unavoidable but the energy density at which it takes place does not need to be close to the one of the bounce.

B. Perturbation equations

Quite a few results were derived both for the background and the perturbations in [18]. However, the equation for perturbations was not fully consistent. This is what we correct here.

The perturbed Einstein equations for a flat Universe filled with a scalar field in the polymerization framework has already been derived in the deformed algebra approach [23]. The Hamilton equation of motion for background variables is written thanks to the elementary Poisson brackets, as previously explained. Following the standard procedure, the equations of motions for the perturbed variables are decomposed in scalar, vector, and tensor modes. The physical part is then extracted by

considering terms invariant under both Gauss and diffeomorphism transformations. This results in

$$v''_{S/T} - \mathcal{G}^{(2)}(c,p)\nabla^2 v_{S/T} - \frac{z''_{S/T}}{z_{S/T}} v_{S/T} = 0, \tag{34}$$

where the prime denotes a derivation with respect to the conformal time $d\eta = p^{-1/2}dt$ and $v_{S/T}$ is the Mukhanov variable for, respectively, scalar and tensor modes. We have also defined $\mathcal{G}^{(2)}(c,p) \coloneqq \partial_c \mathcal{G}^{(1)}(c,p)$. This is in agreement with what was previously found for the usual correction [35].

By performing a Fourier decomposition on the k modes and introducing the variables $h_k = v_k/z$ and $\tilde{g}_k = \sqrt{p}\dot{h}_k/\mathcal{G}^{(2)}(c,p)$, it is possible to rewrite Eq. (34) as a set of first-order coupled differential equations. For tensor modes, one gets

$$\begin{aligned} \dot{h}_k &= \frac{\mathcal{G}^{(2)}(c,p)}{\sqrt{p}}\tilde{g}_k, \\ \dot{\tilde{g}}_k &= -2H\tilde{g}_k - ak^2 h_k, \end{aligned} \tag{35}$$

whereas, for scalar modes, the equations are

$$\begin{aligned} \dot{h}_k &= \frac{1}{\sqrt{p}}\tilde{g}_k, \\ \dot{\tilde{g}}_k &= -2H\tilde{g}_k - a\mathcal{K}(k,t,c,p)h_k, \end{aligned} \tag{36}$$

with $\mathcal{K}(k,t,c,p) = \mathcal{G}^{(2)}(c,p)k^2a^{-2} - H\dot{z}_S z_S^{-1} - \ddot{z}_S z_S^{-1}$.

C. Initial conditions for perturbations

Following the logic of causality (and remaining consistent with the background evolution), the initial conditions for perturbations are set in the prebounce contracting branch. The perturbations are thereafter propagated through the bounce and the Euclidean phase until they exit the horizon during the inflationary stage. This approach is different from the one depicted in [23] in which the authors set initial conditions for the perturbations at the onset of inflation. In this latter case, by construction, the perturbations never feel the high-energy quantum regime and the Euclidean phase. This is why our results are deeply different.

The usual canonical quantization procedure is applied for each mode v_k. In the Heisenberg picture,

$$\hat{v}_k(\eta) = v_k(\eta)\hat{a}_k + v_k^*(\eta)\hat{a}^\dagger_{-k}, \tag{37}$$

where $\hat{a}_k$, $\hat{a}_k^\dagger$ are, respectively, the annihilation and creation operators, satisfying the usual commutation relation. This leads to the so-called Wronskian condition:

$$v_k\frac{dv_k^*}{d\eta} - v_k^*\frac{dv_k}{d\eta} = i, \tag{38}$$

implying restrictions for the mode coefficients. In particular, the Minkowski vacuum can be rewritten (in the case of tensors modes) as

$$\begin{aligned} h_k(t_i) &= (2k)^{-1/2}a^{-1}(t_i), \\ \tilde{g}_k(t_i) &= -i(k/2)^{-1/2}a^{-1}(t_i) - (2k)^{-1/2}H(t_i). \end{aligned} \tag{39}$$

Initial conditions for scalar modes are way harder to derive, in particular due to the shape of the potential in the contracting branch of (all) bouncing models [45]. One can however rely on an appropriate WKB approximation [42]. In this approach, we constrain the mode coefficients from the Wronskian equation and choose the coefficients to describe a wave propagation in the positive time direction. One is then able to derive the initial conditions of h_k and g_k for scalar modes.

IV. NUMERICAL RESULTS AND DISCUSSIONS

For Gaussian perturbations, the full statistical information is given by the 2-point correlation function. In a very standard manner, the scalar and tensor primordial power spectra are expressed as functions of the Mukhanov variable and the associated potentials evaluated at the horizon crossing:

$$P_S(k) = \frac{k^3}{2\pi^2}\left|\frac{v_k}{z_S}\right|^2\Bigg|_{k=aH} \tag{40}$$

and

$$P_T(k) = \frac{4k^3}{\pi^2}\left|\frac{v_k}{z_T}\right|^2\Bigg|_{k=aH}. \tag{41}$$

In order to study the polymerization effects on those primordial power spectra, one needs to choose explicit expressions for the $g(k,p)$ function. There are not many constraints on the shape of g: mainly the low-energy limit and periodicity. Still, following [23], solving the anomalies in the algebra of constraints imposes g to be of the following form:

$$g(c,p) = p^{1/2}\varphi(cp^{-1/2}), \tag{42}$$

$\varphi(cp^{-1/2})$ being an arbitrary function of $(cp^{-1/2})$. For the rest of this work we rewrite

$$g(c,p) \coloneqq \bar{\mu}^{-1}f(x), \tag{43}$$

where $\bar{\mu}$ is the parameter already introduced in Eq. (15). With this notation, one can easily retrieve the usual LQC prescription with $f(x) = \sin(x)$.

Several functions have been considered in [18], for example

$$f_{sqr}(x) = \sin(x)\sqrt{1 + A_1 x^{n_1}(x-\pi)^{n_2}}, \tag{44}$$

with $n_i \geq 1$, $i \in \{1,2\}$ and $A_1 \geq 0$, together with

$$f_{\cos}(x) = \sin(x)\sqrt{(1+C_1)^{-1}\sum_{n=0}^{C_1}\cos^{2n}(x)}, \quad (45)$$

with $C_1 \geq 1$. One can easily show that such parametrizations have the correct behavior in the low-energy limit.

To specifically study the effects of the change of signature, we introduce a new function $f_\zeta(x)$ which has GR as a limit and allows to parametrically control the $G^{(2)} \leq 0$ region. It reads

$$f_\zeta(x) = \sqrt{\frac{\zeta^2+\pi^2}{\zeta^2+4(x-\pi/2)^2}}\sin(x), \quad (46)$$

where ζ is the free parameter associated with the signature change. Figure 1 shows how the parameter changes the shape of the function.

A. Tensor primordial power spectrum

For illustrative purposes, the numerical computation of the time evolution of the tensor mode amplitude squared $|v_k|^2$ is displayed in Fig. 2 for the f_ζ polymerization choice with $\zeta = 1$. As for other plots, Planck units are used. The contraction phase can easily be seen, together with the bounce, close to $t = 2.275 \times 10^7 t_P$, and the inflationary phase on the right side of the plot.

Tensor primordial power spectra for the polymerization choices defined in the previous section are presented in Figs. 3–5. Those spectra are evaluated at the end of the slow-roll inflationary phase, when slow-roll hypotheses break down. We ensure that all modes of interest, i.e., those represented on the spectra, are outside the horizon at that

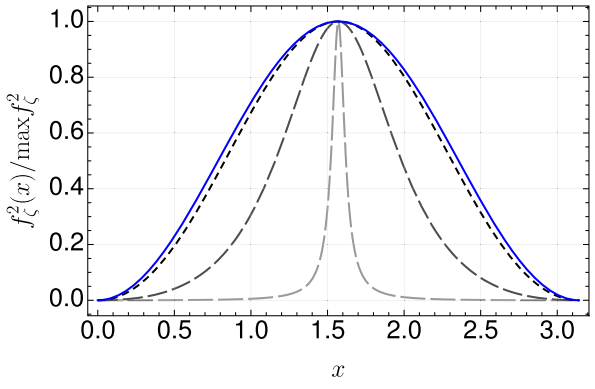


FIG. 1. Graphical representation of the f_ζ polymerization choice for various values of ζ ($\zeta = 5$ in black with short dashes, $\zeta = 1$ in gray with intermediate dashes, and $\zeta = 0.1$ in light gray with big dashes). The usual $\sin^2(x)$ prescription is represented in solid blue line.

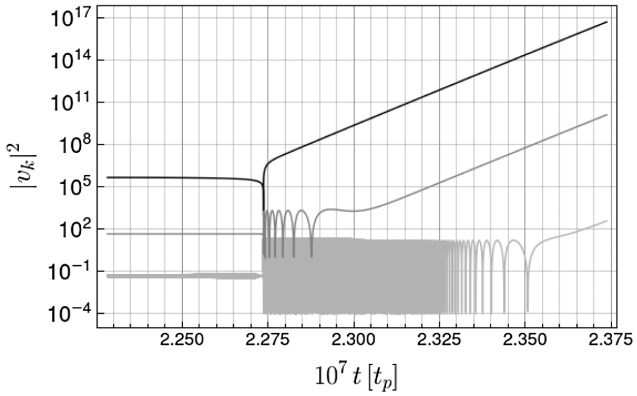


FIG. 2. Time dependence of the tensor mode amplitudes $|v_k|^2$, in the f_ζ polymerization framework, with $\zeta = 1$, for the comobile wave numbers $k = 10^{-6}$, $k = 10^{-2}$, and $k = 10$ from top to bottom.

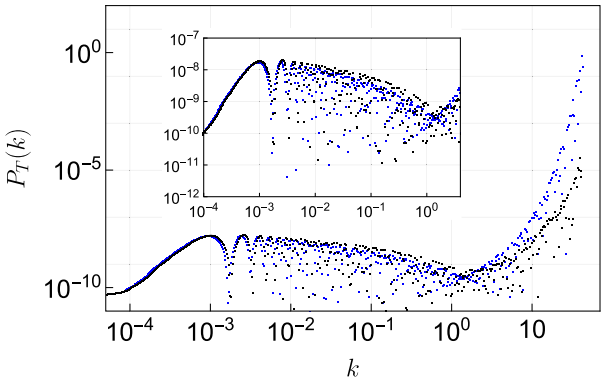


FIG. 3. Tensor primordial power spectrum for the f_{sqr} polymerization (black dots) with $A_1 = 1$ and $n_i = 1$ ($i \in \{1,2\}$), together with the usual $\sin^2(x)$ prescription (smaller blue dots). A zoom on the oscillatory regime is also represented.

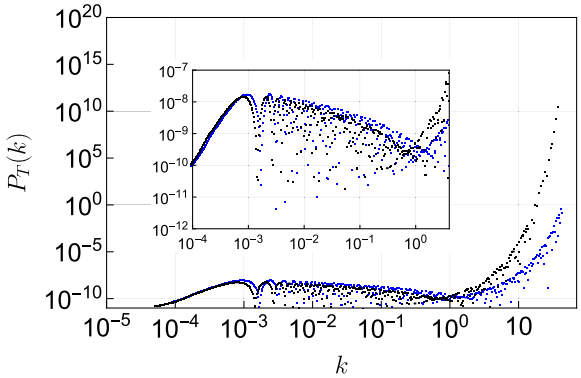


FIG. 4. Tensor primordial power spectrum for the $f_{\cos}$ polymerization (black dots) with $C_1 = 2$, together with the usual $\sin^2(x)$ prescription (smaller blue dots). A zoom on the oscillatory regime is also represented.

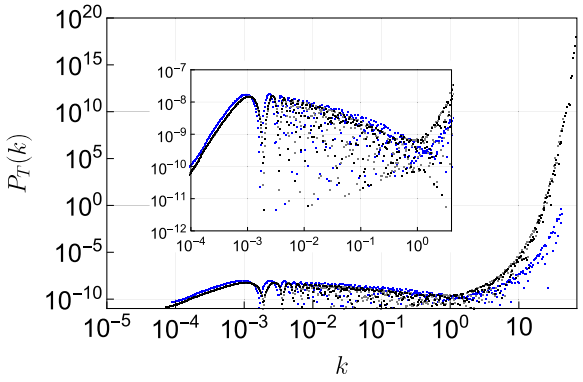


FIG. 5. Tensor primordial power spectrum for the f_ζ polymerization with $\zeta = 1$ (black dots) and $\zeta = 0.1$ (gray dots), together with the usual $\sin^2(x)$ prescription (blue dots). A zoom on the oscillatory regime is also represented.

moment. Whatever the polymerization choice, the spectra exhibit three regimes:

(1) A scale-invariant behavior in the infrared limit $(k \in]-\infty; 10^{-4}])$;
(2) An oscillatory behavior (associated with the bounce) for $(k \in [10^{-4}; 2])$; and
(3) An exponential divergence in the ultraviolet limit $(k \in [2; +\infty[)$.

The infrared (IR) regime corresponds to the largest wavelengths. Those modes do exit the horizon during the contracting phase (before slow-roll inflation takes places) and remain frozen during the bounce and the subsequent phase of inflation. However, during contraction, the comoving Hubble radius (hence the tensor potential) behaves similarly as it does during the slow-roll phase $aH \sim 2/\eta$ and $z_T''/z_T \sim 2/\eta^2$. Moreover, the tensor potential converges towards 0 when going backward in time in the contracting branch and those modes can initially be normalized using the usual Bunch-Davies vacuum. This situation therefore corresponds to the classical solution of inflation and the associated tensor power spectrum is scale invariant. Those perturbations are not impacted by the presence of the bounce. As will be discussed further in this paper, the situation is more complicated for scalar perturbations as the scalar potential does not vanish when going backward in time in the contracting branch, leading to a scale-dependent behavior of the scalar spectra in the IR. This issue has been investigated in [45], but is however not of high importance as, for the vast majority of the parameter space, those modes cannot be observed in the cosmological microwave background (CMB). A more exhaustive interpretation of those results has been widely studied in previous articles [12,39,43,44,49]. It basically means that depending on the number of inflationary e-folds, the model is either indiscernible from GR[1] (brief inflation), marginally compatible with GR, or fully different from GR (long inflation). We insist once more that the UV increase is not in itself inconsistent as the power spectrum does anyway *not* describe the real world in the $k \to \infty$ limit. Nonlinear local effects rule in this regime. In addition, it should be pointed out that modes are propagated in the Euclidean regime using their Fourier expansion, which remains conceptually unclear.

The results displayed in the previously mentioned figures are not difficult to interpret (within the assumptions of the model). For tensor modes, the potential z_T''/z_T depends only on the scale factor a and its derivatives. In other words, the potential depends only upon background variables. Even with quite exotic generalized holonomy corrections, those variables are mostly equivalent to the usual loop quantum cosmology ones (see [18]). Nevertheless, some deviations from the standard behavior can be observed in the ultraviolet. This is due to the maximum value of f: if it differs from unity, the bounce-energy density is not exactly the same than the one of the usual $\sin^2(x)$ bounce. This can be explicitly seen in Fig. 6 for the f_ζ polymerization choice.

B. Scalar primordial power spectrum

The numerical results for the scalar primordial power spectra are given in Figs. 7–9 for the same polymerization choices. Three regimes can still be identified:

(1) A power law $(\propto k^3)$ in the infrared $(k \in]-\infty; 10^{-3}])$;
(2) Oscillations for $(k \in [10^{-3}; 2])$; and
(3) A divergence in the ultraviolet $(k \in [2; +\infty[)$.

Once again, the meanings of the main features have already been studied (see, in particular, [42]).

Starting from the definition of $z_S := a^2\dot{a}^{-1}\dot{\phi}$, one obtains

$$\frac{z_S''}{z_S} = -a^2(\partial_\phi^2 V) + 2\dot{a}^2 - 2\kappa f''\dot{\phi}(\partial_\phi V)a^4\dot{a}^{-2} - \frac{7}{2}a^2\kappa f''\dot{\phi}^2 + 3a^2\kappa\dot{\phi}^4 + \frac{1}{2}a^4\dot{a}^{-2}\kappa^2 f''^2\dot{\phi}^4. \quad (47)$$

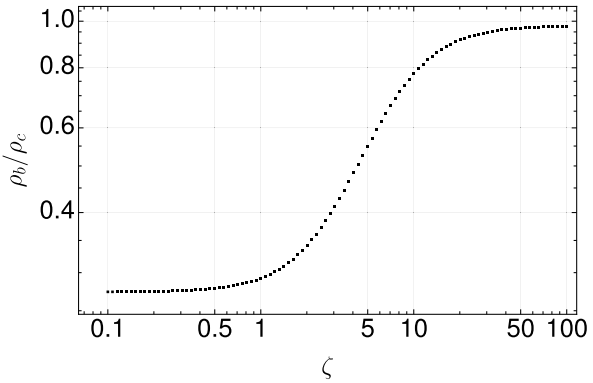


FIG. 6. ζ dependence of the bounce-energy density for a background described by the f_ζ polymerization choice.

[1]We do not consider here the subtle normalization effects associated with the preceding deflation.

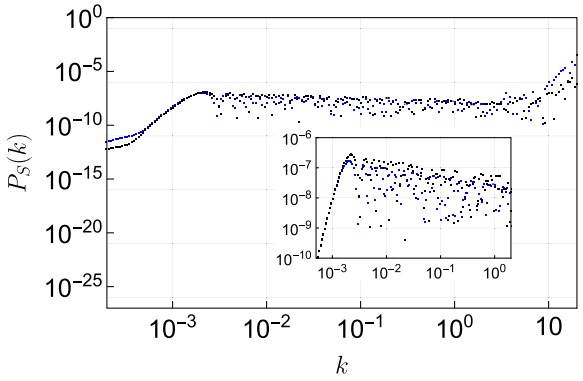


FIG. 7. Scalar primordial power spectrum for the f_{sqr} polymerization (black dots) with $A_1 = 1$ and $n_i = 1$ $(i \in \{1, 2\})$, together with the usual $\sin^2(x)$ prescription (blue dots). A zoom on the oscillatory regime is also represented.

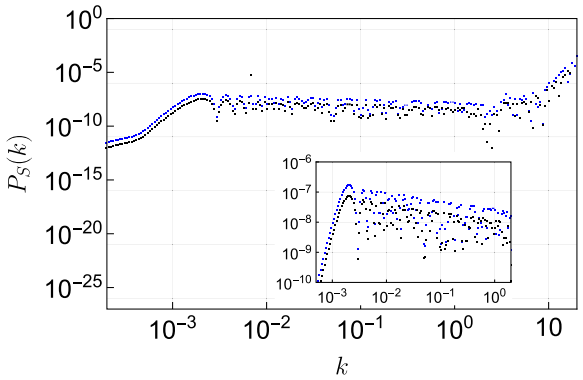


FIG. 8. Scalar primordial power spectrum for the $f_{\cos}$ polymerization (black dots) with $C_1 = 2$, together with the usual $\sin^2(x)$ prescription (smaller blue dots). A zoom on the oscillatory regime is also represented.

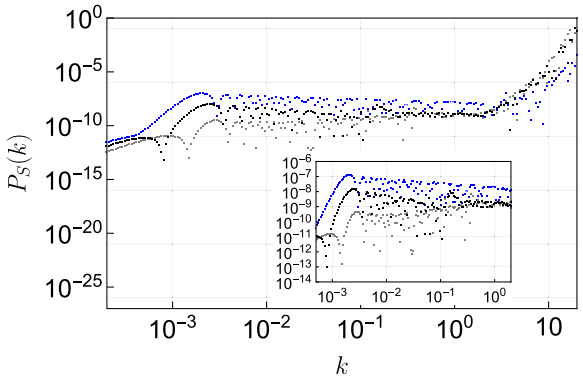


FIG. 9. Scalar primordial power spectrum for the f_ζ polymerization with $\zeta = 1$ (black dots) and $\zeta = 0.1$ (gray dots), together with the usual $\sin^2(x)$ prescription (blue dots). A zoom on the oscillatory regime is also represented.

This complicated expression is what makes the case of scalar perturbations specific. In addition to the previously explained issues with initial conditions (that are not related with this specific model but inherent to all bouncing models), the fact that the polymerization choice also appears in the z_S''/z_S term of the propagation equation is what makes the study of scalar modes subtle. This is however the most interesting part of the game as they are directly related with CMB measurements.

To allow a direct comparison with data, one needs to convert comobile values into physical ones. In this work, as usual and useful when studying bouncing models, we normalized the scale factor to unity at the bounce time. The conversion therefore requires to know the number of e-folds between the bounce and the decoupling. In particular, it requires the knowledge of the number of inflationary e-folds N_{inf}, which cannot be fully fixed by the model but depends on contingent parameters (such as the phase of the scalar field during the contraction phase). Extensive discussions can be found in [49] and [51,52]. In practice, the physical wave number k_{phys} is related to the comobile wave number k used in the different plots of this article by

$$k_{\mathrm{phys}} = k\left(e^{N_{\mathrm{inf}}}\frac{T_{\mathrm{RH}}}{T_{\mathrm{dec}}}\right)^{-1}, \tag{48}$$

where T_{RH} and T_{dec} are, respectively, the reheating and decoupling temperatures.

The main conclusion that can be drawn from all the plots is that the spectra remain remarkably close one to the other, and similar to the "standard" deformed algebra $\sin^2(x)$ one. In the scalar case, this was not an *a priori* expected result. This shows that the precise shape of the holonomy correction has a very weak influence on the details of the observables, even if initial conditions are set in the contracting branch, and the perturbations propagated through the bounce and the Euclidean phase. This is an important point for the reliability of the model.

This work focuses on the *shape* of the power spectrum as this is precisely where effects of generalized holonomy corrections are expected to play a significant role. General considerations on the amplitude of the spectrum and on the scalar-tensor ratio in this framework can be found in [43,49].

V. FALSIFIABILITY

In principle, it could be that measurements allow to constrain the inflaton potential. In this case, the duration of inflation would somehow be predicted by the model (see, in [18], the extension to generalized holonomy correction of the results from [25]). Should the latter be high enough so that the observational window falls in the UV part of the spectrum, the model would be discarded. What conclusions could then be drawn? Obviously, the situation is intricate as

quite a few explicit and implicit hypotheses are always at play in a cosmological scenario. At the heuristic level, this would discard a specific approach to generalizing gravity through a periodization of the Ashtekar connection (with the suitable low-curvature limit). More deeply, the delicate and important point would be to clarify the link between this specific approach and LQG in general.

On the one hand, it is true that the deformed algebra framework does not retain much of the complicated structure of loop quantum gravity. It particular, it is obviously "less quantum" than the dressed metric approach.

On the other hand, it could be argued that it actually captures the core ingredients. Gauge fixing before quantization is often harmless. However, the constraints considered here are not of the kind of those encountered in Yang-Mills theories. When quantum corrected, the gauge transformation they generate are not of the usual form. Therefore, gauge fixing before quantization might lead to choose the gauge according to transformations that need to be modified; hence, the inconsistency.

In addition, in the case of gravity, the dynamics is part of the gauge system [53,54]. Consistency therefore imposes to quantize gauge transformations and the dynamics simultaneously. It is not correct to fix the gauge in order to derive the dynamics. The deformed algebra approach solves both of those issues [31] and should therefore be taken seriously.

The relation between this approach and the full theory is still unclear. If it was, in the future, shown to be reliably related to LQG, possible conflicts with data would rule out the main theory. On the other hand, if this framework was demonstrated to miss key features of the full theory, it would discard only this specific way to describe the cosmological dynamics.

It is fair to underline that this remains an open question at this stage. It should however be stressed that the unforeseen link between the deformed algebra approach and the disappearance of time, as predicted by the Hartle-Hawking proposal [55], is quite remarkable. Even more impressive is the way it might cure the weaknesses of the original proposal [47,48].

Finally, as previously reminded, it could also be that modes are not correctly propagated in the Euclidean phase. In this work, we make minimalist assumptions and work in Fourier space to avoid obvious problems with the definition of a wave in a timeless space. Another interesting view was suggested in [50].

VI. CONCLUSION

In this work we have considered generalized holonomy corrections, as the usual harmonic choice made in loop quantum cosmology is far from being the only possible one. It has even been recently argued that there is no fundamental reason for focusing on this specific shape [17].

We have studied three different generic functions having general relativity as their low-energy limit and satisfying the basic loop gravity requirements. One of them is specifically parametrized so that the position of the Euclidean region, corresponding to a change of concavity, can be easily varied and probed.

The generalized holonomy correction appears both at the background level and in the propagation equation for perturbations. In addition, for scalar modes, it also enters the z_S''/z_S term. This leads to an intricate situation which cannot be fully understood intuitively.

To clarify the situation, we have numerically calculated the primordial power spectra in all cases, setting initial conditions in the prebounce contracting branch. Since in this setting (motivated by general arguments), cosmological perturbations are propagated through the bounce and the Euclidean phase, a bigger sensitivity of the spectra to the shape of the holonomy correction than the one established in [23] could have been expected. However, we have shown that whatever the (reasonable) form of the function and values of the parameters, the overall shape of the spectra remains unchanged with respect to the usual deformed algebra LQC results. This shows that the known conclusions are robust.

Obviously, the actual content of the Universe in the contraction phase is not known and this constitutes a weakness for all bouncing models. As pointed out in [56,57] this might raise some interesting paradoxes. In this work, the only assumption required is that a scalar field dominates over all the other possible contents at high energy before the bounce. Although speculative, this assumption makes sense as it both leads to the desired phase on inflation and seems favored by grand unified models of particle physics [58]. Obviously, a detailed description on an "inverse-reheating" process is still missing. More important than the actual content is the question of anisotropies, extensively discussed, e.g., in [14].

In the future, it would be interesting to generalize this investigation to the dressed metric approach. In this case, the way the new holonomy correction might alter the propagation equation is, however, less clear and requires further investigations.

[1] A. Ashtekar and E. Bianchi, Rep. Prog. Phys. **84**, 042001 (2021).
[2] C. Rovelli, *Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2004).
[3] C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2014), ISBN 1107069629, 9781107069626, 9781316147290.
[4] C. Kiefer, in *Time and Matter 2007* (University of Nova Gorica Press, Oxford, 2007).
[5] C. Rovelli and F. Vidotto, arXiv:2211.06718.
[6] A. Ashtekar, A. Corichi, and P. Singh, Phys. Rev. D **77**, 024046 (2008).
[7] A. Ashtekar and P. Singh, Classical Quantum Gravity **28**, 213001 (2011).
[8] I. Agullo, A. Ashtekar, and W. Nelson, Phys. Rev. Lett. **109**, 251301 (2012).
[9] A. Barrau, T. Cailleteau, J. Grain, and J. Mielczarek, Classical Quantum Gravity **31**, 053001 (2014).
[10] P. Diener, B. Gupt, and P. Singh, Classical Quantum Gravity **31**, 105015 (2014).
[11] A. Ashtekar and A. Barrau, Classical Quantum Gravity **32**, 234001 (2015).
[12] B. Bolliet, J. Grain, C. Stahl, L. Linsefors, and A. Barrau, Phys. Rev. D **91**, 084035 (2015).
[13] E. Alesci and F. Cianfrani, Phys. Rev. D **92**, 084065 (2015).
[14] K. Martineau, A. Barrau, and S. Schander, Phys. Rev. D **95**, 083507 (2017).
[15] S. Gielen and D. Oriti, Phys. Rev. D **98**, 106019 (2018).
[16] A. Perez, Phys. Rev. D **73**, 044007 (2006).
[17] L. Amadei, A. Perez, and S. Ribisi, Phys. Rev. D **107**, 086007 (2023).
[18] C. Renevey, K. Martineau, and A. Barrau, Phys. Rev. D **105**, 063521 (2022).
[19] A. Ashtekar and P. Singh, Classical Quantum Gravity **28**, 213001 (2011).
[20] M. Bojowald, Living Rev. Relativity **8**, 11 (2005).
[21] J. Ben Achour, S. Brahma, and M. Geiller, Phys. Rev. D **95**, 086015 (2017).
[22] J. Yang, Y. Ding, and Y. Ma, arXiv:0902.1913.
[23] Y. Han and M. Liu, Classical Quantum Gravity **35**, 105017 (2018).
[24] L. Linsefors and A. Barrau, Phys. Rev. D **87**, 123509 (2013).
[25] L. Linsefors and A. Barrau, Classical Quantum Gravity **32**, 035010 (2015).
[26] Y. Akrami *et al.* (Planck Collaboration), Astron. Astrophys. **641**, A10 (2020).
[27] L. Linsefors and A. Barrau, Phys. Rev. D **87**, 123509 (2013).
[28] I. Agullo, A. Ashtekar, and W. Nelson, Phys. Rev. D **87**, 043507 (2013).
[29] I. Agullo, A. Ashtekar, and W. Nelson, Classical Quantum Gravity **30**, 085014 (2013).
[30] B.-F. Li and P. Singh, Phys. Rev. D **106**, 086015 (2022).
[31] A. Barrau, M. Bojowald, G. Calcagni, J. Grain, and M. Kagan, J. Cosmol. Astropart. Phys. 05 (2015) 051.
[32] M. Bojowald, G. M. Hossain, M. Kagan, and S. Shankaranarayanan, Phys. Rev. D **78**, 063547 (2008).
[33] J. Mielczarek, T. Cailleteau, A. Barrau, and J. Grain, Classical Quantum Gravity **29**, 085009 (2012).
[34] T. Cailleteau, J. Mielczarek, A. Barrau, and J. Grain, Classical Quantum Gravity **29**, 095010 (2012).
[35] T. Cailleteau, A. Barrau, J. Grain, and F. Vidotto, Phys. Rev. D **86**, 087301 (2012).
[36] J. Mielczarek, Springer Proc. Phys. **157**, 555 (2014).
[37] J. Mielczarek, AIP Conf. Proc. **1514**, 81 (2013).
[38] J. Mielczarek, Europhys. Lett. **108**, 40003 (2014).
[39] L. Linsefors, T. Cailleteau, A. Barrau, and J. Grain, Phys. Rev. D **87**, 107503 (2013).
[40] A. Barrau and J. Grain, *Loop Quantum Gravity and Observations* (World Scientific, London, 2017), pp. 281–306.
[41] J. Mielczarek, L. Linsefors, and A. Barrau, Int. J. Mod. Phys. D **27**, 05 (2018).
[42] S. Schander, A. Barrau, B. Bolliet, L. Linsefors, J. Mielczarek, and J. Grain, Phys. Rev. D **93**, 023531 (2016).
[43] A. Barrau and J. Grain, arXiv:1607.07589.
[44] K. Martineau, A. Barrau, and J. Grain, Int. J. Mod. Phys. D **27**, 1850067 (2018).
[45] A. Barrau, P. Jamet, K. Martineau, and F. Moulin, Phys. Rev. D **98**, 086003 (2018).
[46] E. Wilson-Ewing, C.R. Phys. **18**, 207 (2017).
[47] M. Bojowald and S. Brahma, Phys. Rev. Lett. **121**, 201301 (2018).
[48] M. Bojowald and S. Brahma, Phys. Rev. D **102**, 106023 (2020).
[49] B. Bolliet, A. Barrau, J. Grain, and S. Schander, Phys. Rev. D **93**, 124011 (2016).
[50] M. Bojowald and J. Mielczarek, J. Cosmol. Astropart. Phys. 08 (2015) 052.
[51] A. Barrau, Eur. Phys. J. C **80**, 579 (2020).
[52] C. Renevey, A. Barrau, K. Martineau, and S. Touati, J. Cosmol. Astropart. Phys. 01 (**2021**) 018.
[53] C. Rovelli, Found. Phys. **44**, 91 (2014).
[54] C. Rovelli, Fundam. Theor. Phys. **199**, 107 (2020).
[55] J. B. Hartle and S. W. Hawking, Phys. Rev. D **28**, 2960 (1983).
[56] A. Barrau and L. Linsefors, J. Cosmol. Astropart. Phys. 12 (2014) 037.
[57] A. Barrau, K. Martineau, and F. Moulin, Phys. Rev. D **96**, 123520 (2017).
[58] S. Mohanty and A. Nautiyal, Phys. Rev. D **78**, 123515 (2008).

Conclusion

General Relativity has been a very successful theory of gravity for over a century and there is so far no sign of its decline. Nonetheless, its lack of consistency with the quantum sector together with some of the issues that have emerged from its application to cosmology and black hole physics, led fundamental physicists to find alternative approaches to the standard description of the gravitational sector. In this thesis, we focused our attention on the phenomenology of several of these issues in the context of black holes, dark energy and early universe cosmology.

We started with the study of the evolution of black holes when surrounded by an isotropic fluid. As a first attempt, we used the Schwarzschild solution outside of its domain of validity to better understand the evolution of event horizons when a black hole is bathed in a fluid at thermal equilibrium. We showed in particular that the behaviour of the Schwarzschild black hole is catastrophic in the context of cosmology, meaning that either its mass diverges in a finite time or it falls down to zero, due to Hawking radiation. We expect the divergent trajectory to be due to the inappropriate use of the Schwarzschild solution, at least from the time when the black hole horizon reaches the cosmological horizon. To go beyond this simple solution, we relaxed the vacuum assumption and assumed an isotropic metric in Painlevé-Gullstrand coordinates. We showed that two different apparent horizons emerge, a cosmological and a black hole horizons. Furthermore, we demonstrated for the most general case that a black hole horizon never shrinks, which is in contradiction with the generalized McVittie metric. We then tried to solve the Einstein field equations for the simpler case of an Eulerian perfect fluid and concluded that this cannot describe a fluid around a central mass, except in the case of a cosmological constant. In the future, the connection between the Schwarzschild approach and the isotropic solution should be made with simplifying assumptions applied to the Nielsen Visser metric. It would also be very interesting to show, in the general approach, that the cosmological and black hole horizons indeed merge in extreme conditions.

Cosmology, as derived from the homogeneous and isotropic solutions of General Relativity, has been greatly successful to describe the history of the universe. However, the description of the late time accelerated expansion with a cosmological constant suffers some conceptual issues. String theory is sometimes coined as a non-falsifiable theory, due to the very large set of low energy effective theories that can be derived from it, but recently its experimental fallibility has been put forward with the swampland program and cosmological surveys measuring the equation of state of dark energy. In this thesis, we saw that thanks to the Euclid satellite, the Vera Rubin observatory and the SKA, the de Sitter conjecture will be in strong tension with observations, if the constant nature of dark energy is confirmed. It is remarkable that string theory, which is believed to be almost unfalsifiable, might be excluded, or at least disfavored, using cosmological surveys. Nonetheless, this will not be the

case until the de Sitter conjecture has been proven to be correct and the upper value for the de Sitter condition has been found. While such future works are left to string theorists, this thesis focused on developing new tests for other dynamical dark energy theories, in particular screened Horndeski models, through gravitational waves. The effect of the screened scalar field on the waveform of the gravitational waves sourced by a binary system of compact objects were derived. Interestingly, it has been found that even a fully screened scalar field, having no direct effect on the binary system and near zone metric, would still influence the gravitational waveform through long range effects. The direct effect of the screened scalar field on the behaviour of the system has been shown to be completely negligible for the cubic Galileon theory. The long range effects, entering at the 1.5PN order in the waveform, are so far unobservable using current detectors. However, with multi-band observations, which will be made possible with LISA, together with pulsar constraints on the radiative power of gravitational waves, one could hope to observe the accumulative effect of the long range correction. Hence, further work in this direction should be made as well as the calculation of the waveform of the scalar waves, which are also detectable using interferometers.

On the other side of the cosmological time scale, the concordance model is plagued with the big bang singularity. However, together with the theory of inflation, we saw that it actually contains all the necessary ingredients to resolve the singularity with a curvature bounce. Even though the bouncing solution of a closed de Sitter space-time has been know for a while, its application to inflation has been so far neglected. It is rather fascinating that no exotic physics or quantum gravity is needed to solve the big bang singularity problem and the hypothesis of a positive spatial curvature remains highly plausible with current observations. Such potential resolution of the big bang singularity is a strong motivation to consider a non-zero curvature density parameter and settle the question of the shape of the universe that has been revived with Planck's CMB data that are in need of an anomalous lensing amplitude. Nevertheless, the spatial curvature is so far still consistent with a flat universe and we looked at the possible signature of the bounce in the CMB. We saw that potential future detection of the tensor power spectrum would not be able to highlight the curvature bounce, but a small signature might be present in the temperature anisotropies, if the fluctuations of the CMB were initiated prior to the bounce. A better understanding of the origin of the perturbations in bouncing scenarii is needed to confirm the small effects of the pre-bounce universe on the CMB, even though these effects are not statistically significant. Beyond the significance of these effects, the fact that the length of the deflationary period of the pre-bounce universe can have an effect on the expanding phase is also quite remarkable. Even though single field inflation predicts Gaussian fluctuations, so far favored by Planck's data, a bounce can be the origin of non-Gaussianities and further work should be made in this direction. Additionally, we also believe that our work has an indirect consequence on quantum cosmology theories that use the Wheeler-DeWitt approach. Such theories often require a closed universe to be consistent, but as we saw, the bounce would then happen well above the Planck regime, before the need of quantum gravity effects.

Nevertheless, a non-zero curvature is still unobserved and it is sadly likely to stay that way if inflation lasted more than 75 e-folds. Therefore, other scenarii for the origin of the universe, such as the bounce predicted by LQC, are also plausible and their phenomenology must be studied. Inherent to all quantum theories, ambiguities can affect the predictions of LQC and a consistency check should be made. This was done in the last two papers included in this thesis and the predictions of LQC has been proven to only be insignificantly

altered by different holonomy corrections. While this is a first step in the the proof of a consistent LQC, a more general approach should be pursued for that purpose. In particular, the holonomy corrections emerging from the use of different representations of $su(2)$ might lead to a pre-bounce universe that does not behave as the post-bounce universe, i.e. the correction $g(c,p)$ does not tend to c^2 for small energy density. This type of corrections were not included in our work and should be investigated. While the study of modified Friedmann equations for different holonomy corrections was conducted in the context of LQC, such modified equations could emerge from very different assumptions and therefore also be used in other models of modified or quantum gravity.

Beyond some conceptual and theoretical issues, the concordance model of cosmology together with the theory of inflation, are extremely successful to describe the observed cosmological phenomena. The lack of experimental signature of beyond standard model physics gives a hard time to fundamental theoretical physicists and thousands of different models trying to explain dark energy or the origin of the universe have be invented. With such a rise of alternative theories to General Relativity, the need for phenomenological studies has never been greater and this thesis is intended as a step in this direction. The hope to discover the signature of exotic physics in a recent future is still far from lost. The era of precision cosmology is at its beginning and already several tensions between the concordance model and observation have been identified. Some of the most famous are the Hubble tension, the S_8 tension, the Lithium problem or the lensing anomaly. Impressive future cosmological surveys are expected to bring crucial new data for astrophysics and cosmology, starting with the recently launched Euclid satellite, but also LISA, SKA, the Vera Rubin observatory or the ELT. The future of cosmology appears bright, full of excitements and hopefully full of new discoveries.

Bibliography

[1] Albert Einstein. On the electrodynamics of moving bodies. *Annalen Phys.*, 17:891–921, 1905.

[2] Albert Einstein. The Foundation of the General Theory of Relativity. *Annalen Phys.*, 49(7):769–822, 1916.

[3] David Hilbert. Die Grundlagen der Physik. 1. *Gott. Nachr.*, 27:395–407, 1915.

[4] D. Lovelock. The Einstein tensor and its generalizations. *J. Math. Phys.*, 12:498–501, 1971.

[5] Ruben Aldrovandi and José Geraldo Pereira. *Teleparallel Gravity: An Introduction.* Springer, 2013.

[6] Jose Beltrán Jiménez, Lavinia Heisenberg, and Tomi Koivisto. Coincident General Relativity. *Phys. Rev. D*, 98(4):044048, 2018.

[7] Lavinia Heisenberg. A systematic approach to generalisations of General Relativity and their cosmological implications. *Phys. Rept.*, 796:1–113, 2019.

[8] Norbert Straumann. *General Relativity.* Graduate Texts in Physics. Springer, Dordrecht, 2013.

[9] Sean M. Carroll. Lecture notes on general relativity. 12 1997.

[10] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley and Sons, New York, 1972.

[11] Steven Weinberg. *Cosmology.* 2008.

[12] A. Zee. *Einstein Gravity in a Nutshell.* Princeton University Press, New Jersey, 5 2013.

[13] Prakash Sarkar, Jaswant Yadav, Biswajit Pandey, and Somnath Bharadwaj. The scale of homogeneity of the galaxy distribution in SDSS DR6. *Mon. Not. Roy. Astron. Soc.*, 399:L128–L131, 2009.

[14] C. Marinoni, J. Bel, and A. Buzzi. The Scale of Cosmic Isotropy. *JCAP*, 10:036, 2012.

[15] Daniel Baumann. Inflation. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pages 523–686, 2011.

[16] Daniel Baumann. Primordial Cosmology. *PoS*, TASI2017:009, 2018.

[17] Daniel Baumann. *Cosmology*. Cambridge University Press, 7 2022.

[18] R. Gharechahi, M. Nouri-Zonoz, and A. Tavanfar. A tale of two velocities: Threading versus slicing. *Int. J. Geom. Meth. Mod. Phys.*, 15(03):1850047, 2017.

[19] Rudeep Gaur and Matt Visser. Cosmology in Painlevé-Gullstrand coordinates. *JCAP*, 09:030, 2022.

[20] Georges Lemaitre. The Expanding Universe. *Mon. Not. Roy. Astron. Soc.*, 91:490–501, 1931.

[21] Georges Lemaitre. A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae. *Annales Soc. Sci. Bruxelles A*, 47:49–59, 1927.

[22] J. Silk et al. *Particle Dark Matter: Observations, Models and Searches*. Cambridge Univ. Press, Cambridge, 2010.

[23] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.

[24] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[25] T. M. C. Abbott et al. Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D*, 98(4):043526, 2018.

[26] Valerio Faraoni. *Cosmological and Black Hole Apparent Horizons*, volume 907. 2015.

[27] Ivan Booth. Black hole boundaries. *Can. J. Phys.*, 83:1073–1099, 2005.

[28] Roger Penrose. Gravitational collapse and space-time singularities. *Phys. Rev. Lett.*, 14:57–59, 1965.

[29] S. W. Hawking and R. Penrose. The Singularities of gravitational collapse and cosmology. *Proc. Roy. Soc. Lond. A*, 314:529–548, 1970.

[30] Richard L. Arnowitt, Stanley Deser, and Charles W. Misner. Dynamical Structure and Definition of Energy in General Relativity. *Phys. Rev.*, 116:1322–1330, 1959.

[31] Karl Schwarzschild. On the gravitational field of a mass point according to Einstein's theory. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1916:189–196, 1916.

[32] Kazunori Akiyama et al. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *Astrophys. J. Lett.*, 875:L1, 2019.

[33] Valerio Faraoni. Evolving black hole horizons in General Relativity and alternative gravity. *Galaxies*, 1(3):114–179, 2013.

[34] Changjun Gao, Xuelei Chen, Valerio Faraoni, and You-Gen Shen. Does the mass of a black hole decrease due to the accretion of phantom energy. *Phys. Rev. D*, 78:024008, 2008.

[35] Alex B. Nielsen and Matt Visser. Production and decay of evolving horizons. *Class. Quant. Grav.*, 23:4637–4658, 2006.

[36] Valerio Faraoni, Changjun Gao, Xuelei Chen, and You-Gen Shen. What is the fate of a black hole embedded in an expanding universe? *Phys. Lett. B*, 671:7–9, 2009.

[37] Valerio Faraoni. Cosmological apparent and trapping horizons. *Phys. Rev. D*, 84:024003, 2011.

[38] Valerio Faraoni. Embedding black holes and other inhomogeneities in the universe in various theories of gravity: a short review. *Universe*, 4(10):109, 2018.

[39] G. C. McVittie. The mass-particle in an expanding universe. *Mon. Not. Roy. Astron. Soc.*, 93:325–339, 1933.

[40] Nemanja Kaloper, Matthew Kleban, and Damien Martin. McVittie's Legacy: Black Holes in an Expanding Universe. *Phys. Rev. D*, 81:104044, 2010.

[41] Sean M. Carroll. *Spacetime and Geometry.* Cambridge University Press, 7 2019.

[42] Walter C. Hernandez and Charles W. Misner. Observer Time as a Coordinate in Relativistic Spherical Hydrodynamics. *Astrophys. J.*, 143:452, 1966.

[43] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[44] E. Babichev, V. Dokuchaev, and Yu. Eroshenko. Black hole mass decreasing due to phantom energy accretion. *Phys. Rev. Lett.*, 93:021102, 2004.

[45] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proc. Nat. Acad. Sci.*, 15:168–173, 1929.

[46] R. A. Alpher, H. Bethe, and G. Gamow. The origin of chemical elements. *Phys. Rev.*, 73:803–804, 1948.

[47] Ralph A. Alpher and Robert Herman. Evolution of the Universe. *Nature*, 162(4124):774–775, 1948.

[48] G. Gamow. The Evolution of the Universe. *Nature*, 162(4122):680–682, 1948.

[49] Arno A. Penzias and Robert Woodrow Wilson. A Measurement of excess antenna temperature at 4080-Mc/s. *Astrophys. J.*, 142:419–421, 1965.

[50] Katherine Freese. Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars. *EAS Publ. Ser.*, 36:113–126, 2009.

[51] Gary Steigman. Primordial Nucleosynthesis in the Precision Cosmology Era. *Ann. Rev. Nucl. Part. Sci.*, 57:463–491, 2007.

[52] Jean-Philippe Uzan. The Big-Bang Theory: Construction, Evolution and Status. *Prog. Math. Phys.*, 76:1–72, 2021.

[53] Keith A. Olive. TASI lectures on dark matter. In *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2002): Particle Physics and Cosmology: The Quest for Physics Beyond the Standard Model(s)*, pages 797–851, 1 2003.

[54] Varun Sahni. Dark matter and dark energy. *Lect. Notes Phys.*, 653:141–180, 2004.

[55] Juan Garcia-Bellido. Cosmology and astrophysics. In *2004 European School of High-Energy Physics*, pages 267–342, 2 2005.

[56] Paolo Salucci. The distribution of dark matter in galaxies. *Astron. Astrophys. Rev.*, 27(1):2, 2019.

[57] Dragan Huterer and Daniel L Shafer. Dark energy two decades after: Observables, probes, consistency tests. *Rept. Prog. Phys.*, 81(1):016901, 2018.

[58] Antonio Padilla. Lectures on the Cosmological Constant Problem. 2 2015.

[59] D. J. Fixsen, E. S. Cheng, J. M. Gales, John C. Mather, R. A. Shafer, and E. L. Wright. The Cosmic Microwave Background spectrum from the full COBE FIRAS data set. *Astrophys. J.*, 473:576, 1996.

[60] Mordehai Milgrom. Dynamics with a non-standard inertia-acceleration relation: an alternative to dark matter. *Annals Phys.*, 229:384–415, 1994.

[61] Jacob D. Bekenstein. Relativistic gravitation theory for the MOND paradigm. *Phys. Rev. D*, 70:083509, 2004. [Erratum: Phys.Rev.D 71, 069901 (2005)].

[62] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933.

[63] Benoit Famaey and Stacy McGaugh. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Rev. Rel.*, 15:10, 2012.

[64] Maxim Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman, C. Jones, S. Murray, and W. Tucker. Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56. *Astrophys. J.*, 606:819–824, 2004.

[65] Vera C. Rubin and W. Kent Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophys. J.*, 159:379–403, 1970.

[66] Pieter van Dokkum et al. A galaxy lacking dark matter. *Nature*, 555(7698):629–632, 2018.

[67] Pavel E. Mancera Piña et al. Off the Baryonic Tully–Fisher Relation: A Population of Baryon-dominated Ultra-diffuse Galaxies. *Astrophys. J. Lett.*, 883(2):L33, 2019.

[68] D. Scolnic et al. SUPERCAL: CROSS-CALIBRATION OF MULTIPLE PHOTOMETRIC SYSTEMS TO IMPROVE COSMOLOGICAL MEASUREMENTS WITH TYPE Ia SUPERNOVAE. *Astrophys. J.*, 815(2):117, 2015.

[69] Shadab Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 470(3):2617–2652, 2017.

[70] Kevork N. Abazajian et al. The Seventh Data Release of the Sloan Digital Sky Survey. *Astrophys. J. Suppl.*, 182:543–558, 2009.

[71] Daniel J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005.

[72] Henk Hoekstra and Bhuvnesh Jain. Weak Gravitational Lensing and its Cosmological Applications. *Ann. Rev. Nucl. Part. Sci.*, 58:99–123, 2008.

[73] M. Betoule et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *Astron. Astrophys.*, 568:A22, 2014.

[74] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

[75] Steven Weinberg. The Cosmological Constant Problem. *Rev. Mod. Phys.*, 61:1–23, 1989.

[76] Lucas Lombriser. On the cosmological constant problem. *Phys. Lett. B*, 797:134804, 2019.

[77] Willis E. Lamb and Robert C. Retherford. Fine Structure of the Hydrogen Atom by a Microwave Method. *Phys. Rev.*, 72:241–243, 1947.

[78] H. B. G. Casimir. On the Attraction Between Two Perfectly Conducting Plates. *Indag. Math.*, 10:261–263, 1948.

[79] Claudia de Rham, Gregory Gabadadze, and Andrew J. Tolley. Resummation of Massive Gravity. *Phys. Rev. Lett.*, 106:231101, 2011.

[80] Claudia de Rham and Lavinia Heisenberg. Cosmology of the Galileon from Massive Gravity. *Phys. Rev. D*, 84:043503, 2011.

[81] Paolo Creminelli, Guido D'Amico, Jorge Norena, and Filippo Vernizzi. The Effective Theory of Quintessence: the w<-1 Side Unveiled. *JCAP*, 02:018, 2009.

[82] Minjoon Park, Kathryn M. Zurek, and Scott Watson. A Unified Approach to Cosmic Acceleration. *Phys. Rev. D*, 81:124008, 2010.

[83] Giulia Gubitosi, Federico Piazza, and Filippo Vernizzi. The Effective Field Theory of Dark Energy. *JCAP*, 02:032, 2013.

[84] Jolyon K. Bloomfield, Éanna É. Flanagan, Minjoon Park, and Scott Watson. Dark energy or modified gravity? An effective field theory approach. *JCAP*, 08:010, 2013.

[85] Cumrun Vafa. The String landscape and the swampland. 9 2005.

[86] Hirosi Ooguri and Cumrun Vafa. On the Geometry of the String Landscape and the Swampland. *Nucl. Phys. B*, 766:21–33, 2007.

[87] Eran Palti. The swampland and string theory. *Contemp. Phys.*, 62(3):165–179, 2022.

[88] A. Weltman et al. Fundamental physics with the Square Kilometre Array. *Publ. Astron. Soc. Austral.*, 37:e002, 2020.

[89] Luca Amendola et al. Cosmology and fundamental physics with the Euclid satellite. *Living Rev. Rel.*, 16:6, 2013.

[90] Paul A. Abell et al. LSST Science Book, Version 2.0. 12 2009.

[91] C. Brans and R. H. Dicke. Mach's principle and a relativistic theory of gravitation. *Phys. Rev.*, 124:925–935, 1961.

[92] Joe Kennedy, Lucas Lombriser, and Andy Taylor. Reconstructing Horndeski models from the effective field theory of dark energy. *Phys. Rev. D*, 96(8):084051, 2017.

[93] E. C. G. Stueckelberg. Interaction energy in electrodynamics and in the field theory of nuclear forces. *Helv. Phys. Acta*, 11:225–244, 1938.

[94] M. Ostrogradsky. Mémoires sur les équations différentielles, relatives au problème des isopérimètres. *Mem. Acad. St. Petersbourg*, 6(4):385–517, 1850.

[95] Gregory Walter Horndeski. Second-order scalar-tensor field equations in a four-dimensional space. *Int. J. Theor. Phys.*, 10:363–384, 1974.

[96] B. P. Abbott et al. Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. *Astrophys. J. Lett.*, 848(2):L13, 2017.

[97] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J. Lett.*, 848(2):L12, 2017.

[98] Lucas Lombriser and Andy Taylor. Breaking a Dark Degeneracy with Gravitational Waves. *JCAP*, 03:031, 2016.

[99] Clifford M. Will. The Confrontation between General Relativity and Experiment. *Living Rev. Rel.*, 17:4, 2014.

[100] Cyril Renevey. Review on tests of General Relativity And Modified Gravity Using Pulsar Timing. 5 2019.

[101] Austin Joyce, Bhuvnesh Jain, Justin Khoury, and Mark Trodden. Beyond the Cosmological Standard Model. *Phys. Rept.*, 568:1–98, 2015.

[102] Austin Joyce, Lucas Lombriser, and Fabian Schmidt. Dark Energy Versus Modified Gravity. *Ann. Rev. Nucl. Part. Sci.*, 66:95–122, 2016.

[103] Claudia de Rham. Massive Gravity. *Living Rev. Rel.*, 17:7, 2014.

[104] Clare Burrage and Jeremy Sakstein. Tests of Chameleon Gravity. *Living Rev. Rel.*, 21(1):1, 2018.

[105] Justin Khoury and Amanda Weltman. Chameleon cosmology. *Phys. Rev. D*, 69:044026, 2004.

[106] G. R. Dvali, Gregory Gabadadze, and Massimo Porrati. 4-D gravity on a brane in 5-D Minkowski space. *Phys. Lett. B*, 485:208–214, 2000.

[107] Claudia de Rham, Andrew J. Tolley, and Daniel H. Wesley. Vainshtein Mechanism in Binary Pulsars. *Phys. Rev. D*, 87(4):044025, 2013.

[108] A. I. Vainshtein. To the problem of nonvanishing gravitation mass. *Phys. Lett. B*, 39:393–394, 1972.

[109] Ana Avilez-Lopez, Antonio Padilla, Paul M. Saffin, and Constantinos Skordis. The Parametrized Post-Newtonian-Vainshteinian Formalism. *JCAP*, 06:044, 2015.

[110] Xing Zhang, Wen Zhao, He Huang, and Yifu Cai. Post-Newtonian parameters and cosmological constant of screened modified gravity. *Phys. Rev. D*, 93(12):124003, 2016.

[111] Gregory Gabadadze, Kurt Hinterbichler, and David Pirtskhalava. Classical Duals of Derivatively Self-Coupled Theories. *Phys. Rev. D*, 85:125007, 2012.

[112] Ryan McManus, Lucas Lombriser, and Jorge Peñarrubia. Finding Horndeski theories with Einstein gravity limits. *JCAP*, 11:006, 2016.

[113] Ryan McManus, Lucas Lombriser, and Jorge Peñarrubia. Parameterised Post-Newtonian Expansion in Screened Regions. *JCAP*, 12:031, 2017.

[114] Cyril Renevey, Joe Kennedy, and Lucas Lombriser. Parameterised post-Newtonian formalism for the effective field theory of dark energy via screened reconstructed Horndeski theories. *JCAP*, 12:032, 2020.

[115] Clifford M. Will and Alan G. Wiseman. Gravitational radiation from compact binary systems: Gravitational wave forms and energy loss to second postNewtonian order. *Phys. Rev. D*, 54:4813–4848, 1996.

[116] Ryan N. Lang. Compact binary systems in scalar-tensor gravity. II. Tensor gravitational waves to second post-Newtonian order. *Phys. Rev. D*, 89(8):084014, 2014.

[117] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Phys. Rev. D*, 23:347–356, 1981.

[118] Andrei D. Linde. A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems. *Phys. Lett. B*, 108:389–393, 1982.

[119] Jerome Martin, Christophe Ringeval, and Vincent Vennin. Encyclopædia Inflationaris. *Phys. Dark Univ.*, 5-6:75–235, 2014.

[120] Jerome Martin and Robert H. Brandenberger. The TransPlanckian problem of inflationary cosmology. *Phys. Rev. D*, 63:123501, 2001.

[121] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. Planck evidence for a closed Universe and a possible crisis for cosmology. *Nature Astron.*, 4(2):196–203, 2019.

[122] Leandros Perivolaropoulos and Foteini Skara. Challenges for ΛCDM: An update. *New Astron. Rev.*, 95:101659, 2022.

[123] Roya Mohayaee, Mohamed Rameez, and Subir Sarkar. Do supernovae indicate an accelerating universe? *Eur. Phys. J. ST*, 230(9):2067–2076, 2021.

[124] Nathan J. Secrest, Sebastian von Hausegger, Mohamed Rameez, Roya Mohayaee, and Subir Sarkar. A Challenge to the Standard Cosmological Model. *Astrophys. J. Lett.*, 937(2):L31, 2022.

[125] Will Handley. Curvature tension: evidence for a closed universe. *Phys. Rev. D*, 103(4):L041301, 2021.

[126] Aurélien Barrau. A pure general relativistic non-singular bouncing origin for the Universe. *Eur. Phys. J. C*, 80(6):579, 2020.

[127] Viatcheslav F. Mukhanov, H. A. Feldman, and Robert H. Brandenberger. Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions. *Phys. Rept.*, 215:203–333, 1992.

[128] Béatrice Bonga, Brajesh Gupt, and Nelson Yokomizo. Tensor perturbations during inflation in a spatially closed Universe. *JCAP*, 05:021, 2017.

[129] Béatrice Bonga, Brajesh Gupt, and Nelson Yokomizo. Inflation in the closed FLRW model and the CMB. *JCAP*, 10:031, 2016.

[130] Ali A. Asgari, Amir H. Abbassi, and Jafar Khodagholizadeh. On the perturbation theory in spatially closed background. *Eur. Phys. J. C*, 74:2917, 2014.

[131] Ivan Agullo and Parampreet Singh. *Loop Quantum Cosmology*, pages 183–240. WSP, 2017.

[132] Killian Martineau. *Quelques aspects de cosmologie et de physique des trous noirs en gravitation quantique à boucles*. PhD thesis, LPSC, Grenoble, 2019.

[133] Bryce S. DeWitt. Quantum Theory of Gravity. 1. The Canonical Theory. *Phys. Rev.*, 160:1113–1148, 1967.

[134] J. B. Hartle and S. W. Hawking. Wave Function of the Universe. *Phys. Rev. D*, 28:2960–2975, 1983.

[135] Edward W. Kolb and Michael S. Turner. *The Early Universe*, volume 69. 1990.

[136] Alexander Vilenkin. The Birth of Inflationary Universes. *Phys. Rev. D*, 27:2848, 1983.

[137] R. Giles. The Reconstruction of Gauge Potentials From Wilson Loops. *Phys. Rev. D*, 24:2160, 1981.

[138] Abhay Ashtekar and Parampreet Singh. Loop Quantum Cosmology: A Status Report. *Class. Quant. Grav.*, 28:213001, 2011.

[139] S. W. Hawking. Particle Creation by Black Holes. *Commun. Math. Phys.*, 43:199–220, 1975. [Erratum: Commun.Math.Phys. 46, 206 (1976)].

[140] Thomas Cailleteau, Jakub Mielczarek, Aurelien Barrau, and Julien Grain. Anomaly-free scalar perturbations with holonomy corrections in loop quantum cosmology. *Class. Quant. Grav.*, 29:095010, 2012.

[141] Alejandro Perez. On the regularization ambiguities in loop quantum gravity. *Phys. Rev. D*, 73:044007, 2006.

[142] Lautaro Amadei, Alejandro Perez, and Salvatore Ribisi. Landscape of polymer quantum cosmology. *Phys. Rev. D*, 107(8):086007, 2023.