Primordial black holes from inflation and their gravitational wave signals

Memoria de tesis doctoral realizada por

Julián Leonardo Rey Idler

presentada ante el Departamento de Física Teórica
de la Universidad Autónoma de Madrid
para la obtención del Título de Doctor en Ciencias

Tesis doctoral dirigida por

Dr. Guillermo Ballesteros Martínez

Departamento de Física Teórica
de la Universidad Autónoma de Madrid

Universidad Autónoma de Madrid

Instituto de Física Teórica UAM-CSIC

Madrid, 1 de septiembre de 2022

Articles

This thesis is based on the following articles,

1. **Detuning primordial black hole dark matter with matter domination and axion monodromy** [1]
   G. Ballesteros, **J. Rey** and F. Rompineve
   JCAP **06** (2020) 014 [arXiv:1912.01638]

2. **Primordial black holes as dark matter and gravitational waves from single-field polynomial inflation** [2]
   G. Ballesteros, **J. Rey**, M. Taoso and A. Urbano
   JCAP **07** (2020) 025 [arXiv:2001.08220]

3. **Stochastic inflationary dynamics beyond slow-roll and consequences for primordial black hole formation** [3]
   G. Ballesteros, **J. Rey**, M. Taoso and A. Urbano
   JCAP **08** (2020) 043 [arXiv:2006.14597]

4. **Primordial black holes and gravitational waves from dissipation during inflation** [4]
   G. Ballesteros, M. A. G. García, A. Pérez Rodríguez, M. Pierre and **J. Rey**
   Submitted to *JCAP* [arXiv:2208.14978]

Agradecimientos

Al momento de entregar esta tesis se habrán cumplido diez años desde la primera vez que pisé un salón de clases universitario, tanto tiempo como el transcurrido en la Odisea. La alegoría no es coincidencia, pues el viaje ha estado plagado de problemas y tribulaciones. Desde los cierres de la USB como consecuencia de las protestas, pasando por las interminables horas en autobuses con olor a gasolina atascados en la redoma de Hoyo de la Puerta y mi encuentro con el temido ogro de la Coordinación de Servicio Comunitario (que nada tiene que envidiarle al cíclope Polifemo) y concluyendo con los problemas burocráticos de inmigración en España. Pero es aquí donde termina la metáfora, porque Ulises, después de diez años de viaje, consiguió regresar a casa. Nunca hubiera pensado que la parte más fácil de una carrera en física teórica sería la física.

Esta sección está dedicada a todas aquellas personas que de un modo u otro me han ayudado a llegar a donde estoy. En primer lugar, me gustaría agradecer a mis padres, Julián y Mery, pues fue su sacrificio y esfuerzo porque jamás me faltara nada lo que me permitió estudiar durante tantos años y espero poder devolverles algún día aunque sea una décima parte de lo que me dieron. En segundo lugar quisiera darle las gracias a mi supervisor, Guillermo, por darme una oportunidad y guiarme por el campo de batalla que es la academia. Sus consejos sobre física han sido invaluables, por no mencionar su apoyo en la búsqueda de becas y postdocs. Pero sobre todo, le agradezco por haberme apoyado siempre durante mi tránsito por España y haberme ofrecido su ayuda cuando lo necesité. Le debo mucho también a mis colaboradores, Mathias, Marcos, Alejandro, Alfredo, Marco y Fabrizio. Sobra decir que la mayor parte de esta tesis no habría visto nunca la luz de no ser por su duro trabajo y largas discusiones.

Finalmente, me gustaría agradecer a mis amigos, sin ningún orden particular. Adriana (y Moka), por haberme recibido tantas veces en su hogar (y en su carro), en tres países distintos, pero, sobre todo, por siempre haber estado dispuesta a hablar y reírse conmigo. Mary, por haber hecho la pandemia un poco más soportable y haber compartido conmigo durante un período importante de su vida, siempre tendrá un lugar especial en mi mente. Anabel, por haber estado ahí para mí siempre que la necesité, por haberme empujado a ser ambicioso y haberme apoyado y aconsejado durante la carrera y durante mis aplicaciones a doctorado. Iván, por haberme enseñado todo lo que sé de matemática, haberme apoyado con mis aplicaciones y haber pasado tantas horas resolviendo problemas conmigo en el Ampere, la sala de lectura de matemática en la USB y tantos sitios más. Sabrina, por haberme acompañado y apoyado durante nuestros últimos meses en la USB (que fueron los más duros) y a quien debo en gran medida mi bienestar mental durante ese período. Diego, por

haber sido mi compañero de lucha en las trincheras durante tantos años de carrera, por haber tomado el camino difícil conmigo (adelantando cursos en verano, yendo los fines de semana a la oficina de Lázaro a aprender álgebra y descifrando las demostraciones del Spivak), así como las incontables partidas de ludo y dominó. Guillermo, por haberme recibido tantas veces en su casa, haber descifrado conmigo los acertijos de Mehrdad y haber sufrido conmigo las aplicaciones a doctorado y postdocs. Y por supuesto, a Paola, por haber estado dispuesta siempre a hablar conmigo y hacerme reír. Todas estas personas han estado ahí siempre que las necesité y de no ser por su ayuda y consejos, tal vez esta tesis no existiría.

Espero poder continuar haciendo ciencia durante mucho tiempo, porque, a pesar de las dificultades, no hay un sentimiento comparable en el mundo a haber tenido el privilegio de estudiar el trabajo de los grandes y poder entender un pedacito del universo. Hay más cosas en el cielo y en la tierra, Horacio, de las que han sido soñadas en tu filosofía.

Abstract

We consider the possibility that the majority of dark matter in the Universe consists of black holes of primordial origin, and study the prospects of generating them in single-field models of inflation. Three different scenarios are presented. The first two rely on the presence of an ultra-slow-roll phase in inflation due to an inflection point in the potential. One of these scenarios is characterized by a quartic polynomial potential and is (arguably) the simplest model of inflation able to produce a large population of primordial black holes. The second scenario is aimed at ameliorating the tuning problems present in inflection-point models, and involves a setup that employs the advantages of gravitational collapse in a long epoch of early matter domination, as well as a potential based on a string-inspired class of models in which the inflaton is identified with a non-compact axion field. The third scenario we consider is fundamentally different from the inflection-point models, and consists on obtaining the large peak in the power spectrum of curvature perturbations necessary for black hole formation from a transient dissipative phase during inflation. In this case the enhancement of the power spectrum occurs due to the presence of a stochastic thermal noise source in the equation of motion for the fluctuations.

We consider the impact of quantum diffusion on the inflationary dynamics during an ultra-slow-roll phase and show, by means of a fully analytical approach, that the power spectrum of comoving curvature perturbations computed in stochastic inflation matches precisely, at the linear level, the result obtained by solving the Mukhanov-Sasaki equation. Finally, we compute the stochastic background of gravitational waves generated in each scenario. In particular, we study the gravitational waves induced during an early matter-dominated era and determine how much of the parameter space remains available after taking into account the bounds on the gravitational wave energy density arising from the abundance of light elements produced during Big-Bang nucleosynthesis and cosmic microwave background experiments. We examine the gauge dependence of the resulting signal and, by using a heuristic argument based on symmetry properties and dimensional analysis, determine the full gauge-invariant expression for the energy density of gravitational waves at next-to-leading order in perturbations. We discuss the prospects of detecting the resulting signal with the LISA experiment.

Resumen

Consideramos la posibilidad de que la mayor parte de la materia oscura del universo consista de agujeros negros de origen primordial y estudiamos la capacidad de los modelos de inflación con un único campo de producirlos. Presentamos tres escenarios distintos. Los primeros dos hacen uso de una fase en la cual el campo rueda muy lentamente gracias a la presencia de un punto de inflexión en el potencial. Uno de estos escenarios está caracterizado por un potencial polinómico de cuarto orden y es (posiblemente) el modelo más simple de inflación que es capaz de producir una población considerable de agujeros negros primordiales. El segundo escenario tiene como objetivo aminorar los problemas de ajuste fino presentes en los modelos con puntos de inflexión e involucra una configuración que emplea las ventajas del colapso gravitacional en una época temprana en la que el universo estaría dominado por materia no relativista, así como un potencial basado en una clase de modelos inspirados por la teoría de cuerdas en los que el inflatón es identificado con un campo axiónico no compacto. El tercer escenario que consideramos es fundamentalmente distinto de los modelos con puntos de inflexión y consiste en obtener el incremento del espectro de perturbaciones de curvatura necesario para la formación de agujeros negros a partir de una fase disipativa de corta duración durante la inflación. En este caso el incremento en la amplitud del espectro ocurre gracias a la presencia de una fuente estocástica de ruido térmico en la ecuación de movimiento de las fluctuaciones.

Estudiamos también el impacto que tiene la difusión cuántica sobre la dinámica inflacionaria durante una fase en la cual el campo rueda muy lentamente y demostramos, mediante un enfoque totalmente analítico, que el espectro de las perturbaciones calculado en el formalismo de inflación estocástica es exactamente igual al obtenido al resolver la ecuación de Mukhanov y Sasaki. Finalmente, calculamos el fondo estocástico de ondas gravitacionales generado en cada escenario y, en particular, estudiamos las ondas gravitacionales inducidas durante una fase temprana en la cual el universo estaría dominado por materia no relativista y determinamos qué región del espacio de parámetros es eliminada después de tomar en cuenta las cotas sobre la densidad de energía de las ondas gravitacionales que provienen de la abundancia de elementos ligeros producidos durante la nucleosíntesis y experimentos del fondo cósmico de radiación. Examinamos la dependencia del resultado bajo transformaciones de coordenadas y, utilizando un argumento heurístico basado en propiedades de simetría y análisis dimensional, determinamos la expresión completa y manifiestamente invariante bajo cambios de coordenadas de la densidad de energía de las ondas gravitacionales a orden subdominante en las perturbaciones. Discutimos la posibilidad de detectar la señal resultante utilizando el experimento LISA.

Table of contents

Introduction **7**

Chapter 1. Black holes from cold inflation **28**

1.1 Cold inflation . . . 28
1.2 Primordial black holes . . . 32
1.3 Perturbative reheating . . . 41
1.4 The curvature perturbation . . . 44

Chapter 2. Inflationary potentials **50**

2.1 Polynomial potential . . . 50
2.2 Higher-dimensional operators . . . 58
2.3 Monodromy-inspired potential . . . 64

Chapter 3. Quantum backreaction **73**

3.1 Stochastic inflation . . . 73
3.2 Classicalization of the modes . . . 80
3.3 The noise matrix . . . 85
3.4 Numerical analysis . . . 90

Chapter 4. Black holes from warm inflation **94**

4.1 Dissipation during inflation . . . 94
4.2 The matrix formalism . . . 100
4.3 Solving the stochastic equations . . . 104
4.4 Analytical approximation . . . 106

Chapter 5. Gravitational wave signals **114**

5.1 Induced gravitational waves . . . 114
5.2 Gauge dependence of the signal . . . 120
5.3 Energy density bounds . . . 126
5.4 Gravitational waves in warm inflation . . . 132

Conclusions **142**

Appendices **151**

A Cosmological perturbation theory . . . 151
B Power counting and effective theories . . . 154
C Stochastic differential equations . . . 156
D Microphysics of the dissipative coefficient . . . 161
E Stochastic and quantum expectation values . . . 164
F Analytical gravitational wave integrals . . . 167
G Induced gravitational waves in different gauges . . . 168

References **173**

Introduction

The dark matter problem

At present, there is overwhelming evidence that the matter content of our Universe cannot be entirely accounted for. Dark matter makes up around 85% of the amount of matter consistent with observations, whereas the standard, baryonic matter accounts for the rest. The earliest evidence of dark matter came in the 1930s [5], when the virial theorem was used to estimate the velocity dispersion of galaxies in the Coma Cluster, and it was determined that baryonic matter alone would yield a result about two orders of magnitude smaller than the observed one [6]. The idea of dark matter, however, did not take off until the 1970s, when it was determined that galaxy rotation curves flatten as the distance from the galactic center increases [7], a fact that cannot be described by using baryonic matter alone. Since then, the amount of evidence in favor of some as-yet-unknown, electromagnetically non-interacting form of matter has vastly grown, and ranges from the gravitational lensing of distant sources produced by dark matter [8] and the measurement of the dark matter energy density from the peaks in the angular power spectrum of the cosmic microwave background (CMB) anisotropies [9] to the measurement of the rate of expansion of the Universe using type Ia supernovae [10, 11].

From the point of view of particle physics, a series of well-motivated dark matter candidates have been proposed [12], although none of them have been detected so far. These include, for instance, weakly interacting massive particles (WIMPs) proposed due to the observation that a particle with a self-interaction cross section characteristic of the electroweak force would produce the correct abundance of dark matter today, particles arising from supersymmetric extensions to the Standard Model, sterile neutrinos which would interact with the rest of the Standard Model only through gravity, and axions, proposed as a possible solution to the so-called strong CP problem of quantum chromodynamics. On the side of astrophysics, objects typically far less luminous than stars were also considered as viable dark matter candidates, such as planets, brown dwarfs, neutron stars, and black holes. These were collectively labeled as massive compact halo objects, or MACHOs, and most were ultimately ruled out due to gravitational microlensing surveys, and the determination of the amount of baryonic matter from measurements of the abundance of light elements produced during Big Bang nucleosynthesis [5].

A compelling alternative which has so far stood the test of time is the possibility that dark matter is comprised of black holes formed before nucleosynthesis (and therefore via mechanisms different from the usual stellar collapse) and with masses that cannot be probed

by gravitational microlensing. These objects are known as primordial black holes (PBHs). Since these black holes do not interact electromagnetically, they are viable cold dark matter candidates, where the attribute cold refers to the fact that, being compact objects with non-relativistic velocities, a fluid of PBHs would behave as presureless dust. The possibility that black holes could have formed in the early Universe due to its higher energy density was first introduced in [13–15], but they were first proposed as dark matter candidates in [16]. Assuming the majority of dark matter to be comprised of PBHs, in this thesis we will investigate different aspects about their formation and observable astrophysical signatures. The most studied mechanism of PBH formation relies on gravitational collapse triggered by large density fluctuations in the early Universe. These density fluctuations are often assumed to originate during a period of accelerated expansion known as cosmic inflation.

Cosmic inflation

The discovery of the cosmic microwave background [17] was one of the most important scientific findings of the last century, and solidified the Big Bang model of cosmology. However, it also brought with itself a series of puzzles that we still struggle to explain today. Arguably, the most important of these issues are the horizon and flatness problems. The first one refers to the fact that the CMB temperature is nearly uniform across the entire visible Universe, even though different regions across this observable patch were causally disconnected at the time photons decoupled from the primordial plasma. The second one refers to the fact that the spatial curvature of the Universe, which is a free parameter of the theory, is strongly constrained to be a very small number, leading to a fine-tuning problem. Although one could always dismiss these issues by appealing to the fine-tuning of initial conditions, a dynamical explanation would be highly desirable. The inflationary paradigm [18–20] aims to deliver such an explanation and solve both of these issues in a single stroke by postulating that the Universe went through an early phase of accelerated expansion.

The appeal of cosmic inflation is not restricted to the solution of the CMB puzzles, however. Soon after the discovery of the CMB and motivated by the idea that the formation of structure in the Universe must have been seeded by small density fluctuations, cosmologists began looking for anisotropies in the CMB radiation [21]. The first measurements of this anisotropy were performed by the COBE satellite [22], which found a nearly Gaussian and scale-invariant distribution of temperature fluctuations of order $\delta T/T \sim \mathcal{O}(10^{-5})$. A distribution with precisely these characteristics can be obtained within the inflationary framework, adding to the success of the idea. In the simplest models of inflation, a phase of accelerated expansion can occur if the energy density of the Universe is dominated by a scalar field that slowly rolls down a potential. Quantum fluctuations of this field then translate into the anisotropies we observe in the CMB, and provide the seeds for the formation of the large scale structure in the Universe. These perturbations can be shown to freeze once their associated wavenumber becomes comparable to the Hubble rate, and are therefore blind to

subhorizon dynamics until their horizon re-entry, which allows us to easily connect them to CMB observations.

There are many CMB experiments active at present, such as Planck and BICEP, which have managed to either set strong bounds on inflationary observables or measure them with high precision. The tilt of the scalar power spectrum, for instance, has been measured at the 0.1% level at the time of writing [23], whereas the tensor-to-scalar ratio r that measures the difference between the amplitude of the scalar power spectrum and that of the primordial stochastic gravitational wave background is constrained at the percent level [24]. Experiments such as the ground-based CMB-S4 telescopes [25] and the LiteBIRD satellite [26] are set to begin operating within the next decade and will further tighten these measurements and potentially detect other not-yet-seen features of the CMB, such as the imprint left by primordial gravitational waves in the CMB polarization in the form of B-modes.

Primordial black holes

PBHs are very compelling dark matter candidates. On the theoretical side, the simplest models of formation do not require any additions to the Standard Model of particle physics other than inflation. On the other hand, being compact objects, they also exhibit a wide variety of astrophysical signatures [27]. At the time of writing, PBHs can account for all the dark matter provided their masses lie in the window

$$10^{-16}\ M_\odot \lesssim M_{\rm PBH} \lesssim 10^{-11} M_\odot. \tag{1}$$

The upper limit in this range of masses is due to microlensing observations using the Subaru Telescope's HSC camera [28], and the lower end is due to the fact that black holes with very small masses would be evaporating today (or have evaporated already if the masses are sufficiently small), emitting Hawking radiation and leading to the observation of extragalactic [29, 30] and galactic [31, 32] γ-rays, the tightest bound being the latter, imposed by the INTEGRAL satellite. We remark that there are other bounds on the lower end of this range, such as the ones arising from the study of electromagnetic energy injection in the CMB [33, 34], the injection of electron-positron cosmic rays into the galaxy due to Hawking radiation, which would be observed by the Voyager 1 probe [35], and the annihilation of positrons contributing to the Galactic 511 keV line [36–38]. Throughout this thesis we will be concerned only with the mass range shown in (1), but let us point out that PBHs with higher masses are constrained by microlensing observations from EROS/MACHO [39] and OGLE [40], by the effect their accretion would have on the CMB [41, 42], and the non-observation of the stochastic gravitational wave background by LIGO [43].

A bound due to gamma-ray burst femtolensing [44] existed previously in the mass range (1), but was disputed in [45], where it was argued that most gamma-ray bursts are too large to be used for femtolensing observations. The possibility that a subset of bursts with appropriate sizes could be used was also entertained in [45], although a large number of them

would have to be detected by future experiments for this to be a viable observational tool. Other previously existing bounds in this mass range include the explosion of white dwarfs [46] due to runaway thermonuclear fusion caused by the friction generated by the transit of a sufficiently massive black hole through the star, and the accretion and subsequent destruction of a neutron star by black holes [47]. These bounds were disputed in [48], where the capture rate of black holes by neutron stars was derived[1] and hydrodynamic simulations were performed to estimate the temperature and timescale of the shock produced by a black hole that passes through a white dwarf. These results determined the constraints to be less effective than previously thought.

At this point, the relation between the inflationary paradigm and the formation of PBHs should not come as a surprise. In the simplest and most studied PBH formation mechanisms, inflation provides the seeds for the large density fluctuations that trigger the gravitational collapse of matter into black holes [50] at later stages. The density fluctuations observed in the CMB turn out to be far too small to produce a significant amount of black holes, so for PBHs to constitute a sizable fraction of the observed dark matter, said fluctuations must be enhanced at distance scales that are much smaller than the ones constrained by CMB experiments. This enhancement of the density fluctuations can be accomplished in single-field models of inflation if the inflaton traverses through a region of its potential with an approximate inflection point, which would then trigger a so-called ultra-slow-roll (USR) phase [51] which increases the primordial spectrum $\mathcal{P}_{\mathcal{R}}$ of curvature perturbations. As we will discuss later on, large density fluctuations can also be produced from a transient dissipative phase during inflation due to the interaction between the inflaton field and an underlying bath of thermalized radiation, which leads to a source of stochastic noise in the equation of motion for the fluctuations that enhances the primordial spectrum. The enhancement of the power spectrum can also be obtained in two-field hybrid models of inflation, see [52] for the earliest implementation of this idea. Throughout this thesis we will be concerned only with inflationary mechanisms of PBH formation, but let us remark that other, non-inflationary alternatives have also been proposed, such as the collapse of topological defects [53–57], the collision of vacuum bubbles [58, 59], the collapse of Fermi balls [60], the increased probability of gravitational collapse during a phase transition due to a change in the equation of state of the Universe [61], and the collapse of a large buildup of particles in a shrinking false vacuum bubble [62].

Let us turn our attention to the aforementioned inflection-point models of PBH production. The first attempt to implement this mechanism was put forward in [63],[2] but interest in the idea remained largely dormant until the first detection of gravitational waves by LIGO [65]. Since then, the field has seen a revitalized interest and a large number of proposed models have emerged. One of the first modern takes on the inflection-point idea was performed

[1] See also [49] for further work in this direction.

[2] See also [64], where a linear potential with a smooth change in slope able to generate an enhancement in the power spectrum was studied, although not with the purpose of producing PBHs.

in [66], where a potential based on the ratio of two polynomials was proposed. Shortly afterwards, the possibility that the feature in the potential could have a radiative origin was explored in [67, 68]. In [68] it was assumed that the inflaton had a polynomial potential dominated by the quartic term, and the inflection point then arises from the precise tuning of the coefficients of the logarithmic quantum corrections. In addition, to fit the CMB data, [68] considered a completely general but non-minimal coupling between the inflaton and the Ricci scalar, which flattens the potential at large field values. The possibility of obtaining the feature by carefully adjusting the coefficients of a polynomial potential, without the need for quantum corrections, was also studied in [68]. More exotic examples of models are those motivated by string theory and supergravity, such as axion-like potentials [69], or potentials constructed within the framework of supersymmetric α-attractor models [70].

These models do not come without disadvantages, however. In general, the inflection point is not a feature that arises naturally in these potentials, and their parameters must instead be engineered specifically for this purpose. On the other hand, to accomplish the enhancement of the primordial spectrum, a severe tuning of the parameters in the potential is usually required. There are two reasons for this. The first is that the size of the power spectrum is exponentially sensitive to the decrease in the speed of the inflaton when it reaches the inflection point. If the parameters are not carefully adjusted, the inflaton either overshoots the local minimum that usually precedes the inflection point at high speed (so that fluctuations are barely enhanced and PBHs are underproduced), or it gets stuck in said local minimum. Another issue is that the speed necessary to produce a significant amount of PBHs in a given model is often at odds with the speed necessary to yield enough e-folds of inflation to solve the horizon and flatness problems of the CMB. The region in parameter space that can satisfy all of these constraints is usually very narrow. The second reason for the tuning is that PBHs are usually assumed to be formed during the radiation-dominated (RD) era preceding Big Bang nucleosynthesis, so that the current PBH abundance depends exponentially on the power spectrum, and small deviations from the sweet spot in which the fraction of PBHs as dark matter is $\mathcal{O}(1)$ therefore result in PBH abundances orders of magnitude too small to be of interest for the dark matter problem. In this thesis we will see how some of these problems can be ameliorated in the context of a specific inflationary potential inspired by axion monodromy inflation.

One effect that was not taken into account in the early treatments of inflection-point models of PBH production is that of quantum diffusion. In single-field inflation, we typically assume the homogeneous part of the field to behave classically, and only treat the perturbations quantum mechanically. In the standard picture, these perturbations become constant and classical once they cross the Hubble horizon, determining the power spectrum. Quantum fluctuations could, however, backreact on the classical trajectory of the inflaton, a possibility that cannot be handled by the above description. This effect must instead be understood in the framework of stochastic inflation [71], in which long-wavelength fluctuations

are sourced by small-wavelength ones. The latter behave as classical random variables and their effect can be described via a set of stochastic differential equations. It is known that in the slow-roll regime this procedure leads to a $\mathcal{P}_\mathcal{R}$ in agreement with the result of standard perturbation theory at the linear level (see e.g. [72]), but whether or not this result holds if the inflaton goes through a USR phase (as in the aforementioned inflection-point models) is less clear. In the presence of a USR phase, it has been speculated that the backreaction could potentially modify the peak height of $\mathcal{P}_\mathcal{R}$, the location of the peak itself, or even broaden the PBH mass function [68]. Studying the effect of quantum diffusion is therefore important for the correct determination of the PBH abundance, and is a subject we will treat in this thesis. Another advantage of the stochastic inflation formalism is that it also allows the calculation of the non-Gaussian corrections to the probability distribution function of the primordial curvature fluctuations, which can have an effect on the PBH abundance [73–76].

Gravitational waves

The primordial stochastic background of gravitational waves, which is potentially within the reach of detectors that will begin operating in the near future (such as LISA [77]) and is comparable to the CMB in the implications it would have for cosmology, has long been considered a powerful tool to probe the physics of the early Universe. There are many reasons for this, not the least of which is the fact that gravitational waves can probe processes and energy scales inaccessible by other means, such as phase transitions, topological defects and reheating after inflation. Another reason is that by using the techniques of cosmological perturbation theory, it is possible to show that at leading order in perturbations the tensor modes of the metric (that is, the gravitational waves) decouple from other degrees of freedom and propagate freely, affected only by the expansion of the Universe and carrying information about the processes that generated them [78].

The fact that PBH formation requires a large scalar power spectrum leads to an enhancement of the second-order tensor modes of the metric, which are sourced by terms quadratic in first-order scalars in Einstein's equations [79–81]. Moreover, if PBHs constitute all of the dark matter, then it should be possible to detect the corresponding induced stochastic background of gravitational waves if they form during a RD era and, in particular, if their masses lie within the mass window (1), then this stochastic background would be observable[3] by the LISA experiment [77, 82, 83]. This is due to the fact that, assuming a monochromatic mass distribution, the mass of a PBH that forms during a RD era due to collapse induced by large scalar fluctuations is related to the peak frequency of the corresponding gravitational wave spectrum via [82, 83]

$$M_{\rm PBH} \simeq \gamma\big(3\times 10^{-15} M_\odot\big)\left(\frac{0.1{\rm Hz}}{f_{\rm peak}}\right)^2, \tag{2}$$

[3]However, it is not clear at present that a PBH signal can be distinguished from other sources.

where γ is an $\mathcal{O}(1)$ efficiency factor. It can be seen from this equation that the frequencies that LISA can test correspond to unconstrained PBH masses in the window (1). Not only this, but the amplitude of the signal also turns out to lie roughly four orders of magnitude above the LISA sensitivity curve, as we will show.

The possibility that gravitational waves can be induced at second order in perturbations was, to the best of our knowledge, first suggested in [79], although the idea did not gain traction until it started to be considered as a possible way to probe the existence of PBHs. The reason for this is that the amplitude of these induced gravitational waves is typically suppressed with respect to the first-order contribution, and they therefore only become relevant when there is a large peak in the scalar power spectrum, as is the case in the aforementioned inflationary models of PBH production. These gravitational waves are induced not only during inflation, but also afterwards, once the perturbations re-enter the horizon, and therefore contain information about both the mechanism that produced them and the subsequent cosmological history.

If we want to use gravitational waves as an observable to probe the existence of PBHs, it is necessary to carefully examine the calculation of the signal in order to determine whether the results can be trusted. It has been pointed out recently that the usual calculation of the stochastic background of gravitational waves induced at second order from scalar fluctuations is not entirely solid, since the second-order tensor piece of the metric is not a gauge-invariant variable [84–89]. This gauge dependence then propagates to the observable quantity of interest: the energy density of gravitational waves. It has been recently claimed [86, 90] that despite the above remarks, the usual calculation of the energy density of gravitational waves is valid beyond leading order in perturbations, and this quantity is approximately gauge-invariant as long as it is measured late in the radiation era, since the scalars that source the second-order tensor modes decay quickly with time in this case, so that the latter propagate linearly and effectively become first-order quantities [86]. However, the energy density of gravitational waves should be a gauge-invariant quantity which is well-defined at all times, independently of when it is measured. The determination of the correct expression for this quantity beyond leading order in perturbations remains an open problem, one that we shall explore in more detail in this thesis.

Overview of the thesis

In this thesis we consider two different implementations of the inflection-point mechanism of PBH formation. The first consists on making the (arguably) simplest possible choice for the potential of the inflaton, namely, a quartic polynomial, together with a non-minimal coupling to the Ricci scalar [2, 68]. If the coefficients in the polynomial are chosen appropriately, this leads to the presence of an inflection point in the potential. As we will see, in this model the main issue is that the tilt of the power spectrum at CMB scales is in mild tension with the experimentally measured values if we require that enough PBHs with masses in the window

(1) are produced to account for all the dark matter. We will show that this issue can be resolved by either considering very simple, well-motivated extensions of the cosmological ΛCDM model (such as the addition of neutrino masses), or by adding higher-dimensional operators (the presence of which is to be expected anyway) to the potential. The second possibility we consider is more elaborate and is aimed at alleviating the aforementioned downsides of inflection-point models of PBH production. The key aspects in this second setup are an inflationary potential which naturally features the existence of several local minima, and an early matter-dominated (eMD) epoch right after inflation (which can be obtained, for instance, from perturbative reheating), during which the likeliness of gravitational collapse is augmented [50, 91, 92]. We find that these two ingredients can also lead to the formation of a sizable fraction of PBHs with masses in the range (1) [1].

The reason that an eMD era ameliorates some of the aforementioned tuning problems is that, due to the absence of radiation pressure, it is easier for the matter cloud to collapse, so the enhancement of the power spectrum required for the PBHs to form turns out to be orders of magnitude smaller than in the standard radiation scenario, even after taking into account the fact that the angular momentum and non-sphericity of the cloud inhibit collapse [92, 93]. Moreover, although the dependence of the PBH abundance on the power spectrum is still exponential during the eMD era, its sensitivity to small changes in the amplitude of the fluctuations that induce collapse turns out to be much milder than in the RD scenario. The formation of PBHs during an eMD phase has been well-studied in the literature, see e.g. [91] for early work and [92–94] for recent updates. As we will show, in order to take full advantage of the eMD scenario (in the sense that a sizeable PBH abundance is obtained with a power spectrum that is as close as possible to its value at CMB scales), the transition temperature between the eMD and the RD eras must be of order $\lesssim 10^5$ GeV. Similar conclusions were reached in previous works. Here, we expand with respect to estimates previously presented in [70, 93].

So far we have focused exclusively on inflection-point models of PBH production. As we mentioned earlier, however, there are other possibilities. In particular, black holes could be generated from a transient dissipative phase during inflation, a mechanism that is fundamentally different from the inflection-point models discussed above. The effect of dissipation on the dynamics of the background inflaton field can be described by introducing a friction term $\propto \Gamma\dot{\phi}$ in its equation of motion [95, 96]. One might conjecture that this term could slow down the inflaton in the same way that an inflection point in the potential does, leading to a large enhancement of the power spectrum. As we will show, this argument is not correct, since, although the background dynamics of the inflaton are similar in both scenarios, the equations of motion for the fluctuations are quite different. However, as we mentioned earlier, the power spectrum can still be enhanced due to the presence of a stochastic source of thermal noise for the fluctuations that originates from the interaction between the inflaton and an underlying thermalized radiation bath. We assume this source to be active only for

a short period in order to produce PBHs with adequate masses in the range (1). We model our scenario using a phenomenological approach, assuming that the dissipative coefficient Γ has a localized peak and a well-motivated temperature dependence. We perform a full numerical analysis of the dynamics by directly solving the stochastic differential equations for the perturbations. We integrate the system for multiple realizations of the thermal noise, thereby obtaining a histogram for $\mathcal{P}_{\mathcal{R}}(k)$ at each k, allowing us to determine its statistical properties. The most interesting quantity is not necessarily the full probability distribution for the power spectrum, however, but rather its expectation value, which is a deterministic quantity. We show that this expectation value can be calculated by reformulating the stochastic system of differential equations for the fluctuations as a single deterministic matrix differential equation for their two-point functions.

In addition to the exploration of the specific models of PBH formation discussed above, in this thesis we also reassess the computation of the primordial power spectrum, in the presence of a USR phase, using the framework of stochastic inflation, both analytically and numerically [3]. We find that (at linear order) the power spectrum computed using the full machinery of stochastic inflation agrees with the result of standard perturbation theory, and we confirm our results numerically using the aforementioned polynomial model. We find that the choice of the coarse-graining scale is critical, since in the presence of a USR phase perturbations take longer to classicalize, and so the standard choice made in slow-roll inflation is no longer valid. We do not study higher-order correlators for the perturbations in this thesis, but we remark that the fact that the power spectrum remains unchanged does not necessarily imply that stochastic effects are irrelevant for the black hole abundance, since large non-Gaussianities are still expected to modify their formation probability [73, 74, 97].

Since scenarios in which PBHs form during an eMD era have significant advantages, it is worth asking whether the gravitational wave signal in this case changes with respect to the usual radiation scenario. Specifically, we would like to know whether the resulting stochastic background can be detected by future experiments, such as LISA and BBO/DECIGO. One might conjecture that this is not the case, since, as discussed earlier, the enhancement of the scalar power spectrum required for PBHs to be able to account for all dark matter is much smaller in this scenario. However, we will show that due to the fact that scalar fluctuations do not decay with time during a matter-dominated era, the gravitational wave spectrum increases in amplitude as the transition temperature between the eMD and RD eras decreases. This additional enhancement counters the relatively small amplitude of the scalar power spectrum, rendering the stochastic gravitational wave background observable by LISA. We also examine whether the resulting spectrum satisfies the bound on the present abundance of gravitational waves Ω_{GW} arising from measurements of the CMB [98] and the abundance of light elements produced during Big Bang nucleosynthesis [99], and find that this leads to a severe constraint in the eMD scenario, ruling out part of the available parameter space. How much of the parameter space is ruled out depends on the transition

temperature, but as we will see, for certain values of this quantity this bound can be even stronger than the ones coming from Hawking radiation experiments.

The gauge dependence of the second-order tensor modes of the metric has been extensively studied for both matter- and radiation-dominated Universes (see e.g. [84, 85, 89]). In the eMD-to-RD transition scenario, however, the Universe goes through both phases, so it is important to include the two in the computation. This calculation is presented here for the first time, and we find that because the eMD era has a finite duration, the set of problematic gauges coincides with the one in the pure RD case [85]. In particular, we examine the issue by considering four different gauges (Newtonian, uniform curvature, uniform expansion and comoving), and find that the results coincide at late times (that is, deep into the RD era) for the first three gauges. We postulate that the issue of the gauge dependence stems from the incorrect assumption that the corrections at subleading order in perturbations to the abundance of gravitational waves $\Omega_{\rm GW}$ can be obtained by simply substituting[4] $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ in the usual formula (the derivation of which can be found in e.g. [100, 101]),

$$\Omega_{\rm GW}(\tau, k) = \frac{1}{6}\left(\frac{k}{\mathcal{H}}\right)^2 \langle \mathcal{P}_h(\tau, k)\rangle_{\rm W}, \tag{3}$$

where $\mathcal{P}_h$ denotes the power spectrum of $h_{ij}^{\rm TT}$, τ denotes conformal time, $\mathcal{H} = aH$ the conformal Hubble parameter, and the brackets denote an average over many wavelengths. Indeed, $\Omega_{\rm GW}$ must be a gauge-invariant, observable quantity, and cannot be expressed in terms of the gauge-dependent $\boldsymbol{h}_{ij}^{\rm TT}$. In this thesis, we provide a heuristic derivation of the full expression for this quantity, which is valid in any gauge and independent of the energy content of the Universe, by using an argument based on symmetry properties and dimensional analysis.

The thesis is structured as follows. In Chapter 1 we review the basic aspects of inflation and primordial black hole formation in matter- and radiation-dominated eras. In Chapter 2 we implement the idea of black hole production through an inflection point in the potential by using two different inflationary models, one based on a polynomial potential and another one inspired by axion monodromy. In Chapter 3 we explore the effect that the quantum diffusion of the inflaton field has on the power spectrum of comoving cuvature perturbations. In Chapter 4 we discuss how black holes can form due to a transient dissipative phase during inflation. Finally, in Chapter 5 we calculate the gravitational wave signals in each scenario, and discuss their observational prospects.

[4] We use $h_{ij}^{\rm TT}$ to denote the first-order, transverse, traceless tensor modes, and the bold symbol $\boldsymbol{h}_{ij}^{\rm TT}$ to denote second-order modes.

Introducción

El problema de la materia oscura

En la actualidad, existe una gran cantidad de evidencia que apunta al hecho de que el contenido de materia del universo no se puede explicar por completo. La materia oscura constituye alrededor del 85% de la cantidad de materia consistente con las observaciones, mientras que la materia bariónica estándar representa el resto. La evidencia más temprana de materia oscura se produjo en la década de 1930 [5], cuando se usó el teorema virial para estimar la dispersión de velocidades de las galaxias en el cúmulo de Coma y se determinó que la materia bariónica por sí sola daría un resultado aproximadamente dos órdenes de magnitud menor que el observado [6]. La idea de la materia oscura, sin embargo, no despegó sino hasta la década de 1970, cuando se determinó que las curvas de rotación de las galaxias se aplanan a medida que aumenta la distancia al centro galáctico [7], un hecho que no se puede describir usando solo materia bariónica. Desde entonces, la cantidad de evidencia a favor de alguna forma de materia que aún no se conoce y que no interactúa electromagnéticamente ha crecido enormemente y va desde la desviación de la luz proveniente de fuentes distantes cuando la materia oscura actúa como lente gravitacional [8] y la medición de la densidad de energía de la materia oscura utilizando los picos en el espectro angular de potencias de las anisotropías del fondo cósmico de microondas (CMB, por sus siglas en inglés) [9] hasta la medición de la tasa de expansión del universo utilizando supernovas de tipo Ia [10, 11].

Desde el punto de vista de la física de partículas han sido propuestos una serie de candidatos con una motivación física sólida para explicar la materia oscura [12], aunque ninguno de ellos ha sido detectado hasta ahora. Estos incluyen, por ejemplo, partículas masivas que interactúan débilmente (WIMPs, por sus siglas en inglés), propuestas debido a la observación de que una partícula con una tasa de autointeracción característica de la fuerza electrodébil produciría la abundancia correcta de materia oscura observada hoy en día, partículas que surgen de extensiones supersimétricas al Modelo Estándar, neutrinos estériles que interactuarían con el resto del Modelo Estándar solo a través de la gravedad y axiones, propuestos como una posible solución al problema de CP fuerte de la cromodinámica cuántica. Del lado de la astrofísica, objetos típicamente mucho menos luminosos que estrellas también fueron considerados durante un tiempo como candidatos viables a la materia oscura, tales como los planetas, las enanas marrones, las estrellas de neutrones y los agujeros negros. Estos objetos de halos compactos masivos (MACHOs, por sus siglas en inglés) fueron, en su mayoría, descartados debido a observaciones de microlentes gravitacionales y la determinación de la cantidad de materia bariónica a partir de mediciones de la abundancia de elementos ligeros

producidos durante la nucleosíntesis del Big Bang [5].

Una alternativa viable que hasta ahora ha superado la prueba del tiempo es la posibilidad de que la materia oscura esté compuesta por agujeros negros formados antes de la nucleosíntesis (y, por lo tanto, a través de mecanismos distintos al colapso estelar habitual) y con masas que no pueden ser probadas por microlentes gravitacionales. Estos objetos se conocen como agujeros negros primordiales (PBHs, por sus siglas en inglés). Dado que estos agujeros negros no interactúan electromagnéticamente, son candidatos viables como materia oscura fría, donde el atributo frío se refiere a que, al ser objetos compactos con velocidades no relativistas, un fluido de PBHs se comportaría como polvo sin presión. La posibilidad de que los agujeros negros se hayan formado en el universo primitivo debido a la mayor densidad de energía se introdujo por primera vez en [13–15], aunque la primera vez que fueron propuestos como candidatos a la materia oscura fue en [16]. Suponiendo que la mayoría de la materia oscura esté compuesta por PBHs, en esta tesis investigaremos diferentes aspectos sobre su formación y señales astrofísicas observables. El mecanismo más estudiado de formación de PBHs se basa en el colapso gravitatorio provocado por grandes fluctuaciones de densidad en el universo primitivo. Estas fluctuaciones de densidad podrían generarse durante un período de expansión acelerada conocido como inflación cósmica.

Inflación cósmica

El descubrimiento del fondo cósmico de microondas [17] fue uno de los hallazgos científicos más importantes del siglo pasado y solidificó el modelo cosmológico del Big Bang. Sin embargo, también trajo consigo una serie de acertijos que aún hoy nos cuesta explicar. Podría decirse que los más importantes de estos problemas son los problemas de horizonte y planitud. El primero se refiere al hecho de que la temperatura del CMB es casi uniforme en todo el universo visible, aunque las diferentes regiones de este parche observable eran causalmente disconexas en el momento en que los fotones se desacoplaron del plasma primordial. El segundo se refiere al hecho de que la curvatura espacial del universo, que es un parámetro libre de la teoría, debe de ser un número muy pequeño para ser consistente con las observaciones, lo cual genera un problema de ajuste fino. Aunque uno siempre podría descartar estos problemas apelando al ajuste fino de las condiciones iniciales, tener una explicación dinámica sería lo ideal. El paradigma inflacionario [18–20] tiene como objetivo ofrecer tal explicación y resolver ambos problemas simultáneamente postulando que el universo pasó por una fase temprana de expansión acelerada.

Sin embargo, el atractivo de la inflación cósmica no está limitado a la solución de los acertijos del CMB. Poco después del descubrimiento del CMB y motivados por la idea de que la formación de estructura en el universo debía de tener origen en pequeñas fluctuaciones de densidad producidas en el universo temprano, los cosmólogos comenzaron a buscar anisotropías en la radiación del CMB [21]. Las primeras mediciones de esta anisotropía fueron realizadas por el satélite COBE [22], que encontró una distribución de fluctuaciones

de temperatura de orden $\delta T/T \sim \mathcal{O}(10^{-5})$. Una distribución con precisamente estas características se puede obtener en el paradigma inflacionario, lo cuál contribuyó al éxito de la idea. En los modelos más simples de inflación puede ocurrir una fase de expansión acelerada si la densidad de energía del universo está dominada por un campo escalar que rueda lentamente por su potencial. Las fluctuaciones cuánticas de este campo se traducen luego en las anisotropías que observamos en el CMB y proporcionan las semillas para la formación de la estructura a gran escala en el universo. Se puede demostrar que estas perturbaciones se congelan una vez que su número de onda asociado se vuelve comparable al parámetro de Hubble y, por lo tanto, son ciegas a la dinámica subhorizonte hasta que entran de nuevo en el mismo, lo cual nos permite conectarlas fácilmente con las observaciones del CMB.

En la actualidad, existen una gran cantidad de experimentos del CMB activos, tales como Planck y BICEP, que han logrado o bien establecer fuertes límites sobre los observables inflacionarios, o medirlos con alta precisión. La inclinación del espectro de potencias escalar, por ejemplo, ha sido determinada con una precisión relativa del 0.1% al momento de escribir esta tesis [23]. Similarmente, existe una cota al nivel porcentual sobre el cociente tensor-escalar r que mide la diferencia entre la amplitud del espectro de potencias escalar y la del fondo primordial estocástico de ondas gravitacionales [24]. Experimentos tales como los telescopios terrestres CMB-S4 [25] y el satélite LiteBIRD [26] empezarán a operar durante la próxima década, reforzando aún más estas mediciones y posiblemente detectando otras características aún no vistas del CMB, como la huella dejada por las ondas gravitacionales primordiales en la polarización del mismo en forma de modos B.

Agujeros negros primordiales

Los PBHs son candidatos intrigantes para explicar la materia oscura. Desde el punto de vista teórico, los modelos de formación más simples no requieren ninguna adición al Modelo Estándar de física de partículas además de la inflación. Por otro lado, al ser objetos compactos, los mismos exhiben también una amplia variedad de señales astrofísicas [27]. Al momento de escribir esta tesis, los PBHs pueden constituir toda la materia oscura siempre que sus masas se encuentren en la ventana

$$10^{-16}\ M_\odot \lesssim M_{\rm PBH} \lesssim 10^{-11} M_\odot. \tag{4}$$

El límite superior de este rango de masas se debe a las observaciones de microlentes gravitacionales utilizando la cámara HSC del Telescopio Subaru [28], mientras que el límite inferior se debe al hecho de que los agujeros negros con masas muy pequeñas se estarían evaporando hoy (o se habrían evaporado ya si sus masas fueran lo suficientemente pequeñas), emitiendo radiación de Hawking y conduciendo a la observación de rayos γ extragalácticos [29, 30] y galácticos [31, 32], siendo la cota más fuerte ésta última, impuesta por el satélite INTEGRAL. Existen también otros límites en el extremo inferior de este rango, como los que surgen del estudio de la inyección de energía electromagnética en el CMB [33, 34], la

inyección de rayos cósmicos de electrones y positrones en la galaxia debido a la radiación de Hawking, que sería observada por la sonda Voyager 1 [35] y la aniquilación de los positrones que contribuyen a la línea galáctica de 511 keV [36–38]. En esta tesis estaremos interesados solo en el rango de masas mostrado en (4), pero señalamos que los PBHs con masas más altas están limitados por las observaciones de microlentes de EROS/MACHO [39] y OGLE [40], por el efecto que tendría su acreción en el CMB [41, 42] y por la no observación del fondo estocástico de ondas gravitacionales por el interferómetro LIGO [43].

Anteriormente existía un límite debido a observaciones de femtolentes gravitacionales producidos por ráfagas de rayos gamma [44] en el rango de masas (4), pero el mismo fue disputado en [45], donde se argumentó que la mayoría de las ráfagas de rayos gamma son demasiado grandes para ser utilizadas como lentes. En [45] también se consideró la posibilidad de utilizar un subconjunto de ráfagas con tamaños apropiados, aunque un gran número de ellas tendría que ser detectado por experimentos futuros para hacer de esta una herramienta de observación viable. Otros límites previamente existentes en este rango de masas incluyen la explosión de enanas blancas [46] debido a la fusión termonuclear causada por la fricción generada por el tránsito de un agujero negro suficientemente masivo a través de la estrella y la acreción y posterior destrucción de una estrella de neutrones por agujeros negros [47]. Estos límites fueron discutidos en [48], donde se derivó la tasa de captura de agujeros negros por estrellas de neutrones[5] y se realizaron simulaciones hidrodinámicas para estimar la temperatura y la escala de tiempo del choque producido por un agujero negro que pasa a través de una enana blanca. Estos resultados determinaron que las cotas eran menos efectivas de lo que se pensaba anteriormente.

La relación entre el paradigma inflacionario y la formación de PBHs no debería de sorprender al lector ahora. En los mecanismos de formación de PBHs más simples y estudiados, la inflación proporciona las semillas para las grandes fluctuaciones de densidad que desencadenan el colapso gravitatorio de la materia en agujeros negros [50] en etapas posteriores. Las fluctuaciones de densidad observadas en el CMB resultan ser demasiado pequeñas para producir una cantidad significativa de agujeros negros, por lo que para que los PBHs constituyan una fracción considerable de la materia oscura observada, el tamaño de dichas fluctuaciones debe de incrementar a escalas de distancia que son mucho más pequeñas que las observadas por los experimentos del CMB. Este incremento de las fluctuaciones de densidad se puede lograr en modelos de inflación de un solo campo si el inflatón atraviesa una región de su potencial con un punto de inflexión aproximado, lo cual desencadena una fase de rodamiento muy lento (USR, por sus siglas en inglés) [51] durante la cual aumenta el espectro primordial $\mathcal{P}_{\mathcal{R}}$ de perturbaciones de curvatura. Como veremos más adelante, también se pueden producir grandes fluctuaciones de densidad a partir de una fase disipativa transitoria durante la inflación debido a la interacción entre el inflatón y un baño subyacente de radiación en equilibrio térmico que a su vez conduce a una fuente de ruido estocástico en la ecuación

[5] Véase también [49] para más trabajo en esta dirección.

de movimiento para las fluctuaciones. El incremento del espectro de potencias también se puede obtener en modelos híbridos de inflación de dos campos, véase [52] para la primera implementación de esta idea. A lo largo de esta tesis nos ocuparemos únicamente de los mecanismos inflacionarios de formación de PBHs, pero recalcamos que también se han propuesto otras alternativas no inflacionarias, tales como el colapso de los defectos topológicos [53–57], la colisión de burbujas de vacío [58, 59], el colapso de bolas de Fermi [60], el aumento en la probabilidad de colapso gravitacional durante una transición de fase debido a un cambio en la ecuación de estado del universo [61] y el colapso de una gran acumulación de partículas en una burbuja de vacío falso que se encoge [62].

Desviemos ahora nuestra atención a los modelos antes mencionados de producción de PBHs con un punto de inflexión. La primera implementación de este mecanismo fue presentada en [63],[6] pero el interés en la idea permaneció en gran medida latente hasta la primera detección de ondas gravitacionales por parte de LIGO [65]. Desde entonces, el campo ha experimentado un interés renovado y ha surgido una gran cantidad de modelos propuestos. Una de las primeras versiones modernas de la idea del punto de inflexión se encuentra en [66], donde se propuso un potencial basado en el cociente de dos polinomios. Poco después se exploró en [67, 68] la posibilidad de que el punto de inflexión en el potencial pudiera tener un origen radiativo. En [68] se asumió que el inflatón tenía un potencial polinómico dominado por el término cuártico y el punto de inflexión surge entonces del ajuste preciso de los coeficientes de las correcciones cuánticas logarítmicas. Además, para ajustar los datos del CMB, [68] consideró un acoplamiento no mínimo completamente general entre el inflatón y el escalar de Ricci, que aplana el potencial a valores grandes del campo. La posibilidad de obtener el punto de inflexión ajustando cuidadosamente los coeficientes de un potencial polinómico sin necesidad de recurrir a correcciones cuánticas también fue estudiada en [68]. Algunos ejemplos más exóticos de modelos son aquellos motivados por la teoría de cuerdas y la supergravedad, tales como los potenciales axiónicos [69], o los potenciales construidos dentro del marco de modelos supersimétricos de atractores α [70].

Sin embargo, estos modelos no vienen sin desventajas. En general, el punto de inflexión no es una característica que surja naturalmente en estos potenciales, de modo que los mismos deben de ser construidos específicamente con este propósito. Por otro lado, para lograr el incremento del espectro primordial, generalmente se requiere un fuerte ajuste de los parámetros en el potencial. Existen dos razones para esto. La primera es que el tamaño del espectro de potencias es exponencialmente sensible a la disminución en la velocidad del inflatón cuando el mismo alcanza el punto de inflexión. Si los parámetros no se ajustan con cuidado, el inflatón o bien pasa sobre el mínimo local que suele preceder al punto de inflexión a alta velocidad (de modo que las fluctuaciones apenas incrementan y no se produce una cantidad significativa de PBHs), o se atasca en dicho mínimo local. Otro problema es que la velocidad

[6] Véase también [64], donde fue propuesto un potencial lineal con un cambio suave en la pendiente capaz de generar un incremento en el espectro, aunque no con el propósito de producir PBHs.

necesaria para producir una cantidad significativa de PBHs en un modelo dado a menudo está en conflicto con la velocidad necesaria para producir suficiente inflación para resolver los problemas de horizonte y planitud del CMB. La región en el espacio de parámetros que puede satisfacer todas estas restricciones suele ser muy estrecha. La segunda razón para el ajuste es que generalmente se supone que los PBHs se forman durante la era dominada por la radiación (RD, por sus siglas en inglés) que precede a la nucleosíntesis del Big Bang, de modo que la abundancia actual de PBHs depende exponencialmente del espectro de potencias y las pequeñas desviaciones del punto óptimo en en el que la fracción de PBHs como materia oscura es $\mathcal{O}(1)$, por lo tanto, dan como resultado abundancias órdenes de magnitud más pequeñas que las de interés para el problema de la materia oscura. En esta tesis veremos cómo algunos de estos problemas pueden aliviarse en el contexto de un potencial inflacionario específico inspirado en la monodromía axiónica.

Un efecto que no se tuvo en cuenta en los primeros modelos de producción de PBHs con un punto de inflexión es el de la difusión cuántica. En la inflación de un solo campo, generalmente asumimos que la parte homogénea del campo se comporta de manera clásica y solo tratamos las perturbaciones de forma cuántica. En el escenario estándar, estas perturbaciones se vuelven constantes y clásicas una vez que cruzan el horizonte de Hubble, determinando el espectro de potencias. Sin embargo, las fluctuaciones cuánticas podrían reaccionar sobre la trayectoria clásica del inflatón, una posibilidad que no puede ser manejada por la descripción anterior. En cambio, este efecto debe entenderse en el marco de la inflación estocástica [71], en el que las fluctuaciones de longitud de onda corta actúan como fuente para las de longitud de onda grande. Estas últimas se comportan como variables aleatorias clásicas y su efecto puede describirse mediante un conjunto de ecuaciones diferenciales estocásticas. Se sabe que en el régimen de rodamiento lento este procedimiento conduce a un $\mathcal{P}_\mathcal{R}$ igual al obtenido utilizando la teoría de perturbaciones estándar a nivel lineal (véase [72]), pero no está claro si este resultado se cumple o no si el inflatón pasa por una fase de USR (como en los modelos de punto de inflexión antes mencionados). En presencia de una fase USR, se ha especulado que la reacción de las perturbaciones podría potencialmente modificar la altura del pico de $\mathcal{P}_\mathcal{R}$, la ubicación del pico en sí, o incluso hacer más ancha la función de masa de los PBHs [68]. Por lo tanto, estudiar el efecto de la difusión cuántica es importante para la correcta determinación de la abundancia de los PBHs y es un tema que trataremos en esta tesis. Otra ventaja del formalismo de la inflación estocástica es que también permite el cálculo de las correcciones no gaussianas a la función de distribución de probabilidad de las fluctuaciones de curvatura primordiales, que pueden tener un efecto sobre la abundancia de los PBHs [73–76].

Ondas gravitacionales

El fondo estocástico primordial de ondas gravitacionales, que está potencialmente al alcance de detectores que comenzarán a operar en el futuro cercano (tales como LISA [77]) y es

comparable al CMB en las implicaciones que tendría para la cosmología, se ha considerado durante mucho tiempo como una poderosa herramienta para sondear la física del universo primitivo. Hay muchas razones para esto, una de las cuales es el hecho de que las ondas gravitacionales pueden sondear procesos y escalas de energía inaccesibles por otros medios, tales como transiciones de fase, defectos topológicos y el recalentamiento después de la inflación. Otra razón es que, usando las técnicas de la teoría de perturbaciones cosmológica, es posible mostrar que, a primer orden en las perturbaciones, los modos tensoriales de la métrica (es decir, las ondas gravitacionales) se desacoplan de otros grados de libertad y se propagan libremente, afectados solo por la expansión del universo y transmitiendo información sobre los procesos que las generaron [78].

El hecho de que la formación de PBHs requiera de un espectro de potencias escalar relativamente grande conduce a un incremento en los modos tensoriales de segundo orden de la métrica que surge como consecuencia de la presencia de términos cuadráticos en perturbaciones escalares de primer orden en las ecuaciones de Einstein [79–81] para las perturbaciones. Además, si los PBHs constituyen toda la materia oscura, entonces debería ser posible detectar el fondo estocástico de ondas gravitacionales inducido si los mismos se forman durante una era de radiación y, en particular, si sus masas se encuentran dentro de la ventana (4), entonces este fondo estocástico sería observable[7] por el experimento LISA [77, 82, 83]. Esto se debe al hecho de que, suponiendo una distribución de masas monocromática, la masa de un PBH que se forma durante una era de radiación debido al colapso inducido por grandes fluctuaciones escalares está relacionada con la frecuencia máxima del espectro de ondas gravitacionales correspondiente a través de la expresión [82, 83]

$$M_{\rm PBH} \simeq \gamma\left(3\times 10^{-15} M_{\odot}\right)\left(\frac{0.1{\rm Hz}}{f_{\rm peak}}\right)^{2}, \tag{5}$$

donde γ es un factor de eficiencia $\mathcal{O}(1)$. A partir de esta ecuación se puede ver que las frecuencias que LISA puede sondear corresponden a masas de PBHs en la ventana (4). No solo esto, sino que la amplitud de la señal también se encuentra aproximadamente cuatro órdenes de magnitud por encima de la curva de sensibilidad de LISA, como mostraremos luego.

La posibilidad de que las ondas gravitacionales pudieran ser inducidas a segundo orden en teoría de perturbaciones fue, hasta donde sabemos, sugerida por primera vez en [79], aunque la idea no ganó fuerza sino hasta que comenzó a considerarse como una posible forma de sondear la existencia de PBHs. La razón de esto es que la amplitud de estas ondas gravitacionales inducidas suele estar suprimida con respecto a la contribución de primer orden, por lo que solo cobran relevancia cuando existe un gran pico en el espectro de potencias escalar, como es el caso de los modelos inflacionarios de producción de PBHs antes mencionados.

[7] Sin embargo, actualmente no está claro que una señal producida por PBHs se pueda distinguir de otras fuentes.

Estas ondas gravitacionales se inducen no solo durante la inflación, sino también después, una vez que las perturbaciones vuelven a entrar en el horizonte, y por lo tanto contienen información sobre el mecanismo que las produjo y la historia cosmológica posterior.

Si queremos utilizar las ondas gravitacionales como un posible observable para sondear la existencia de PBHs, es necesario examinar cuidadosamente el cálculo de la señal para determinar si se puede confiar en los resultados. Se ha señalado recientemente que el cálculo habitual del fondo estocástico de ondas gravitacionales inducidas a segundo orden a partir de fluctuaciones escalares no es del todo sólido, ya que la perturbación tensorial de segundo orden de la métrica no es una variable invariante bajo cambios de coordenadas [84–89]. Esta dependencia del sistema de coordenadas se propaga a la cantidad observable de interés: la densidad de energía de las ondas gravitacionales. Recientemente se afirmó [86, 90] que, a pesar de los comentarios anteriores, el cálculo habitual de la densidad de energía de las ondas gravitacionales es válido más allá del orden dominante en las perturbaciones y esta cantidad es aproximadamente invariante bajo cambios de coordenadas siempre que se mida tarde en la era de radiación, ya que los escalares que generan los modos tensoriales de segundo orden decaen rápidamente con el tiempo en este caso, de modo que estos últimos se propagan linealmente y efectivamente se convierten en cantidades de primer orden [86]. Sin embargo, la densidad de energía de las ondas gravitacionales debe de ser una cantidad invariante bajo cambios de coordenadas que esté bien definida en todo momento, independientemente de cuándo se mida. La determinación de la expresión correcta para esta cantidad más allá del orden dominante en las perturbaciones sigue siendo un problema abierto, que exploraremos con más detalle en esta tesis.

Visión general de la tesis

En esta tesis consideramos dos implementaciones diferentes del mecanismo de producción de PBHs con un punto de inflexión. El primero consiste en hacer la elección más simple posible para el potencial del inflatón, esto es, un polinomio cuártico, junto con un acoplamiento no mínimo al escalar de Ricci [2, 68]. Si los coeficientes en el polinomio se eligen apropiadamente, esto conduce a la presencia de un punto de inflexión en el potencial. Como veremos, en este modelo el problema principal es que la inclinación del espectro de potencias a escalas del CMB se encuentra en tensión con los valores medidos experimentalmente si requerimos que se produzcan suficientes PBHs con masas en la ventana (4) para explicar toda la materia oscura. Mostraremos que este problema se puede resolver considerando extensiones muy simples y bien motivadas del modelo cosmológico ΛCDM (como la adición de las masas de los neutrinos), o agregando operadores de mayor dimensión (cuya presencia es de esperar de todos modos) al potencial. La segunda posibilidad que consideramos es más complicada y tiene como objetivo aliviar las desventajas antes mencionadas de los modelos de producción de PBHs con puntos de inflexión. Los aspectos clave en esta segunda configuración son un potencial inflacionario que naturalmente contiene varios mínimos locales y una época

temprana dominada por materia no relativista (eMD, por sus siglas en inglés) justo después de la inflación (que se puede obtener, por ejemplo, del recalentamiento perturbativo), durante la cual la probabilidad de colapso gravitatorio aumenta [50, 91, 92]. Encontramos que estos dos ingredientes también pueden conducir a la formación de una fracción considerable de PBH con masas en el rango (4) [1].

La razón por la que una era de eMD alivia algunos de los problemas de ajuste fino antes mencionados es que, debido a la ausencia de presión de radiación, es más fácil que una nube de materia colapse, por lo que el incremento del espectro de potencias requerido para que se formen los PBHs resulta ser órdenes de magnitud más pequeño que en el escenario de radiación estándar, incluso después de tener en cuenta el hecho de que el momento angular y la no esfericidad de la nube suprimen el colapso [92, 93]. Además, aunque la dependencia de la abundancia de los PBHs del espectro de potencias sigue siendo exponencial durante la era de eMD, su sensibilidad a pequeños cambios en la amplitud de las fluctuaciones que inducen el colapso resulta ser mucho más leve que en el escenario de RD. La formación de PBHs durante una fase de eMD ha sido ampliamente estudiada en la literatura, véase, por ejemplo, [91] (uno de los primeros trabajos), o [92–94] para actualizaciones recientes. Como mostraremos, para aprovechar al máximo el escenario de eMD (en el sentido de que se pueda conseguir una abundancia considerable de PBHs con un espectro de potencias lo más cercano posible a su valor a escalas del CMB), la temperatura de transición entre las eras de eMD y RD debe ser de orden $\lesssim 10^5$ GeV. En trabajos anteriores se ha llegado a conclusiones similares; en esta tesis mejoramos las estimaciones presentadas en [70, 93].

Hasta ahora nos hemos centrado exclusivamente en los modelos de producción de PBHs con un punto de inflexión. Sin embargo, como mencionamos anteriormente, existen otras posibilidades. En particular, los agujeros negros podrían generarse durante una fase disipativa transitoria durante la inflación, un mecanismo que es fundamentalmente diferente de los modelos de punto de inflexión discutidos anteriormente. El efecto de la disipación en la dinámica del inflatón se puede describir introduciendo un término de fricción $\propto \Gamma\dot{\phi}$ en su ecuación de movimiento [95, 96]. Uno podría conjeturar que este término podría ralentizar al inflatón de la misma manera que lo hace un punto de inflexión en el potencial, conllevando a su vez un incremento del espectro. Como mostraremos, este argumento no es correcto, ya que, aunque la dinámica del inflatón es similar en ambos escenarios, las ecuaciones de movimiento de las fluctuaciones son bastante diferentes. Sin embargo, como mencionamos anteriormente, el espectro de potencias puede incrementar debido a la presencia de una fuente estocástica de ruido térmico para las fluctuaciones cuyo origen yace en la interacción entre el inflatón y un baño de radiación en equilibrio térmico subyacente. Asumimos que esta fuente está activa solo por un período corto para producir PBHs con masas adecuadas en el rango (4). Modelamos nuestro escenario utilizando un enfoque fenomenológico, suponiendo que el coeficiente disipativo Γ tiene un pico localizado y una dependencia de la temperatura con una motivación teórica sólida. Realizamos un análisis numérico completo de la dinámica resolviendo

directamente las ecuaciones diferenciales estocásticas para las perturbaciones. Integramos el sistema para múltiples realizaciones del ruido térmico, obteniendo así un histograma para $\mathcal{P}_{\mathcal{R}}(k)$ en cada k, permitiéndonos determinar sus propiedades estadísticas. Sin embargo, la cantidad más interesante no es necesariamente la distribución de probabilidad completa para el espectro de potencias, sino su valor esperado, que es una cantidad determinista. Mostramos que este valor esperado se puede calcular reformulando el sistema estocástico de ecuaciones diferenciales para las fluctuaciones como una única ecuación diferencial matricial determinista para sus funciones de correlación a dos puntos.

Además de la exploración de los modelos específicos de formación de PBHs discutidos anteriormente, en esta tesis también reevaluamos el cálculo del espectro de potencias primordial, en presencia de una fase de USR, usando el marco de la inflación estocástica, tanto analítica como numéricamente [3]. Encontramos que (a orden lineal) el espectro de potencias calculado usando la maquinaria del formalismo de inflación estocástica concuerda con el resultado obtenido utilizando la teoría de perturbaciones estándar, y confirmamos nuestros resultados numéricamente usando el modelo polinómico antes mencionado. Encontramos que la elección de la escala que separa a las fluctuaciones clásicas de las cuánticas es fundamental, ya que en presencia de una fase de USR, las perturbaciones tardan más en clasicalizarse y, por lo tanto, la elección estándar realizada en el escenario inflacionario de rodamiento lento usual ya no es válida. No estudiamos correladores de orden superior para las perturbaciones en esta tesis, pero señalamos que el hecho de que el espectro de potencias no cambie a orden lineal no implica necesariamente que los efectos estocásticos sean irrelevantes para la abundancia de agujeros negros, ya que se espera que las no-gausianidades modifiquen también su probabilidad de formación [73, 74, 97].

Dado que los escenarios en los que se forman PBHs durante una era de eMD tienen ventajas significativas, vale la pena preguntarse si la señal de ondas gravitacionales en este caso cambia con respecto al escenario de radiación habitual. Específicamente, nos gustaría saber si el fondo estocástico resultante puede detectarse mediante experimentos futuros, como LISA y BBO/DECIGO. Uno podría conjeturar que este no es el caso, ya que, como se discutió anteriormente, el incremento del espectro de potencias escalar requerido para que los PBHs puedan dar cuenta de toda la materia oscura es mucho menor en este escenario. Sin embargo, mostraremos que debido al hecho de que las fluctuaciones escalares no decaen con el tiempo durante una era de eMD, el espectro de ondas gravitacionales aumenta en amplitud a medida que disminuye la temperatura de transición entre las eras de eMD y RD. Este incremento adicional contrarresta la amplitud relativamente pequeña del espectro de potencias escalar, lo que hace que el fondo estocástico de ondas gravitacionales sea observable por LISA. También examinamos si el espectro resultante satisface los límites sobre la abundancia actual de ondas gravitacionales $\Omega_{\rm GW}$ que surgen de las mediciones del CMB [98] y la abundancia de elementos ligeros producidos durante la nucleosíntesis del Big Bang. [99] y encontramos que esto restringe el escenario de eMD eliminando parte del espacio de parámetros disponible. La

porción del espacio de parámetros que se descarta depende de la temperatura de transición, pero como veremos, para ciertos valores de esta cantidad, este límite puede ser incluso más fuerte que los que provienen de los experimentos de radiación de Hawking.

La dependencia de los modos tensoriales de segundo orden de la métrica del sistema de coordenadas se ha estudiado ampliamente tanto para universos dominados por materia no relativista como por radiación (véase, por ejemplo, [84, 85, 89]). Sin embargo, en el escenario de transición de eMD a RD el universo pasa por ambas fases, por lo que es importante incluir las dos en el cálculo. Este cálculo se presenta aquí por primera vez y encontramos que, debido al hecho de que la época de eMD tiene una duración finita, el conjunto de sistemas de coordenadas problemáticos coincide con el del caso de RD [85]. En particular, examinamos el problema considerando cuatro sistemas de coordenadas diferentes (newtoniano, curvatura uniforme, expansión uniforme y comovimiento), y encontramos que los resultados coinciden en a tiempos tardíos (es decir, al final de la era de RD) para los primeros tres. Postulamos que el problema de la dependencia de coordenadas se deriva de la suposición incorrecta de que las correcciones a orden subdominante en perturbaciones a la abundancia de ondas gravitacionales $\Omega_{\rm GW}$ pueden obtenerse simplemente sustituyendo[8] $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ en la fórmula habitual (cuya derivación puede ser encontrada en [100, 101], por ejemplo),

$$\Omega_{\rm GW}(\tau, k) = \frac{1}{6}\left(\frac{k}{\mathcal{H}}\right)^2 \langle \mathcal{P}_h(\tau, k)\rangle_{\rm W}, \tag{6}$$

donde $\mathcal{P}_h$ denota el espectro de potencias de $h_{ij}^{\rm TT}$, τ denota el tiempo conforme, $\mathcal{H} = aH$ el parámetro de Hubble conforme y los corchetes denotan un promedio sobre muchas longitudes de onda. De hecho, $\Omega_{\rm GW}$ debe ser una cantidad observable invariante bajo transformaciones de coordenadas, de modo que no puede expresarse en términos de $\boldsymbol{h}_{ij}^{\rm TT}$ solamente. En esta tesis proporcionamos una derivación heurística de la expresión completa para esta cantidad, válida en cualquier sistema de coordenadas e independiente del contenido energético del universo, usando un argumento basado en propiedades de simetría y análisis dimensional.

La tesis se estructura de la siguiente manera. En el Capítulo 1 revisamos los aspectos básicos de la inflación y la formación de agujeros negros primordiales en eras dominadas por materia no relativista y radiación. En el Capítulo 2 implementamos la idea de producir agujeros negros a través de un punto de inflexión en el potencial usando dos modelos inflacionarios diferentes, uno basado en un potencial polinómico y otro inspirado en la monodromía axiónica. En el Capítulo 3 exploramos el efecto que tiene la difusión cuántica del inflatón sobre el espectro de potencias de perturbaciones de curvatura. En el Capítulo 4 discutimos cómo se pueden formar agujeros negros debido a una fase disipativa transitoria durante la inflación. Finalmente, en el Capítulo 5 calculamos las señales de ondas gravitacionales en cada escenario y discutimos las posibilidades de observarlas en el futuro.

[8] Usamos $h_{ij}^{\rm TT}$ para denotar los modos tensoriales de primer orden, transversales y sin traza, y el símbolo en negrita $\boldsymbol{h}_{ij}^{\rm TT}$ para denotar modos de segundo orden.

CHAPTER 1

Black holes from cold inflation

In this introductory chapter we review the basic facts about inflation, reheating and the formation of primordial black holes. In Section 1.1 we briefly review how inflation provides a solution to the horizon and flatness problems of the cosmic microwave background and establish the conventions and notation we will use throughout the rest of the thesis. In Section 1.2 we compute the mass and abundance of these black holes depending on whether they form during a radiation or early matter-dominated era, and we determine the size of the fluctuations necessary for these primordial black holes to be able to account for the entirety of the observed dark matter. In Section 1.3 we show how an early period of matter domination can be obtained via perturbative reheating, and determine the number of e-folds of inflation necessary to solve the horizon and flatness problems. Finally, in Section 1.4 we solve the equation of motion for the comoving curvature perturbation analytically (by making a suitable set of assumptions), and review the mechanism by which the power spectrum is enhanced in the presence of an inflection point in the inflationary potential.

1.1 ■ Cold inflation

In this section we briefly review the basic facts about inflation. An extensive review can be found in [102]. We begin by assuming that the Universe is dominated by a perfect fluid with equation of state $p = w\rho$ and that spacetime is described by the Friedmann-Lemaître-Robertson-Walker metric, which is given in conformal time $d\tau = adt$ and in spherical coordinates by

$$ds^2 = a^2\left(-d\tau^2 + \frac{dr^2}{1-k_c r^2} + r^2 d\Omega^2 \right). \tag{7}$$

We refer to the constant k_c as the spatial curvature of the Universe. This is a number that must be determined experimentally, and is not fixed by any fundamental principles. The scale factor $a(\tau)$ is a function of time that describes how the size of a spatial patch changes with the expansion of the Universe. The evolution of a is determined by the Friedmann equations

$$H^2 = \frac{\rho}{3M_p^2} - \frac{k_c}{a^2}, \tag{8}$$

$$\dot{H} + H^2 = -\frac{1}{6M_p^2}(\rho + 3p), \tag{9}$$

which can be obtained from Einstein's equations and where $H = \dot{a}/a$ is the Hubble parameter. Throughout this thesis we use dots to denote derivatives with respect to cosmic time ($\cdot \equiv d/dt$) and primes to denote derivatives with respect to conformal time ($' \equiv d/d\tau$). The first of these equations can be recast as

$$\Omega - 1 = \Omega_k, \qquad \text{with} \qquad \Omega = \frac{\rho}{3M_p^2 H^2}, \qquad \Omega_k = \frac{k_c}{a^2 H^2}. \tag{10}$$

The quantity Ω_k is constrained to be $\Omega_k \ll 1$ by CMB experiments (see e.g. [23, 103]). This is a puzzling fact, since, as we will see momentarily, the quantity $(aH)^{-1}$, known as the comoving Hubble radius, increases in time for conventional matter sources (e.g. dust and radiation), so in order to satisfy the constraint we must have $k_c \ll 1$. This is known as the flatness problem. From now on we set $k_c = 0$, which makes the Friedmann equations invariant under a rescaling of a and therefore allows us to normalize the scale factor to unity today, $a_0 = 1$.

The particle horizon is the maximum distance a photon can travel from the Big Bang at $t = 0$ to some later time t (and therefore determines the size of the observable Universe),

$$\Delta r = \Delta\tau = \int_0^t \frac{dt'}{a(t')} \propto \frac{1}{aH} \propto a^q, \qquad q = \frac{1}{2}(1 + 3w). \tag{11}$$

The fact that $(aH)^{-1} \propto a^q$ can be shown by solving the Friedmann equations. As mentioned earlier, the scale factor a describes the evolution of the size of a spatial patch. A patch with size $R_p(\tau_0) = R_0$ today will have size $R_p(\tau) = a(\tau)R_0$ at some other time τ. We see from the above equation, however, that the particle horizon does not grow linearly with a, but rather as $\Delta r \propto a^q$. Since the particle horizon represents the size of a causally connected patch, we see that for conventional matter sources (e.g. radiation with $w = 1/3$ or dust with $w = 0$), a spatial patch that is causally connected today must have been disconnected at some point in the past. Nevertheless, measurements show that the CMB temperature is nearly uniform across the entire visible Universe (on sufficiently large scales). The question of how this equilibrium was achieved if the patch was causally disconnected at some point is known as the horizon problem.

The key observation now is that if the Universe had gone through a phase in which $(aH)^{-1}$ decreases, then at some point the spatial patch corresponding to the observable Universe would have been larger than the particle horizon (provided this phase was long enough), solving the horizon problem. Similarly, the quantity $\Omega_k \propto (aH)^{-1}$ is driven to a small value, solving the flatness problem. We refer to such a phase as *inflation*. From eq. (11) we see that such a period requires $w < -1/3$. The simplest way in which this condition can be satisfied is if the energy budget of the Universe is dominated by a scalar field slowly rolling down a potential. To see this, let us consider the corresponding Lagrangian for such

a field assuming it is minimally coupled to gravity

$$\mathcal{L} = \frac{1}{2}M_p^2 R + \frac{1}{2}(\partial_\mu \phi)^2 - V(\phi). \tag{12}$$

The pressure and energy density of the field are, assuming that the field is homogeneous,

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{13}$$

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{14}$$

Thus, if the kinetic energy is negligible compared to the potential, we have $w = p/\rho \simeq -1 < -1/3$, and therefore a period of inflation takes place. To quantify how negligible the kinetic energy is with respect to the potential, it is convenient to define

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \qquad \eta \equiv -\frac{1}{2}\frac{\dot{\epsilon}}{H\epsilon}. \tag{15}$$

These are known as the slow-roll parameters. It can be shown from Friedmann's equations that inflation takes place when $\epsilon < 1$. The regime in which $\epsilon \ll 1$ and $|\eta| \ll 1$ is known as slow-roll inflation. This inflationary period only lasts a finite amount of time, since eventually the field reaches the minimum of the potential and the potential and kinetic energy become comparable, leading to $\epsilon > 1$ and ending inflation. The field then proceeds to oscillate around the minimum of the potential, decaying into other fields and populating the Universe with radiation. This process is known as reheating.[9]

One of the core features of inflation is that it allows us to understand the CMB anisotropies. For this we need to make use of cosmological perturbation theory, reviewed in detail with our conventions in Appendix A. The field is expanded into a homogeneous background piece $\phi(t)$ and a space-dependent perturbation $\delta\phi(t, \boldsymbol{x})$. Similarly, we can expand the metric around the homogeneous FLRW background as[10]

$$ds^2 = -a^2\,(1+2\varphi)\,d\tau^2 + 2a^2\partial_i B dx^i d\tau + a^2\big[\,(1-2\psi)\,\delta_{ij} + 2\partial_i\partial_j E\big]dx^i dx^j. \tag{16}$$

If we work in Newtonian gauge ($B = E = 0$) and in the absence of anisotropic stress, we can use one of Einstein's equations to set $\varphi = \psi$. The energy density, pressure, and momentum

[9] This is not the only way in which inflation can end, however. Another possibility is that the inflation gradually transfers its energy to a thermal bath of radiation throughout its evolution. Eventually, the radiation energy density begins to dominate and inflation ends. This scenario is known as warm inflation, and will be studied in Chapter 4.

[10] Vector perturbations are not typically produced in single-field inflation (and in any case decay quickly with the expansion of the Universe), so we ignore them. We also neglect tensor perturbations for the time being, these will be treated in detail in Chapter 5.

perturbations for the Lagrangian (12) are, respectively,

$$\delta\rho = \dot\phi\dot{\delta\phi} - \dot\phi^2\psi + V_\phi\delta\phi, \tag{17}$$

$$\delta p = \dot\phi\dot{\delta\phi} - \dot\phi^2\psi - V_\phi\delta\phi, \tag{18}$$

$$\delta q = -\dot\phi\delta\phi, \tag{19}$$

and the remaining Einstein's equations are then, in Fourier space,

$$3H(\dot\psi_k + H\psi_k) + \frac{k^2}{a^2}\psi_k = -\frac{1}{2M_p^2}\Big(\dot\phi\dot{\delta\phi}_k - \dot\phi^2\psi_k + V_\phi\delta\phi_k\Big), \tag{20}$$

$$\dot\psi_k + H\psi_k = \frac{1}{2M_p^2}\dot\phi\delta\phi_k, \tag{21}$$

$$\ddot\psi_k + 4H\dot\psi_k + (2\dot H + 3H^2)\psi_k = \frac{1}{2M_p^2}\Big(\dot\phi\dot{\delta\phi}_k - \dot\phi^2\psi_k - V_\phi\delta\phi_k\Big). \tag{22}$$

The connection to CMB observables is most easily made by making use of the following gauge-invariant variable, known as the comoving curvature perturbation,

$$-\mathcal{R}_k \equiv \frac{H}{\rho+p}\dot\phi\delta\phi_k + \psi_k. \tag{23}$$

By combining Einstein's equations (20–22), we can find a second-order differential equation for this quantity

$$\ddot{\mathcal{R}}_k + (3-2\eta)H\dot{\mathcal{R}}_k + \frac{k^2}{a^2}\mathcal{R}_k = 0. \tag{24}$$

This is known as the Mukhanov-Sasaki equation, and will be studied in detail in Section 1.4. This equation is usually solved by imposing the so-called Bunch-Davies initial conditions,

$$\mathcal{R}_k = -\frac{e^{-ik\tau}}{2M_p a\sqrt{k\epsilon}}, \tag{25}$$

which can be obtained by asking that the state that minimizes the expectation value of the Hamiltonian in the far past coincides with the Minkowski vacuum [102]. The main quantity of interest for us will be the power spectrum of $\mathcal{R}$, defined via

$$\mathcal{P}_\mathcal{R}(k) \equiv \frac{k^3}{2\pi^2}|\mathcal{R}_k|^2\Big|_{k\ll aH}. \tag{26}$$

The power spectrum of $\mathcal{R}$ can be related to the temperature and density fluctuations observed in the CMB, so its behaviour on small scales $10^{-4}\text{Mpc}^{-1} \lesssim k \lesssim 0.5\text{Mpc}^{-1}$ is well-known. On these scales, the power spectrum is well described by the following nearly scale-invariant function,

$$\mathcal{P}_\mathcal{R}(k) \simeq A_s\left(\frac{k}{k_\star}\right)^{n_s-1}, \tag{27}$$

where $k_\star$ is some arbitrary reference scale, usually taken to be $k_\star = 0.05$ Mpc^{-1}. The amplitude and spectral index have been determined to be $A_s \simeq 10^{-9}$ and $n_s \simeq 0.965$ at this scale [23, 103]. As we will show in Section 1.4, a spectrum with these characteristics can be obtained within the framework of slow-roll inflation if the potential is chosen appropriately. This is one of the great successes of the inflationary paradigm. As we will see momentarily, the power spectrum is also the quantity that determines whether black holes can form.

1.2 ■ Primordial black holes

As discussed in the previous section, inflation provides the seeds for the density fluctuations we observe today. Had these fluctuations been large enough at a particular scale, they could have induced gravitational collapse of the matter in the surrounding region, leading to the formation of a black hole [50]. The masses of the resulting black holes, as we will show, would then be determined by the scale at which these fluctuations are large. In this section we review the calculation of the mass and abundance of these black holes in radiation and matter-dominated eras. A schematic representation of the evolution of the scale factor across the various phases of the cosmological history in the scenario we consider is shown in Fig. 1. To be clear, we consider PBH formation only during an eMD era, and not during the late MD era that follows Big Bang Nucleosynthesis.

The mass and abundance of the PBHs depend on the equation of state of the Universe when the wavenumber of the fluctuations that induce collapse becomes comparable to the Hubble radius after inflation. During RD the radiation pressure opposes the gravitational collapse, whereas during MD any overdensity grows since the pressure is zero. We use $t_m = 2/(3H_m)$ to denote the time at which the eMD era ends. For simplicity, we consider that the Universe thermalizes instantaneously at t_m and becomes radiation dominated. Given that in RD the Hubble expansion rate is $H = 1/(2t)$ and the energy density during this period is therefore $\rho = 3M_p^2/(4t^2)$, we can define the temperature T_m of the radiation bath at thermalization through $\rho = (\pi^2/30) g_\star(T) T^4$ as

$$T_m = \left(\frac{M_p}{t_m}\right)^{1/2} \left(\frac{4\pi^2 g_\star(T_m)}{90}\right)^{-1/4}, \tag{28}$$

where $g_\star = \sum_b g_{\star b} + (7/8) \sum_f g_{\star f}$ counts the effective number of degrees of freedom and the sums over b and f run, respectively, over the baryonic and fermionic species whose masses are below T at any given time. In the Standard Model, for temperatures above the electroweak phase transition one has $g_\star = 106.75$.

An early phase of MD after inflation can occur during perturbative reheating if the inflaton oscillates rapidly in a quadratic minimum [104]. In this case, we can identify the temperature T_m with the reheating temperature and we will often refer to it in this way. We will elaborate more on the connection between T_m and inflation in Section 1.3. However, it

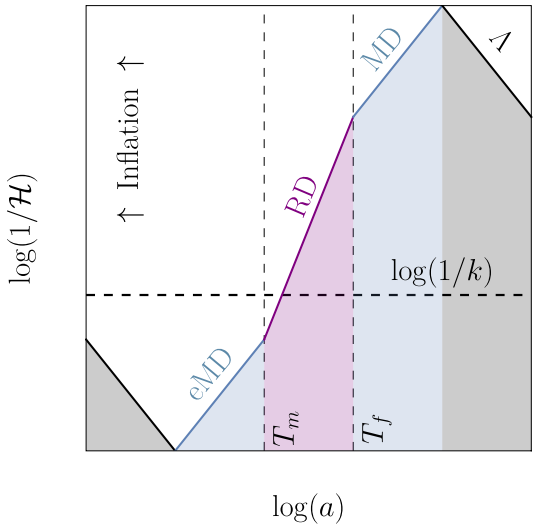

Figure 1: *Schematic depiction of the evolution of the conformal Hubble factor $\mathcal{H} \equiv aH$ in the scenarios we consider. T_m denotes the temperature at which the instantaneous transition between the early matter era (eMD) and the radiation era (RD) occurs. T_f denotes the temperature at the end of the RD era.*

is worth noting that there are other possibilities to realize a phase of eMD (e.g. oscillations of other heavy scalars) and our discussion until Section 1.3 does not depend on the origin of this phase.

As mentioned above, PBHs form from the collapse of regions with large density fluctuations when their spatial extension, characterized by some scale k, becomes comparable to the size of a Hubble patch. The mass of the individual PBHs is mainly given by the mass contained in this Hubble patch at the time of horizon re-entry for the scale k. This scale cannot be defined unambiguously; see [105] for a recent discussion. A common approximation for peaked spectra identifies k with the location of the peak of the primordial spectrum in linear perturbation theory. We will use this approximation, which is sufficient for our purposes. Then, the mass $M_{\rm PBH}$ of individual PBHs is

$$M_{\rm PBH} = 4\pi\gamma \frac{M_p^2}{H}, \tag{29}$$

where H is equal to $1/(2t)$ for RD after inflation and $2/(3t)$ for an eMD phase. The coefficient γ quantifies the efficiency of the collapse. Numerical analyses in the case of RD indicate that γ depends on the spectral shape of the density fluctuations and that the actual mass depends mildly on the density threshold that triggers the formation of a PBH [106]. We will neglect these dependencies and use $\gamma = 0.2$ for RD (see [50, 106]), which is accurate enough for our purposes. The actual efficiency of the collapse in MD is uncertain but may be expected to be

higher than in RD due to the absence of radiation pressure. For concreteness, we take $\gamma = 1$ for MD throughout this thesis, although we keep γ unspecified in most of the expressions below.

An overdensity of comoving scale k re-enters the Hubble horizon at time t_k, when the condition $k = a(t_k)H(t_k) \equiv a_k H_k$ is satisfied. If this occurs during an eMD phase which ends at time t_m, we can write: $a_k = (a_k/a_m)\,(a_m/a_0)$, where $a_m \equiv a(t_m)$ and a_0 is the scale factor today, which we normalize to one. Entropy conservation from t_m until today implies that $a_m/a_0 = (T_0/T_m)\,(g_{\star s}(T_0)/g_{\star s}(T_m))^{1/3}$, where T_0 is the current CMB temperature. During the phase of eMD there is no thermal equilibrium, but the scaling $a_k/a_m = (H_m/H_k)^{2/3}$ can be used. Combining these results with the condition for horizon crossing and using that $3M_p^2 H_m^2 = (\pi^2/30) g_\star(T_m) T_m^4$ to eliminate H_m, we obtain

$$a_k = \frac{\pi^2}{90} g_\star(T_m) \frac{g_{\star s}(T_0)}{g_{\star s}(T_m)} \frac{T_0^3\, T_m}{k^2 M_p^2}, \tag{30}$$

where we are keeping the number of effective entropy ($g_{\star s}$) and temperature ($g_\star$) relativistic degrees of freedom distinct and T_0 is the temperature of radiation today. This expression allows us to write the PBH mass of eq. (29) as

$$M_{\rm PBH} = \gamma \frac{2\pi^3}{45} \left(\frac{T_0}{k}\right)^3 \frac{g_{\star s}(T_0)}{g_{\star s}(T_m)} g_\star(T_m) T_m \quad \text{for eMD.} \tag{31}$$

If the PBHs form during RD, the expression for their mass can be obtained following a similar logic. In this case

$$a_k = \frac{\pi}{3\sqrt{10}} \left(\frac{g_{\star s}(T_0)}{g_{\star s}(T_k)}\right)^{2/3} \frac{T_0^2}{k M_p} \sqrt{g_\star(T_k)}, \tag{32}$$

and therefore

$$M_{\rm PBH} = \gamma \frac{4\pi^2}{3\sqrt{10}} \left(\frac{T_0}{k}\right)^2 \left(\frac{g_{\star s}(T_0)}{g_{\star s}(T_k)}\right)^{2/3} \sqrt{g_\star(T_k)} M_p \quad \text{for RD.} \tag{33}$$

The PBH mass thus scales as k^{-2} if PBHs form during RD and as k^{-3} if they form during eMD. In the latter case the PBH mass depends also on the duration of the eMD phase through the reheating temperature T_m, see eq. (28). For the purpose of comparison, it is useful to write both mass expressions in terms of some benchmark values for k, T_m, and the mass of the Sun ($M_\odot$) [1],

$$M_{\rm PBH} \simeq 2.8 \times 10^{-16} \left(\frac{\gamma}{0.2}\right) \left(\frac{g_\star(T_k)}{g_{\star s}(T_k)}\right)^{2/3} \left(\frac{106.75}{g_\star(T_k)}\right)^{1/6} \left(\frac{10^{14}\ \text{Mpc}^{-1}}{k}\right)^2 M_\odot \quad \text{for RD,} \tag{34}$$

$$M_{\rm PBH} \simeq 2.4 \times 10^{-17} \gamma \left(\frac{g_\star(T_m)}{g_{\star s}(T_m)}\right) \left(\frac{10^{14}\ \text{Mpc}^{-1}}{k}\right)^3 \left(\frac{T_m}{10^5\ \text{GeV}}\right) M_\odot \quad \text{for eMD.} \tag{35}$$

The expressions above have been obtained by setting $T_0 = 2.7255$ K [107], $g_\star(T_0) = 2.00$, and $g_{\star s}(T_0) = 3.91$. These values for the entropy and temperature degrees of freedom correspond to assuming that all three neutrinos are non-relativistic today, see [105].

We are interested in the current abundance of PBHs with respect to that of dark matter:

$$f_{\rm PBH} = \frac{\Omega^0_{\rm PBH}}{\Omega^0_{\rm DM}}. \tag{36}$$

In the approximation of rapid collapse of the overdensity, $f_{\rm PBH}$ can be written in terms of β, i.e. the ratio of the collapsing energy density to the total energy density at the time of the collapse:

$$\beta = \frac{1}{\gamma}\frac{\rho_{\rm PBH}(t_k)}{\rho(t_k)}. \tag{37}$$

Here $\rho_{\rm PBH}$ and ρ are the PBH and total energy densities, respectively.

In the case of PBHs formed during an eMD phase, by means of entropy conservation and using $\rho_m = (\pi^2 g(T_m)/30)T_m^4$, we obtain

$$f_{\rm PBH} = \gamma\beta\frac{\Omega^0_\gamma}{\Omega^0_{\rm DM}}\frac{g_\star(T_m)g_{\star s}(T_0)}{g_\star(T_0)g_{\star s}(T_m)}\frac{T_m}{T_0}, \tag{38}$$

where we have assumed that the transition between the different epochs depicted in Fig. 1 is instantaneous.

The analogous expression for RD is obtained from eq. (38) by simply setting $T_m = T_k$, where T_k is the temperature of the radiation at the time of formation. In this case we can write T_k as a function of the Hubble rate and relate this to the PBH mass through eq. (29). Then, the expressions for the PBH abundance as a function of the quantity β in the RD and eMD cases are, respectively [1]

$$f_{\rm PBH} \simeq \left(\frac{\gamma}{0.2}\right)^{3/2}\left(\frac{\beta}{8.9\times10^{-16}}\right)\left(\frac{g_\star(T_k)}{106.75}\right)^{-1/4}\left(\frac{g_\star(T_k)}{g_{\star s}(T_k)}\right)\left(\frac{M_{\rm PBH}}{10^{-15}\ M_\odot}\right)^{-1/2} \quad \text{(RD)}, \tag{39}$$

$$f_{\rm PBH} \simeq \gamma\left(\frac{\beta}{5.5\times10^{-15}}\right)\left(\frac{g_\star(T_m)}{g_{\star s}(T_m)}\right)\left(\frac{T_m}{10^5\ \text{GeV}}\right) \quad \text{(eMD)}. \tag{40}$$

The temperature dependence of the PBH abundance in the eMD case implies that a shorter duration of this phase (i.e. a higher reheating temperature) implies a larger abundance, see eq. (28). This is simply due to the fact that PBHs, being cold dark matter, dilute slower than radiation as the Universe expands. Therefore, the longer the duration of the RD phase is (i.e. the shorter the eMD phase is), the higher the abundance of PBHs. Notice that we could also write $f_{\rm PBH}$ in the eMD case as a function of the PBH mass, using eq. (35). However, unlike in the RD case of eq. (39) this introduces explicitly the wavenumber k, which makes the formula more cumbersome. We often assume that the distribution of PBH masses at formation time

is monochromatic, which is a good approximation for sufficiently peaked primordial spectra. If this condition is not satisfied, the total PBH abundance can be obtained by integrating $\int f_{\rm PBH}(M_{\rm PBH}) d\log(M_{\rm PBH})$.

So far we have obtained general expressions for the abundance as functions of the collapsing energy density fraction β, but we have not discussed how this quantity is computed in each case. In the RD case we use the approximation

$$\beta(k) = \frac{1}{\sqrt{2\pi\sigma^2(k)}} \int_{\delta_c}^{\infty} \exp\left[-\frac{\delta^2}{2\sigma^2(k)}\right] d\delta, \qquad \text{for RD.} \tag{41}$$

In this expression $\delta = \delta\rho/\rho$ is the density contrast in the total matter gauge (since this is the quantity that can be extracted from numerical simulations [108, 109]), δ_c is the estimate for the δ threshold for gravitational collapse during RD and $\sigma^2(k)$ is the variance of the density contrast smoothed over a comoving distance scale $\sim 1/k$, given by

$$\sigma^2(k) = \frac{4(1+w)^2}{(5+3w)^2} \int \frac{dq}{q} \left(\frac{q}{k}\right)^4 \mathcal{P}_{\mathcal{R}}(q) W^2(q/k)\,, \tag{42}$$

where $w = 1/3$ for RD. In this expression $\mathcal{P}_{\mathcal{R}}(k)$ is the dimensionless power spectrum of the comoving curvature perturbation $\mathcal{R}$ and $W(x)$ is a window function which we will take to be Gaussian. The expression (41) is obtained by applying the Press-Schechter formalism [110] and assuming that the primordial fluctuations leading to PBH formation are Gaussian. The Gaussian approximation will be sufficient for our purposes in this thesis, but we will have more to say on non-Gaussianities at the end of Section 3.3. The value of δ_c is known to depend on the profile of the collapsing overdensities, but should be between 0.4 and 0.5, see [111–113].

The physical interpretation of eq. (41) is transparent: only overdensities above the threshold δ_c can collapse into a black hole. For PBHs formed during eMD, the situation is very different, due to the absence of radiation pressure. In this case, eq. (42) still applies (now with $w = 0$) but, for sufficiently large variances, the collapsing energy fraction has been estimated to be[11] [91, 92]

$$\beta(k) \simeq 0.056\, \sigma^5(k), \quad \text{for eMD and } \sigma \gtrsim \sigma_{\rm ang}. \tag{43}$$

This formula was derived in [92] by applying the so-called Zel'dovich approximation [114] to model the nonlinear growth of density perturbations in the framework of Newtonian cosmology. The perturbations are assumed to follow a Gaussian distribution, and deviations from spherical symmetry are encoded in a deformation tensor whose entries are parameterized

[11] Inhomogeneities of the collapsing overdensity may further suppress PBH formation during eMD [94]. However, such extra suppression depends on certain assumptions on the final stages of the collapse, which might be evaded in realistic setups; see the discussion in [94]. For these reasons, we neglect inhomogeneities in our estimates and use eqs. (43) and (44) throughout this thesis.

by using the Doroshkevich distribution [115]. The non-sphericity of the collapsing cloud leads to a pancake collapse [114], so that the Schwarzschild criteria can no longer be used to determine whether a black hole forms, and it must instead by replaced by the hoop conjecture [116]. These ingredients lead to a complicated integral for the collapse probability distribution that must be performed numerically. The above expression is the result of a numerical fit to this integral and should therefore be regarded as a semi-analytical formula. This expression for β, which accounts for the effect of asphericities in the collapsing region, is valid only if σ is larger than a certain value[12] $\sigma_{\rm ang} \simeq 0.005$ [93]. Below this value, the effect of the angular momentum of the collapsing region becomes relevant, and the expression above must be replaced by [93]

$$\beta(k) \simeq 1.9 \times 10^{-7} f_q(q_c) \mathcal{I}^6 \sigma^2(k) \exp\left[-0.147 \left(\frac{\mathcal{I}^2}{\sigma(k)}\right)^{2/3}\right], \quad \text{for eMD, and } \sigma \lesssim \sigma_{\rm ang}. \tag{44}$$

This formula is obtained by applying the theory of angular momentum in structure formation developed in [117] to PBH collapse. The Kerr bound, which relates the mass of a spinning black hole to its angular momentum and determines the condition under which the black hole singularity is hidden by an event horizon, is adopted as the formation criteria, and the fluctuations are once again assumed to follow a Gaussian distribution. In the above formula, $\mathcal{I}$ is an $\mathcal{O}(1)$ parameter[13][93] and $f_q(q_c)$ is the fraction of mass with a level of quadrupolar asphericity q smaller than a threshold $q_c \simeq 2.4(\mathcal{I}\,\sigma)^{1/3}$. Following the estimates of [93] we will assume $f_q(q_c) = 1$ in our numerical examples.

From eqs. (41) and (44), it is clear that in order to obtain a sizeable abundance at a specific mass, the power spectrum $\mathcal{P}_\mathcal{R}$ of curvature fluctuations produced during inflation must be enhanced at the scales of interest (i.e. those unconstrained by current experiments). The simplest way to model the peaked power spectrum resulting from the kind of models we will consider is by using a Dirac delta,

$$\mathcal{P}_\mathcal{R}(k) = \mathcal{A}_\delta k_\delta \delta(k - k_\delta). \tag{45}$$

The variance of the density contrast, given in eq. (42), can be computed directly in this case. For a Gaussian window function, we obtain

$$\sigma^2(k) = \mathcal{A}_\delta \frac{8(1+w)^2}{\pi(5+3w)^2} \left(\frac{k_\delta}{k}\right)^4 \exp\left(-\frac{k_\delta^2}{k^2}\right). \tag{46}$$

This can be plugged directly into eqs. (39) and (40) to obtain the PBH abundance. The

[12] We remark that the numerical fits in eqs. (43, 44) fail around $\sigma_{\rm ang}$, thereby making β discontinuous at this value.

[13] The variance of the angular momentum $\langle \mathbf{L}^2 \rangle$ and σ are related through $\mathcal{I}$ via $45t\langle \mathbf{L}^2 \rangle^{1/2} \simeq 8\pi \mathcal{I}(ar)^5 \rho\sigma$, where r is the initial comoving radius of the overdensity and ρ is the homogeneous energy density of the eMD universe.

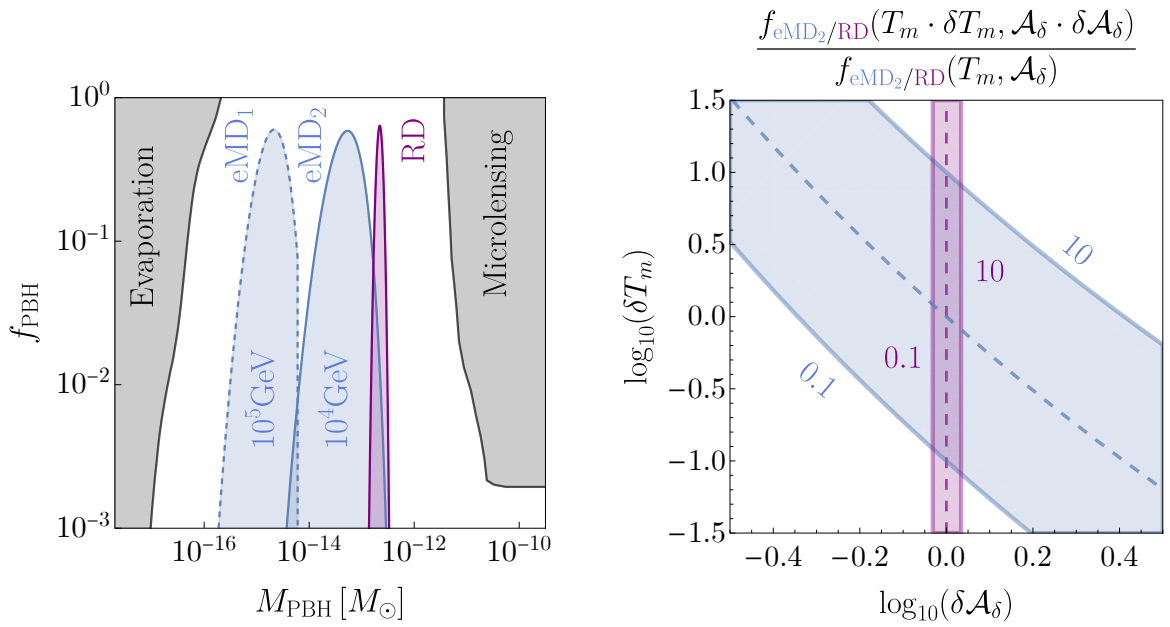

Figure 2: *Left panel: PBH abundance for the three examples shown in Table 1, together with galactic γ-ray constraint in [32] and HSC lensing [48, 118] constraints. Right panel: sensitivity of $f_{\rm PBH}$ to the transition temperature and the amplitude of the power spectrum in eq. (45) for the central values in the RD and eMD$_2$ examples. The shaded regions depict an order of magnitude change in the abundance, and it remains unchanged on the dashed lines.*

result is depicted in the left panel of Fig. 2 for three different parameter choices, shown in Table 1.

Example	A_δ	k_δ	T_m
eMD$_1$	0.00021	$10^{13.5}\text{Mpc}^{-1}$	10^5GeV
eMD$_2$	0.00046	$10^{12.7}\text{Mpc}^{-1}$	10^4GeV
RD	0.041	$10^{12.7}\text{Mpc}^{-1}$	∞

Table 1: *Parameter choices for three examples that yield an abundance $f_{\rm PBH}$ of $\mathcal{O}(1)$, depicted in Fig. 2. In RD we need $\mathcal{P}_\mathcal{R}\sim 10^{-2}$ to obtain $f_{\rm PBH}\sim\mathcal{O}(1)$, whereas in eMD we need $\mathcal{P}_\mathcal{R}\sim 10^{-4}$.*

The previous equations also summarize a key difference between PBH formation in eMD and RD. Whereas in RD the PBH abundance is exponentially sensitive to the primordial power spectrum $\mathcal{P}_\mathcal{R}$, in the eMD case this dependence is a power-law for fluctuations larger than $\sigma_{\rm ang}$. The reason for this difference lies in the threshold for gravitational collapse: in RD this is given by δ_c, while in eMD essentially any overdensity undergoes gravitational collapse, due to the very small Jeans length of non-relativistic matter.[14] Angular momentum effects need to be taken into account in eMD if σ is small enough. In that case, the approximate

[14]The Jeans length determines the critical radius above which an overdensity collapses and is proportional to the speed of sound of the fluctuations.

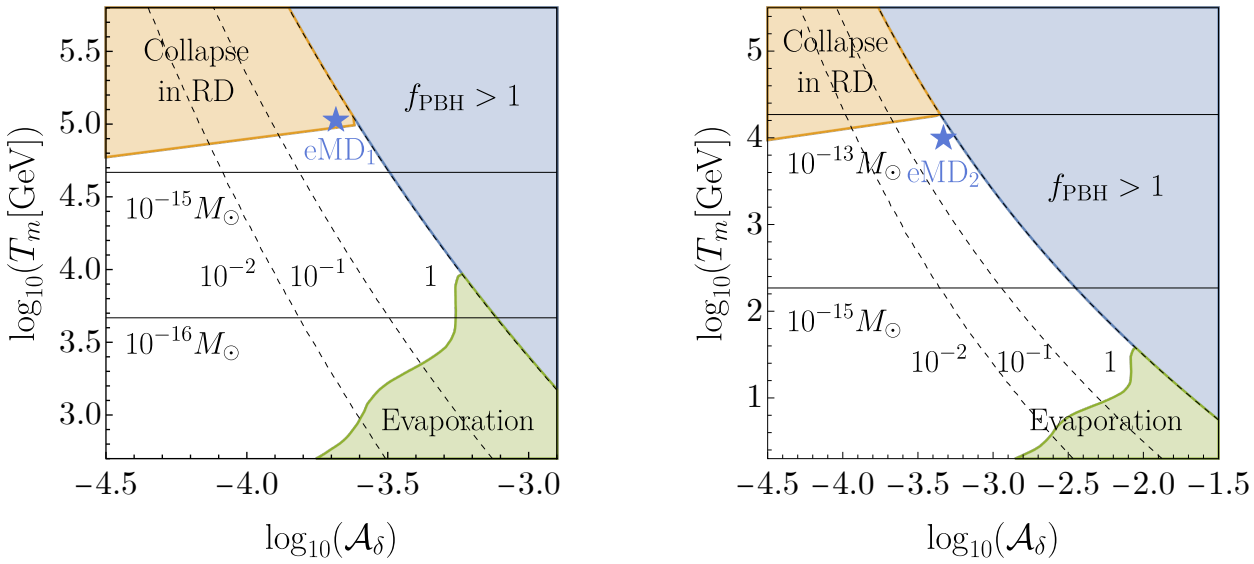

Figure 3: *Left panel: PBH abundances and masses (at the peak of the distribution), together with the constraint of eq. (47), and the galactic γ-ray constraint in [32], with k_δ fixed as in the eMD$_1$ example of Table 1. Right panel: same as left panel, for the eMD$_2$ example of Table 1. The values of T_m and $\mathcal{A}_\delta$ corresponding to each example are shown with a star. The eMD$_1$ example lies right at the edge of the constraint in eq. (47), which we remark should be understood as an order of magnitude estimate.*

power-law behavior of β is lost, but its sensitivity to the primordial power spectrum is still much milder than in RD. In the right panel of Fig. 2 we compare the sensitivity of β to changes in $\mathcal{P}_\mathcal{R}$. The PBH fraction changes much more dramatically with $\mathcal{P}_\mathcal{R}$ for PBHs formed during RD than in eMD. This translates into a higher level of tuning in the parameters of the inflationary potential if PBHs with $f_{\rm PBH} \sim 1$ form during RD.

Eqs. (41) and (43) should be supplemented with an additional constraint which arises from requiring that the collapsing fluctuation reaches the non-linear regime during the eMD era. This requirement significantly limits the available parameter space, which we show in Fig. 3. Using that the linear density contrast grows as the scale factor during matter domination, one finds that only fluctuations larger than $\sigma_{\rm nl} \simeq (H_m/H_k)^{2/3}$ reach non-linearity before the end of the eMD epoch. Smaller fluctuations take longer to reach the non-linear regime and thus do not complete the collapse during eMD. This constraint can be conveniently expressed as follows [92]

$$\sigma \gtrsim 1.9 \times 10^{-4} \left(\frac{g_\star(T_m)}{106.75}\right) \left(\frac{106.75}{g_{\star s}(T_m)}\right)^{2/3} \left(\frac{T_m}{10^5 \mathrm{GeV}}\right)^2 \left(\frac{10^{14} \mathrm{Mpc}^{-1}}{k}\right)^2 . \tag{47}$$

When the rotation of the collapsing fluctuation plays a role (i.e. for $\sigma \lesssim \sigma_{\rm ang}$) a different constraint arises. This can be understood as follows: during the linear evolution of an overdensity, its angular momentum grows; in particular, the longer the duration of the linear evolution, the stronger will be the effect of angular momentum on the gravitational collapse. Therefore, a different lower bound on σ arises from requiring that the growth of angular momentum does not prevent PBH formation. The resulting constraint is $\sigma \gtrsim 5H_m/(2\mathcal{I}H_k)$ –we refer the reader to [93] for its derivation– and can be rewritten as

$$\sigma \gtrsim 10^{-5} \left(\frac{1}{\mathcal{I}}\right) \left(\frac{g_\star(T_m)}{106.75}\right)^{3/2} \left(\frac{106.75}{g_{\star s}(T_m)}\right) \left(\frac{T_m}{10^5 \text{GeV}}\right)^3 \left(\frac{10^{14}\text{Mpc}^{-1}}{k}\right)^3 . \tag{48}$$

We remark, however, that these bounds (especially the latter) should be understood as an order of magnitude estimate and not a hard cutoff. The precise determination of the effect of angular momentum on the collapse of PBHs requires further study, in particular from the point of view of numerical simulations. In what follows we restrict our attention to the constraint in eq. (47), which is slightly stronger than the one above for the parameters we are interested in.

The advantage of considering an eMD phase for PBH formation is reflected in Fig. 3, where we plot the PBH masses according to (35) as well as the abundance $f_{\rm PBH}$ obtained by combining eqs. (40) and (44), which is the appropriate expression for β in the region of parameter space where $\sigma \lesssim \sigma_{\rm ang}$.[15] We also show the observational bounds on PBH masses from Hawking evaporation [32] as well as the constraint (47). In the viable region of parameter space, the minimal amplitude of the primordial power spectrum which is needed to obtain $f_{\rm PBH} = 1$ is $\mathcal{P}_\mathcal{R} \gtrsim 10^{-4}$ (for reheating temperatures $T_m \lesssim 10^5$ GeV). This is five orders of magnitude above the value at CMB scales and two orders of magnitude smaller than the benchmark value $\mathcal{P}_\mathcal{R} \gtrsim 10^{-2}$ required in RD. We remark that the reason that the most interesting range of reheating temperatures is $T_m \lesssim 10^5$ GeV is that for higher reheating temperatures the constraint of eq. (47) comes into effect and the collapse occurs at least partially during RD, so that the probability that a black hole will form becomes strongly suppressed. The slightly weaker constraint of eq. (48) also kicks in around the same region, so the probability is further suppressed by the high angular momentum of the collapsing matter cloud. Black hole formation is still possible for higher reheating temperatures, but then the amplitude of the power spectrum required to obtain $f_{\rm PBH} \simeq 1$ would be higher than the $\mathcal{P}_\mathcal{R} \gtrsim 10^{-4}$ value quoted above and shown in Fig. 3. Only for reheating temperatures of order $T_m \lesssim 10^5$ GeV can we take full advantage of the eMD era. In addition, we remark that

[15] The analysis of [93] suggests that the value $\sigma_{\rm ang} \simeq 0.005$ should be taken as an order of magnitude estimate, rather than as a sharp threshold. In particular, effects of order higher than second in angular momentum may lower this value of $\sigma_{\rm ang}$ and extend the validity of (43). This would lead to slightly smaller values of $\mathcal{P}_\mathcal{R}$ being required for $f_{\rm PBH}$ smaller than one, which are more advantageous for PBH formation, but the constraint (47) would then disfavor a larger region of parameter space, leading to slightly lower reheating temperatures.

eqs. (34, 35), make clear the fact that the most interesting scales for the formation of PBHs of interest for the dark matter problem are around $k \sim 10^{14}$ Mpc^{-1}, since only for these scales do we obtain masses in the unconstrained window (1). As can be seen from the eMD$_1$ and eMD$_2$ examples of Table 1 and Fig. 3, the only effect of varying this scale is that the reheating temperature T_m required to take full advantage of the eMD era changes slightly.

1.3 ■ Perturbative reheating

In the previous section we found that intermediate reheating temperatures $T_m \lesssim 10^5$ GeV are required to take full advantage of an eMD phase and produce PBHs with masses in the unconstrained window (1) and $f_{\rm PBH} \sim 1$ (see Fig. 3 and the relevant discussion). While we did not specify the origin of the eMD phase, reheating provides a straightforward option. In this section we briefly describe how an early period of matter domination can be obtained from perturbative reheating (or, more generally, from having a scalar field oscillate around the minimum of a quadratic potential [104]). We consider the Lagrangian for a minimally coupled scalar in eq. (12). A potential is generically quadratic if expanded around its minimum,[16]

$$V \simeq \frac{1}{2} m^2 \phi^2. \tag{49}$$

The field ϕ obeys the equation of motion

$$\ddot{\phi} + 3H\dot{\phi} + V_\phi = 0. \tag{50}$$

During this oscillating phase the field trajectory can be parameterized as [119]

$$\phi(t) = \Phi(t) \sin(mt), \tag{51}$$

where the envelope function $\Phi(t)$ encodes the information about the redshift and decay of the inflaton. By averaging the energy density (13) and pressure (14) of the field over one oscillation, we obtain

$$\langle \rho \rangle = \frac{1}{2}\langle \dot{\phi}^2 \rangle + \frac{1}{2} m^2 \langle \phi^2 \rangle \simeq \frac{1}{2} m^2 \Phi(t)^2, \tag{52}$$

$$\langle p \rangle = \frac{1}{2}\langle \dot{\phi}^2 \rangle - \frac{1}{2} m^2 \langle \phi^2 \rangle \simeq 0, \tag{53}$$

where we have assumed that the timescale over which $\Phi(t)$ varies is much smaller than the period of one oscillation. We therefore find that a field oscillating in a quadratic potential indeed behaves as presureless dust. If the main contribution to the energy budget of the

[16] There are of course exceptions to this (for instance, if the potential is a monomial ϕ^k with $k \neq 2$, or if there is some symmetry preventing the appearance of a quadratic term), but generically this is the leading term in a series expansion around the minimum.

Universe during this period of reheating is due to ϕ, the Universe is effectively matter-dominated.

A perturbative description of reheating such as the one presented above is not always appropriate due to the occurrence of preheating [119]. This is the process by which explosive particle production takes place as a consequence of parametric resonance driven by the inflaton oscillations. The occurrence of preheating typically quenches the existence of a prolonged phase of eMD [120]. Therefore, it is worth stating the conditions under which preheating is prevented. We illustrate this by considering a coupling of the form $\mu\phi\chi^2$ between the inflaton ϕ and another scalar field χ. The envelope function $\Phi(t)$ in eq. (51) is given by $\Phi(t) = M_p/(\sqrt{3\pi}mt)$, with $mt \simeq n/2\pi$, where n is the number of accumulated oscillations at time t [119]. By studying the growth of χ fluctuations, efficient preheating can be shown to occur as long as both of the following conditions are fulfilled [119]:

$$\mu \lesssim 32\Phi(t), \quad \text{and} \quad 4\mu\Phi(t) \gtrsim \sqrt{m^3 H}. \tag{54}$$

Violating one of these two conditions is enough to ensure that preheating does not happen. The first of these inequalities will be satisfied initially (for small n) provided that $\mu \ll M_p$. However, the second condition can be violated if μ is sufficiently small, preventing preheating from occurring at this stage. Once n is large enough, both conditions are violated and thus preheating never occurs.

Let us therefore impose that preheating does not occur and derive the value of μ such that $T_m \lesssim 10^6$ GeV can be achieved. For the reheating channel under consideration, the perturbative decay rate of ϕ into χ is given by $\Gamma = \mu^2/(8\pi m)$. In the absence of preheating, perturbative reheating proceeds until $H \sim \Gamma$, when the energy density stored in the inflaton is approximately $3M_p^2\Gamma^2$. Equating this to the energy density of the radiation bath (under our assumption of instantaneous transition between eMD and RD), we get

$$T_m = \sqrt{\Gamma M_p}\left(\frac{\pi^2 g_\star(T_m)}{90}\right)^{-1/4} \simeq 10^5 \text{ GeV}\left(\frac{\mu}{10^3 \text{ GeV}}\right)\left(\frac{10^{13} \text{ GeV}}{m}\right)^{1/2}. \tag{55}$$

Therefore, $T_m \sim \mathcal{O}(10^5)$ GeV is achieved if $\mu \sim 10^{-16} M_p$ for typical values of the inflaton mass in high scale models, $m \sim 10^{13}$ GeV. It is straightforward to check that for such small values of μ the second of the conditions (54) is violated, which shows that the estimate (55) is consistent with the assumption of inefficient preheating.

It remains to be seen whether such small values of μ are feasible in concrete scenarios of inflation and reheating. Interestingly, in inflationary models inspired by string compactifications, such as the one we will consider in Section 2.3, the inflaton can be a modulus and thus couple only gravitationally to light degrees of freedom. In this case, the decay rate Γ is Planck-suppressed, its typical form for decay into a scalar pair being $\Gamma \sim m^3/(48\pi M_p^2)$ (see e.g. [121]). For $m \sim 5 \times 10^{-6} M_p$ this translates into $\mu \sim 10^{-11} M_p$. For these values of

μ, preheating is again avoided. However, reaching $T_m \lesssim 10^6$ GeV still requires an extra suppression of the decay rate. Model building possibilities in this direction can be found e.g. in [121], in the framework of the large volume scenario [122, 123] of moduli stabilization.

Let us relate the reheating temperature to the number of e-folds ($dN = Hdt$) between the time at which the largest observable scales left the horizon and the end of inflation. The number of e-folds elapsed from the moment a scale k satisfies $k = aH$ during inflation until its end is[17]

$$N(k) = 63.55 + \frac{1}{4}\log\frac{\Omega_r^0}{h^2} - \log\frac{k}{a_0 H_0} - \frac{1}{12}\log\frac{\rho_{end}}{\rho_m} + \frac{1}{4}\log\frac{\rho_k}{\rho_{end}} + \frac{1}{4}\log\frac{\rho_k}{(10^{16}\text{GeV})^4}, \quad (56)$$

where $H_0 = 100h$ km s^{-1}Mpc^{-1} and ρ_k is the energy density of the universe at $k = aH$ during inflation. The subscripts $_{end}$ and $_m$ refer to the end of inflation and the end of the period of eMD. The quantity Ω_r^0 is the current radiation density. We can relate ρ_m to T_m simply through

$$\rho_m = \frac{\pi^2}{30} g_\star(T_m) T_m^4. \quad (57)$$

Given a model of inflation, we can determine ρ_k and ρ_{end} and then use the last two equations to find out the required reheating temperature T_m for a specific value of $N(k)$. For instance, if we assume $10H_{end} \sim H_k \simeq 10^{13}$GeV, so that $10^2\rho_{end} \sim \rho_k \sim (10^{16}\text{GeV})^4$, we get

$$N(k) \simeq 49 - \log\left(\frac{k}{0.05\ \text{Mpc}^{-1}}\right) + \frac{1}{3}\log\left(\frac{T_m}{10^5\ \text{GeV}}\right), \quad (58)$$

where we have used the values of the cosmological parameters in [23, 103]. If the reheating temperature is $T_m \sim 10^5$ GeV, inflation lasts approximately 50 e-folds after fluctuations of wavenumbers comparable to the Planck fiducial scale ($k_\star = 0.05$ Mpc^{-1}) exit the horizon. As can be seen from eqs. (34, 35), the most interesting scales for PBH formation (regardless of whether they form during an eMD or RD era) are around $k \sim 10^{14}$ Mpc^{-1}, in the sense that it is only for these scales that we obtain masses in the unconstrained range (1) of interest for dark matter. According to eq. (58), these fluctuations should exit the horizon during inflation approximately 15 e-folds before the end of inflation. Before moving on, let us remark that the ~ 50 e-folds number quoted above refers only to the time elapsed between the moment $k_\star$ leaves the horizon and the end of inflation. The total duration of inflation is unconstrained.

[17] A similar expression was first given in [124]. We have followed an analogous derivation and chosen to write the numerical factors differently.

We have seen that to produce a population of PBHs of interest for dark matter we need a power spectrum of order[18] $\mathcal{P}_\mathcal{R} \sim 10^{-4} - 10^{-2}$ (depending on whether the black holes form during eMD or RD) at scales $k \sim 10^{14}$ Mpc^{-1}. As we discussed in Section 1.1, however, CMB observations have determined the power spectrum to have an amplitude $\mathcal{P}_\mathcal{R} \sim 10^{-9}$ on large distance scales ($k \lesssim 0.5$ Mpc^{-1}). This means that we need to introduce some mechanism to enhance the power spectrum on small scales. In this section we discuss how this can be accomplished in single-field inflation by solving the Mukhanov-Sasaki equation (24) analytically. Similar analytical models have also been studied elsewhere, e.g. in [126]. See also [109, 127, 128] for in-depth discussions of the dynamics leading to an enhancement of the power spectrum.

The Mukhanov-Sasaki equation (24) for $\mathcal{R}_k$ is, in terms of the number of e-folds,

$$\frac{d^2\mathcal{R}_k}{dN^2} + (3 - \epsilon - 2\eta)\frac{d\mathcal{R}_k}{dN} + \frac{k^2}{a^2H^2}\mathcal{R}_k = 0. \tag{59}$$

In the models we will consider, the condition $\epsilon \ll 1$ is always satisfied, so we can set $\epsilon \simeq 0$ for simplicity. We see from eq. (15) that this implies $\dot{H} \simeq 0$, so that H can be considered constant. The last simplifying assumption we make is to take η as a piecewise constant function. In particular this should reproduce the correct spectrum in the slow-roll (SR) regime, where $\eta \simeq 0$. As we will see, in the models we will consider in Chapter 2, the slow-roll regime holds initially, but then η reaches a large, positive value once the field encounters the inflection point in the potential. The reason for this will be explained at the end of this section. We refer to this period in which $\eta \gg 1$, which lasts a few e-folds at most, as an ultra-slow-roll (USR) phase. Finally, once this period is over, η relaxes back to a negative value, so that ϵ starts growing and inflation eventually ends. The evolution of η we consider is illustrated in the left panel of Fig. 4. The value of η on each phase is denoted by η_i. Since H can be considered constant and $a \propto e^N$ by definition, we have

$$aH = a_\star e^{N-N_\star} H, \tag{60}$$

where $N_\star$ is an arbitrary reference time, which we choose as the time at which the SR phase ends and the USR phase begins. We can set $N_\star = 0$ and normalize $a_\star = 1$. The solution to

[18] One of the main assumption underlying these estimates is that the primordial perturbations follow a Gaussian distribution. As we have mentioned already, non-Gaussianities could potentially alter these numbers. However, we remark that, as pointed out in [125], due to the exponential sensitivity of the PBH abundance on the power spectrum, the effect of non-Gaussianities can be countered by simply multiplying the power spectrum by a factor of $\mathcal{O}(1)$. We therefore expect that these numbers are robust as an order of magnitude estimate.

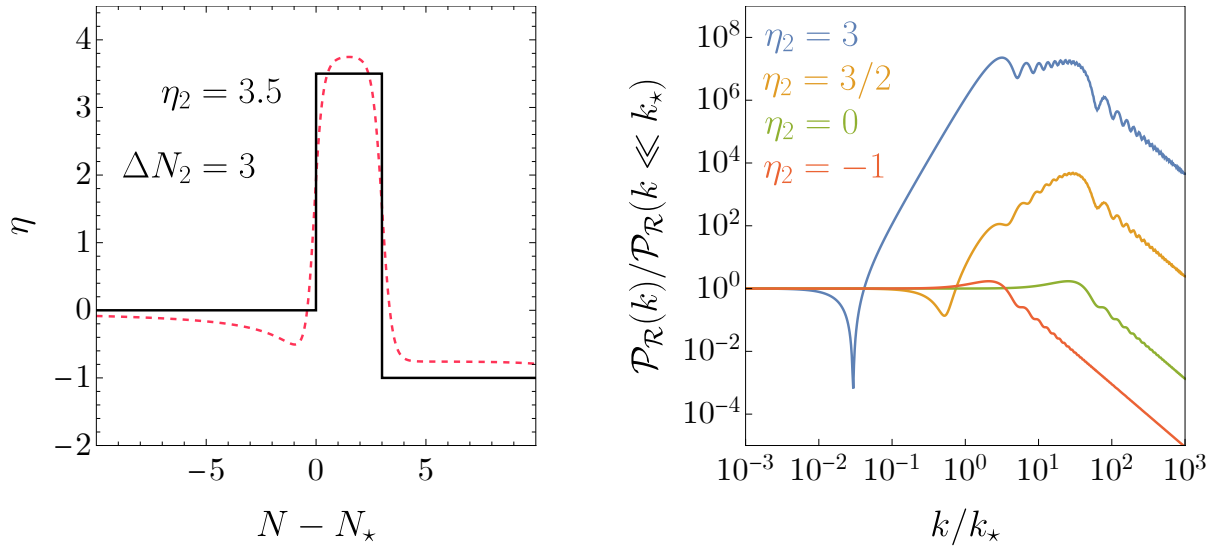

Figure 4: *Left panel: simplified evolution of the slow-roll parameter η we consider in this section (black, solid) compared to the typical evolution of η in numerical models of PBH production from an inflection point (red, dashed). Right panel: examples of different spectra computed with eq. (67). The reference scale $k_\star$ has been chosen as the scale that crosses the horizon at the time at which the USR phase begins. We set $\eta_1 = 0$, $\eta_3 = -1$, and the duration of the USR phase as $\Delta N_2 = 2.5$ for all examples in this panel.*

eq. (59) in the i-th phase is then, in terms of Bessel functions of the first kind,

$$\mathcal{R}_k^{(i)} = \frac{1}{M_p}\sqrt{\frac{\pi}{8H\epsilon_1}}e^{-(3-2\eta_i)N/2}\left[\alpha_i J_{(3-2\eta_i)/2}\left(ke^{-N}/H\right) + \beta_i J_{-(3-2\eta_i)/2}\left(ke^{-N}/H\right)\right], \tag{61}$$

where we have assumed that $\eta_1 = 0$ so that ϵ is constant in this phase, and we denote its value by ϵ_1. If modes are assumed to be in the Bunch-Davies vacuum in the far past –see eq. (25)– then the integration constants are, in the first phase,

$$\alpha_1 = -1, \qquad \beta_1 = -i. \tag{62}$$

The constants α_i and β_i in the subsequent phases can be found by imposing continuity of the solutions and their derivatives at the beginning of each phase.

An important fact is that, because aH grows exponentially during inflation, the last term in eq. (59) is negligible at late times. At this stage one of the two solutions to the equation becomes $\mathcal{R}_k \simeq$ constant and the other one decays quickly in time, so that the perturbation freezes once it crosses the horizon (a time defined by the condition $k = aH$) and is therefore unaffected by the subhorizon dynamics. This is the reason that, as we anticipated in Section 1.1, $\mathcal{R}_k$ is the preferred variable to make the connection between inflation and the CMB observables. As we will see momentarily, in the presence of a USR phase this fact remains

true, although perturbations no longer freeze shortly after crossing the horizon, but rather after the USR phase ends (a fact that will have important consequences when we discuss quantum diffusion in Chapter 3). The power spectrum of $\mathcal{R}_k$ is obtained by evaluating the curvature perturbation at some late time during the third phase, after it has frozen

$$\mathcal{P}_{\mathcal{R}}(k) = \lim_{N\to\infty} \frac{k^3}{2\pi^2} |\mathcal{R}_k^{(3)}(N)|^2. \tag{63}$$

The limit can be found by using the following identity

$$\lim_{y\to 0} y^{-m} J_m(y) = \frac{2^{-m}}{\Gamma_{\rm E}(m+1)}, \tag{64}$$

where $\Gamma_{\rm E}$ denotes Euler's Gamma function. Combining this with eq. (61), we have

$$\lim_{N\to\infty} \mathcal{R}_k^{(3)} \propto \lim_{N\to\infty} e^{-(3-2\eta_3)N/2} \Big[\alpha_3(k)\big(ke^{-N}/H\big)^{(3-2\eta_3)/2} + \beta_3(k)\big(ke^{-N}/H\big)^{-(3-2\eta_3)/2}\Big] \tag{65}$$

$$\propto \big(k/H\big)^{-(3-2\eta_3)/2} \beta_3(k), \tag{66}$$

up to some overall normalization factor which does not depend on k and is not important for our current purpose. Therefore,

$$\mathcal{P}_{\mathcal{R}}(k) \propto |\beta_3(k)|^2 \big(k/H\big)^{2\eta_3}. \tag{67}$$

In particular, if $\eta_1 = \eta_2 = \eta_3 = 0$, then $\beta_3 = i$ and the power spectrum is exactly scale invariant, as we anticipated at the end of Section 1.1. The explicit analytical expression for β_3 is quite cumbersome and not particularly illuminating, so we do not write it explicitly. Examples of spectra computed using eq. (67) for different parameter choices are shown in the right panel of Fig. 4. In these examples the power spectrum is approximately scale-invariant for small values of k, whereas for $k \simeq k_\star$ (where $k_\star = H$ is the reference scale, chosen as the scale that crosses the horizon at the time at which the USR phase begins) the spectrum is enhanced by several orders of magnitude.

When the power spectrum is enhanced, there is often a dip (which we define as the lowest point in the power spectrum) present before the peak. We will now show that there is a simple way to estimate both the size of the peak and the position of the dip in the power spectrum. To estimate the size of the peak notice that the solution to eq. (59) is, on superhorizon scales $k \ll aH$ where the last term can be neglected,

$$\mathcal{R}_k \simeq \mathcal{R}_{\rm cross} + \frac{d\mathcal{R}}{dN}\bigg|_{N_{\rm cross}} \int_{N_{\rm cross}}^{N} \exp\bigg[-\int_{N_{\rm cross}}^{N'} (3-2\eta) dN''\bigg] dN', \tag{68}$$

where we have assumed that the solution becomes valid around the time of horizon crossing $N_{\rm cross}$. In standard slow-roll inflation we have $\eta = 0$ and the second term in this expression

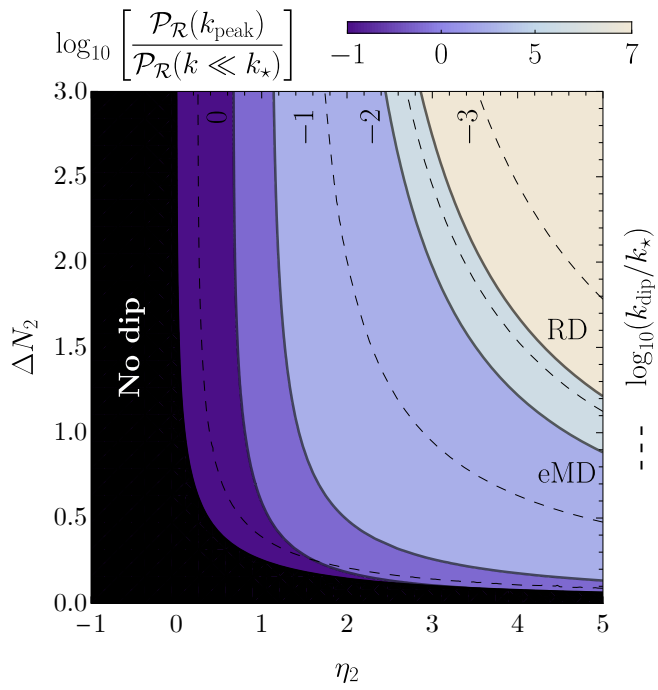

Figure 5: *Size of the peak (colored contours) and position of the dip (dashed lines) in the power spectrum, obtained by using the approximate expressions in eqs. (69) and (72), respectively, for the benchmark values $\eta_1 = 0$ and $\eta_3 = -1$. The enhancement of seven (five) orders of magnitude required to account for the observed dark matter abundance with primordial black holes if they form during a radiation-dominated (early matter-dominated) era is denoted by the label RD (eMD).*

is exponentially suppressed, so that only the first term, which is constant, contributes, and $\mathcal{R}_k$ freezes outside the horizon as we saw earlier. If $\eta > 3/2$, the second term will grow exponentially and quickly overcome the first one. This is the origin of the enhancement of the power spectrum. This enhancement therefore occurs only after the mode has left the horizon (on subhorizon scales, the solution in eq. (68) is not valid, and the mode simply oscillates in the Bunch-Davies vacuum). For constant η the integral can be computed explicitly and the size of the peak can be estimated via

$$\frac{\mathcal{P}_\mathcal{R}(k_{\rm peak})}{\mathcal{P}_\mathcal{R}(k \ll k_\star)} \simeq \left|\frac{\mathcal{R}_{\rm peak}}{\mathcal{R}_{\rm cross}}\right|^2 \simeq \left|\frac{2\eta_2 - 1 - i}{6 - 4\eta_2}\left(1 - e^{-(3-2\eta_2)\Delta N_2}\right)\right|^2, \tag{69}$$

where the prefactor comes from assuming that the Bunch-Davies solution (25) is still valid at horizon crossing and using it to determine the ratio $(d\mathcal{R}/dN)_{\rm cross}/\mathcal{R}_{\rm cross}$ in eq. (68). We have also used the fact that the mode $k_{\rm peak}$ that undergoes the greatest enhancement is the one that crosses the horizon just as the USR phase begins. The reason for this is that if the mode leaves the horizon before the first phase is over, it will be exponentially suppressed due to the second term of eq. (68). Modes that leave after the second phase has begun will

spend less time outside of the horizon, and the enhancement will therefore be smaller. This means that the peak will always be located roughly at

$$k_{\rm peak} \simeq k_\star, \tag{70}$$

where we remind the reader that we chose the reference scale $k_\star$ as the one that crosses the horizon as the second phase begins. The dip in the power spectrum occurs because as the second term in eq. (68) grows and overcomes the first one, a partial cancellation occurs between both. Thus, if the second term is never large enough (that is, whenever $\eta \lesssim 3/2$ or if the phase is too short), no dip will be present. Since the dip is always located before the peak (which is roughly at $k_\star$) its position $k_{\rm dip}$ can be obtained by expanding eq. (67) as a power series in $k/k_\star$ around $k = 0$ and finding the value $k_{\rm dip}$ at which the resulting expression vanishes at leading order in $k/k_\star$. We have

$$\left[\frac{\mathcal{P}_\mathcal{R}(k)}{\mathcal{P}_\mathcal{R}(k \ll k_\star)}\right]^{1/2} \propto \left|1 - \frac{k^2}{k_{\rm dip}^2} + \mathcal{O}\left[(k/k_\star)^3\right]\right|. \tag{71}$$

The resulting expression is,[19] for the benchmark parameters $\eta_1 = 0$ and $\eta_3 = -1$,

$$\frac{k_{\rm dip}^2}{k_\star^2} \simeq \frac{e^{(3-2\eta_2)\Delta N_2}}{a_{\rm dip} + b_{\rm dip} e^{(3-2\eta_2)\Delta N_2} + c_{\rm dip} e^{(1-2\eta_2)\Delta N_2}}, \tag{72}$$

where we have defined the coefficients,

$$a_{\rm dip} \equiv \frac{8\eta_2(\eta_2+1)}{3+4\eta_2(\eta_2-2)}, \qquad b_{\rm dip} \equiv \frac{10\eta_2}{3-2\eta_2}, \qquad c_{\rm dip} \equiv \frac{2\eta_2+2}{2\eta_2-1}. \tag{73}$$

Eq.(72) is plotted in Fig. 5. For parameters in the shaded region of this figure the right-hand side of eq. (72) becomes negative and thus there is no dip present in the spectrum.

As we have seen, a brief phase of $\eta > 3/2$ enhances the power spectrum. We mentioned earlier that such a phase can be achieved if the inflaton reaches a relatively flat region in the potential (in particular, it can happen if the potential has a near-inflection point). Let us make the connection between these two statements. If the potential is flat, the last term in the equation of motion (50) for the inflaton can be neglected, leading to

$$\ddot{\phi} + 3H\dot{\phi} \simeq 0. \tag{74}$$

On the other hand, by combining the Friedmann equations (8) and (9) we can show that

$$\dot{H} = -\frac{\dot{\phi}^2}{2M_p^2}. \tag{75}$$

[19] A similar expression was obtained in [126] for the particular case $\eta_2 = 3$.

This equation does not involve any approximations. By combining these two equations we find $M_p^2 \ddot{H} \simeq 3H\dot{\phi}^2$, which in turn leads to

$$\dot{\epsilon} = 2\frac{\dot{H}^2}{H^3} - \frac{\ddot{H}}{H^2} \simeq 2H\epsilon^2 - \frac{6}{H}\frac{\dot{\phi}^2}{2M_p^2} \simeq -6H\epsilon \qquad \longrightarrow \qquad \eta = -\frac{1}{2}\frac{\dot{\epsilon}}{H\epsilon} \simeq 3. \tag{76}$$

where we have neglected the small ϵ^2 term. We therefore find that if $V_\phi \simeq 0$, then $\eta \simeq 3$, leading to an enhancement of the power spectrum. This is often referred to as an ultra-slow-roll phase. As we have seen, however, the spectrum will be enhanced whenever $\eta > 3/2$, so throughout this thesis we use the term ultra-slow-roll to refer to this regime, even if $\eta \neq 3$. This implies, in particular, that the potential does not need to be exactly flat for the spectrum to be enhanced, i.e. an approximate inflection point suffices.

CHAPTER 2

Inflationary potentials

In this chapter we will study two different implementations of the inflection point mechanism of PBH production. The first model is studied in Section 2.1 and consists on making the simplest possible choice for the inflaton potential; namely, a polynomial, together with a non-minimal coupling to the Ricci scalar. We find that the main issue in this model is the mild tension that arises between the predicted tilt of the power spectrum and the current CMB measurements. We will show that this tension can be alleviated if simple extensions to the base ΛCDM model are considered, such as the addition of neutrino masses. In Section 2.2, we show that another way of ameliorating this tension is by adding higher-dimensional operators to the potential, the presence of which cannot be avoided. In Section 2.3 we consider a model inspired by axion monodromy inflation aimed at alleviating some of the fine-tuning issues present in most models of PBH production from an inflection point by considering a potential that naturally features several local minima, as well as the addition of an early matter-dominated era after inflation during which gravitational collapse is enhanced.

2.1 ■ Polynomial potential

As we saw in Section 1.4, a brief period of ultra-slow-roll in which the slow-roll parameter η reaches a large, positive value can enhance the power spectrum by several orders of magnitude. Such a period can be achieved if the inflaton reaches a relatively flat region in the potential, such as an inflection point. The first attempt to implement this mechanism was put forward in [63], where a potential with a plateau able to generate a large peak in the curvature power spectrum was considered. However, in subsequent years the idea remained largely dormant until the first detection of gravitational waves by LIGO [65], which revitalized interest in the field. Many proposals have been explored since then, from the possibility of obtaining the feature in the potential from quantum corrections [67, 68] to more exotic models such as axion-like potentials [69], or potentials arising from supersymmetric α-attractor models [70].

In this section we will be interested in the simplest inflationary model with this feature. The model we study here was first proposed in [68], where a potential characterized by a quartic polynomial was considered, together with a non-minimal coupling to the Ricci scalar. The non-minimal coupling can be eliminated by performing a field redefinition, which translates into the appearance of a polynomial dividing the potential. This model is very similar to one of the first modern takes on the inflection-point idea, namely, the one in [66],

where a potential based on the ratio of two polynomials was proposed. There are two crucial differences between both models. The first one is the origin of the dividing polynomial in the potential, which in [68] appears due to the field redefinition, whereas in [66] it was postulated from the start. The second one is the fact that, due to the aforementioned redefinition, the canonically normalized field that is identified with the inflaton is different in both models. These differences have a relevant phenomenological consequence: the examples provided in [66] lead to PBHs several orders of magnitude heavier, which are allowed only at the level of $f_{\rm PBH} \lesssim 0.1$ due to the microlensing bounds from the EROS project [39]. Another related work is [129], which differs from the one presented here in two respects: a running of the quartic coupling of the inflaton above a mass threshold is introduced, and the power spectrum is computed only in the slow-roll approximation. The radiative potential presented in [68] also resembles the one studied here, although in that case the polynomial is assumed to be dominated by the quartic term, and the inflection point arises from the tuning of logarithmic quantum corrections, rather than from the coefficients of the polynomial. A flattening of the potential at large field values is accomplished in that case also by considering a non-minimal coupling to the Ricci scalar.

Following the discussion in [68], let us consider the Lagrangian for a scalar field non-minimally coupled to gravity in the Jordan frame,

$$\mathcal{L} = -\frac{1}{2}(M_p^2 + \xi\phi^2)R + \frac{1}{2}g_{\mu\nu}\partial^\mu\phi\partial^\nu\phi - V(\phi). \tag{77}$$

There is no symmetry that forbids the coupling $\xi\phi^2 R$ from appearing in the Lagrangian, and thus its presence cannot be avoided.[20] We can get rid of this coupling by performing a field-dependent conformal transformation of the metric,

$$g_{\mu\nu} \to \left(1 + \frac{\xi\phi^2}{M_p^2}\right) g_{\mu\nu}. \tag{78}$$

The Lagrangian then becomes, in the so-called Einstein frame,

$$\mathcal{L} = -\frac{1}{2}M_p^2 R + \frac{1}{2}\left[\frac{1 + \xi(1+6\xi)\phi^2/M_p^2}{(1+\xi\phi^2/M_p^2)^2}\right] g_{\mu\nu}\partial^\mu\phi\partial^\nu\phi - \frac{V(\phi)}{(1+\xi\phi^2/M_p^2)^2}. \tag{79}$$

The new kinetic term of ϕ can be canonically normalized by means of the field redefinition

$$\frac{dh}{d\phi} = \frac{\sqrt{1 + \xi(1+6\xi)\phi^2/M_p^2}}{1+\xi\phi^2/M_p^2}. \tag{80}$$

[20] A non-minimal coupling to gravity of the form ϕR which is linear in the field ϕ can be eliminated by a field redefinition at the prize of redefining M_p in the Jordan frame.

The solution to this differential equation is

$$h = \sqrt{\frac{1+6\xi}{\xi}}\sinh^{-1}\left[(\phi/M_p)\sqrt{\xi(1+6\xi)}\right] - \sqrt{6}\tanh^{-1}\left[\frac{\sqrt{6}\xi(\phi/M_p)}{\sqrt{1+\xi(1+6\xi)(\phi/M_p)^2}}\right], \quad (81)$$

where we have used the boundary condition $h(\phi = 0) = 0$. The equation of motion for the new, canonically normalized field is

$$\frac{d^2h}{dN^2} + 3\frac{dh}{dN} - \frac{1}{2}\left(\frac{dh}{dN}\right)^3 + \left[3 - \frac{1}{2}\left(\frac{dh}{dN}\right)^2\right]\frac{d\log U}{dh} = 0, \quad (82)$$

where we denote by $U[\phi(h)] \equiv V[\phi(h)]/[1+\xi\phi(h)^2/M_p^2]^2$ the potential of the canonically normalized field h. Suppose that in eq. (77) we choose a generic polynomial potential for ϕ,

$$V(\phi) = \sum a_n\phi^n. \quad (83)$$

The corresponding physical potential in the Einstein frame is then [68]

$$U[\phi(h)] = \frac{a_4}{(1+\xi\phi^2/M_p^2)^2}\left(\tilde{a}_2\phi^2 + \tilde{a}_3\phi^3 + \phi^4 + \sum_{n\geq 5}\tilde{a}_n\phi^n\right)\Big|_{\phi(h)}, \quad (84)$$

where $\tilde{a}_n \equiv a_n/a_4$. If we restrict ourselves to the case $a_{n\geq 5} = 0$, this potential has the right properties to provide a working inflationary model, since it has a minimum at $\phi = 0$ where inflation ends and it flattens, thanks to the presence of the non-minimal coupling, at large field values. This fact remains approximately true if $a_{n\geq 5} \neq 0$ as long as these coefficients are sufficiently small. We will have more to say about this case in the next section.

We look for viable inflationary solutions in the special case in which the values of the parameters a_i and ξ are such that an approximate stationary inflection point is present a few e-folds before the end of inflation at the field value $\phi = \phi_0$. To this end it is convenient to rewrite the above potential with a different set of parameters. A stationary inflection point is defined by the two conditions $U_\phi(\phi_0) = U_{\phi\phi}(\phi_0) = 0$. If we impose these conditions we can eliminate the coefficients $\tilde{a}_{(2,3)}$. Some trivial algebra then gives the following potential,

$$U(\phi) = \frac{\lambda\phi^4}{4!(1+\xi\phi^2/M_p^2)^2}\left\{\frac{\phi_0^2}{\phi^2}(1+c_2)\left[2\left(3+\xi\frac{\phi_0^2}{M_p^2}\right) + \sum_{n\geq 5}\tilde{a}_n\mathcal{F}_n(\phi_0,\xi)\right]\right.$$
$$\left. - \frac{\phi_0}{\phi}(1+c_3)\left[8 - \sum_{n\geq 5}\tilde{a}_n\mathcal{G}_n(\phi_0,\xi)\right] + \left(3+\xi^2\frac{\phi_0^4}{M_p^4}\right)\left[1+\sum_{n\geq 5}\tilde{a}_n\phi^{n-4}\right]\right\}, \quad (85)$$

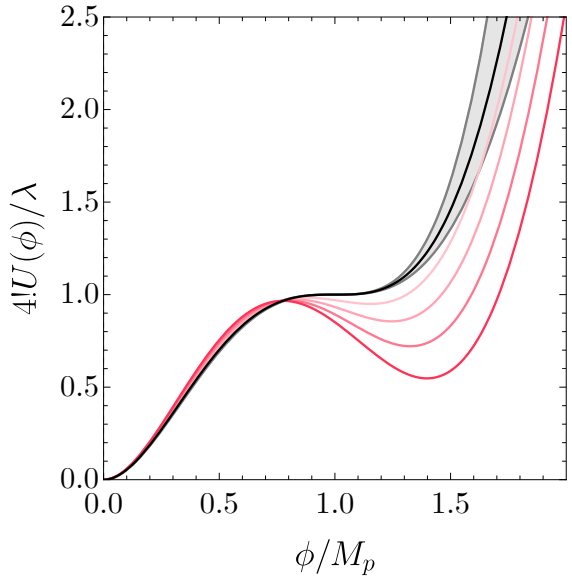

Figure 6: *Shape of the potential in eq. (88) for* $\phi_0 = 1$ *and* $\xi = 0.1$ *in the presence of an exact stationary inflection point* ($c_2 = c_3 = 0$, *solid black line) and for a slight deformation of it* ($0.02 \leqslant c_2 = c_3 \leqslant 0.1$, *red lines with increasing tonality of red). The region shaded in light gray illustrates the effect of changing* ξ *in the range* $\xi = 0.1 \pm 0.075$ *(while keeping* $c_2 = c_3 = 0$*), with larger values of* ξ *corresponding to milder slopes of the potential. Increasing the value of* ξ*, therefore, has the effect of reducing the classical velocity with which the inflaton field reaches the approximate stationary inflection point. Too large values of* ξ *eventually trap the inflaton field in the local minimum.*

where, without loss of generality, we have defined $\lambda \equiv 24a_4/(3 + \xi^2\phi_0^4/M_p^4)$, and

$$\mathcal{F}_n(\phi_0, \xi) \equiv \frac{3}{2}n(n-3)\phi_0^{n-4} + (6 - 5n + n^2)\xi\frac{\phi_0^{n-2}}{M_p^2} - \frac{1}{2}(n-4)(n-3)\xi^2\frac{\phi_0^n}{M_p^4}, \tag{86}$$

$$\mathcal{G}_n(\phi_0, \xi) \equiv (2-n)n\phi_0^{n-4} + (8 - 6n + n^2)\xi^2\frac{\phi_0^n}{M_p^4}. \tag{87}$$

We have also introduced two small parameters $c_{(2,3)}$ to replace $\tilde{a}_{(2,3)}$ in such a way that the potential in eq. (85) has an exact stationary inflection point when $c_{(2,3)} = 0$. If we restrict this expression to the fourth-order case $\tilde{a}_{n\geq 5} = 0$, we have

$$U(\phi) = \frac{\lambda\phi^4}{4!(1 + \xi\phi^2/M_p^2)^2}\left[3 + \xi^2\frac{\phi_0^4}{M_p^4} - 8(1+c_3)\frac{\phi_0}{\phi} + 2(1+c_2)\left(3 + \xi\frac{\phi_0^2}{M_p^2}\right)\frac{\phi_0^2}{\phi^2}\right]. \tag{88}$$

We remark that all we have done is replace the parameters $a_{(2,3,4)}$ in eq. (83) for the more useful λ and $c_{(2,3)}$, but the potentials are completely equivalent. The shape of this potential is shown in Fig. 6 for different values of the parameters.

Let us restrict our attention to the potential in eq. (88). Higher-dimensional operators will be considered in the next section. In the left panel of Fig. 7 we show the inflationary

dynamics corresponding to the solution of eq. (82) for the following parameters

$$c_2 = 0.011, \quad c_3 = 0.0089, \quad \phi_0 = M_p, \quad \text{and} \quad \tilde{a}_{n\geq 5} = 0, \tag{89}$$

while the values of λ and ξ are tuned in order to obtain, respectively, the correct normalization of the power spectrum A_s at CMB scales and the maximum amount of PBH abundance compatible with observations. We will present the result of a scan over different values of $c_{(2,3)}$, with λ and ξ tuned accordingly, later in this section, but the solution for the parameters in eq. (89) is representative of the entire parameter space. Typical values of λ and ξ for the solutions that we find are of order $\lambda \sim \mathcal{O}(10^{-9})$ and $\xi \sim \mathcal{O}(0.1)$. We remark that in this section we only consider PBH formation during a radiation-dominated era. Formation during an early matter-dominated era will be considered in the model of Section 2.3.

The blue line in Fig. 7 represents the physical potential $U[\phi(h)]$ as a function of the canonically normalized field h. Inflation starts at large field values and ends at the absolute minimum of the potential, at $h = 0$. Before the end of inflation, the potential features the presence of an approximate stationary inflection point with a local minimum and a subsequent local maximum. The solid red line represents the inflaton velocity dh/dN as a function of the canonically normalized field h. From the time at which the reference scale $k_\star = 0.05\,\mathrm{Mpc}^{-1}$, where we fit the CMB observables, exits the horizon until the end of inflation we count $\Delta N \simeq 51.4$ e-folds. This is enough to solve the horizon and flatness problems, as discussed in Section 1.3.[21] When the inflaton draws near the approximate stationary inflection point its velocity suddenly decreases to nearly zero, so that it almost stops, but has just enough inertia to overcome the barrier. As discussed in the previous section, this part of the dynamics is known as an ultra-slow-roll (USR) phase [51], and corresponds to the vertical region shaded in pink, which lasts for approximately $\Delta N \simeq 3.08$ e-folds (we define it as the phase in which $\eta > 3/2$, following the discussion in Section 1.4). The power spectrum corresponding to this solution and calculated by solving the Mukhanov-Sasaki equation (24) numerically is shown in the right panel of Fig. 7.

In order to make contact with CMB observables, at scales $10^{-4}\mathrm{Mpc}^{-1} \lesssim k \lesssim 0.5\mathrm{Mpc}^{-1}$, we fit our power spectrum with the parametric function

$$\log \mathcal{P}_\mathcal{R}(k) = \log A_s + \left(n_s - 1 + \frac{\alpha}{2} \log \frac{k}{k_\star} + \frac{\vartheta}{6} \log^2 \frac{k}{k_\star} + \dots \right) \log \frac{k}{k_\star}, \tag{90}$$

with $\alpha = dn_s/d\log k$, $\vartheta = d^2 n_s/d\log k^2$. At the pivot scale $k_\star = 0.05\mathrm{Mpc}^{-1}$, we find[22]

$$\log(10^{10} A_s) \simeq 3.06, \qquad n_s \simeq 0.9491, \qquad \alpha \simeq -10^{-3}, \qquad \vartheta \simeq 2 \times 10^{-4}. \tag{91}$$

[21] Whether or not the horizon and flatness problems are solved really depends on the elapsed number of e-folds between the time at which the largest observable scale $k \simeq 10^{-3}\mathrm{Mpc}^{-1}$ crosses the horizon and the end of inflation, but the difference between both numbers is of a few e-folds at most, so in practice using either is okay, given the uncertainty on e.g. the duration of reheating.

[22] We always refer to CMB parameters (such as n_s) at the pivot scale $k_\star = 0.05\mathrm{Mpc}^{-1}$.

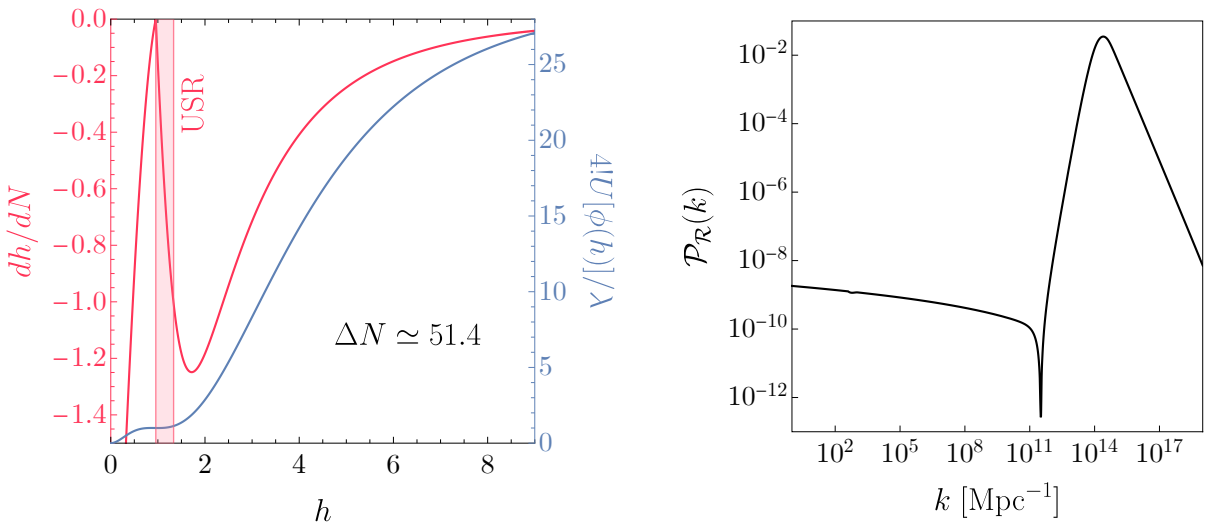

Figure 7: *Left panel: inflaton velocity (red line, units given on the left-side y-axis) as a function of the (canonically normalized) inflaton field value (bottom x-axis). Physical potential (blue line, units given on the right-side y-axis) as a function of the (canonically normalized) inflaton field value. Right panel: power spectrum of comoving curvature perturbations as a function of the comoving scale k obtained by solving numerically the Mukhanov-Sasaki equation (24).*

For the tensor-to-scalar ratio, we find, by means of the slow-roll approximation, $r \simeq 0.03$. All these values but the spectral index, n_s, are in good agreement with observations.[23] The fit of the spectral index n_s results in a 3σ tension with the latest Planck constraints if one takes the analysis obtained assuming the 6-parameter base ΛCDM model or the extension in which the running of the spectral index is added as an additional free parameter[24] [23],

$$\Lambda\text{CDM}: \quad n_s = 0.9649 \pm 0.0042, \tag{92}$$

$$\Lambda\text{CDM} + \frac{dn_s}{d\log k}: \quad n_s = 0.9641 \pm 0.0044. \tag{93}$$

A scan over the parameters of the potential[25] seems to suggest that values of n_s slightly smaller than the one expected on the basis of eqs. (92, 93) is a general result, and not just a vice of the specific numerical solution analyzed in this section. More in detail, we find that increasing the value of n_s in order to reduce the tension with eqs. (92, 93) results in a shift of the peak of the PBH mass distribution towards smaller values of $M_{\rm PBH}$. Even though we find

[23] We remark that including the BICEP and Keck Array data tightens the bound on r quoted in [23, 103] from $r < 0.11$ at $k_\star = 0.05\,\text{Mpc}^{-1}$ to $r < 0.035$, see [24, 130]. Our polynomial model still satisfies this stronger constraint.

[24] The addition of BAO data increases the best fit value of n_s in both cases just at the level of 0.2%, see [23].

[25] A less intensive scan was already performed for [68], with the same qualitative result, which we now confirm.

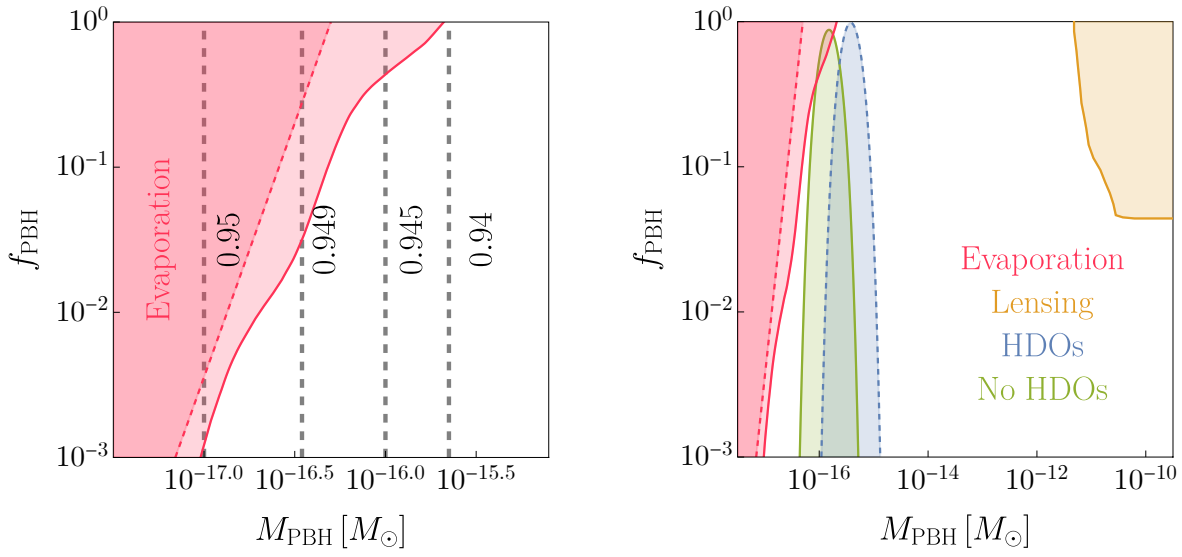

Figure 8: *Left panel: fraction of the dark matter in the form of PBHs as a function of the PBH masses. The values of n_s corresponding to solutions whose mass distribution is peaked at each particular mass are shown as dashed vertical lines. Solutions with acceptable values for the cosmological parameters A_s, r, α, ϑ, ΔN and satisfying the Hawking evaporation constraint populate the white region. The extra-galactic γ-ray constraints in [29, 30] are shown in dashed red, and the bound of [32] in solid red. Right panel: fractional abundance of PBHs with respect to the dark matter abundance as a function of the PBH mass for the parameter values discussed in the text, with and without higher-dimensional operators (HDOs), assuming the PBHs form during an RD era. The solution in green satisfies the extra-galactic γ-ray constraints in [29, 30], but there is a small overlap if the stronger bound of [32] is considered. This overlap can be eliminated by considering slightly lower $\mathcal{O}(1)$ abundances or by shifting the peak of the distribution to higher masses at the price of having lower values of n_s (see left panel).*

that in these cases it is still possible to obtain a peak in the power spectrum of order $\sim 10^{-2}$, the Hawking evaporation constraints kick in and very rapidly forbids sizable abundances of PBHs. On the contrary, moving the peak of the PBH mass distribution towards larger values of $M_{\rm PBH}$ implies a decrease of the spectral index below $n_s \simeq 0.94 - 0.95$, thus exacerbating the tension with CMB observables.

The above facts are illustrated in the left panel of Fig. 8, where we show the result of a scan over the parameter space of the model. All solutions found are characterized by acceptable values for the CMB observables A_s, r, α, ϑ in eq. (90) and by the condition that the time elapsed from the moment that the pivot scale leaves the horizon until the end of inflation is $\Delta N > 50$ e-folds. The plot confirms that values of the spectral index $n_s > 0.95$ are compatible with a fraction of dark matter in the form of PBHs smaller than 10^{-3} while a sizeable fraction can be obtained only if $n_s \simeq 0.948$ or smaller.

In light of the above discussion, it is important to interpret the actual relevance of

the spectral index tension. The value of n_s quoted in eq. (92) refers to the 6-parameter base ΛCDM model while the one in eq. (93) to the extension in which also the running of the spectral index is added as a free parameter. We can however consider other popular extensions of the 6-parameter base ΛCDM model [103]. Indeed, there are valid motivations to believe that the base ΛCDM model does not capture all the relevant physics throughout the evolution of the Universe. For instance, in the base ΛCDM model the neutrinos are massless, and although this is a reasonable first-order approximation, neutrino oscillations indicate that neutrinos have a small but non-zero mass. It is, therefore, more than legitimate to extend the base ΛCDM model by including the sum over the active neutrino masses, $\sum m_\nu$. Another plausible extension includes the effective number of relativistic degrees of freedom, $N_{\rm eff}$. Quoting from the official Planck 2018 release [103]:

$$\Lambda\text{CDM} + N_{\rm eff}: \quad n_s = 0.9589 \pm 0.0084, \tag{94}$$

$$\Lambda\text{CDM} + N_{\rm eff} + \sum m_\nu: \quad n_s = 0.9587 \pm 0.0086, \tag{95}$$

$$\Lambda\text{CDM} + N_{\rm eff} + \frac{dn_s}{d\log k}: \quad n_s = 0.950 \pm 0.011, \tag{96}$$

and in all these motivated extensions the tension with respect to the value $n_s \simeq 0.95$ is significantly reduced. It is indeed known that including a marginalization over the total neutrino mass or the number of relativistic degrees of freedom could induce a shift towards lower values in the determination of n_s[26] [132]. On a more speculative ground, the increasing statistical significance of the so-called "Hubble tension" (the discrepancy between the values of the present day Hubble expansion rate H_0 derived from the local distance ladder and the CMB, see [133]) further motivates the need of some new physics beyond the base ΛCDM model. Following this line of reasoning, in [134] a global analysis of current cosmological data in a cosmological scenario that is significantly more extended than the one provided by the base ΛCDM model was considered. They included as free parameters α, $N_{\rm eff}$, Σm_ν and the equation of state of dark energy, finding a preferred value of the spectral index of order $n_s \simeq 0.95$. In light of these results, we argue that values of the spectral index of order $n_s \simeq 0.94 - 0.95$ favored by our analysis (assuming the majority of dark matter to be comprised of PBHs) could consistently fit in the context of a cosmological model that extends the standard base ΛCDM one.

Before closing this section, let us mention how the different uncertainties in the calculation of the mass and abundance of PBHs could affect the results presented here. In this section, we have assumed that the mass of the PBHs is proportional to the mass contained in one Hubble volume at horizon crossing. Numerical simulations, in combination with the application of peaks theory, suggest that the mass of the PBHs depends on the shape of the perturbation from which it is formed [135]. Moreover it is related to the mass in a Hubble volume at a

[26] See also [131], where the cosmological parameters were derived in the context of a model with a strongly self-interacting massive neutrino, resulting in lower values of n_s.

time slightly different from the horizon crossing time. The effect of these issues is encoded in the efficiency factor γ in eq. (34). Unfortunately, at present the calculation of this factor carries $\mathcal{O}(1)$ uncertainties that do not allow to determine precisely the relation between the mass $M_{\rm PBH}$ and the scale k at which the enhancement of the power spectrum occurs. This is important because, as shown in Fig. 8, shifting the mass by a factor $\gtrsim 3$ would push solutions with $n_s = 0.95$ away from the Hawking evaporation constraints. In this way it would be possible to explain all the dark matter in terms of PBHs with these solutions. These uncertainties therefore prevent a reliable determination of the fate of the models with low n_s.

Another issue is the threshold δ_c in eq. (41), the value of which is also uncertain [111–113] and seems to depend on the shape of the power spectrum [135] (see, however, [136], where a universal threshold is proposed). The exact value of this quantity only determines the size of the PBH abundance and is therefore unlikely to affect the results of this section, which rely only on the relation between the mass of the PBHs and the comoving scale at which the peak in the spectrum is located. A similar remark can be made about the effect of non-Gaussianities. We have derived our results by using the Gaussian approximation in eq. (41). However, as we have mentioned already, the value of the power spectrum required to obtain $f_{\rm PBH} \sim \mathcal{O}(1)$ could be altered in the presence of the non-Gaussian corrections to this equation, which would appear as additional factors inside the integrand. Nonetheless, as explained in [125], since this equation is exponentially sensitive to the amplitude of the power spectrum, the effect of non-Gaussianities can be countered by simply changing the power spectrum by an $\mathcal{O}(1)$ factor. Such a change can be easily accomplished in the examples we have discussed in this section by slightly changing the value of the ξ parameter in the potential, which has no bearing on n_s. In other words, the n_s problem is essentially independent of the non-Gaussian corrections to the PBH abundance, since the issue in question is the position of the PBH mass distribution, which is controlled by eq. (34), an expression that is independent of the β function in eq. (41).

2.2 ■ Higher-dimensional operators

In this section we consider a simple, and arguably natural, way of circumventing the n_s tension. Instead of restricting ourselves to the case $\tilde{a}_{n\geq 5} = 0$, let us consider higher-dimensional operators (HDOs) in the potential. The HDOs considered in this section can be generated in the context of a toy ultraviolet completion to the theory discussed in Appendix B. We shall argue that a natural organization of the series of HDOs leads to good inflationary solutions with a value of the spectral index in perfect agreement with Planck data. In fact, a single five-dimensional operator with a naturally small coefficient and negligible higher-order terms is sufficient.

In models of large-field inflation one should generically take into consideration HDOs in the inflaton potential. It is then natural to question whether these corrections might spoil solutions that lead to a considerable abundance of PBHs at the renormalizable level, either by lowering the abundance, or by changing the power spectrum parameters at CMB scales. A precise characterization of the HDOs is of course possible only if one knows the ultraviolet completion of gravity. Nevertheless, it is still possible to gain an interesting insight from a minimal set of assumptions. To validate our construction, we shall follow two simple rules:

i) The HDOs must be subdominant compared to the leading renormalizable terms.

ii) The HDOs must be organized in the form of a "convergent" series. The meaning of this will be immediately clear.

Let us explain our rationale in more detail. If we rewrite eq. (84) as follows

$$U(\phi) = \frac{a_4}{(1+\xi\phi^2/M_p^2)^2}\left[\tilde{a}_2\phi^2 + \tilde{a}_3\phi^3 + \phi^4\left(1 + \sum_{n\geq 5}\tilde{a}_n\phi^{n-4}\right)\right], \tag{97}$$

the conditions *i)* and *ii)* translate into the order relation

$$\cdots < \tilde{a}_n\phi^{n-4} < \cdots < \tilde{a}_5\phi < 1, \tag{98}$$

meaning that, at each order $n \geqslant 5$ in the expansion, the coefficient $\tilde{a}_n$ has to be small enough compared to the previous one to compensate the additional power of ϕ, which can easily be $\mathcal{O}(10)$ at CMB scales. The description in terms of the effective operators breaks down at large field values where eq. (98) ceases to be valid. Driven by a pure phenomenological approach, one can, for instance, fix the coefficient $\tilde{a}_5$ to a very small number such that the condition $\tilde{a}_5\phi < 1$ is satisfied all along the inflationary trajectory while setting the remaining coefficients $\tilde{a}_{n\geqslant 6}$ to zero. It can indeed be checked that, if we take for simplicity the same values for $c_{(2,3)}$ and ϕ_0 chosen in eq. (89), the presence of a dimension-five HDO with coefficient $\tilde{a}_5 \sim \mathcal{O}(10^{-3})$ (together with $\tilde{a}_{n\geqslant 6} = 0$) is enough to give acceptable solutions with $n_s \simeq 0.960$ and the correct mass and abundance of PBHs. However, it is better to follow some organization principle that may help elucidate the physical interpretation of eq. (98).

Broadly speaking, the ultraviolet theory that generates the HDOs in eq. (97) will be described, at least, by a mass scale M and a dimensionless coupling g. Let us discuss how these fundamental quantities enter in our construction. We rewrite each HDO as

$$\mathcal{O}_n = a_n\phi^n = \frac{\phi^n}{\Lambda_n^{n-4}}, \quad n \geq 5, \tag{99}$$

where, for each operator, we introduce a suppression scale Λ_n (which is not necessarily equal to M_p). The scale Λ_n defines the strength of the effective interaction $\mathcal{O}_n$, and it is given by the ratio between a mass scale and a certain power of couplings. A simple but

pertinent example is that of the electroweak scale v. Its inverse squared, the Fermi constant $G_{\rm F} = 1/v^2$, defines the strength of the dimension-six four-fermion operator in the Fermi theory, and can be defined by means of the ratio between the W mass and the weak gauge coupling. As anticipated, we make the simplified assumption that the ultraviolet completion that is responsible for the generation of the effective operators in eq. (99) is characterized by a single coupling g and a single mass scale M. In such a case, by means of dimensional analysis [137], we expect in the weak coupling limit the scaling

$$\frac{1}{\Lambda_n} = \frac{g^{\frac{n-2}{n-4}}}{M} \qquad \longrightarrow \qquad \frac{1}{\Lambda_n^{n-4}} = g^2 \left(\frac{g}{M}\right)^{n-4} \equiv \frac{g^2}{\Lambda^{n-4}}, \tag{100}$$

where $\Lambda \equiv M/g$. We refer the reader to Appendix B for a detailed derivation of the scaling in eq. (100). The mass scale M can be considered as the mass associated to new degrees of freedom populating the ultraviolet theory, while g characterizes their self-coupling as well as their coupling with ϕ. Consequently, if we compare, in the spirit of eq. (97), the HDO $\mathcal{O}_n$ with the renormalizable term $a_4\phi^4$, we can write (keeping only track of powers of M and g, and neglecting $\mathcal{O}(1)$ proportionality coefficients)

$$a_4\phi^4 + \mathcal{O}_n = a_4\phi^4\left[1 + \underbrace{\left(\frac{g^2}{a_4}\right)\left(\frac{g}{M}\right)^{n-4}}_{\tilde{a}_n}\phi^{n-4}\right]. \tag{101}$$

From dimensional analysis, we know that the coefficient a_4 has the dimension of a coupling squared. This implies that the ratio g^2/a_4 is a genuine dimensionless number (see Appendix B for details). The hierarchy among the coefficients $\tilde{a}_n$ for $n \geq 5$ can be obtained if $\Lambda > \phi$. More precisely, the conditions in eq. (98) translate into

$$\frac{1}{\Lambda} < \min\left\{\frac{a_4}{g^2}\frac{1}{\phi_{\rm in}}, \frac{1}{\phi_{\rm in}}\right\}, \tag{102}$$

where $\phi_{\rm in}$ corresponds to the field value at the time at which the largest observable comoving scale $k \simeq 10^{-3}{\rm Mpc}^{-1}$ crosses the horizon (since we need to trust our theoretical description at least up to such field values). Clearly, as we already mentioned, for sufficiently large values of ϕ the hierarchy in eq. (98) will break down. This is just a manifestation of the old problem of initial conditions in large-field inflation. We stress that this problem is by no means unique to our model, but completely generic for large field inflation, and we do not aim to solve it here. We simply assume that the slow-roll approximation is valid at $\phi = \phi_{\rm in}$. One further condition needs to be satisfied. As it is clear from the previous discussion, we expect new states associated to the ultraviolet completion of our effective theory to lie at the mass scale M. We have to check, therefore, that the energy density during inflation is not high enough to excite these states (which could alter our effective inflationary potential). This means that the relation $H < M = g\Lambda$ must hold.

Bearing in mind these conceptual limitations, let us investigate some numerical consequences of the potential of eq. (97). If we set for simplicity $g^2 = a_4$, the HDOs are controlled by one single dimensionful free parameter, the inverse scale $1/\Lambda$, and we have

$$U(\phi) = \frac{a_4}{(1+\xi\phi^2/M_p^2)^2}\left[\tilde{a}_2\phi^2 + \tilde{a}_3\phi^3 + \phi^4\left(1 + \sum_{n\geq 5} c_n \frac{\phi^{n-4}}{\Lambda^{n-4}}\right)\right], \tag{103}$$

where the coefficients c_n are $\mathcal{O}(1)$ dimensionless numbers whose exact values cannot be computed with dimensional analysis alone. Qualitatively, the effect of the HDOs is shown in the left panel of Fig. 9 where we consider for illustration $\Lambda^{-1} = 10^{-3} M_p^{-1}$, $\xi = 0.3$, $\phi_0 = M_p$ and $c_{(2,3)} = 0$, together with $c_n = 1$ for $n \neq 2, 3$ (dashed red line). The solid black line is obtained by setting $c_n = 0$ for all n. At small field values, the stationary inflection point is not affected by the presence of the HDOs since it is controlled by the quartic and cubic coefficients. At large field values, on the contrary, the presence of the HDOs introduces a small deviation with respect to the renormalizable case (solid versus dashed line), and alters the first and second derivatives of the potential, thus changing the slow-roll parameters at the CMB pivot scale $k_\star$.

We will now consider two benchmark examples of solutions. In the first example, we assume $c_n = 1$ for $n \neq (2, 3)$. For each value of Λ^{-1} we consider inflationary solutions which give rise to $f_{\rm PBH} \simeq 1$ and consistently fit CMB observables at large scales. In order to facilitate the comparison with the renormalizable case, in Fig. 9 we show solutions with fixed number of e-folds $\Delta N \simeq 51$ from the time at which the pivot scale $k_\star$ crosses the horizon until the end of inflation. The values of $c_{(2,3)}$ and ϕ_0 are the same used in eq. (89) for the renormalizable case while λ and ξ are tuned, for each value of Λ, in such a way to obtain, respectively, the correct normalization of the power spectrum at CMB scales and the condition $f_{\rm PBH} \simeq 1$ on the abundance of PBHs. Furthermore, it is important to remark that all solutions shown in the right panel of Fig. 9 have, by construction, the position of the peak of the power spectrum, $k_{\rm peak}$, fixed at the value $k_{\rm peak} \simeq 1.5 \times 10^{14}$ Mpc^{-1}. This choice gives an abundance of PBHs peaked at $M_{\rm PBH} \simeq 5 \times 10^{17}$ g, which is compatible with the possibility of having 100% of dark matter in the form of PBHs. Moreover, it eliminates all those solutions, like the ones in the left panel of Fig. 8, in which larger values of n_s are obtained at the expense of a larger $k_{\rm peak}$ (and larger ΔN). We consider HDOs up to $n = 8$, and check that our results remain stable if further higher-order terms are added. In the right panel of Fig. 9 we show, for each one of these solutions, the corresponding value of n_s. If Λ^{-1} is too small, the impact of the HDOs is negligible and it is possible to have 100% of dark matter in the form of PBHs only for values of the spectral index that are $3\,\sigma$ away from the central value of Planck, as discussed earlier and shown in the left panel of Fig. 8. However, by increasing the value of Λ^{-1} without clashing against eq. (98) (region shaded in gray), the small correction introduced at large ϕ gives values of the spectral index that are in perfect agreement with the current observational bounds. This is shown by the red solid

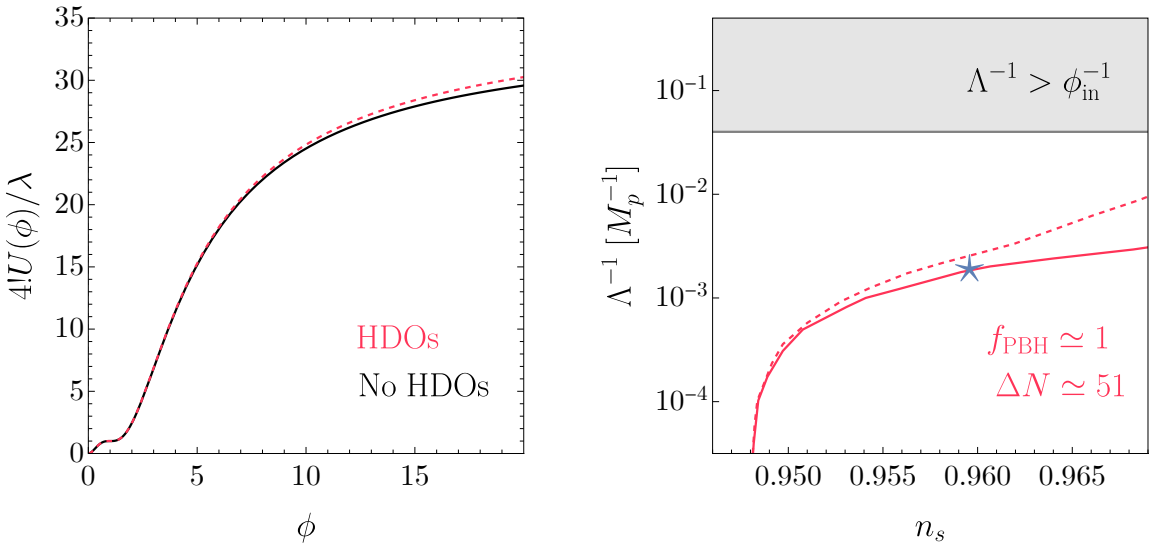

Figure 9: *Left panel. Corrections to the potential in eq. (97) due to the presence of HDOs. Right panel. Inflationary solutions that give 100% of dark matter in the form of PBHs in the plane* $(1/\Lambda, n_s)$ *where* $\Lambda = M/g$ *defines the suppression scale that controls the impact of HDOs (with large values of* Λ *that correspond to more and more suppressed HDOs). All remaining cosmological observables respect their corresponding CMB constraints. We show solutions with fixed number of e-folds* $\Delta N \simeq 51$*. We show two representative cases. The first case corresponds to the solid red line, and represents the impact of the HDOs in eq. (103) with* $g^2 = a_4$ *and* $c_n = 1$*. The second case corresponds to the dashed red line, and represents the impact of the HDOs in eq. (103) with* $g^2 = a_4$ *and* $c_n = (-1)^{n+1}$*. The blue star marks the solution whose PBH abundance is plotted in the right panel of Fig. 8 (blue region with dashed boundary).*

line in Fig. 9.

Let us now consider specifically the solution marked by the blue star which has $n_s \simeq 0.96$ and $\Lambda^{-1} \simeq 2 \times 10^{-3} M_p^{-1}$. As specified before, we include in our analysis HDOs up to $n = 8$ but for $\Lambda^{-1} \simeq 2 \times 10^{-3} M_p^{-1}$ it is possible to see that the first two with $n = (5, 6)$ dominate over the remaining ones. The corresponding population of PBHs is shown in the right panel of Fig. 8 (blue region with dashed boundary). We find that having 100% of dark matter in the form of PBHs is in excellent agreement with CMB observations (though as already mentioned, the tensor to scalar ratio $r = 0.03$ is fairly close to the bound obtained by adding the BICEP and Keck Array data to the Planck analysis [24, 130]). The solution also satisfies the strongest Hawking evaporation constraints imposed by the INTEGRAL satellite [31, 32]. The value of a_4 is fixed by the amplitude of the power spectrum at CMB scales, and we find $a_4 \simeq 10^{-10}$. Since we assumed $g^2 = a_4$, we have $g \simeq 10^{-5}$. From our discussion, it follows that for the mass scale $M = g\Lambda$ we have $M \simeq 10^{-2} M_p$. The condition $H < M$ is therefore verified, since $H \sim \sqrt{a_4}/\xi \simeq 10^{-5} M_p$. The same conclusion holds true for all solutions in the right panel of Fig. 9.

In the second example that we consider we take $c_n = (-1)^{n+1}$. This choice can be justified, for instance, in the context of toy ultraviolet completions such as the one discussed in Appendix B. We perform the same analysis discussed before, and our result is shown by the red dashed line in Fig. 9. Also in this case the series of HDOs can fix the n_s tension. For large values of Λ, the impact of HDOs is negligible, as in the case with $c_n = 1$. Larger values of Λ^{-1}, in comparison to the case $c_n = 1$, are needed in order to obtain the same n_s. The reason is that in the case $c_n = (-1)^{n+1}$ the alternating signs cause a partial cancellation of the HDOs.

To better understand how the presence of HDOs solves the n_s tension and gives an abundance of PBHs in the right mass window, it is instructive to emphasize three points. The first is that the modification of the potential induced by HDOs increases the value of the slow-roll parameter η towards less negative values. This $\mathcal{O}(1)$ change in η is enough to modify, at the percent level, the value of n_s, from $n_s \simeq 0.948$ to $n_s \simeq 0.970$. The second is that the presence of HDOs makes the power spectrum considerably more flat at large scales, going from $n_s \simeq 0.948$ to $n_s \simeq 0.970$. As a consequence, the dip of the power spectrum shifts towards larger values of k. The shift in the position of the dip of $\mathcal{P}_\mathcal{R}(k)$, for an ultra-slow-roll phase of the same duration, would also shift the position of the peak of $\mathcal{P}_\mathcal{R}(k)$ towards larger k, producing PBHs with unacceptably small mass. Finally, note that without including HDOs, we do not have the freedom to change the potential at large field values. Consequently, the simplest way to get a larger n_s is to fit CMB observables at larger field values. If we assume the renormalizable potential to be dominated by the quartic term, it is indeed possible to obtain, by means of the slow-roll approximation, the qualitative scaling $n_s = 1 - 16/[(1+6\xi)\phi^2_{\rm CMB}] + O(1/\phi^3_{\rm CMB})$. However, increasing $\phi_{\rm CMB}$ increases, in turn, the value of ΔN and shifts the peak of the power spectrum towards smaller scales. This qualitative behavior characterizes all solutions in Fig. 8 which produce a sizable abundance of PBHs. Although we do not have a rock-solid mathematical proof, our numerical scan suggests that this is a generic feature of the fourth-order polynomial model. These simple points make clear that the n_s tension can be solved with HDOs while obtaining PBHs as the totality of dark matter in perfect agreement with observational constraints.

It is worth noting now that the perturbative unitarity breaking scale for the potential of eq. (85) is $\Lambda_{\rm U} = M_p/\xi$ [138, 139]. New dynamics must arise at a scale lower or equal than $\Lambda_{\rm U}$ in order to restore unitarity. The values of $\xi \sim \mathcal{O}(0.1)$ that we find in our solutions push $\Lambda_{\rm U}$ above M_p. It is thus tantalizing that $M \sim 10^{-2} M_p \ll \Lambda_{\rm U}$, as we can speculate with the possibility that the new states arising at the mass scale M (and the corresponding HDOs) may harbinger a UV completion ensuring unitarity beyond $\Lambda_{\rm U}$.

The numerical analysis carried out in this section shows that small corrections to the inflaton potential generated by HDOs –whose presence, in particular in the context of large-field models of inflation, has no reason to be neglected– have the capability to fix the n_s tension pointed out earlier. Before concluding, it is worth mentioning some possible ex-

tensions of our analysis. First, the condition $g^2 = a_4$, that we imposed for simplicity, has no fundamental reason to be true, and relaxing it would open an additional direction in the parameter space (g, Λ) that would be interesting to explore. Second, one can in principle add HDOs that include field derivatives; for instance an effective operator of this kind with canonical mass-dimension d would have the general form $\mathcal{O}_d^{\partial} = (1/\Lambda_{\partial}^{d-4})\partial^{n_\partial}\phi^{n_\phi}$ with $d = n_\partial + n_\phi$. These operators can arise, for instance, in the context of the toy UV theory discussed in Appendix B. Although we expect these operators to be slow-roll suppressed at large field values, where HDOs are most relevant, it would be interesting to include them in the analysis.

2.3 ■ Monodromy-inspired potential

In this section we move away from the simple polynomial model we studied earlier to present a more elaborate scenario aimed at alleviating some of the downsides of inflection-point models of PBH production; specifically, the fact that the inflection point often has to be introduced in the potential in an ad hoc manner, and the fact that PBH formation in a radiation-dominated (RD) era leads to a severe fine-tuning of the parameters in the model. Our setup is characterized by a string-inspired inflationary potential which naturally features several local minima, and an early matter dominated (eMD) epoch after inflation.

Our scenario is inspired by axion monodromy inflation (AMI) [140, 141]. In this framework, the inflaton is a pseudo-scalar field with a discrete shift symmetry which is broken by a non-periodic potential term. The inflaton potential features axionic oscillations, superimposed on a monodromic term. When the amplitude of these modulations is large enough, near-inflection points and local minima appear in the inflationary trajectory. AMI can arise from string compactifications, where the inflaton field is generically accompanied by other heavy scalar fields, called moduli. As the inflaton travels $\gtrsim M_p$ distances during inflation, the heavy moduli tend to shift from their VEVs and backreact on the inflationary trajectory, leading to a flattening of the inflaton potential [142, 143] at large field values. The non-periodic part of the potential is generally quadratic for $\phi \lesssim M_p$ and behaves as a monomial for $\phi \gtrsim M_p$. In certain realizations of AMI, which will be of particular interest for us, the amplitude of the axionic potential oscillations is suppressed at large field values [144]. Therefore, AMI can provide inflationary potentials which exhibit two distinct regions (see Fig. 10): the first one, close to the global minimum, can feature the critical points that are desired for PBH formation; the second region, at large field values, does not display oscillations and is instead ideal to realize large field inflation. This setup can naturally accommodate a long eMD epoch after the end of inflation, before the inflaton decays completely and reheats the Universe. The minima of the potential for small field values is approximately quadratic and can support small oscillations of the inflaton. As per the discussion in Section 1.3, if these oscillations dominate the energy budget of the Universe, the latter enters into a phase of eMD

after inflation. Besides, in the framework of moduli stabilization in string compactifications [122, 123], the inflaton is often identified with a modulus field, meaning that it couples only gravitationally to the visible and hidden sectors. Therefore, reheating occurs at a slow pace via Planck-suppressed operators and the eMD phase can have a prolonged duration. AMI potentials typically feature the following functional form

$$V(\phi) = V_{\rm mon} + V_{\rm cos} = \frac{m^2 F^2}{2p}\left[\left(1+\frac{\phi^2}{F^2}\right)^p - 1\right] + \Lambda(\phi)^4 \cos\left(\frac{\phi}{f}+\delta\right). \tag{104}$$

Let us first focus on $V_{\rm mon}$. This term features three parameters: two energy scales, F and m, and an exponent p, which may be either positive or negative. For $\phi \ll F$, this part of the potential is well approximated by a parabola. For $\phi \gg F$, the potential either grows as ϕ^{2p} (if $p > 0$), or saturates to a plateau (if $p < 0$). In most constructions $p \leq 1$, meaning that the parabola tends to flatten at large field values. Typically, these flattening effects kick in at $\phi \sim F \lesssim M_p$, possibly fitting nicely with general arguments against the validity of single-field EFT descriptions of large field inflation, such as the weak gravity conjecture [145] (see [146, 147] for applications to axion inflation, and [148–150] for related conjectures). Observationally, this feature is essential for the viability of (104) as an inflationary potential. Indeed, power-like potentials with $p \geq 1$ are strongly constrained by CMB data, since they predict a large amplitude of primordial B-modes [23]. Models with $p = 1/3, 1/2$ [143] are still marginally compatible with CMB data. Here we will focus on $p < 1$, while allowing also for concrete values of p (such as $1/6$) beyond the ones that have been obtained so far in concrete stringy setups. An explicit realization of AMI with $p < 0$ –a possibility that we will consider– has been provided in [151].

Let us now discuss the second part of the potential (104), $V_{\rm cos}$. This contains the distinctive axionic oscillations, superimposed on $V_{\rm mon}$. Crucially, their amplitude $\Lambda(\phi)^4$ depends on the inflaton value. We follow [143] and parameterize this dependence as follows

$$\Lambda(\phi)^4 = \Lambda_0^4\, e^{-\left(\frac{\phi}{\phi_\Lambda}\right)^{p_\Lambda}}, \tag{105}$$

with $\phi_\Lambda \gtrsim M_p$ and p_Λ either positive or negative. Putting (104) and (105) together, we can write [1]

$$V(\phi) = m^2 f^2 \left\{ \frac{1}{2p}\frac{F^2}{f^2}\left[\left(1+\frac{\phi^2}{F^2}\right)^p - 1\right] + \kappa e^{-\left(\frac{\phi}{\phi_\Lambda}\right)^{p_\Lambda}} \cos\left(\frac{\phi}{f}+\delta\right)\right\} + V_0, \tag{106}$$

where we have added a constant V_0, which ensures $V = 0$ at the reheating minimum. The implications of $V_{\rm cos}$ then depend on p_Λ, p and the rescaled amplitude of the oscillations $\kappa \equiv \Lambda_0^4/(m^2 f^2)$. Let us consider the impact of κ and neglect the exponential prefactor for

now. Close to $\phi = 0$ we can approximate eq. (106) by

$$V(\phi) \approx m^2 f^2 \left[\frac{1}{2} \frac{\phi^2}{f^2} + \kappa \cos\left(\frac{\phi}{f} + \delta \right) \right] . \tag{107}$$

It is then straightforward to see that the potential (107) exhibits local minima for $\kappa \geq 1$, whereas for $\kappa < 1$ the oscillating part of the potential only gives rise to small bumps in the axion potential. We are interested in local minima which appear close to the bottom of the inflationary potential (i.e. for $\phi/M_p \ll 10$) and we will thus consider $\kappa \geq 1$.

Let us now return to the full potential. Depending on the sign of p_Λ the amplitude of the oscillations is exponentially suppressed or enhanced at large field values. The value of p_Λ is determined by the source of the non-perturbative effects that induce $V_{\cos}$ and by moduli stabilization. See [144] for examples with both $p_\Lambda > 0$ and $p_\Lambda < 0$. We are interested in $p_\Lambda > 0$ since then oscillations are absent at $\phi \gg \phi_\Lambda$ and the flatness of the potential allows to fit the CMB without tuning, while still featuring local minima at smaller field values. This particular behavior of the inflationary potential is also somewhat similar to what has been used in the relaxion mechanism [152]. Finally, let us discuss the parameter δ, which should be included on general grounds, since $V_{\rm mon}$ and $V_{\cos}$ have a priori no reason to be aligned. Furthermore, the choice $\delta = 0$ leads to the presence of two degenerate minima at the bottom of the potential, which may lead to stable domain walls during the reheating phase, when the field can oscillate along the full potential. For these reasons, in what follows we take $\delta \sim 1$.

The potential in eq. (106) is shown in Fig. 10 for different parameter choices. The figure illustrates the key feature of our inflationary potentials: beyond $\phi \sim 2M_p$, the potential is essentially indistinguishable from a standard monomial with power controlled by p, while at small field values the periodic axionic oscillations lead to a rich structure of local minima, the depth of which is controlled by κ. This parameter can be changed without affecting the inflationary potential at large field values. Inflation along our potential proceeds as follows. First, for large field values ($\phi \gg M_p$), the inflaton slowly rolls down the potential. This phase is the one responsible for the small CMB temperature anisotropies. Then, an ultra-slow-roll phase can be achieved as the inflaton traverses one of the local minima at $\phi \sim M_p$. In this regime, as we have seen, super-horizon curvature fluctuations are exponentially enhanced, leading to PBH formation upon horizon re-entry. Interestingly, in this scenario the two phases are decoupled from one another.

Due to the presence of local minima, the inflaton does not necessarily end up in the global minimum of the potential. In fact, in the regime $\kappa \gg 1$ the field typically gets classically stuck in one of the local minima closest to the global minimum. Let us estimate the tunneling rate to the global minimum from one of the nearest neighbouring local minima. This is proportional to e^{-S_t}, where the tunneling action S_t can be easily estimated in the

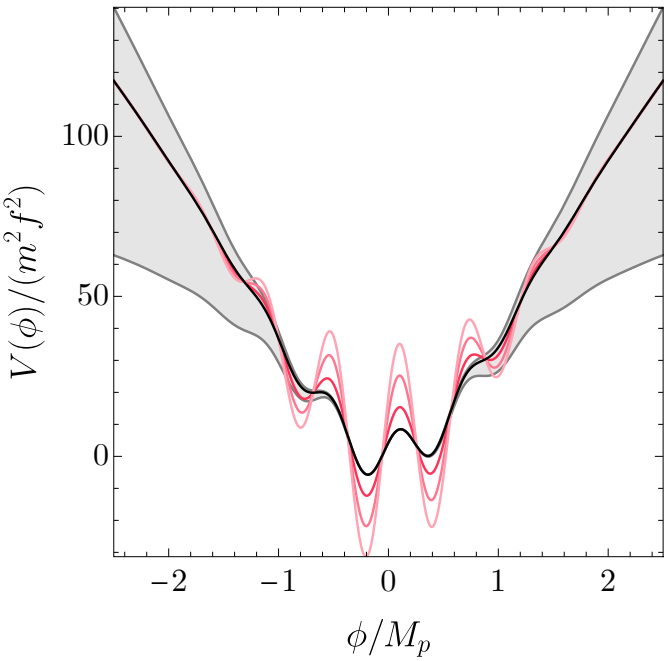

Figure 10: *Shape of the potential in eq. (106) for* $V_0 = 0$, $F = M_p$, $f = 0.1M_p$, $p_\Lambda = 2$, $\phi_\Lambda = M_p$, *and* $\delta = -1$. *The solid black line corresponds to* $p = 1/6$ *and* $\kappa = 8$. *The region shaded in light gray illustrates the effect of changing* p *in the range* $-1/2 \leq p \leq 1/3$, *with larger values of* p *corresponding to steeper potentials at large field values. The effect of varying* κ *in the range* $8 \leq \kappa \leq 35$ *is shown as red lines with increasing brightness.*

thin-wall approximation [153] as follows:

$$S_t \simeq \frac{27\pi^2}{2}\frac{\sigma^4}{(\Delta V)^3} \sim \frac{27\pi^2}{2}\frac{(\Delta\phi)^4(\kappa m^2 f^2)^2}{(m^2 f^2)^3} \sim \frac{27\pi^2}{2}\kappa^2\left(\frac{f}{m}\right)^2 . \tag{108}$$

Here we have approximated the tension of the bubble wall as $\sigma \sim \Delta\phi\sqrt{\Lambda_0^4}$, with $\Delta\phi \sim f$ the distance in field space between minima and ΔV the difference in height between the minima. For $p = 1$, the amplitude of the temperature anisotropies in the CMB implies $m \sim 10^{-6} M_p$. For $|p| < 1$, the CMB normalization depends also on F and larger values of m are allowed. Typical values of f in string compactifications are $10^{-3} M_p \lesssim f \lesssim M_p$. Therefore, for $\kappa \sim \mathcal{O}(10)$ (as it will be in our examples) S_t is very large and tunneling to the global minimum is an extremely suppressed process, regardless of the prefactor. The constant V_0 is chosen in such a way that the minimum where the inflaton stops has $V_0 = 0$. After inflation ends, the inflaton then oscillates around an approximately quadratic minimum and gives rise to the desired epoch of eMD. The Universe may then have a neighbouring AdS vacuum (the global minimum of the full potential in eq. (106), if the inflaton gets stuck in a preceding local minimum). As we have shown, this does not pose any cosmological threat

to the stability of the Universe.

Let us remark now that a potential (also inspired by AMI) with multiple approximate inflection points was considered in [69] in the context of PBH formation (during radiation domination). It differs from ours in several respects that are worth mentioning. The most important difference is that the potential of [69] features oscillations also at large field values. This implies that the CMB is fit in this case by tuning very finely the parameters of two trigonometric functions, in such a way that the CMB scales coincide with a sufficiently flat region of the potential for large field values. The axion decay constant in the case of [69] takes larger values: $f \gtrsim 0.6\ M_p$. In our case the CMB observables are essentially independent of the axion decay constant, which means f takes somewhat smaller values: $f \gtrsim 0.2 M_p$. Another difference is that the examples of [69] do not display multiple minima, but rather a successions of approximate plateaus. In that case inflation ends at the absolute minimum of the potential. Instead, we consider the possibility of several local minima where the inflaton may get stuck. Finally, [69] focused on PBH formation during RD, while we take into account the possibility of a long eMD epoch, which allows for a significant tuning reduction in the inflationary parameters.

We now proceed to solve the Mukhanov-Sasaki equation (24) numerically to find the power spectrum $\mathcal{P}_\mathcal{R}$ and the PBH abundance for concrete examples that satisfy the most recent Planck constraints [23, 103]. The potential in eq. (106) can feature many local minima whose depth grows as the inflaton travels from larger to smaller values. During its trajectory the inflaton passes through several of these minima, and the enhancement of the primordial spectrum becomes most pronounced when it goes through the next-to-last minimum, before stopping definitively; since it is in this region that it slows down the most. The depth of the minima is controlled by the parameter κ, as we mentioned earlier. Any large enough value of κ ensures the existence of minima which may lead to abundant PBH formation. The actual value of the abundance is determined by the speed of the inflaton as it climbs out of the next-to-last minimum before reheating. This speed is, in turn, fixed by the precise value of κ. In our examples, we adjust the parameter κ to obtain $f_{\rm PBH} \sim \mathcal{O}(1)$, and find $\kappa \sim \mathcal{O}(10)$. The smaller amplitude of $\mathcal{P}_\mathcal{R}$ required to account for all the dark matter with PBHs formed in eMD tends to reduce the number of e-folds that the inflaton field spends traversing the local minimum responsible for PBH formation with respect to the case of RD. Therefore, imposing $f_{\rm PBH} \sim \mathcal{O}(1)$ with masses in the window (1), some examples of potentials that are ruled out for PBH formation during RD due to an excess of inflation –see [124] and eq. (56)– may become viable changing κ appropriately if the PBHs form during an eMD phase. The same can be expected to occur for other models with an approximate inflection point.

Let us now discuss the effects of the rest of the parameters of the potential. We start with F and ϕ_Λ, which control the location in field space at which the flattening effects kick. We take them to be of order M_p. The parameter $f \lesssim M_p$ governs the width of the local minima. We can distinguish between different scenarios depending on its value. For values of f close

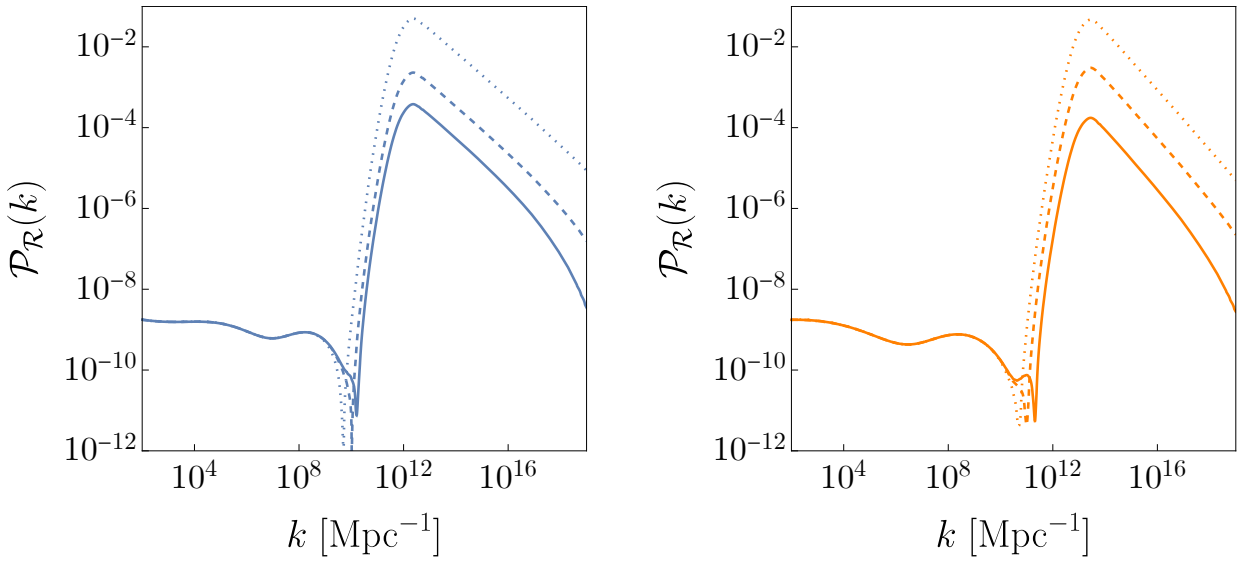

Figure 11: *Primordial spectra for Examples 1 (left) and 2 (right). The dotted and dashed lines correspond to the parameters choices in Table 2 labeled RD and eMD, respectively. The solid lines (not included on the table) are shown for ease of comparison and correspond to setting* $\mathcal{I} = 1$, *together with* $\kappa = 8.253$ *for Example 1, and* $\kappa = 14.984$ *for Example 2.*

to M_p, the inflaton encounters at most one local minimum before inflation ends. In this limit the model is essentially an implementation of the standard mechanism of PBH production from a quasi-inflection point. We will not consider this situation here, but we remark that it is a possibility capable of producing an interesting population of PBHs for the dark matter problem, if $\kappa \sim \mathcal{O}(1)$. In the opposite limit, for $f \ll 1$, the inflaton may roll all the way down to the global minimum and oscillate in a region of the potential which can encompass several local minima. This case is not relevant for PBH formation and we will not consider it either.[27] We thus focus on intermediate values of f for which the inflaton still travels over several minima before inflation ends. We find that if $f \lesssim 0.1 M_p$ and $\mathcal{P}_\mathcal{R} \sim \mathcal{O}(10^{-4})$ is imposed at its maximum, the field generically does not spend enough time (typically at most ~ 5 e-folds) on the local minimum previous to the end of inflation. In this case inflation does not last long enough to solve the horizon problem. We focus on $f = 0.2 M_p$ in our examples.

We will consider two concrete choices for the parameter p, one with $p > 0$ and the other one with $p < 0$. In particular, we find that the largest positive value of p which is compatible with the latest Planck data and at the same time leads to a significant amount of light PBHs is smaller than $p = 1/3$, which is the smallest positive exponent for which an explicit string construction currently exists [143]. We choose $p = 1/6$ to produce our first example and $p = -1/2$ for our second example. We focus on the case in which the inflaton gets stuck in the local minimum closest to the global minimum, which corresponds to rather large values

[27] However, it may present interesting consequences for reheating [154].

of κ. The scenario with the inflaton rolling all the way down to the global minimum can also be easily realized for both $p > 0$ and $p < 0$, by taking smaller values of κ. In both cases, we set $V = 0$ at the minimum of the potential where the field ends its trajectory.

	f	δ	p_Λ	ϕ_Λ	p	F	$m \cdot 10^5$	$V_0 \cdot 10^{11}$	κ (eMD/RD)
• Example 1	$0.2M_p$	-1	1	$1.15M_p$	1/6	$0.75M_p$	$3M_p$	$-5.1M_p^4$	8.254/8.254538
• Example 2	$0.2M_p$	-1	1	$1.15M_p$	-1/2	$1.85M_p$	$2M_p$	$2.0M_p^4$	14.986/14.986471

	$\phi_\star$	n_s	r	α	ΔN(eMD/RD)	$f_{\rm PBH}$(eMD/RD)	$M_{\rm PBH}$
• Example 1	$6.85M_p$	0.970	0.068	-0.009	51/57	0.9/0.4	$10^{-13}M_\odot$
• Example 2	$6.04M_p$	0.970	0.036	0.02	52/56	0.2/0.6	$10^{-15}M_\odot$

Table 2: *Parameters and predictions for the two examples we consider. The parameters give the correct normalization of the spectra at $k_\star = 0.05$ Mpc^{-1}, corresponding to $\phi = \phi_\star$. The CMB parameters are given at this scale, and ΔN denotes the number of e-folds from $\phi_\star$ to the end of inflation. The values of κ are given with the precision needed to attain the corresponding $f_{\rm PBH}$ in each case. We have set $T_m = 10^4$ GeV and $\mathcal{I} = 3.1$ for Example 1, and $T_m = 2 \cdot 10^5$ GeV and $\mathcal{I} = 4.4$ for Example 2.*

We show the resulting curvature power spectra in Fig. 11 for $p = 1/6$ and $p = -1/2$. The parameters of these two examples are given in Table 2. Additionally, for each one of these values of p, we obtain two different power spectra by considering two choices of κ. The spectra shown with continuous lines in Fig. 11 lead to $f_{\rm PBH} \sim \mathcal{O}(0.1-1)$ if the collapse occurs during the eMD epoch with reheating temperature $T_m \lesssim 10^6$ GeV. For these examples, we use the expression for β which takes angular momentum into account, (44) with $f_q(q_c) = 1$. We have chosen $\mathcal{I} = 3.1$ for Example 1, and $\mathcal{I} = 4.4$ for Example 2. In contrast, the spectra in dashed lines lead to a large $f_{\rm PBH}$ if they form during RD, and require further tuning of κ. The predictions for the inflationary observables and the PBH masses and fractions in our examples are also reported in Table 2.[28] We remark that although the examples presented in Table 2 satisfy the most recent Planck constraints, they do not satisfy the bound on r obtained if the BICEP and Keck Array data are included in the analysis [24, 130], which was published after [1] and is $r < 0.035$ (although Example 2 is within $\sim 1\,\sigma$ of this value). It might be possible to find examples that satisfy this stronger constraint if larger negative values of p are considered. We also point out that our examples satisfy the strongest Hawking evaporation constraints imposed by the INTEGRAL satellite [31, 32].

Since the suppression of the PBH formation probability away from the peak of the spec-

[28]The abundance in the case of RD has been computed using eq. (41), which assumes Gaussian fluctuations. However, due to the large change in the inflaton velocity in the region of the potential responsible for PBH formation, large non-Gaussianities may be produced. The impact of non-Gaussianities on the PBH abundance is not entirely understood yet but, as shown in [125], it is possible to compensate for their effect by slightly changing the size of the power spectrum, or, alternatively, the threshold δ_c in eq. (41) for the case of PBH formation during RD, so taking them into account will not significantly alter our results. The effect of non-Gaussianities on PBH formation during an MD era in models of inflation with an inflection point remains to be studied.

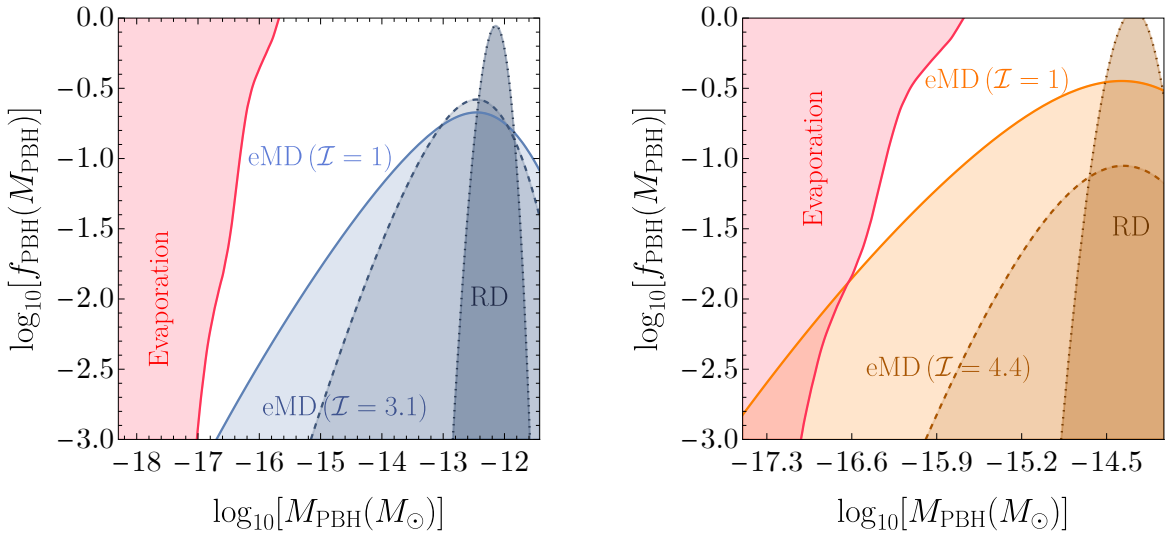

Figure 12: *Distribution of PBH masses for Examples 1 (left) and 2 (right). The largest masses shown in each panel correspond to the last modes that undergo collapse during the eMD era. The red region corresponds to the bound in [32]. We have also included for comparison the curves obtained by setting* $\mathcal{I}=1$*, with* $\kappa=8.253$ *for Example 1, and* $\kappa=14.984$ *for Example 2.*

trum is much milder in eMD than in RD, the mass distribution functions (depicted in Fig. 12) decay much more quickly in the latter case. This means that the extragalactic γ-ray bounds coming from PBH evaporation become more difficult to evade (for a given k value at the peak of the spectrum) if the PBHs form during an eMD era. This is a generic aspect of the eMD scenario for light PBHs. We find that the the bounds can be evaded in our examples, provided that the parameter $\mathcal{I}$ in eq. (44) is around $\mathcal{I}\sim 3$. For $\mathcal{I}=1$ only a very narrow primordial power spectrum can take full advantage of the eMD phase without clashing with the evaporation bounds. The largest masses shown in Fig. 12 correspond to the last modes that undergo collapse during the eMD era. The shape of the distribution for larger masses depends (unless one assumes pure RD right after inflation) on the details of the transition between eMD and RD, but we expect a rapid decay of the abundance due to the fact that the collapse occurs at least partially during the RD stage.

From these examples, we are able to extract the amount of tuning on the parameter κ which is required in order to obtain $f_{\rm PBH}\sim 1$ in our setup. We focus on this parameter because it is the one that controls the depth of the minima in the potential and therefore the enhancement of the spectrum, as illustrated in the examples of Table 2. We quantify the tuning [155] by calculating $\left|\Delta\kappa/\kappa\right|$, where $\Delta\kappa$ is the difference in κ between a successful example with $f_{\rm PBH}\sim 1$ and the closest κ which invalidates the example, with the rest of the parameters kept fixed. In other words, $\Delta\kappa$ is given by the minimal precision with which κ needs to be specified to obtain $f_{\rm PBH}\sim 1$. In the eMD case, we find that κ has to be chosen

with a relative precision (tuning) of order 10^{-2} % for $p = 1/6$ and $6 \cdot 10^{-3}$ % for $p = -1/2$. For RD, we instead find that the required tuning is increased to order 10^{-5} % for $p = 1/6$ and $6 \cdot 10^{-6}$ % for $p = -1/2$. Thus, an eMD phase alleviates the tuning in the potential parameters by three orders of magnitude. We expect this conclusion to remain valid for other models, since they are mainly a consequence of the discussion in Section 1.2.

CHAPTER 3

Quantum backreaction

In this chapter we reasses the calculation of the primordial power spectrum in the presence of an ultra-slow-roll phase by using the stochastic inflation formalism. In Section 3.1 we review the basic facts about the framework and describe the role of the coarse-graining scale. In Section 3.2 we derive the classicalization condition for the curvature perturbation in presence of a USR phase, using the occupation number density as a proxy. In Section 3.3 we use the analytical solution to the Mukhanov-Sasaki equation presented in Section 1.4 to find analytical expressions for the correlation functions of the noise terms that appear in the background equations of motion for the inflaton and show that the power spectrum in the stochastic inflation formalism coincides, at the linear level, with the result obtained by means of standard perturbation theory. Finally, in Section 3.4 we confirm our results numerically in the context of the polynomial model introduced in the previous chapter.

3.1 ■ Stochastic inflation

In the standard inflationary picture, the inflaton field is split into a homogeneous background piece and a space-dependent perturbation. The background piece is treated classically, and only the perturbation is treated quantum-mechanically. Soon after horizon crossing, the Fourier modes of the perturbation are placed in highly squeezed states which undergo a quantum-to-classical transition [156, 157], leading to the classical distribution of temperature fluctuations that we observe in the CMB. In this formalism, the classical background field follows a set of deterministic equations of motion and its evolution determines the behaviour of the perturbations. This treatment, however, does not take into account the possibility that perturbations might backreact on the classical inflaton trajectory, a shortcoming that the framework of stochastic inflation aims to remedy.

In the stochastic inflation formalism [71], the inflaton field is instead split into a coarse-grained, long-wavelength piece, and a short-wavelength perturbation. This splitting is performed by choosing a suitable coarse-graining scale k_σ, which must be introduced into the formalism by hand. Motivated by the fact that during slow-roll perturbations classicalize soon after horizon crossing, the most appropriate choice for this scale turns out to be $k_\sigma = \sigma a H$, where the coarse-graining parameter σ can be set to $\sigma \lesssim 1$ in the slow-roll regime. Notice that k_σ is a time-dependent function. The physical picture is then that short-wavelength modes gradually leave the horizon as time passes and thus the coarse-grained part of the field changes dynamically. After classicalizing, the short-wavelength perturbations behave as clas-

sical stochastic variables which act as sources of stochastic noise for the coarse-grained fields, which now follow non-deterministic, Langevin equations of motion. During slow-roll, this procedure is known to yield a power spectrum of curvature perturbations in agreement with the usual perturbation theory approach at the linear level (see e.g. [72]). However, whether the same result holds in the presence of an ultra-slow-roll phase is a delicate issue. As we will show, since perturbations in this case take longer to classicalize after horizon crossing, the choice of the coarse-graining scale is critical, and only for $k_\sigma = \sigma a H$ with $\sigma \sim \mathcal{O}(10^{-5})$ can the results of the formalism be trusted.

In [158] it was argued that in the presence of a USR phase, the probability distribution of the fluctuations could be altered significantly, ultimately leading to different results for the mass and abundance of primordial black holes formed in models of inflection-point inflation. The issue was later studied in [159], which made use of a potential based on the ratio of two polynomials first presented in [66] and concluded that in the presence of a USR phase the power spectrum of curvature perturbations computed at the linear level is enhanced by several orders of magnitude with respect to the usual perturbation theory result. These results were later disputed in [160], where it was instead concluded that the formalism leads to a power spectrum in agreement with the standard result at the linear level and at leading order in the slow-roll expansion, claims that were confirmed in [161], where it was concluded that the stochastic corrections to the power spectrum in the presence of a USR phase are negligible even beyond leading order in the slow-roll parameters. The validity of the formalism beyond slow-roll was also addressed in [162], which agreed with the conclusions of [161]. The calculation of the power spectrum was performed in the article that this chapter is based on [3], where we found that, as we shall see, the power spectrum computed in the stochastic formalism and at the linear level agrees with the standard result as long as the coarse-graining scale is appropriately chosen.

Let us remark that, although interesting on their own, these results are not the end of the story. If one of the main motivation for studying the stochastic formalism beyond slow-roll is to determine the correct way to calculate the mass and abundance of PBHs formed in inflection-point models of inflation, then higher-order correlators should also be considered. In fact, within the framework of stochastic inflation it is possible to calculate the full probability distribution of the fluctuations. Analytical results in this direction can be found in [73, 74, 97]. These articles were published shortly after [3] (or around the same time), and all agree that the probability distribution for the primordial fluctuations, although Gaussian at the peak, decays exponentially, a fact that is relevant for the correct determination of the PBH mass distribution. These results were confirmed numerically in [75, 76]. In this thesis we do not study higher-order correlators, but it is important to point out their relevance for PBHs.

We now reassess the computation of the primordial power spectrum in the presence of a USR phase using the framework of stochastic inflation, both analytically and numerically,

following the discussion in [3]. To separate the long- and short-wavelength modes, we can split the quantized inflaton field $\hat{\phi}(t, \boldsymbol{x})$ in two different pieces,

$$\hat{\phi}(t, \boldsymbol{x}) = \hat{\phi}_{\rm C}(t, \boldsymbol{x}) + \hat{\phi}_{\rm Q}(t, \boldsymbol{x}), \tag{109}$$

where

$$\hat{\phi}_{\rm C}(t, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{W}[k_\sigma(t) - k] \left[a_{\boldsymbol{k}} \phi_k(t) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} + \text{h.c.} \right], \tag{110}$$

$$\hat{\phi}_{\rm Q}(t, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{W}[k - k_\sigma(t)] \left[a_{\boldsymbol{k}} \phi_k(t) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} + \text{h.c.} \right]. \tag{111}$$

Here, $\hat{\phi}_{\rm C}(t, \boldsymbol{x})$ represents the coarse-grained part of the field containing the long-wavelength modes, and $\hat{\phi}_{\rm Q}(t, \boldsymbol{x})$ contains the short-wavelength ones. In what follows, we will treat the latter as a perturbation which acts as a source for the coarse-grained piece $\hat{\phi}_{\rm C}$. The quantities $a_{\boldsymbol{k}}$ and its hermitian conjugate $a^\dagger_{\boldsymbol{k}}$ represent creation and annihilation operators. Whether a mode is short- or long-wavelength is determined by the window function $\mathcal{W}$ and the cutoff scale k_σ. As we have anticipated already, in standard slow-roll, perturbations can be shown to classicalize (as we will do in the next section) shortly after they cross the horizon. Therefore, an appropriate choice for the cutoff scale in this case is $k_\sigma = \sigma a H$, with $\sigma \lesssim 1$ a constant parameter. The situation in the presence of a USR phase changes, however, since perturbations take longer to classicalize after horizon crossing. In this case, the cutoff scale can still be chosen as $k_\sigma = \sigma a H$, but the cutoff parameter must be $\sigma \ll 1$. The exact value this quantity must have to correctly capture the dynamics is model-dependent, but we will estimate it to be $\sigma \sim \mathcal{O}(10^{-5})$ in the next section for standard inflection point models. We will have more to say on the role of this parameter at the end of this section. We choose the window function to be the Heaviside step function,

$$\mathcal{W}[k_\sigma(t) - k] = \Theta[k_\sigma(t) - k]. \tag{112}$$

Since the only role of the window function is to determine whether a mode has classicalized or not, results should not depend strongly on its choice.[29] The simple choice shown above has the convenient consequence that the stochastic noise appearing in the equation of motion for the coarse-grained field can be modelled as a Wiener process, as we will see momentarily. It is important to stress that both fields $\hat{\phi}_{\rm C}$ and $\hat{\phi}_{\rm Q}$ in eq. (109) have an intrinsic quantum nature. Since the window function selects modes with $k < k_\sigma$, the spatial dependence in $\hat{\phi}_{\rm C}$

[29] We remark, however, that it was noted in [163] that sharp cutoffs such as the one chosen here lead to issues in the noise correlators at large spatial distances. Smooth window functions have been proposed in [164–166]. As far as we are aware, the effect that varying the window function would have in the presence of a USR phase has not been studied, and would be an interesting direction for future work.

can be neglected on these scales compared to the temporal one, and we can define

$$\hat{\phi}_{\rm C}(t,\boldsymbol{x}) \simeq \hat{\phi}_{\rm C}(t) \equiv \bar{\phi}(t). \tag{113}$$

In other words, the coarse-grained field $\hat{\phi}_{\rm C}$ can be considered homogeneous, at each time t, over a length scale $L_\sigma(t) \equiv (\sigma a H)^{-1}$. Similar considerations apply to the conjugate momentum $\hat{\pi}$ of the field, which can also be decomposed in long- and short-wavelength pieces

$$\hat{\pi}(t,\boldsymbol{x}) = \hat{\pi}_{\rm C}(t,\boldsymbol{x}) + \hat{\pi}_{\rm Q}(t,\boldsymbol{x}), \tag{114}$$

where

$$\hat{\pi}_{\rm C}(t,\boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{W}[k_\sigma(t) - k] \left[a_{\boldsymbol{k}} \pi_k(t) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} + \text{h.c.}\right], \tag{115}$$

$$\hat{\pi}_{\rm Q}(t,\boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{W}[k - k_\sigma(t)] \left[a_{\boldsymbol{k}} \pi_k(t) e^{i\boldsymbol{x}\cdot\boldsymbol{k}} + \text{h.c.}\right], \tag{116}$$

in analogy with eqs. (109, 110, 111).

Let us determine the equation of motion for $\bar{\phi}$ when $\hat{\phi}_{\rm Q}$ is thought of as a perturbation. We work using the spatially flat gauge $\psi = E = 0$. Our convention for the metric perturbations is given in eq. (16) (see also Appendix A for a review of perturbation theory). The equations of motion of the full inflaton field $\hat{\phi}(t,\boldsymbol{x})$ and its conjugate momentum $\hat{\pi}(t,\boldsymbol{x})$ take the form

$$\frac{d\hat{\phi}}{dt} = \frac{1}{a^3}(1+\varphi)\hat{\pi} + \frac{1}{a}(\nabla B)\cdot(\nabla\hat{\phi}), \tag{117}$$

$$\frac{d\hat{\pi}}{dt} = \frac{1}{a}\nabla\cdot(\hat{\pi}\nabla B) + a\nabla^2\hat{\phi} + a\nabla\cdot(\varphi\nabla\hat{\phi}) - a^3(1+\varphi)V_\phi(\hat{\phi}). \tag{118}$$

If we now plug eq. (109) –together with an analogue decomposition for the conjugate momentum $\hat{\pi}(t,\boldsymbol{x})$– into eqs. (117, 118), we find, linearizing in $\hat{\phi}_{\rm Q}$ and $\hat{\pi}_{\rm Q}$, the system of Langevin equations (see Appendix C for a review of stochastic differential equations)

$$\frac{d\bar{\phi}}{dt} = \frac{\bar{\pi}}{a^3} + \xi_\phi, \tag{119}$$

$$\frac{d\bar{\pi}}{dt} = -a^3 V_\phi(\bar{\phi}) + \xi_\pi, \tag{120}$$

where we have defined the so-called noise operators (since $a_{\boldsymbol{k}}$ is an operator)

$$\xi_\phi \equiv -\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{d\mathcal{W}}{dt}\left[a_{\boldsymbol{k}}\phi_k e^{+i\boldsymbol{x}\cdot\boldsymbol{k}} + a^\dagger_{\boldsymbol{k}}\phi^\star_k e^{-i\boldsymbol{x}\cdot\boldsymbol{k}}\right], \tag{121}$$

$$\xi_\pi \equiv -\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{d\mathcal{W}}{dt}\left[a_{\boldsymbol{k}}\pi_k e^{+i\boldsymbol{x}\cdot\boldsymbol{k}} + a^\dagger_{\boldsymbol{k}}\pi^\star_k e^{-i\boldsymbol{x}\cdot\boldsymbol{k}}\right], \tag{122}$$

where the Fourier modes $\phi_k(t)$ and $\pi_k(t)$ defined in (110, 115) satisfy the Hamiltonian system

$$\dot{\phi}_k = \frac{1}{a^3}\left(\bar{\pi}\varphi_k + \pi_k\right), \tag{123}$$

$$\dot{\pi}_k = -\frac{\bar{\pi}}{a}k^2 B_k - ak^2\phi_k - a^3\Big[\varphi_k V_\phi(\bar{\phi}) + \phi_k V_{\phi\phi}(\bar{\phi})\Big]. \tag{124}$$

The time derivative of the window function is, in accordance with eq. (112),

$$\frac{d\mathcal{W}}{dt} = \frac{d}{dt}\Theta[k - k_\sigma(t)] = -\delta[k - k_\sigma(t)]\frac{dk_\sigma(t)}{dt}. \tag{125}$$

In Fourier space, the metric perturbations obey

$$H\varphi_k = \frac{1}{2M_p^2}\phi_k\frac{d\bar{\phi}}{dt}, \tag{126}$$

$$\frac{k^2}{a}B_k = 3H\varphi_k + \frac{1}{2M_p^2}\left[\frac{\phi_k}{H}V_\phi(\bar{\phi}) - \frac{\varphi_k}{H}\left(\frac{d\bar{\phi}}{dt}\right)^2 + \frac{\dot{\phi}_k}{H}\frac{d\bar{\phi}}{dt}\right]. \tag{127}$$

By combining these equations we obtain

$$\ddot{\phi}_k + 3H\dot{\phi}_k + \frac{k^2}{a^2}\phi_k + \left\{V_{\phi\phi}(\bar{\phi}) - \frac{1}{M_p^2}\frac{1}{a^3}\frac{d}{dt}\left[\frac{a^3}{H}\left(\frac{d\bar{\phi}}{dt}\right)^2\right]\right\}\phi_k = 0. \tag{128}$$

If we define $u_k \equiv a\phi_k/M_p$ and change variables to conformal time $ad\tau = dt$ (recall that $' \equiv d/d\tau$), we obtain the stochastic inflation analogue of the Mukhanov-Sasaki equation

$$u_k'' + \left(\frac{z''}{z} - k^2\right)u_k = 0, \qquad \text{with} \qquad z \equiv \frac{\bar{\phi}'}{H}, \tag{129}$$

where the difference with respect to the standard perturbation theory result (24) once we identify $u_k = z\mathcal{R}_k$ is that z is constructed using the coarse-grained field $\bar{\phi}(t)$, and not the classical background field $\phi(t)$. Similarly, the quantity ϕ_k (or u_k) should be understood as a Fourier mode of the short-wavelength perturbation in eq. (111) and not as the standard quantity $\delta\phi_k$ from perturbation theory.

Despite their intrinsic quantum nature, a classical interpretation can be assigned to $\bar{\phi}$ and ξ_ϕ and ξ_π in eqs. (119, 120). If we compute the equal-time commutator $[\xi_\phi(t,\boldsymbol{x}), \pi_\phi(t,\boldsymbol{x}')]$, we find, in the limit $k_\sigma|\boldsymbol{x} - \boldsymbol{x}'| \ll 1$

$$[\xi_\phi(t,\boldsymbol{x}), \xi_\pi(t,\boldsymbol{x}')] \propto \phi_{k_\sigma}(t)\pi^\star_{k_\sigma}(t) - \phi^\star_{k_\sigma}(t)\pi_{k_\sigma}(t) \simeq 0. \tag{130}$$

Since the commutator vanishes, the variables ξ_ϕ and ξ_π can be considered classical (this implies that a similar interpretation can be assigned to $\bar{\phi}$ and $\bar{\pi}$). However, at this stage we cannot assign to them any specific numerical value since they are still defined in terms of operators. This simply means that, from a classical point of view, they must be considered

stochastic variables whose statistical properties are fully determined by computing their correlation functions (that is, by identifying quantum expectation values with statistical moments). It is straightforward to check that $\langle 0|\xi_{\phi,\pi}(t,\boldsymbol{x})|0\rangle = 0$. We define the two-point correlation matrix

$$\Theta(t,\boldsymbol{x};t',\boldsymbol{x}') \equiv \begin{pmatrix} \langle 0|\xi_\phi(t,\boldsymbol{x})\xi_\phi(t',\boldsymbol{x}')|0\rangle & \langle 0|\xi_\phi(t,\boldsymbol{x})\xi_\pi(t',\boldsymbol{x}')|0\rangle \\ \langle 0|\xi_\pi(t,\boldsymbol{x})\xi_\phi(t',\boldsymbol{x}')|0\rangle & \langle 0|\xi_\pi(t,\boldsymbol{x})\xi_\pi(t',\boldsymbol{x}')|0\rangle \end{pmatrix}, \tag{131}$$

with elements

$$\Theta_{fg} \equiv \langle 0|\xi_f(t,\boldsymbol{x})\xi_g(t',\boldsymbol{x}')|0\rangle = \frac{1}{6\pi^2}\frac{dk_\sigma^3}{dt} f_{k_\sigma}(t) g^\star_{k_\sigma}(t) \frac{\sin(k_\sigma|\boldsymbol{x}-\boldsymbol{x}'|)}{k_\sigma|\boldsymbol{x}-\boldsymbol{x}'|}\delta(t-t'). \tag{132}$$

The two-point correlation functions are therefore nonzero only at $t = t'$, a property that defines the so-called white noise. This noise also happens to be Gaussian, since higher-order correlators can be expressed in terms of the two-point functions due to Wick's theorem. These facts, which originate from our choice of window function in eq. (112) justify modelling the noise in the Langevin equations (119, 120) as a Wiener process (see Appendix C for the definition of a Wiener process and a review about stochastic differential equations). Other choices of window functions lead to colored noise sources, for which the standard techniques of stochastic differential equations involving Wiener processes cannot be applied. In the following, we shall restrict our analysis to the simplest case of white noise, although we remark that other choices are also acceptable (see e.g. [164–166]). As we mentioned already, since the only role of this function is to distinguish classicalized modes from quantum modes, the results should not depend heavily on this choice. We are interested in the effect of the noise on length scales $\Delta x \equiv |\boldsymbol{x}-\boldsymbol{x}'|$ over which the coarse-grained field is homogeneous, $\Delta x \ll L_\sigma$. This implies that we can evaluate eq. (132) at the same spatial point, $\boldsymbol{x} \simeq \boldsymbol{x}'$. Eq. (132) becomes

$$\Theta_{fg}(t,\boldsymbol{x};t',\boldsymbol{x}) = \Theta_{fg}(t)\delta(t-t'), \qquad \Theta_{fg}(t) \equiv \frac{d\log k_\sigma}{dt}\underbrace{\frac{k_\sigma^3}{2\pi^2} f_{k_\sigma}(t) g^\star_{k_\sigma}(t)}_{\mathcal{P}_{fg}(t,k_\sigma)}, \tag{133}$$

where $\mathcal{P}_{fg}(t,k_\sigma)$ is the power spectrum of the fluctuations (ϕ,π) evaluated for each time t at the corresponding coarse-graining cutoff wavenumber k_σ.

Let us summarize the above results and highlight a few points. The equations of motion (119, 120) can be interpreted as Langevin equations for the classical stochastic variables $\bar{\phi}$ and $\bar{\pi}$. The quantities ξ_ϕ and ξ_π then behave as sources of stochastic noise, which we model as Wiener processes due to the fact that their correlation functions, given by eq. (133), are non-vanishing only at equal times. These equations of motion can be solved numerically by discretizing the time variable and drawing (from a Gaussian distribution) a different value for the noise sources at each time step. After repeating this procedure for a large number

of realizations, the statistical properties of the system can be determined. This procedure, however, is a very difficult numerical task, since the amplitude of the background noise, given by eq. (133), is determined by the evolution of the fluctuations. The correct way to handle the evolution would then be to solve for the background and perturbations simultaneously, see e.g. [76]. An alternative approach is to expand the coarse-grained fields about their classical counterparts at first order, $\bar{\phi} = \phi_{\rm cl} + \delta\phi_{\rm st}$ and $\bar{\pi} = \pi_{\rm cl} + \delta\pi_{\rm st}$, where $\phi_{\rm cl}$ and $\pi_{\rm cl}$ are defined as the solution of the Langevin equations in the absence of the noise term. This can be thought of as the first step of a recursive strategy [167, 168]. Following this approach, the stochastic noise appears as a source for the classical stochastic variables $\delta\phi_{\rm st}$ and $\delta\pi_{\rm st}$. At first glance, expanding the fields in this way might make it seem like we are just doing standard perturbation theory. However, notice that the variables $\delta\phi_{\rm st}$ and $\delta\pi_{\rm st}$ retain here their statistical meaning precisely because of the presence of the noise terms in the equations governing their evolution. It is, therefore, only by computing the corresponding statistical moments that one can extract information about the distribution of the perturbations. This is the approach that we shall take from now on.

Until now, we have formulated the problem using cosmic time t as the time variable. However, in stochastic inflation using different time variables can lead to different physical results. This is for instance the case if one considers the number of e-folds N instead of the cosmic time t. This is because changing variables from t to N involves H, which is a stochastic variable, since it is a function of $\bar{\phi}$. In [72], it was argued that the number of e-folds N is the time variable which allows one to consistently connect stochastic inflation with results from QFT on curved space-times (see also [169, 170]). This issue is not really relevant for us since all of our results in this chapter will be derived using the analytical model presented in Section 1.4, which assumes that H is constant. Nevertheless, formulating the stochastic dynamics in terms of the number of e-folds can help elucidate the physical interpretation of some of the final equations, in addition to being more convenient for numerical calculations. For this reason, from now on we switch to the description in terms of the number of e-folds. The Langevin equations then take the form

$$\frac{d\bar{\phi}}{dN} = \bar{\pi} + \xi_{\phi}, \tag{134}$$

$$\frac{d\bar{\pi}}{dN} = -(3-\epsilon)\bar{\pi} - \frac{1}{H^2}V_{\phi}(\bar{\phi}) + \xi_{\pi}, \tag{135}$$

where, compared with eqs. (119, 120), we have rescaled the conjugate momentum as $\bar{\pi}/(a^3 H) \to \bar{\pi}$. This rescaling allows a more direct identification of $\bar{\pi}$ with the inflaton velocity. The noise

operators are

$$\xi_\phi = -\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{d\mathcal{W}}{dN}\left[a_k\phi_k e^{+i\boldsymbol{x}\cdot\boldsymbol{k}} + a_k^\dagger\phi_k^\star e^{-i\boldsymbol{x}\cdot\boldsymbol{k}}\right], \tag{136}$$

$$\xi_\phi = -\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{d\mathcal{W}}{dN}\left[a_k\pi_k e^{+i\boldsymbol{x}\cdot\boldsymbol{k}} + a_k^\dagger\pi_k^\star e^{-i\boldsymbol{x}\cdot\boldsymbol{k}}\right], \tag{137}$$

where now ϕ_k and π_k are given by

$$\frac{d\phi_k}{dN} = \bar{\pi}\varphi_k + \pi_k, \tag{138}$$

$$\frac{d\pi_k}{dN} = -(3-\epsilon)\pi_k - \frac{\bar{\pi}k^2}{aH}B_k - \frac{k^2}{(aH)^2}\phi_k - \frac{1}{H^2}\left[\varphi_k V_\phi(\bar{\phi}) + \phi_k V_{\phi\phi}(\bar{\phi})\right]. \tag{139}$$

If we eliminate the metric perturbations by means of the Einstein field equations, we find

$$\frac{d^2\phi_k}{dN^2} + (3-\epsilon)\frac{d\phi_k}{dN} + \left[\frac{k^2}{(aH)^2} + (3-\epsilon)\frac{V_{\phi\phi}(\bar{\phi})}{V(\bar{\phi})}M_p^2 - 2\epsilon(3+\epsilon-2\eta)\right]\phi_k = 0. \tag{140}$$

For the noise correlation function in eq. (133), we have

$$\Theta_{fg}(N) \equiv \frac{d\log k_\sigma}{dN}\frac{k_\sigma^3}{2\pi^2}f_{k_\sigma}(N)g_{k_\sigma}^\star(N), \tag{141}$$

where $k_\sigma = k_\sigma(N)$.

From the discussion presented in this section, it is clear that a proper definition of the coarse-graining cutoff wavenumber k_σ is necessary in order to have a correct interpretation of the stochastic dynamics. In standard slow-roll inflationary models, the choice $k_\sigma = \sigma aH$ is well motivated. However, in the presence of a USR phase the horizon-crossing condition $\sigma \sim 1$ does not offer a correct description of the dynamics of the perturbations and, in particular, it does not always describe appropriately the time after which the curvature perturbation stays constant. To illustrate this, we consider the analytical model developed in Section 1.4. In Fig. 13 we show, by using this model, that modes that leave the horizon during the USR phase take much longer to freeze than modes that leave at early or late times. In the next section we will show that, critically, the freezing condition is equivalent to the classicalization condition. It is therefore important to re-think about the appropriate definition of k_σ in models that feature a USR phase.

3.2 ■ Classicalization of the modes

In this section we present a detailed analysis of the quantum-to-classical transition of inflationary perturbations in the presence of a USR phase. We will argue that the definition $k_\sigma = \sigma aH$ can still be used in this case, provided that the cutoff parameter σ is chosen to

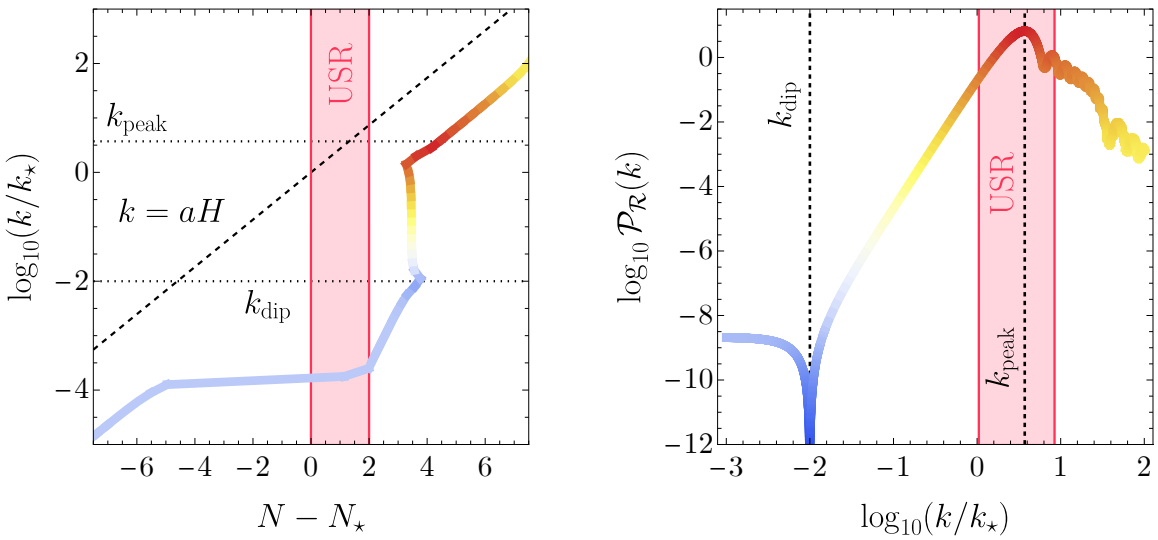

Figure 13: *Left panel: time at which the curvature perturbation freezes outside the horizon for each k. For modes that leave the horizon at early times or after the USR phase ends, the freezing occurs shortly after horizon crossing. Modes that leave the horizon during the USR phase take much longer to freeze. Right panel: power spectrum $\mathcal{P}_\mathcal{R}$ corresponding to the dynamics in the left panel. The reference scale $k_\star$ is chosen as the one that crosses the horizon when the USR phase begins. We have used the analytical model of Section 1.4 with parameters $\eta_1 = 0$, $\eta_2 = 4$, $\eta_3 = -1$, and $\Delta N_2 = 2$.*

be small enough to allow for the classicalization of the relevant modes to occur. To quantize the system, let us begin by noting that the Mukhannov-Sasaki equation (129) in real space can be derived from the following action for $u(\tau, \boldsymbol{x})$

$$\mathcal{S} = \frac{1}{2}\int d\tau d^3\boldsymbol{x} M_p^2 \left[\left(u' - \frac{z'}{z}u\right)^2 - (\nabla u)^2\right]. \tag{142}$$

Defining the conjugate momentum $p = \delta\mathcal{S}/\delta u'$, we obtain $p = u' - (z'/z)u$. Consequently, the Hamiltonian is

$$\mathcal{H}(\tau) = \frac{1}{2}\int d^3\boldsymbol{x} M_p^2 \left[p^2 + (\nabla u)^2 + 2\frac{z'}{z}pu\right]. \tag{143}$$

We promote $u(\tau, \boldsymbol{x})$ and $p(\tau, \boldsymbol{x})$ to quantum operators $\hat{u}(\tau, \boldsymbol{x})$ and $\hat{p}(\tau, \boldsymbol{x})$ with equal-time commutation relations $[\hat{u}(\tau, \boldsymbol{x}), \hat{p}(\tau, \boldsymbol{x}')] = i\delta^3(\boldsymbol{x} - \boldsymbol{x}')$. In Fourier space, we find the following Hamiltonian operator

$$\hat{\mathcal{H}}(\tau) = \frac{1}{2}\int d^3\boldsymbol{k} M_p^2 \left\{\hat{p}_{\boldsymbol{k}}(\tau)\hat{p}^\dagger_{\boldsymbol{k}}(\tau) + k^2\hat{u}_{\boldsymbol{k}}(\tau)\hat{u}^\dagger_{\boldsymbol{k}}(\tau) + \frac{z'}{z}\left[\hat{p}_{\boldsymbol{k}}(\tau)\hat{u}^\dagger_{\boldsymbol{k}}(\tau) + \hat{u}_{\boldsymbol{k}}(\tau)\hat{p}^\dagger_{\boldsymbol{k}}(\tau)\right]\right\}, \tag{144}$$

with commutation relations $[\hat{u}_{\boldsymbol{k}}(\tau), \hat{p}_{\boldsymbol{k}'}(\tau)] = i\delta^3(\boldsymbol{k}+\boldsymbol{k}')$ and $[\hat{u}_{\boldsymbol{k}}(\tau), \hat{p}^\dagger_{\boldsymbol{k}'}(\tau)] = i\delta^3(\boldsymbol{k}-\boldsymbol{k}')$. We study the evolution of the system in the Heisenberg picture. To this end, we introduce the time-dependent ladder operators

$$\hat{u}_{\boldsymbol{k}}(\tau) = \frac{1}{\sqrt{2k}}\left[a_{\boldsymbol{k}}(\tau) + a^\dagger_{-\boldsymbol{k}}(\tau)\right], \tag{145}$$

$$\hat{p}_{\boldsymbol{k}}(\tau) = -i\sqrt{\frac{k}{2}}\left[a_{\boldsymbol{k}}(\tau) - a^\dagger_{-\boldsymbol{k}}(\tau)\right], \tag{146}$$

which have equal-time commutation relations $[a_{\boldsymbol{k}}(\tau), a^\dagger_{\boldsymbol{k}'}(\tau)] = \delta^3(\boldsymbol{k}-\boldsymbol{k}')$. These equations follow from the usual form of the lowering operator

$$a_{\boldsymbol{k}}(\tau) = \sqrt{\frac{k}{2}}\hat{u}_{\boldsymbol{k}}(\tau) + \frac{i}{\sqrt{2k}}\hat{p}_{\boldsymbol{k}}(\tau), \tag{147}$$

after noticing that

$$a^\dagger_{\boldsymbol{k}}(\tau) = \sqrt{\frac{k}{2}}\hat{u}^\dagger_{\boldsymbol{k}}(\tau) - \frac{i}{\sqrt{2k}}\hat{p}^\dagger_{\boldsymbol{k}}(\tau) = \sqrt{\frac{k}{2}}\hat{u}_{-\boldsymbol{k}}(\tau) - \frac{i}{\sqrt{2k}}\hat{p}_{-\boldsymbol{k}}(\tau), \tag{148}$$

where the last step follows from the reality of $u(\tau, \boldsymbol{x})$ and $p(\tau, \boldsymbol{x})$.

The Hamiltonian is

$$\hat{\mathcal{H}}(\tau) = \frac{M_p^2}{2}\int d^3\boldsymbol{k}\Big\{k\underbrace{\left[a_{\boldsymbol{k}}(\tau)a^\dagger_{\boldsymbol{k}}(\tau) + a^\dagger_{-\boldsymbol{k}}(\tau)a_{-\boldsymbol{k}}(\tau)\right]}_{\text{collection of harmonic oscillators}} + i\tfrac{z'}{z}\underbrace{\left[a^\dagger_{-\boldsymbol{k}}(\tau)a^\dagger_{\boldsymbol{k}}(\tau) - a_{\boldsymbol{k}}(\tau)a_{-\boldsymbol{k}}(\tau)\right]}_{\text{interacting term (pair creation)}}\Big\}. \tag{149}$$

The first term in square brackets is the standard part describing a collection of free harmonic oscillators. The second term in square brackets, which vanishes in flat space-time, is an interacting term between the scalar field and the classical gravitational background. The interaction is described by the product of two creation operators for the mode $\boldsymbol{k}$ and $-\boldsymbol{k}$ and it represents the production of pairs of quanta with opposite momentum during the cosmological expansion. In terms of the Hubble parameters, we find

$$\frac{z'}{z} = aH(1+\epsilon-\eta) \simeq \begin{cases} aH & \text{slow-roll phase with } \epsilon \simeq \eta \ll 1, \\ aH(1-\eta) & \text{ultra-slow-roll phase with } \epsilon \ll 1. \end{cases} \tag{150}$$

During a standard phase of slow-roll evolution, the relative importance of the interacting term is controlled by the relation between k and aH. For $k \ll aH$, that is, after the mode with comoving wavenumber k leaves the horizon, the interacting term dominates and a copious pair production enhances exponentially the number of quanta in the original Minkowski vacuum that, consequently, undergoes a quantum-to-classical transition. This justifies the standard definition $k_\sigma = \sigma aH$ for the coarse-graining cutoff wavenumber in the standard

slow-roll case. It also justifies our earlier statement that the classicalization condition is equivalent to asking that the curvature perturbation freezes, since the latter occurs also around $k \simeq aH$. When slow-roll is violated, however, the description of the quantum to classical-transition is more involved. In particular, during an ultra-slow-roll phase η can take large values, flipping the sign of the interacting term in eq. (150).

Let us calculate the occupation number density operator to determine the effect of the interacting term in the Hamiltonian. The Heisenberg equations for the creation and annihilation operators can be written in matrix form as

$$\frac{d}{d\tau}\begin{pmatrix} a_{\boldsymbol{k}}(\tau) \\ a^{\dagger}_{-\boldsymbol{k}}(\tau) \end{pmatrix} = \begin{pmatrix} -ik & z'/z \\ z'/z & ik \end{pmatrix}\begin{pmatrix} a_{\boldsymbol{k}}(\tau) \\ a^{\dagger}_{-\boldsymbol{k}}(\tau) \end{pmatrix}. \tag{151}$$

The off-diagonal terms are responsible for particle creation in curved space-time. To solve this system, we use a Bogoliubov transformation. Starting from some given initial condition at time $\tau_\star$, the ladder operators at time τ can be written as:

$$a_{\boldsymbol{k}}(\tau) = y_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau_\star) + w_{\boldsymbol{k}}(\tau)a^{\dagger}_{-\boldsymbol{k}}(\tau_\star), \tag{152}$$

$$a^{\dagger}_{-\boldsymbol{k}}(\tau) = y^{\star}_{\boldsymbol{k}}(\tau)a^{\dagger}_{-\boldsymbol{k}}(\tau_\star) + w^{\star}_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau_\star), \tag{153}$$

with the condition $|y_{\boldsymbol{k}}(\tau)|^2 - |w_{\boldsymbol{k}}(\tau)|^2 = 1$ which follows from the fact that the commutation relations among ladder operators must be preserved if the evolution is unitary. From eqs. (145, 146), we find

$$\hat{u}_{\boldsymbol{k}}(\tau) = u_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau_\star) + u^{\star}_{\boldsymbol{k}}(\tau)a^{\dagger}_{-\boldsymbol{k}}(\tau_\star), \tag{154}$$

$$i\hat{p}_{\boldsymbol{k}}(\tau) = p_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau_\star) - p^{\star}_{\boldsymbol{k}}(\tau)a^{\dagger}_{-\boldsymbol{k}}(\tau_\star), \tag{155}$$

where we have defined

$$u_{\boldsymbol{k}}(\tau) \equiv \frac{1}{\sqrt{2k}}\Big[y_{\boldsymbol{k}}(\tau) + w^{\star}_{\boldsymbol{k}}(\tau)\Big], \tag{156}$$

$$p_{\boldsymbol{k}}(\tau) \equiv \sqrt{\frac{k}{2}}\Big[y_{\boldsymbol{k}}(\tau) - w^{\star}_{\boldsymbol{k}}(\tau)\Big]. \tag{157}$$

It is easy to see that $u_{\boldsymbol{k}}(\tau)$ and $p_{\boldsymbol{k}}(\tau)$ satisfy the following equations

$$u''_k(\tau) + \left(k^2 - \frac{z''}{z}\right)u_k(\tau) = 0, \tag{158}$$

$$i\left[u'_k(\tau) - \frac{z'}{z}u_k(\tau)\right] - p_k(\tau) = 0, \tag{159}$$

with $u_k(\tau_\star) = 1/\sqrt{2k}$ and $p_k(\tau_\star) = \sqrt{k/2}$. Here we have switched notation $u_{\boldsymbol{k}} \to u_k$ and $p_{\boldsymbol{k}} \to p_k$ because the solutions to these equations depend only on $k = |\boldsymbol{k}|$ (similarly for

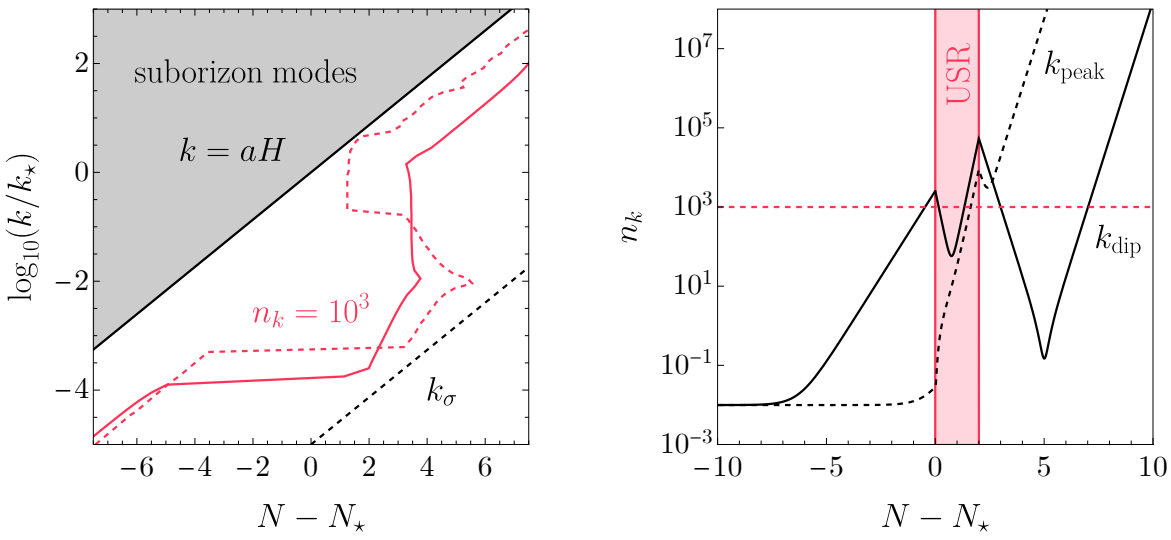

Figure 14: *Left panel: time at which the curvature perturbation freezes outside the horizon for each k (solid red), together with the time at which the occupation number density reaches the value $n_k = 10^3$ dashed red). The cutoff $k_\sigma = \sigma aH$ is also shown for $\sigma = 10^{-5}$ (dashed black). Right panel: evolution of the occupation number density for $k_{\rm peak}$ (dashed black) and $k_{\rm dip}$ (solid black). We have used the analytical model of Section 1.4 with parameters $\eta_1 = 0$, $\eta_2 = 4$, $\eta_3 = -1$, and $\Delta N_2 = 2$.*

y_k and w_k). Notice that from the condition $|y_k(\tau)|^2 - |w_k(\tau)|^2 = 1$ we have, by means of eqs. (156, 157), the condition

$$i\left[\frac{du_k(\tau)}{d\tau}u_k^\star(\tau) - \frac{du_k^\star(\tau)}{d\tau}u_k(\tau)\right] = 1. \tag{160}$$

The strategy now is to solve eqs. (158, 159) and reconstruct, by inverting eqs. (156, 157), the evolution of the ladder operators in eqs. (152, 153). The initial conditions for u_k and p_k follow trivially from eqs. (152, 153). As far as the initial value $\tau_\star$ is concerned, we assume, as customary in the context of inflation, that the system starts in the vacuum state $|0\rangle$ defined by the condition $a_{\boldsymbol{k}}(\tau_\star)|0\rangle = 0$.

The time-dependent occupation number $n_k(\tau)$ is defined, for each mode k, by the expectation value in the original vacuum state of the time-dependent particle number operator $a^\dagger_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau)$,

$$\begin{aligned}\bar{n}_k(\tau) &= \langle 0|a^\dagger_{\boldsymbol{k}}(\tau)a_{\boldsymbol{k}}(\tau)|0\rangle \\ &= \langle 0|\Big[y_k^\star(\tau)a^\dagger_{\boldsymbol{k}}(\tau_\star) + w_k^\star(\tau)a_{-\boldsymbol{k}}(\tau_\star)\Big]\Big[y_k(\tau)a_{\boldsymbol{k}}(\tau_\star) + w_k(\tau)a^\dagger_{-\boldsymbol{k}}(\tau_\star)\Big]|0\rangle \\ &= |w_k(\tau)|^2\delta^3(0),\end{aligned} \tag{161}$$

where the $\delta^3(0)$ has the usual interpretation of spatial volume and arises because we have computed the total number of particles rather than the number density $n_k(\tau)$, which should therefore be defined as $n_k(\tau) \equiv |w_k(\tau)|^2$. By following the aforementioned procedure and using the Wronskian condition (160), the occupation number density can be written as

$$n_k(\tau) = \frac{k}{2}|u_k(\tau)|^2 + \frac{1}{2k}|p_k(\tau)|^2 - \frac{1}{2}. \tag{162}$$

In what follows we will consider the large occupation number condition $n_k(\tau) \gg 1$ as a test for classicality, in accordance with the discussion below eq. (150). By using

$$u_k = \frac{a\dot{\phi}}{M_p H}\mathcal{R}_k, \tag{163}$$

together with the analytical model of Section 1.4 –specifically, eq. (61), together with the fact that $\epsilon \propto e^{-2\eta_i N}$– we obtain, in the i-th region,

$$u_k^{(i)}(\tau) = a\sqrt{2\epsilon}\mathcal{R}_k^{(i)}(\tau), \tag{164}$$

$$p_k^{(i)}(\tau) = ia\sqrt{2\epsilon}\frac{d}{d\tau}\mathcal{R}_k^{(i)}(\tau). \tag{165}$$

The occupation number density is then

$$n_k^{(i)}(\tau) = ka^2\epsilon\left(\left|\mathcal{R}_k^{(i)}(\tau)\right|^2 + \frac{1}{k^2}\left|\frac{d}{d\tau}\mathcal{R}_k^{(i)}(\tau)\right|^2\right) - \frac{1}{2}. \tag{166}$$

The time evolution of this quantity is shown in the right panel of Fig. 14 for two different modes. In the left panel of this figure, we show the time at which curvature perturbations freeze outside the horizon, together with the time at which the occupation number density reaches a value $n_k \gg 1$. Both curves have qualitatively the same shape, confirming the statement that requiring the curvature perturbation to freeze outside the horizon is equivalent to asking that the modes classicalize. As can be seen in the right panel, the exponential growth of n_k that signals the quantum-to-classical transition only occurs after the USR phase has ended, especially for modes close to $k_{\rm dip}$, for which n_k decreases exponentially for a few e-folds after the USR phase. This means that the cutoff $k_\sigma = \sigma aH$ remains a good choice in the presence of a USR phase, but we must be careful to select σ appropriately. As shown in the left panel of Fig. 14, only for $\sigma \lesssim 10^{-5}$ do the modes actually have enough time to classicalize.

3.3 ■ The noise matrix

Now that we have clarified the role of the cutoff choice on the quantum-to-classical transition, we can move on to the calculation of the noise correlation matrix in eq. (141). To this end

we can once again exploit the analytical model of Section 1.4. By combining Einstein's equations and using $u_k = a\phi_k/M_p$, we can show that

$$\phi_k = \sqrt{2\epsilon} M_p \mathcal{R}_k, \tag{167}$$

$$\pi_k = \sqrt{2\epsilon} M_p \left[\frac{d}{dN}\mathcal{R}_k - (\eta + \epsilon)\mathcal{R}_k\right]. \tag{168}$$

Setting $k_\sigma = \sigma a H$ and using $\epsilon \ll 1$, we find the following explicit expressions for the noise correlators in terms of the number of e-folds [3],

$$\Theta^{(i)}_{\phi\phi}(N) = \frac{(\sigma a H)^3}{2\pi^2} 2\epsilon_i M_p^2 |\mathcal{R}^{(i)}_{k_\sigma}(N)|^2, \tag{169}$$

$$\Theta^{(i)}_{\phi\pi}(N) = \frac{(\sigma a H)^3}{2\pi^2} 2\epsilon_i M_p^2 \mathcal{R}^{(i)}_{k_\sigma}(N)^\star \left(\frac{d}{dN}\mathcal{R}^{(i)}_k(N)\Big|_{k_\sigma} - \eta_i \mathcal{R}^{(i)}_{k_\sigma}(N)\right), \tag{170}$$

$$\Theta^{(i)}_{\pi\pi}(N) = \frac{(\sigma a H)^3}{2\pi^2} 2\epsilon_i M_p^2 \left|\frac{d}{dN}\mathcal{R}^{(i)}_k(N)\Big|_{k_\sigma} - \eta_i \mathcal{R}^{(i)}_{k_\sigma}(N)\right|^2. \tag{171}$$

At lowest order in σ the noise correlation functions admit simple analytical expressions if we use eq. (61). The results are shown in Table 3.

$\sigma \to 0$	**Phase 1**	**Phase 2**	**Phase 3**
$\Theta_{\phi\phi}(N)$	$\frac{H^2}{4\pi^2}$	$\frac{H^2}{4\pi^2} e^{-2\eta_2(N-N_\star)}$	$\frac{H^2}{4\pi^2} e^{-2\eta_3(N-\Delta N_2-N_\star)-2\eta_2\Delta N_2}$
$\Theta_{\pi\phi}(N)$	0	$-\frac{H^2}{4\pi^2}\,\eta_2\, e^{-2\eta_2(N-N_\star)}$	$-\frac{H^2}{4\pi^2}\,\eta_3\, e^{-2\eta_3(N-\Delta N_2-N_\star)-2\eta_2\Delta N_2}$
$\Theta_{\pi\pi}(N)$	0	$\frac{H^2}{4\pi^2}\,\eta_2^2\, e^{-2\eta_2(N-N_\star)}$	$\frac{H^2}{4\pi^2}\,\eta_3^2\, e^{-2\eta_3(N-\Delta N_2-N_\star)-2\eta_2\Delta N_2}$

Table 3: *Noise correlation functions in the limit $\sigma \to 0$ [3].*

To solve the stochastic dynamics, we consider, as anticipated in Section 3.1, the expansion of the coarse-grained field about its classical counterpart at first order, namely $\bar{\phi} = \phi_{\rm cl} + \delta\phi_{\rm st}$ and $\bar{\pi} = \pi_{\rm cl} + \delta\pi_{\rm st}$, where $\phi_{\rm cl}$ and $\pi_{\rm cl}$ define the classical trajectory (that is, they are obtained by solving the Langevin equations in the absence of the noise terms, so that they are deterministic quantities and coincide with the background trajectory that would be obtained in standard perturbation theory) and $\delta\phi_{\rm st}$ and $\delta\pi_{\rm st}$ are statistical variables. To interpret $\delta\phi_{\rm st}$ and $\delta\pi_{\rm st}$, therefore, we have to compute their statistical moments. The latter, in full generality, are defined by

$$\langle \delta\phi^n_{\rm st}\delta\pi^m_{\rm st}\rangle_{\rm S}(N) \equiv \int d\bar{\phi} d\bar{\pi}\, [\bar{\phi} - \phi_{\rm cl}(N)]^n [\bar{\pi} - \pi_{\rm cl}(N)]^m P(\bar{\phi}, \bar{\pi}, N), \tag{172}$$

where $P(\Phi, N)$ is the phase-space probability density for the coarse-grained variables $\bar{\phi}$ and $\bar{\pi}$ that solves the Fokker-Planck equation (see e.g. [171–173]. See also Appendix C for a

review of the basic aspects of stochastic differential equations)

$$\frac{\partial P}{\partial N} = -\sum_k \frac{\partial}{\partial \Phi_k}(\mathcal{D}_k P) + \frac{1}{2}\sum_{k\ell} D_{k\ell}\frac{\partial^2 P}{\partial \Phi_k \partial \Phi_\ell}, \tag{173}$$

where $\Phi \equiv (\bar{\phi}, \bar{\pi})^{\mathrm{T}}$. In eq. (173), $\mathcal{D}$ is the drift vector with components

$$\mathcal{D} = \begin{pmatrix} \bar{\phi} \\ -(3-\epsilon_{\bar{\phi}})\left[\bar{\pi} + M_p^2 V_\phi(\bar{\phi})/V(\bar{\phi})\right] \end{pmatrix}, \tag{174}$$

and D is the diffusion matrix

$$D \equiv \frac{\mathbb{1}}{2}\left[\Theta^{(\mathrm{i})}_{\phi\phi}(N) + \Theta^{(\mathrm{i})}_{\pi\pi}(N)\right] + \frac{\sigma_1}{2}\left[\Theta^{(\mathrm{i})}_{\phi\pi}(N) + \Theta^{(\mathrm{i})}_{\pi\phi}(N)\right] + \frac{\sigma_3}{2}\left[\Theta^{(\mathrm{i})}_{\phi\phi}(N) - \Theta^{(\mathrm{i})}_{\pi\pi}(N)\right], \tag{175}$$

where $\sigma_{(1,2,3)}$ are the Pauli matrices. The drift term in the Fokker-Planck equation describes the deterministic part of the dynamics while the diffusion term gives the stochastic one. It is important for what follows to remark that in eq. (174) $\epsilon_{\bar{\phi}}$ indicates the Hubble parameter evaluated on the coarse-grained field $\bar{\phi}$, namely $2M_p^2\epsilon_{\bar{\phi}} = (d\bar{\phi}/dN)^2 = \bar{\pi}^2$.

By using the Fokker-Planck equation, it is possible to write the equations describing the evolution of the statistical moments. Let us focus on the two-point statistical correlators which are relevant for the computation of the power spectrum. We find

$$\frac{d}{dN}\begin{pmatrix} \langle \delta\phi^2_{\mathrm{st}}\rangle_{\mathrm{S}} \\ \langle \delta\phi_{\mathrm{st}}\delta\pi_{\mathrm{st}}\rangle_{\mathrm{S}} \\ \langle \delta\pi^2_{\mathrm{st}}\rangle_{\mathrm{S}} \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ C_\phi & C_\pi & 1 \\ 0 & 2C_\phi & 2C_\pi \end{pmatrix}\begin{pmatrix} \langle \delta\phi^2_{\mathrm{st}}\rangle_{\mathrm{S}} \\ \langle \delta\phi_{\mathrm{st}}\delta\pi_{\mathrm{st}}\rangle_{\mathrm{S}} \\ \langle \delta\pi^2_{\mathrm{st}}\rangle_{\mathrm{S}} \end{pmatrix} + \begin{pmatrix} D_{\phi\phi} \\ D_{\phi\pi} \\ D_{\pi\pi} \end{pmatrix}, \tag{176}$$

where

$$C_\phi = -M_p^2\frac{3-\epsilon_{\mathrm{cl}}}{V^2(\phi_{\mathrm{cl}})}\left[V_{\phi\phi}(\phi_{\mathrm{cl}})V(\phi_{\mathrm{cl}}) - V^2_\phi(\phi_{\mathrm{cl}})\right], \tag{177}$$

$$C_\pi = -3(1-\epsilon_{\mathrm{cl}}) + \pi_{\mathrm{cl}}\frac{V_\phi(\phi_{\mathrm{cl}})}{V(\phi_{\mathrm{cl}})}. \tag{178}$$

where, to be crystal-clear with our notation, we indicate with ϵ_{cl} the Hubble parameter ϵ evaluated on the classical trajectory, namely $2M_p^2\epsilon_{\mathrm{cl}} = (d\phi_{\mathrm{cl}}/dN)^2$. Eq. (176) is of general validity, and is obtained by expanding the components of the drift vector around the classical trajectory and using integration by parts in the Fokker-Planck equation (under the assumption that, by definition, the phase-space probability density P decays fast enough at infinity so that the boundary terms in the integration by parts vanish). Eqs. (177,178) can

be conveniently rewritten in terms of the slow-roll parameters (assuming again that $\epsilon \ll 1$),

$$C_\phi = -(3-\eta)\eta, \tag{179}$$

$$C_\pi = -3. \tag{180}$$

We can once again exploit the analytical model of Section 1.4 to solve this system in each phase by starting from some classical phase-space configuration at some reference initial time. Let us focus on the behavior of the solutions in phase 3. By starting from a classical phase-space configuration at the end of the USR phase $N_{\rm end} = \Delta N_2 + N_\star$ (since in phase 2 the noise decreases exponentially as per Table 3),

$$\langle \delta\phi_{\rm st}^2 \rangle_{\rm S}^{(3)}\big|_{N_{\rm end}} = 0, \qquad \langle \delta\phi_{\rm st}\delta\pi_{\rm st} \rangle_{\rm S}^{(3)}\big|_{N_{\rm end}} = 0, \qquad \langle \delta\pi_{\rm st}^2 \rangle_{\rm S}^{(3)}\big|_{N_{\rm end}} = 0. \tag{181}$$

and using the expressions for the noise at lowest order in σ in Table 3, we find

$$\langle \delta\phi_{\rm st}^2 \rangle_{\rm S}^{(3)} = \frac{H^2}{4\pi^2} e^{2(\eta_3 N_{\rm end} - \eta_2 \Delta N_2)} (N - N_{\rm end}) e^{-2\eta_3 N}, \tag{182}$$

$$\langle \delta\phi_{\rm st}\delta\pi_{\rm st} \rangle_{\rm S}^{(3)} = -\eta_3 \frac{H^2}{4\pi^2} e^{2(\eta_3 N_{\rm end} - \eta_2 \Delta N_2)} (N - N_{\rm end}) e^{-2\eta_3 N}, \tag{183}$$

$$\langle \delta\pi_{\rm st}^2 \rangle_{\rm S}^{(3)} = \eta_3^2 \frac{H^2}{4\pi^2} e^{2(\eta_3 N_{\rm end} - \eta_2 \Delta N_2)} (N - N_{\rm end}) e^{-2\eta_3 N}. \tag{184}$$

One can check that from a generic initial condition the solution will evolve exponentially fast towards eqs. (182, 183, 184). The field diffuses in all three directions in phase-space but, crucially, with very precise relations among the two-point statistical correlators. Before moving on, it is also worth noting that the solution in the first phase is, starting the integration from some arbitrary initial time N_0,

$$\langle \delta\phi_{\rm st}^2 \rangle_{\rm S}^{(1)} = \frac{H^2}{4\pi^2}(N - N_0), \qquad \langle \delta\phi_{\rm st}\delta\pi_{\rm st} \rangle_{\rm S}^{(1)} = 0, \qquad \langle \delta\pi_{\rm st}^2 \rangle_{\rm S}^{(1)} = 0. \tag{185}$$

This is of course nothing but the standard result in slow-roll inflation according to which the inflaton field only diffuses along the $_{\phi\phi}$ direction.

The power spectrum of comoving curvature perturbations is defined by the Fourier transform of the two-point correlation function of $\mathcal{R}$

$$\langle \mathcal{R}(\boldsymbol{x}_1)\mathcal{R}(\boldsymbol{x}_2) \rangle_{\rm S} = \int \frac{d^3\boldsymbol{k}_1}{(2\pi)^3} \frac{d^3\boldsymbol{k}_2}{(2\pi)^3} e^{i\boldsymbol{k}_1\cdot\boldsymbol{x}_1 + i\boldsymbol{k}_2\cdot\boldsymbol{x}_2} \frac{2\pi^2}{k^3} \mathcal{P}_{\mathcal{R}}(k) \delta^3(\boldsymbol{k}_1 + \boldsymbol{k}_2). \tag{186}$$

Since we are only interested in the result on length scales over which the coarse-grained field is homogeneous, we can evaluate the correlators at the same spatial point. The integration

over k will then be limited only to the long wavelength interval $k \in [0, k_\sigma]$. We find

$$\langle \mathcal{R}^2 \rangle_{\rm S} \equiv \langle \mathcal{R}(\boldsymbol{x}) \mathcal{R}(\boldsymbol{x}) \rangle_{\rm S} = \int_0^{k_\sigma} \frac{dk}{k} \mathcal{P}_{\mathcal{R}}(k). \tag{187}$$

If we now change variable from k to the number of e-folds via $k = aH$, so that $dk/k = dN$, and take derivatives on both sides, we find that we can write

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{d}{dN} \langle \mathcal{R}^2 \rangle_{\rm S} = \frac{1}{2M_p^2 \epsilon_{\rm cl}} \left[\frac{d}{dN} \langle \delta\phi_{\rm st}^2 \rangle_{\rm S} + 2\eta_{\rm cl} \langle \delta\phi_{\rm st}^2 \rangle_{\rm S} \right] \Bigg|_{N_\sigma}, \tag{188}$$

where the right-hand side must be evaluated at the time N_σ defined by the condition $k = \sigma aH$ and, as per the discussion in Section 3.2, the cutoff parameter σ must be small enough to allow for classicalization of the modes to occur. A value of $\sigma \lesssim 10^{-5}$ is enough for typical inflection-point models of PBH production, as shown in Fig. 14. If we use eq. (176) to rewrite $d\langle \delta\phi_{\rm st}^2 \rangle / dN$, we find

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{1}{2M_p^2 \epsilon_3} \left[D^{(3)}_{\phi\phi} + 2 \Big(\langle \delta\phi_{\rm st} \delta\pi_{\rm st} \rangle^{(3)}_{\rm S} + \eta_3 \langle \delta\phi_{\rm st}^2 \rangle^{(3)}_{\rm S} \Big) \right] \Big|_{k=k_\sigma}. \tag{189}$$

Let us calculate the power spectrum in eq. (189) analytically in the limit $\sigma \to 0$, in accordance with the above prescription. In phase 3, as a consequence of eqs. (182, 183), we find that $\langle \delta\phi_{\rm st} \delta\pi_{\rm st} \rangle^{(3)}_{\rm S}$ and $\eta_3 \langle \delta\phi_{\rm st}^2 \rangle^{(3)}_{\rm S}$ cancel out. Thus, plugging $\Theta_{\phi\phi}$ from eq. (169) into $D^{(3)}_{\phi\phi}$,

$$\mathcal{P}_{\mathcal{R}}(k) = \frac{(\sigma aH)^3}{2\pi^2} |\mathcal{R}^{(3)}_k(N)|^2 \Big|_{k=k_\sigma}, \tag{190}$$

where, since $k = k_\sigma = \sigma H e^N$,

$$\mathcal{R}^{(3)}_{k_\sigma} = \frac{1}{M_p} \sqrt{\frac{\pi}{8H\epsilon_1}} \left(\frac{k}{\sigma H} \right)^{-(3-2\eta_3)/2} \Big[\alpha_3(k) J_{(3-2\eta_3)/2}(\sigma) + \beta_3(k) J_{-(3-2\eta_3)/2}(\sigma) \Big]. \tag{191}$$

In the $\sigma \to 0$ limit, we therefore have, using eq. (64)

$$\begin{aligned} \lim_{\sigma \to 0} \mathcal{P}_{\mathcal{R}}(k) &= \frac{(\sigma H)^3}{2\pi^2} \frac{1}{M_p^2} \frac{\pi}{8H\epsilon_1} \left(\frac{k}{\sigma H} \right)^{2\eta_3} \left| \beta_3(k) \sigma^{-(3-2\eta_3)/2} \frac{2^{-(3-2\eta_3)/2}}{\Gamma_{\rm E}(\eta_3 - 1/2)} \right|^2 \\ &\propto |\beta_3(k)|^2 \big(k/H \big)^{2\eta_3}, \end{aligned} \tag{192}$$

which coincides exactly with eq. (67). We conclude that the computation of the power spectrum obtained in the context of stochastic inflation matches precisely, even in the presence of an ultra-slow-roll phase, the result obtained by means of the conventional perturbative approach.

The above conclusion is in disagreement with the results of [159], where it was argued that the term proportional to $D^{(3)}_{\phi\phi}$ in eq. (189) reproduces the power spectrum obtained by

solving the Mukhanov-Sasaki equation, and the terms in parentheses yield an additional enhancement to the power spectrum at its peak. The first of these claims is in agreement with the calculation in this section, as long as the cutoff parameter σ is small enough to allow for the modes to classicalize. In the conventional perturbation theory approach, this corresponds to the usual prescription according to which the power spectrum has to be evaluated after the perturbation $\mathcal{R}_k$ freezes to the final constant value that it maintains until its horizon re-entry after the end of inflation. On the other hand, we find that the terms in parentheses in eq. (189) cancel out due to the precise relations among the two-point functions in eqs. (182, 183, 184). One aspect that differentiates the calculation above from that of [159] is that we have used the analytical solution to the Mukhanov-Sasaki equation presented in Section 1.4, whereas a full numerical analysis based on the model of [66] was performed in [159]. To remedy this difference, a numerical analysis based on the polynomial model of Section 2.1 will be performed in the next section.

Before moving on, let us comment on the issue of non-Gaussianities. The above conclusions might naively lead us to think that, since the stochastic inflation approach seems to not have an effect on the power spectrum, the PBH abundance is similarly unaffected. This is not correct, however. The reason is that what enters into the PBH abundance is not the power spectrum per se, but the full probability distribution for the fluctuations, which must be integrated over in order to find the probability of collapse β defined in Section 1.2. One of the advantages of the stochastic approach is that it allows for the calculation of this probability distribution, which can be found by directly solving the Fokker-Planck equation (173). This is the approach that has been pursued in [73, 174], resulting in probability distributions that decay exponentially. Since the integration over the probability distribution is precisely over this exponential tail, these references argue that quantum diffusion heavily impacts the calculation of the PBH abundance. These results have been confirmed numerically in [76]. Although in this thesis we only focus on the calculation of the power spectrum, it would be interesting to extend the analysis of [73] to the class of models discussed here, featuring an epoch of slow-roll violation, using the analytical model in Section 1.4.

3.4 ■ Numerical analysis

In this section we will perform a numerical analysis in the context of the polynomial model discussed in Section 2.1. We remind the reader that the physical, canonically normalized field in this model is h, not ϕ. To be clear about this fact, in this section we use the notation h_k instead of ϕ_k. In the following, we consider the solution which has the abundance shown in green in Fig. 8.

In the left panel of Fig. 15 (compare with the left panel of Fig. 13.) we show for each comoving wavenumber k the transition time (solid black line) after which the corresponding mode freezes to its final constant value. We also show contours of $k_\sigma = \sigma aH$ for different

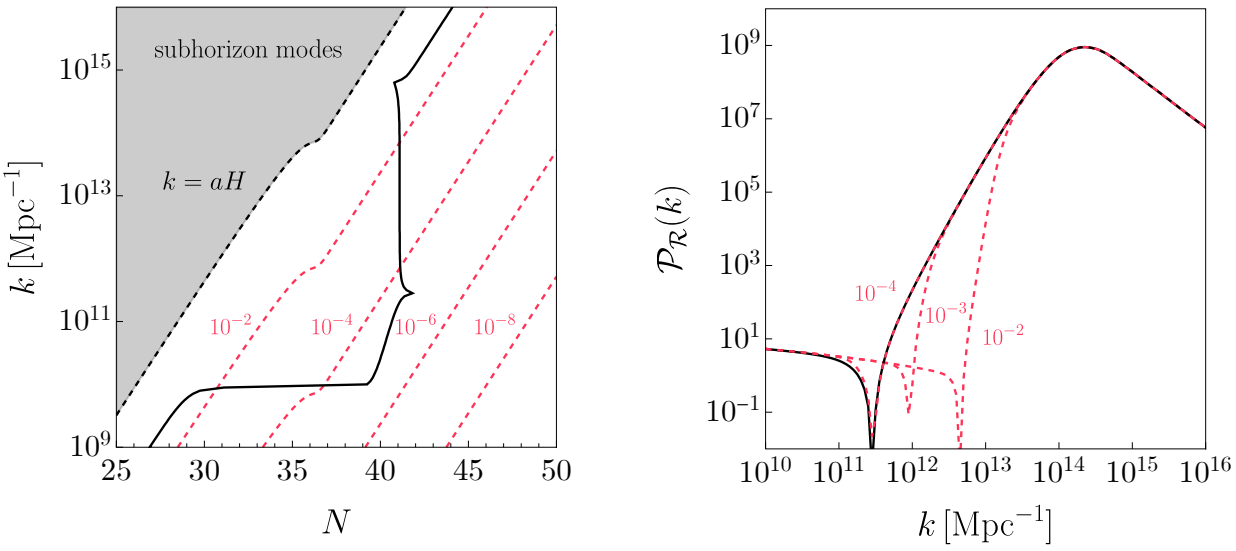

Figure 15: *Left panel: transition time (solid black line) for the comoving wavenumbers k on the y-axis as function of the number of e-folds. This plot refers to the polynomial model discussed in Section 2.1. Diagonal red dashed lines correspond to the condition $k = k_\sigma = \sigma aH$ for different σ. Right panel: power spectrum obtained for the same model by solving numerically the Mukhanov-Sasaki equation (solid black line). We also show the value of $D_{hh}/(2M_p^2\epsilon)$ for different σ (red dashed lines). As discussed in the text, this quantity reproduces, in the limit of small σ, the power spectrum computed by means of the Mukhanov-Sasaki equation. Notice that $D_{hh}/(2M_p^2\epsilon)$ is a function of the number of e-folds, but for fixed σ the dependence on N can be translated into a k-dependence as discussed in eq. (188).*

values of σ with $\sigma = 1$ corresponding to the horizon crossing condition. In the right panel of the same figure we show the power spectrum obtained by solving numerically the Mukhanov-Sasaki equation (solid black line). This plot should be compared with the right panel of Fig. 13 and shows that, as anticipated, the analytical approximation captures all relevant features of the numerical solution.

The elements of the diffusion matrix in eq. (175) are defined in terms of the noise correlators in eq. (141) which take the form

$$\Theta_{fg}(N) = \frac{(1-\epsilon)}{2\pi^2} k_\sigma^3 f_{k_\sigma}(N) g^\star_{k_\sigma}(N), \tag{193}$$

with $f_k, g_k = h_k, \pi_k$ and $k = k_\sigma = \sigma aH$ (where now we no longer consider H as a constant, but rather as a function of time that must be determined numerically by solving the background equations). Part of the complexity of the numerical calculation is that the σ parameter must be kept finite. The fact that the relation $k = \sigma aH$ must hold implies that if the comoving wavenumber k is fixed, then to each N there will correspond a σ such that the relation is verified. The outcome of this procedure is illustrated in Fig. 16, where we plot the

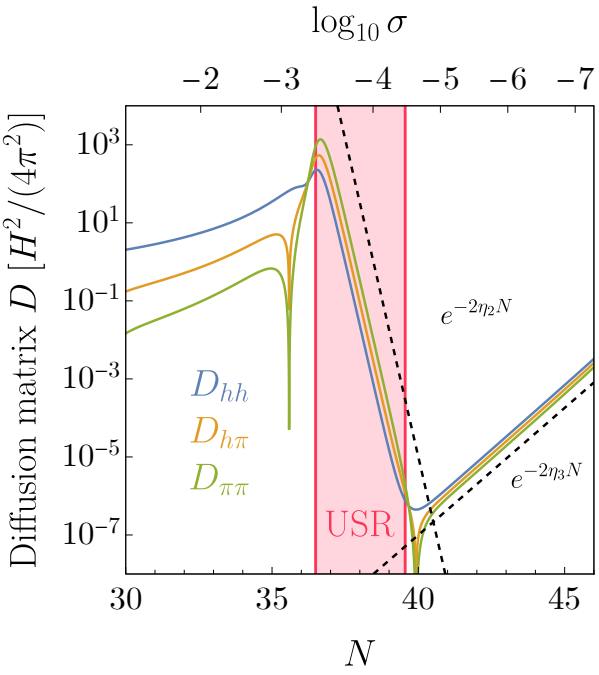

Figure 16: *Elements of the diffusion matrix $D(N)$ in eq. (175) computed for the numerical model discussed in Section 2.1. We fix $k = 5 \times 10^{10}$ Mpc^{-1} and show our results as function of N.*

components of $D(N)$ as a function of N for the representative value $k = 5\times10^{10}$ Mpc^{-1}. This choice corresponds to a mode which crosses the horizon at $N \simeq 28$, before the beginning of the USR phase. The value of σ corresponding to each N is shown on the top x axis. To make the relationship between the top and bottom axes clearer, one can just draw a horizontal line in the left panel of Fig. 15 at the specific value of k analyzed, and see at which N each value of σaH is crossed. In the first phase (that is for $N \lesssim 36$, before the beginning of the USR phase) we have a sizable value of σ. Consequently, the approximations given in Table 3 are not valid, since they are obtained in the $\sigma \to 0$ limit (and for constant H). On the other hand, for $k = 5 \times 10^{10}$ Mpc^{-1} the value of N at which the USR phase begins corresponds to a relatively small value of σ, namely $\sigma \sim 10^{-5}$. This value is so small that, from this point on, our analytical approximations (valid in the $\sigma \to 0$ limit) are now perfectly recovered (see Table 3 and compare with the dotted lines in Fig. 16). These results therefore fully confirm the analytical calculation in the previous section. One of the main messages here is that in the presence of a USR phase, the entries of the noise matrix cannot be taken as $D_{hh} = H^2/4\pi^2$ and $D_{h\pi} = D_{\pi\pi} = 0$. Rather, the noise decreases exponentially as per the results of Table 3 and Fig. 16.

With the entries of the diffusion matrix in hand, we can now proceed to the calculation of the power spectrum in eq. (189). As anticipated in the previous section, the first term

in this equation can be calculated by evaluating the quantity $D_{hh}/(2M_p^2\epsilon)$ at the time N_σ defined by $k = \sigma aH$ for each k. The value of D_{hh} can be read off immediately for any given N (for $k = 5 \times 10^{10}$ Mpc^{-1}) in Fig. 16. As discussed above, for sufficiently small σ (in particular, during phase 3), the matrix elements follow the evolution $\propto e^{-2\eta N}$. Since this is precisely the way in which ϵ evolves as well, the ratio D_{hh}/ϵ consequently settles to a constant value independent of σ, as it should. This fact can be better appreciated in the right panel of Fig. 15, where we show the resulting value of $D_{hh}/(2M_p^2\epsilon)$ for different choices of σ. As expected, only for $\sigma \lesssim 10^{-5}$ do we obtain a power spectrum consistent with the usual perturbation theory result, since it is only for these values of σ that the modes are given enough time to classicalize, as can be confirmed by inspecting the left panel of Fig. 15. These numerical results confirm what we found analytically in eq. (192).

The terms in parentheses in eq. (189) can be found by solving numerically the system in eq. (176). As for the choice of initial conditions, we start the integration from a classical phase space configuration, as we did in eq. (181), although we have checked that solutions derived from different sets of initial conditions settle to the same value within a few e-folds. As before, it is instructive to fix the value of k and consider the solutions as functions of N (or, equivalently, σ). We find that at late times (for sufficiently small σ) $\langle \delta h_{\rm st} \delta\pi_{\rm st}\rangle$ and $\langle \delta h_{\rm st}^2\rangle$ settle to their asymptotic functional forms $\propto N e^{-2\eta N}$ and the combination $\langle \delta h_{\rm st} \delta\pi_{\rm st}\rangle - (\epsilon - \eta)\langle \delta h_{\rm st}^2\rangle$ vanishes asymptotically as in the analytical model. For the mode under consideration ($k = 5 \times 10^{10}$ Mpc^{-1}) this requires $\sigma \lesssim 10^{-5}$, as for the $D_{hh}/(2M_p^2\epsilon)$ term.

We therefore find that the numerical analysis of the model introduced in Section 2.1 confirms our analytical results. The curvature power spectrum computed in stochastic inflation at the linear order in perturbations matches precisely the result obtained by solving the Mukhanov-Sasaki equation using standard perturbation theory. This result is not surprising per se, since it was already well-established in the context of slow-roll inflation, see e.g. [72]. The non-trivial point of our analysis is that we have extended its validity to the case in which an USR phase is present, which is relevant for the formation of PBHs in inflection point models. As a byproduct of our analysis, we have also clarified the role of the stochastic noise and the issue of the quantum-to-classical transition in this scenario.

CHAPTER 4

Black holes from warm inflation

In this chapter we consider the generation of a large power spectrum from a transient dissipative phase during inflation using the warm inflation framework, in which the inflaton gradually transfers its kinetic energy to lighter degrees of freedom throughout its evolution, to describe the dynamics. In Section 4.1 we present the equations of motion for the background and fluctuations in the warm inflation formalism, as well as the phenomenological parameterization for the dissipative coefficient we consider. In Section 4.2 we rewrite the stochastic equations of motion for the fluctuation as a single, deterministic matrix differential equation for their two-point functions in order to determine the stochastic average of the power spectrum. In Section 4.3 we perform a full numerical analysis by directly solving the stochastic differential equations for the fluctuations for many realizations of the noise, allowing us to determine the probability distribution for the power spectrum. Finally, in Section 4.4 we simplify the equations by making a reasonable set of assumptions in order to solve them analytically, and show that the solution reproduces the most important features of the spectrum.

4.1 ■ Dissipation during inflation

In the standard inflationary scenario it is assumed that the coupling between the inflaton and other fields is such that particles produced throughout inflation are quickly diluted by the expansion. The interaction between both sectors only becomes relevant in the reheating stage, during which the inflaton oscillates around the minimum of its potential, decaying into these other fields and populating the Universe with radiation. This is not the only possible way in which inflation can end, however. A gradual transfer of energy can also occur between the inflaton and other fields as it rolls down its potential. If we assume that these other fields comprise a bath of thermalized radiation, eventually their energy density will overcome that of the inflaton and dominate the Universe, ending inflation without the need for a separate stage of reheating. This scenario is known as warm inflation, and is characterized by the presence of a local dissipation term of the form $\Gamma\dot{\phi}$ in the equation of motion for the inflaton. Although the presence of such a term due to particle production had been considered since the early days of inflation, the warm inflation scenario as we know it today was first proposed in [95, 96]. The first models of warm inflation were unsuccessful due to the fact that it seemed difficult to keep temperature corrections to the potential small enough for it to remain relatively flat while still having a large dissipative coefficient,

issues that were first pointed out in [175]. Models that were able to satisfy these constraints required uncompelling ingredients such as non-standard interactions or a large number of fields [176]. It was then realized that evading these issues would be possible if the inflaton did not couple directly to the light fields that comprised the radiation bath, but rather through heavy catalyst fields [177]. The quantum corrections induced by these heavy fields could be cancelled in supersymmetric models while still inducing particle production, and they therefore became a standard ingredient in later warm inflation models. The microphysics of warm inflation in scenarios of this kind are discussed in Appendix D.

The formation of PBHs in warm inflation was studied in [178, 179], where dissipative coefficients that grow monotonically as inflation progresses were considered, leading to spectra that are not peaked, but rather grow towards the end of inflation, producing black holes that are too light to account for the dark matter of the Universe due to the Hawking radiation bounds on the lower end of the mass range (1). The approach we shall take here is to remain agnostic about the origin of the dissipative coefficient. We will model our scenario using a phenomenological approach, assuming that the dissipation is due to a term $\propto \Gamma\dot{\phi}$ entering into the equation of motion of the inflaton, and such that the dissipative coefficient Γ is a peaked function of the inflaton field –so that we can obtain PBH masses in the unconstrained range (1)– and proportional to the third power of the temperature of the radiation bath. This is the main difference between the model in this thesis and the standard warm inflation scenarios, where the dissipation is assumed to be active throughout the duration of inflation, as opposed to only for a few e-folds.

Let us begin by assuming that during inflation the inflaton gradually transfers its energy to a thermalized radiation bath with energy density ρ_r and temperature T related by

$$\rho_r = \frac{\pi^2}{30} g_\star T^4. \tag{194}$$

At the background level, the energy transfer between both components (inflaton and radiation) is modeled via

$$\ddot{\phi} + (3H + \Gamma)\dot{\phi} + V_\phi = 0, \tag{195}$$

$$\dot{\rho}_r + 4H\rho_r = \Gamma\dot{\phi}^2, \tag{196}$$

$$\rho_r + V + \frac{1}{2}\dot{\phi}^2 = 3M_p^2 H^2. \tag{197}$$

These equations ensure that the total energy-momentum tensor is conserved. If friction dominates the background dynamics, the initial conditions for these equations are irrelevant due to the presence of an attractor, much like in the standard slow-roll scenario. This attractor is characterized by the ratios

$$\epsilon_\phi \equiv -\frac{V_\phi}{(3H + \Gamma)\dot{\phi}} \simeq 1, \qquad \epsilon_\rho \equiv \frac{\Gamma\dot{\phi}^2}{4H\rho_r} \simeq 1. \tag{198}$$

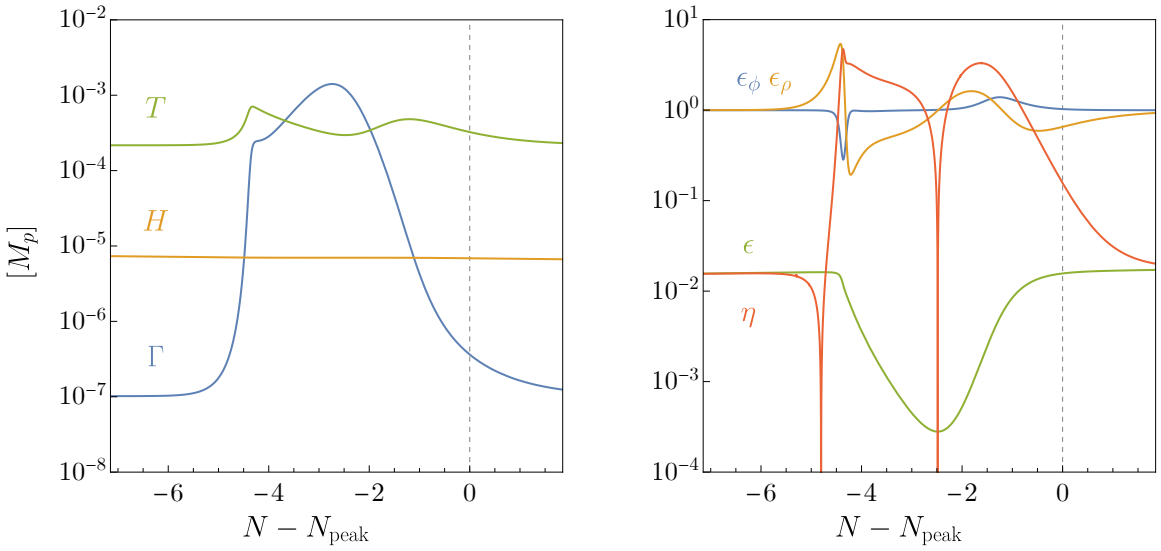

Figure 17: *Left panel: T, H and Γ in units of M_p and as functions of the number of e-folds N, near the value $N = N_{\rm peak}$ at which the peak in the primordial curvature spectrum (shown in Fig. 20) occurs. The analytical expression for Γ as a function of the inflaton background field is given in eq. (211). The parameters chosen for this benchmark example are given in eq. (212). Right panel: slow-roll parameters ϵ and η, and the ratios ϵ_ϕ and ϵ_ρ defined in eq. (198).*

These ratios, along with the slow-roll parameters (15) and the relevant background quantities, are shown in Fig. 17 for a benchmark example of the model we consider here. The figure shows that the dissipative coefficient Γ increases in value, reaching the so-called strongly dissipative regime $\Gamma \gg H$ for a short phase that lasts roughly ~ 4 e-folds. As can be seen in the figure, ϵ remains smaller than 1 throughout this phase and the Hubble parameter remains constant, so that inflation never stops.

Let us move on to the dynamics of the fluctuations. We work in Newtonian gauge and denote by $\delta\phi$ and $\delta\rho^{(r)}$ the perturbations of the inflaton field and the radiation energy density, respectively. We also define

$$\delta q^{(r)} = \frac{4}{3}\rho_r \delta v^{(r)}, \tag{199}$$

where $\delta v^{(r)}$ is the velocity perturbation of the radiation. Einstein's equations are then (see e.g. [180])

$$3H(\dot\psi_k + H\psi_k) + \frac{k^2}{a^2}\psi_k = -\frac{1}{2M_p^2}\left[\delta\rho_k^{(r)} + \dot\phi(\delta\dot\phi_k - \dot\phi\psi_k) + V_\phi\delta\phi_k\right], \tag{200}$$

$$\dot\psi_k + H\psi_k = -\frac{1}{2M_p^2}\left(\delta q_k^{(r)} - \dot\phi\delta\phi_k\right), \tag{201}$$

$$\ddot\psi_k + 4H\dot\psi_k + (2\dot H + 3H^2)\psi_k = \frac{1}{2M_p^2}\left[\frac{1}{3}\delta\rho_k^{(r)} + \dot\phi(\delta\dot\phi_k + \dot\phi\psi_k) - V_\phi\delta\phi_k\right], \tag{202}$$

where the only difference with respect to the standard inflationary scenario in eqs. (20, 21, 20) is the presence of the perturbations of the radiation fluid. The curvature perturbation (23) is now given by

$$\mathcal{R}_k = \frac{H}{\rho + p}\left(\delta q_k^{(r)} - \dot{\phi}\delta\phi_k\right) - \psi_k. \tag{203}$$

The total energy-momentum tensor $T^{\mu\nu} = T^{\mu\nu}_{(\phi)} + T^{\mu\nu}_{(r)}$ is conserved, since [181]

$$\nabla_\mu T^{\mu\nu}_{(\phi)} = Q^\nu, \qquad \nabla_\mu T^{\mu\nu}_{(r)} = -Q^\nu, \tag{204}$$

where the energy transfer vector Q^μ contains a stochastic piece, the form of which is determined by the fluctuation-dissipation theorem (see e.g. [182, 183]),

$$Q_\mu = -\Gamma\, u^\nu \nabla_\nu\phi\, \nabla_\mu\phi + \sqrt{\frac{2\Gamma T}{a^3}}\xi_t \nabla_\mu\phi. \tag{205}$$

In this expression u^ν denotes the 4-velocity of the radiation fluid and $\xi_t \equiv dW_t/dt$, where dW_t is a Wiener increment[30] satisfying $\langle \xi_t(\boldsymbol{x})\xi_{t'}(\boldsymbol{x}')\rangle_{\rm S} = \delta^3(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')$. Here, as in Chapter 3, the brackets $\langle\cdots\rangle_{\rm S}$ denote a stochastic average over different realizations. The linearized equation for $\delta\phi$ is, in Fourier space [183],

$$\begin{aligned}\delta\ddot{\phi}_k + (3H+\Gamma)\delta\dot{\phi}_k + \left(\frac{k^2}{a^2} + V_{\phi\phi} + \dot{\phi}\Gamma_\phi\right)\delta\phi_k \\ + \Gamma_T\frac{\dot{\phi}T}{4\rho_r}\delta\rho_k^{(r)} - 4\dot{\phi}\dot{\psi}_k + (2V_\phi + \Gamma\dot{\phi})\psi_k = \sqrt{\frac{2\Gamma T}{a^3}}\xi_t.\end{aligned} \tag{206}$$

The equations for the radiations perturbations $\delta\rho^{(r)}$ and $\delta q^{(r)}$ are

$$\begin{aligned}\delta\dot{\rho}_k^{(r)} + \left(4H - \Gamma_T\frac{\dot{\phi}^2 T}{4\rho_r}\right)\delta\rho_k^{(r)} - \frac{k^2}{a^2}\delta q_k^{(r)} + \Gamma\dot{\phi}^2\psi_k \\ -4\rho_r\dot{\psi}_k - (\Gamma_\phi\delta\phi_k - 2\Gamma\delta\dot{\phi}_k)\dot{\phi} = -\sqrt{\frac{2\Gamma T}{a^3}}\dot{\phi}\xi_t,\end{aligned} \tag{207}$$

$$\delta\dot{q}_k^{(r)} + 3H\delta q_k^{(r)} + \frac{4}{3}\rho_r\psi_k + \frac{1}{3}\delta\rho_k^{(r)} + \Gamma\dot{\phi}\delta\phi_k = 0. \tag{208}$$

Eqs. (206)–(208), together with one of Einstein's equations, for instance eq. (201), form a complete set for the four variables $\delta\phi$, $\delta\rho^{(r)}$, $\delta q^{(r)}$ and ψ. These equations can be further simplified using the following combination of Einstein's equations:

$$\left(2M_p^2\frac{k^2}{a^2} - \dot{\phi}^2\right)\psi_k + \delta\rho_k^{(r)} + \dot{\phi}\delta\dot{\phi}_k + (V_\phi + 3H\dot{\phi})\delta\phi_k - 3H\delta q_k^{(r)} = 0. \tag{209}$$

Imposing this constraint allows to reduce the number of equations by one, so we can elimi-

[30] See Appendix C for the definition of a Wiener process and a brief review on stochastic differential equations.

nate, for instance, eq. (208). However, we find that not imposing this constraint can be more stable numerically. We use the initial conditions

$$\delta q_k^{(r)} = 0, \qquad \delta \rho_k^{(r)} = 0, \qquad \psi_k = 0, \qquad \delta\phi_k = -\frac{\dot{\phi}}{2M_p\, aH\sqrt{k\epsilon}} e^{-ik\tau}, \tag{210}$$

where we have assumed Bunch-Davies initial conditions for $\delta\phi$. As we will show later, the choice of initial conditions is not very relevant, since the noise term leads to an attractor behaviour for the evolution of the perturbations.

Our main goal is the description of a peak in the curvature power spectrum arising from transient dissipation. The perturbation equations are driven by a source of noise with amplitude $\sim \sqrt{\Gamma T}$, and so they are significantly enhanced whenever Γ is sufficiently large. If the peak of the spectrum of the curvature perturbation is localized around an adequate scale, the PBH mass function will be narrow enough so that their masses fit into the window (1) of interest for dark matter. Therefore, we focus on modeling a dissipative coefficient Γ that satisfies $\Gamma \gg H$ only for a few e-folds. Rather than building a full model of the complete inflationary history we focus on the local description of the dynamics around the relevant region. Although we remain agnostic about the details of the microphysics that gives rise to such a peaked dissipative coefficient, in Appendix D we present a toy example of a Lagrangian with the necessary features that could potentially serve as a basis for future models.

We assume the following parameterization of the dissipative coefficient

$$\Gamma(\phi, T) = \frac{T^3}{m^2 + M^2 \tanh^2\left[(\phi - \phi_\star)/\Lambda\right]}, \tag{211}$$

where, as discussed in Appendix D, the T^3 dependence of Γ arises naturally in a specific low-temperature limit (in which the temperature is much smaller than the mass of the light fields that comprise the radiation bath), which is common in warm inflation. The temperature dependence of Γ is not crucial for the stochastic noise to generate a peak in the primordial power spectrum. A temperature-independent Γ that is peaked as a function of ϕ also produces a similar effect, but the parameterization (211) resembles more closely the actual Γ that may be expected from a concrete Lagrangian in which ϕ couples to other fields.

For our benchmark example of Fig. 17 we choose the following set of parameters:

$$g_\star = 8, \qquad \phi_\star = 22 M_p, \qquad M = 10^{-2} M_p, \qquad m = 1.4 \times 10^{-4} M_p, \qquad \Lambda = 0.1 M_p. \tag{212}$$

We also consider the following inflaton potential

$$V(\phi) = \frac{\lambda}{4}\phi^4, \tag{213}$$

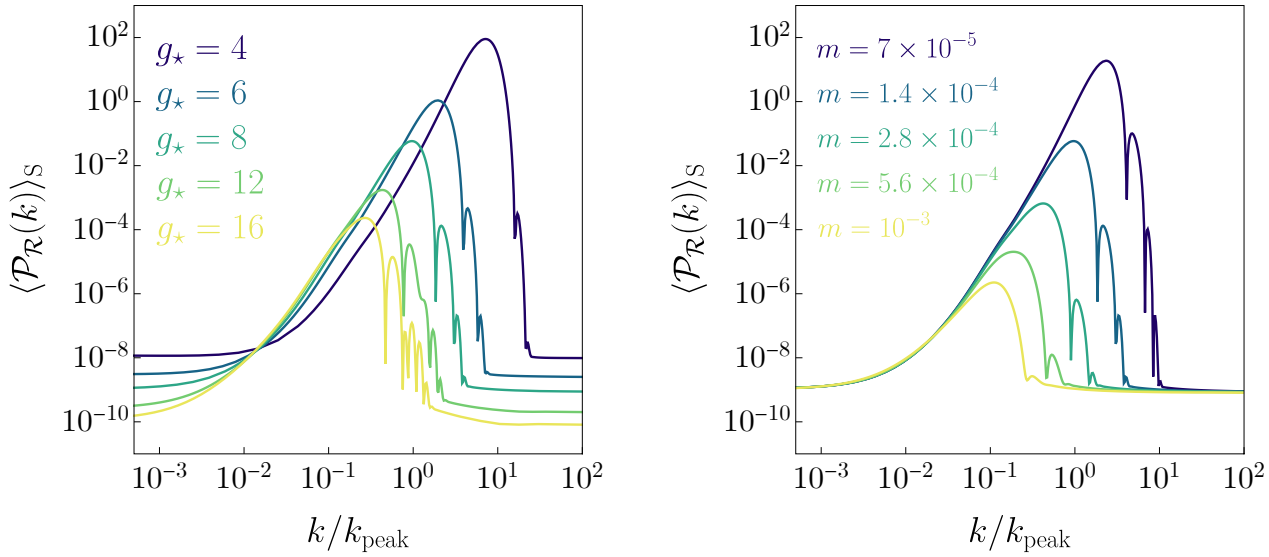

Figure 18: *Stochastic average of the power spectrum computed for the benchmark parameters of eq. (212) by varying $g_\star$ (left panel) and m in units of M_p (right panel). The horizontal axis is normalized at the scale $k = k_{\rm peak}$ at which the peak in the benchmark spectrum (with $g_\star = 8$ and $m = 1.4 \times 10^{-4} M_p$) occurs.*

with $\lambda = 2.5 \times 10^{-15}$. Since our focus is on studying the phenomenology of the dynamics generated by dissipation alone, we could choose any other potential compatible with slow-roll inflation as long as it does not have any peculiar features at the scales we want to analyze. In addition, the potential $V(\phi)$ only needs to be valid a few e-folds before and after the region where $\Gamma \gg H$ because we are only concerned with describing the appearance of a large peak in the primordial power spectrum, which is a local feature. Nevertheless, the value of λ is chosen in such a way that we obtain the correct amplitude for the power spectrum, $A_s \simeq 2 \times 10^{-9}$ [23]. Our choice of parameters leads to a $\mathcal{P}_\mathcal{R}$ with a peak value of $\mathcal{O}(10^{-2})$, which we expect to yield PBHs with an abundance $f_{\rm PBH} \sim \mathcal{O}(1)$ if they form during radiation domination, as per the discussion in Section 1.2.

For the initial conditions of the background variables we choose

$$\phi(N=0) = 26 M_p, \quad \left.\frac{d\phi}{dN}\right|_{N=0} = -\frac{2\sqrt{6}}{\phi_0} M_p^2, \quad \rho_r(N=0) = 10^{-5} M_p^4, \tag{214}$$

although the last two are essentially irrelevant due to the presence of the background attractor discussed in the previous section. This choice makes the background quantities converge quickly to their attractor values. The time at which the localized growth in Γ occurs (and therefore the scale at which the peak in the power spectrum is located) can be controlled by varying $\phi_\star$. Decreasing m or M makes the peak of $\mathcal{P}_\mathcal{R}$ larger. In particular, since we choose $m \ll M$, decreasing m makes $\mathcal{P}_\mathcal{R}$ increase without changing the value of Γ far away from the wavenumbers associated to $\phi_\star$, so that $\mathcal{P}_\mathcal{R}$ retains its normalization at small distance scales. Similarly, increasing Λ makes $\mathcal{P}_\mathcal{R}$ larger. Finally, decreasing $g_\star$ makes the peak of $\mathcal{P}_\mathcal{R}$

larger. To understand why this is the case, let us determine how the coefficient in front of the thermal noise in the equation of motion for the perturbations (206, 207) scales with $g_\star$. This can be done by isolating the temperature dependence of this quantity. Let us define $\gamma(\phi)$ via $\Gamma(\phi, T) = \gamma(\phi)T^3$. Then, by assuming the system is in the attractor solution (198) and using that $\Gamma \gg H$ at the peak, we find

$$\Gamma T \propto \gamma(\phi)\left[\frac{1}{g_\star}\frac{V_\phi^2}{H}\frac{1}{\gamma(\phi)}\right]^{4/7}. \tag{215}$$

We therefore find that decreasing $g_\star$ makes the amplitude of the stochastic noise increase, thereby increasing the curvature power spectrum. The effect of varying m and $g_\star$ on the spectrum is shown in Fig. 18.

4.2 ■ The matrix formalism

Due to the presence of the stochastic thermal noise, the main quantity of interest for us is the expectation value of the power spectrum at a given comoving scale, which we denote by $\langle \mathcal{P}_\mathcal{R}(k)\rangle_\mathrm{S}$. The most straightforward way to compute this quantity (though not necessarily the most economical) is to solve the stochastic equations (201, 206, 207) for a large sample of stochastic realizations, and then calculate their average. Alternatively, one can bypass this procedure by noting that $\langle \mathcal{P}_\mathcal{R}(k)\rangle_\mathrm{S}$ is a deterministic quantity, so that the system of stochastic differential equations can then be rephrased as a single, deterministic, matrix differential equation for the correlators of the scalar fluctuations, in analogy with eq. (176) in the context of stochastic inflation.

Let us begin by noting that the equations of motion can be written, in Fourier space, as a system of linear first-order complex stochastic differential equations. Throughout this section we will work with the number of e-folds as time variable and we define the column vector

$$\Phi \equiv \left(\psi_k, \delta\rho_k^{(r)}, \frac{d\delta\phi_k}{dN}, \delta\phi_k\right)^\mathrm{T}. \tag{216}$$

The equations of motion (201, 206, 207) can then be conveniently written as a system of four first-order stochastic differential equations

$$\frac{d\Phi}{dN} + \boldsymbol{A}\Phi = \boldsymbol{B}\xi_N, \tag{217}$$

where the matrix $\boldsymbol{A}$ and the column vector $\boldsymbol{B}$ are real and independent of Φ. Explicit expressions will be given at the end of this section. We also assume that the constraint in eq. (209) has been imposed to eliminate $\delta q^{(r)}$ from the system. In this equation, ξ_N denotes the Wiener increment from eqs. (206, 207) written in terms of the number of e-folds[31] and

[31] The rule for changing the time variable in the Wiener process is derived in Appendix C.

satisfying, in Fourier space,

$$\langle \xi_N(\boldsymbol{k}) \xi_{N'}(\boldsymbol{k}') \rangle_{\rm S} = (2\pi)^3 \delta(N - N') \delta^3(\boldsymbol{k} - \boldsymbol{k}'). \tag{218}$$

We are interested in the curvature perturbation, which can be written as

$$\mathcal{R} = \boldsymbol{C}^{\rm T} \Phi, \tag{219}$$

where the vector $\boldsymbol{C}$ is can be read off eq. (203). The corresponding power spectrum, averaged over stochastic realizations, can be expressed in terms of the correlation function matrix $\langle \Phi \Phi^\dagger \rangle_{\rm S}$ as

$$\langle \mathcal{P}_\mathcal{R}(k) \rangle_{\rm S} = \frac{k^3}{2\pi^2} \boldsymbol{C}^{\rm T} \langle \Phi \Phi^\dagger \rangle_{\rm S} \boldsymbol{C} \Big|_{k \ll aH}. \tag{220}$$

It makes no difference whether we work with the real and imaginary parts of Φ, or with Φ and its complex conjugate $\Phi^\star$. We now choose the latter option. The probability density $P(\Phi, \Phi^\star, N)$ for the system to be in state $\{\Phi, \Phi^\star\}$ at time N can be obtained by solving the Fokker-Planck equation[32]

$$\frac{\partial P}{\partial N} = \sum_{k\ell} \left[\boldsymbol{A}_{k\ell} \frac{\partial}{\partial \Phi_k} (\Phi_\ell P) + \boldsymbol{A}_{k\ell} \frac{\partial}{\partial \Phi_k^\star} (\Phi_\ell^\star P) + (\boldsymbol{B}\boldsymbol{B}^{\rm T})_{k\ell} \frac{\partial^2 P}{\partial \Phi_k \partial \Phi_\ell^\star} \right]. \tag{221}$$

The two-point statistical moments are defined as

$$\boldsymbol{Q} \equiv \langle \Phi \Phi^\dagger \rangle_{\rm S}(N) \equiv \int \prod_i d\Phi_i \int \prod_j d\Phi_j^\star \; P(\Phi, \Phi^\star, N) \; \Phi \Phi^\dagger. \tag{222}$$

The equation of motion for $\boldsymbol{Q}$ can be found differentiating this equation and using the Fokker-Planck equation.[33] The resulting deterministic differential equation for the matrix $\boldsymbol{Q}$ is

$$\frac{d\boldsymbol{Q}}{dN} = -\boldsymbol{A}\boldsymbol{Q} - \boldsymbol{Q}\boldsymbol{A}^{\rm T} + \boldsymbol{B}\boldsymbol{B}^{\rm T}. \tag{223}$$

By solving this deterministic differential equation we can bypass solving the full system of stochastic differential equations for the perturbations as long as we are only interested in the stochastic average of the power spectrum, which is given by eq. (220).

Let us give explicit expressions for each one of the matrices used in these equations with the number of e-folds as the time variable. The matrix of initial conditions $\boldsymbol{Q}_i \equiv \boldsymbol{Q}(N_{\rm ini})$ is,

[32] The probability density P is a function of two variables (Φ and $\Phi^\star$) which do not obey independent equations of motion (since the noises ξ_N and $\xi_N^\star$ are correlated), so the fact that we can use the Fokker-Planck equation in its canonical form is not obvious. A derivation is performed in Appendix C.

[33] It is also necessary to integrate by parts and use the fact that the probability distribution vanishes on the integration boundaries.

in accordance with eq. (210),

$$\boldsymbol{Q}_i = \frac{1}{2ka^2(N_{\rm ini})}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1+(k/k_i)^2 & -1-i(k/k_i) \\ 0 & 0 & -1+i(k/k_i) & 1 \end{pmatrix}, \tag{224}$$

where k_i is the scale that crosses the horizon at some initial e-fold value $N_{\rm ini}$. In practice, we can start integrating at some time $N_{\rm ini}$ such that $k/k_i \simeq 100$, and terminate the integration a few e-folds after the strongly dissipative phase (in which $\Gamma \gg H$) ends and the mode being computed satisfies $k \ll aH$.

The matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \begin{pmatrix} f_\psi & f_\rho & f_{d\phi} & f_\phi \\ g_\psi + 4\rho_r f_\psi & g_\rho + 4\rho_r f_\rho & g_{d\phi} + 4\rho_r f_{d\phi} & g_\phi + 4\rho_r f_\phi \\ h_\psi + 4(d\phi/dN) f_\psi & h_\rho + 4(d\phi/dN) f_\rho & h_{d\phi} + 4(d\phi/dN) f_{d\phi} & h_\phi + 4(d\phi/dN) f_\phi \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{225}$$

where

$$\begin{aligned}
f_\psi &= 1 + \frac{k^2}{3a^2H^2} - \frac{1}{6M_p^2}\left(\frac{d\phi}{dN}\right)^2, & g_\psi &= \Gamma H\left(\frac{d\phi}{dN}\right)^2 - \frac{k^2}{3a^2}\left[2M_p^2\frac{k^2}{a^2H^2} - \left(\frac{d\phi}{dN}\right)^2\right], \\
f_\rho &= \frac{1}{6M_p^2H^2}, & g_\rho &= 4 - \Gamma_T\frac{HT}{4\rho_r}\left(\frac{d\phi}{dN}\right)^2 - \frac{k^2}{3a^2H^2}, \\
f_{d\phi} &= \frac{1}{6M_p^2}\frac{d\phi}{dN}, & g_{d\phi} &= -\left(\frac{k^2}{3a^2} + 2\Gamma H\right)\frac{d\phi}{dN}, \\
f_\phi &= \frac{V_\phi}{6M_p^2H^2}, & g_\phi &= -\frac{k^2}{3a^2H^2}\left(3H^2\frac{d\phi}{dN} + V_\phi\right) - H\Gamma_\phi\left(\frac{d\phi}{dN}\right)^2, \\
h_\psi &= 2\frac{V_\phi}{H^2} + \frac{\Gamma}{H}\frac{d\phi}{dN}, & h_\rho &= \frac{T\,\Gamma_T}{4H\rho_r}\frac{d\phi}{dN}, \\
h_{d\phi} &= 3 + \frac{\Gamma}{H} + \frac{1}{H}\frac{dH}{dN}, & h_\phi &= \frac{k^2}{a^2H^2} + \frac{V_{\phi\phi}}{H^2} + \frac{\Gamma_\phi}{H}\frac{d\phi}{dN}.
\end{aligned} \tag{226}$$

The vectors $\boldsymbol{B}$ and $\boldsymbol{C}$ are

$$\boldsymbol{B} = \begin{pmatrix} 0 \\ -\sqrt{2\Gamma TH/a^3}\left(\frac{d\phi}{dN}\right) \\ \sqrt{2\Gamma T/(aH)^3} \\ 0 \end{pmatrix}, \qquad \boldsymbol{C} = \frac{1}{3H^2(d\phi/dN)^2 + 4\rho_r}\begin{pmatrix} 2M_p^2k^2/a^2 - 4H^2\left(\frac{d\phi}{dN}\right)^2 - 4\rho_r \\ 1 \\ H^2\left(\frac{d\phi}{dN}\right) \\ V_\phi \end{pmatrix}. \tag{227}$$

It is worth stressing that since the stochastic source ends up dominating the evolution of the perturbations, the choice of initial conditions is in fact not particularly relevant. This fact will be made more clear in Section 4.4, but for now let us illustrate it through a numerical

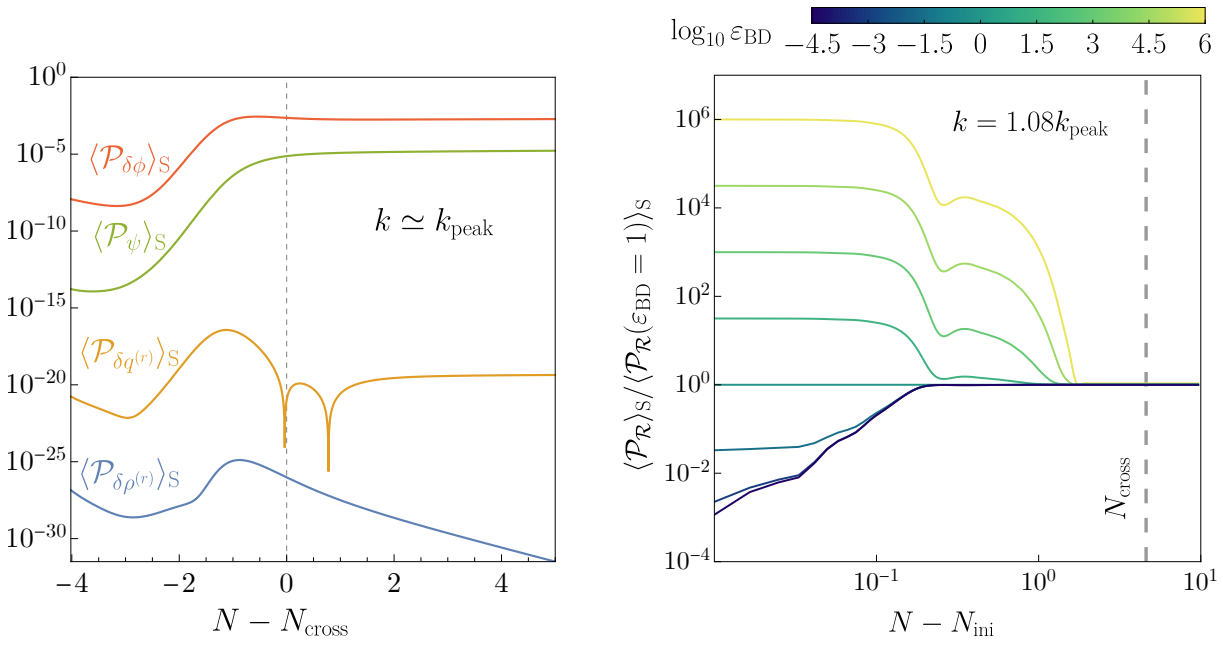

Figure 19: *Left panel: time evolution of the averaged power spectrum of each perturbation as a function of the number of e-folds. Right panel: effect of varying the initial conditions; see the discussion below eq. (228). $N_{\rm cross}$ denotes the time at which the scale k indicated in each panel crosses the horizon ($k = aH$).*

example. Let us parameterize the deviation from the Bunch-Davies initial conditions by multiplying the initial conditions matrix by a small number $\varepsilon_{\rm BD}$,

$$\boldsymbol{Q}_i = \frac{\varepsilon_{\rm BD}}{2ka^2(N_{\rm ini})} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1+(k/k_i)^2 & -1-i(k/k_i) \\ 0 & 0 & -1+i(k/k_i) & 1 \end{pmatrix} . \tag{228}$$

The effect of varying $\varepsilon_{\rm BD}$ with respect to the case $\varepsilon_{\rm BD} = 1$ is shown in the right panel of Fig. 19. Even for very large values of this parameter, $\varepsilon_{\rm BD} \simeq 10^6$, we find that within roughly 1 e-fold (and several e-folds before horizon crossing) the solutions converge to the same value.

As we mentioned earlier, we find that in some cases the system of differential equations is numerically more stable if we do not impose the constraint of eq. (209). This gives rise to an additional equation of motion (for the variable $\delta q^{(r)}$). We have checked that the numerical results using either set of equations are in agreement. The power spectrum obtained from the solution to this matrix differential equation for the benchmark point (212) is shown as a solid line in the left panel of Fig. 20. The evolution of the perturbations for the mode $k_{\rm peak}$ at which the power spectrum peaks is shown in Fig. 19.

4.3 ■ Solving the stochastic equations

In principle, to determine the probability distribution for the stochastic variable Φ, one should solve the Fokker-Planck equation (221), which is a rather difficult task. An alternative consists in estimating numerically the probability distribution by using a frequentist approach, i.e. by solving the system of Langevin equations (217)

$$d\Phi + \boldsymbol{A}\Phi dN = \frac{1}{\sqrt{2}}\boldsymbol{B}\left(dW_N^r + i dW_N^i\right) \tag{229}$$

over many different realizations, where $dW_N^r \equiv \sqrt{2}\mathrm{Re}(\xi_N)dN$ and $dW_N^i \equiv \sqrt{2}\mathrm{Im}(\xi_N)dN$ are real-valued, independent Wiener processes.[34] This is the approach that we adopt in this section.

We impose the following initial conditions, in accordance with eqs. (210, 224)

$$\Phi(N_{\rm ini}) = \left(0, 0, \frac{1}{a(N_{\rm ini})}\frac{i\sqrt{k}}{\sqrt{2}k_i} + \frac{1}{a(N_{\rm ini})}\frac{1}{\sqrt{2k}}, -\frac{1}{a(N_{\rm ini})}\frac{1}{\sqrt{2k}}\right)^{\rm T}. \tag{230}$$

The limits of integration are discussed below eq. (224). We solve the Langevin system with a fixed time-step Runge-Kutta method.[35] The convergence of the solution was checked by successively decreasing the time-step. We found that decreasing the time step below $\Delta N = 10^{-4}$ produces results for the averaged primordial spectrum that are indistinguishable at the percent level.

The left panel of Fig. 20 shows (as light blue dots) a collection of 2160 stochastic realizations of the power spectrum for twenty different values of the wavenumber k. In the same panel, the dark blue dots represent the arithmetic average of all the realizations for each k. The continuous black curve, in turn, corresponds to the numerical solution of the matrix equation (223). The right panel shows the relative difference between the frequentist approach and the matrix formalism solution. The result is a stochastic average which agrees with the matrix formalism results at the percent level.

The Langevin method provides for us not only the means to determine the first moment of the probability distribution of the power spectrum, but with enough sampling we can also determine the full distribution for $\mathcal{P}_{\mathcal{R}}(k)$ at each value of k. The left panel of Fig. 21 shows the normalized histogram for the 2160 realizations for $\log_{10}\mathcal{P}_{\mathcal{R}}(k)$ at $k = 0.8\,k_{\rm peak}$ for illustration. In this same panel we show as a vertical dashed blue line the corresponding expectation value over realizations, and as the vertical red dashed line the mean computed using the matrix formalism (presented in Section 4.2). The continuous black curve corresponds to a skew-normal fit to the (logarithmic) data. A random variable x is skew-normal distributed if its

[34] The factor $\sqrt{2}$ is necessary for the correlation functions of $\mathrm{Re}(\xi_N)$ and $\mathrm{Im}(\xi_N)$ to be properly normalized, as discussed in Appendix C.

[35] We used Wolfram Mathematica and the ItoProcess command to simulate stochastic realizations with the method "StochasticRungeKutta".

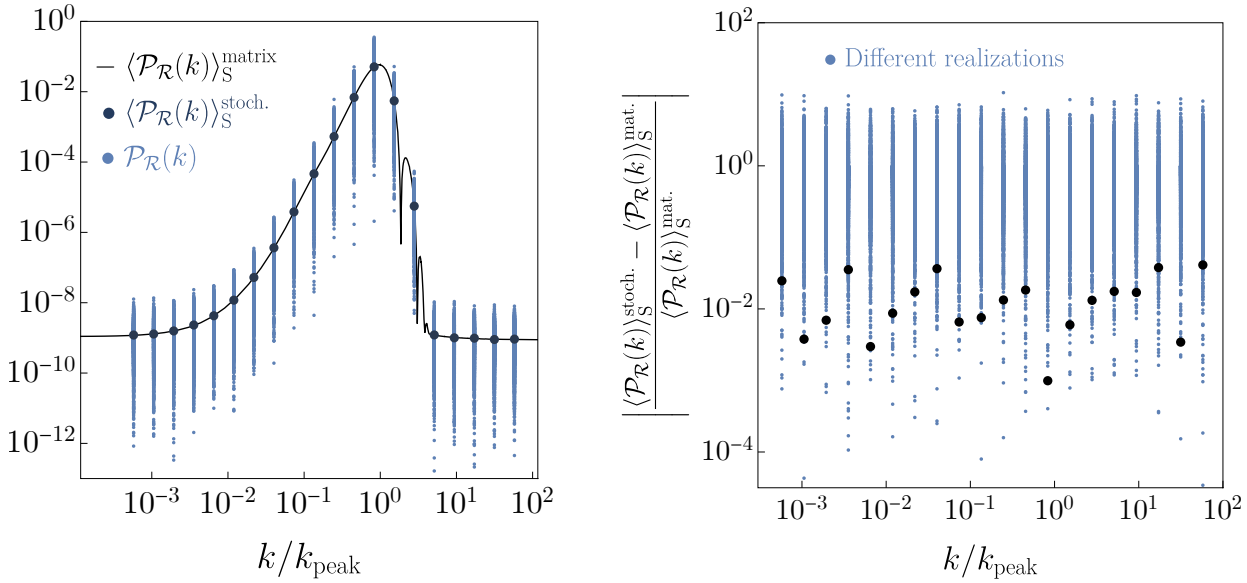

Figure 20: *Left panel: stochastic average of the power spectrum for 20 different values of k (dark blue dots). The number of realizations for each value of k is 2160 (represented as the small, light blue dots). The solid black line represents the average of the power spectrum obtained via the deterministic matrix differential equation derived in Section 4.2. Right panel: absolute value of the relative difference between the stochastic average and the matrix average of the power spectrum. The agreement for each k is at the percent level.*

probability distribution function is given by

$$P_{\text{skew-normal}}(x\,|\,\mu,\sigma,\alpha) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \operatorname{erfc}\left[-\frac{\alpha(x-\mu)}{\sqrt{2}\sigma}\right], \tag{231}$$

where $\operatorname{erfc}(x)$ denotes the complementary error function and $\{\mu,\sigma,\alpha\}$ are free parameters. Therefore, we find that the PDF of $\mathcal{P}_\mathcal{R}$ can be modelled as a *skew-log-normal* distribution. Defining for each k the difference

$$\Delta\log_{10}\mathcal{P}_\mathcal{R} \equiv \log_{10}\mathcal{P}_\mathcal{R} - \log_{10}\langle\mathcal{P}_\mathcal{R}\rangle_\mathrm{S} \tag{232}$$

we find that its probability distribution is very well approximated by a *k-independent* skew-normal distribution. The right panel of Fig. 21 shows the frequentist histogram for the full set of realizations for eq. (232). Together with it we show the corresponding universal skew-normal fit (shown in solid black), with parameters

$$\{\mu,\sigma,\alpha\} = \{0.42, 0.87, -4.15\}. \tag{233}$$

A similar histogram can be created separately for each k, and we find that the standard deviations of the parameters $\{\mu,\sigma,\alpha\}$ for each one of these histograms with respect to the corresponding values for the universal fit shown above are of order $\{3\%, 2\%, 9\%\}$.

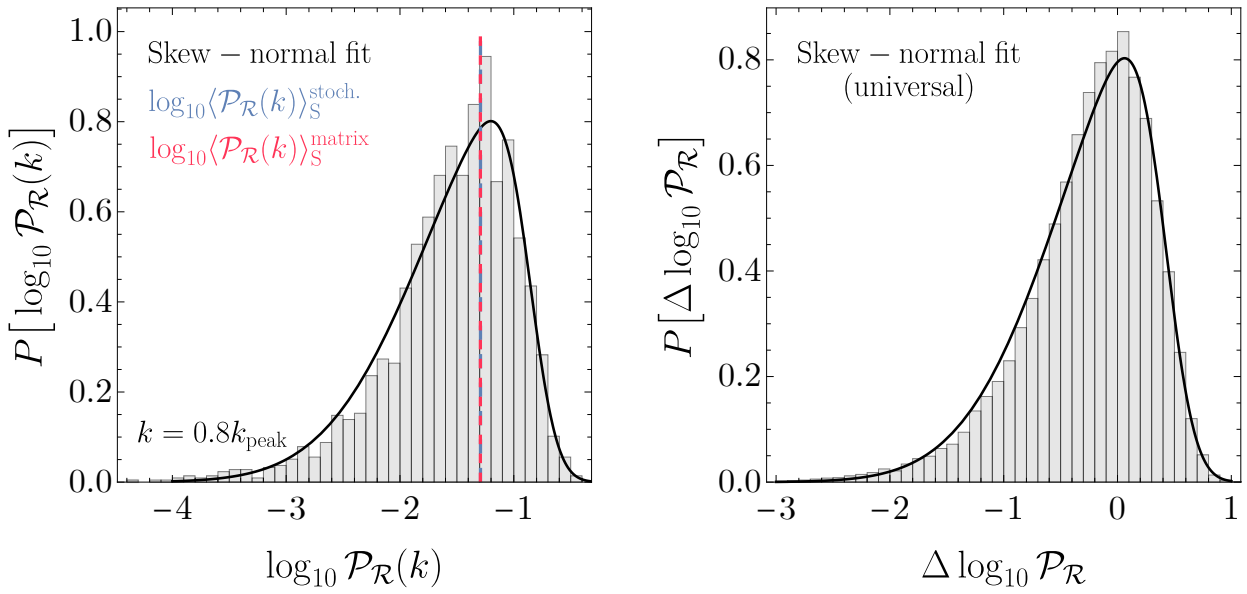

Figure 21: *Left panel: histogram of* $\log_{10}\mathcal{P}_\mathcal{R}(k)$ *for 2160 realizations for* $k = k_\mathrm{peak}$*, together with a skew-normal fit for the probability distribution function. Right panel: histogram of the* k*-independent variable* $\Delta\log_{10}\mathcal{P}_\mathcal{R}$ *defined in eq. (232) using the full set of* 20×2160 *realizations, together with a universal skew-normal fit for the probability distribution function.*

The variance of the probability distribution function for the power spectrum is quite large. From Fig. 21 it is clear that, for a specific realization in a particular Hubble patch the spectrum can reach a value roughly one order of magnitude away from the 10^{-2} value required to obtain $f_\mathrm{PBH} \simeq 1$ if the PBHs form in RD, leading to either overproduction or underproduction of PBHs (according to the Gaussian estimate of the abundance). This effect can always be countered by adjusting any of the parameters in Γ that control the overall size of the average of the power spectrum, as discussed in Section 4.1, as well as the threshold for the collapse (on which the abundance depends exponentially within the Gaussian estimate).

4.4 ■ Analytical approximation

To get a better understanding of the evolution of the perturbations and the shape of the primordial spectrum, it is useful to simplify the equations of motion in such a way that they can be solved analytically. Let us begin by noting that at late times, the only quantity that contributes to the curvature perturbation is $\delta\phi$,

$$\mathcal{R} \simeq -\frac{H\dot{\phi}}{\rho + p}\delta\phi \simeq -\frac{H}{\dot{\phi}}\delta\phi\bigg|_{k\ll aH}. \tag{234}$$

The second observation we make is that, in the equation of motion for $\delta\phi$ (206), we can neglect several terms and still reproduce the most important features of the spectrum,

$$\frac{d^2\delta\phi_k}{dN^2} + \left(3 + \frac{\Gamma}{H}\right)\frac{d\delta\phi_k}{dN} + \left(\frac{k^2}{a^2H^2} + \frac{\Gamma_\phi}{H}\frac{d\phi}{dN}\right)\delta\phi_k + \frac{3}{H^2}\left(\frac{d\phi}{dN}\right)^{-1}\delta\rho_k^{(r)} \simeq 0. \tag{235}$$

This approximation is obtained by discarding terms involving the potential (which are slow-roll suppressed), the metric perturbation (which can be checked numerically to be a good approximation), and assuming that $\Gamma \propto T^3$ and the background remains in the attractor at all times, so that the attractor parameters defined in eq. (198) indeed satisfy $\epsilon_\phi = \epsilon_\rho = 1$. This last approximation is justified in the right panel of Fig. 17, where it can be seen that the background quantities only leave the attractor for very brief periods. In addition, we have found numerically that the stochasticity of the system is encoded in eq. (235) via the $\delta\rho^{(r)}$ term, and therefore the original noise term on the right hand side of eq. (206) can be dropped. Let us explain this last approximation in more detail. If we have an explicit expression for $\delta\rho^{(r)}$ as a function of time, then we can think of the $\delta\rho^{(r)}$ term in eq. (206) as a source term for $\delta\phi$, on the same footing as the ξ_N term. Numerically, we find that the $\delta\rho^{(r)}$ term dominates over the ξ_N term, in the sense that one can set the latter to zero and still correctly reproduce the key features of the spectrum: the location of the peak, and the size of the spectrum; the latter within an order of magnitude of the full numerical result. We remark that this does not mean that the noise ξ_N is irrelevant. In fact, it is precisely the noise in eq. (207) which determines $\delta\rho^{(r)}$, and thus in turn $\delta\phi$. In other words, the enhancement of the power spectrum of the comoving curvature perturbation is due to the source term $\delta\rho^{(r)}$ (the value of which is set by the thermal noise) in the equation of motion for $\delta\phi$.

The strategy we will follow now is to propose a phenomenological parameterization for $\delta\rho^{(r)}$ as a function of time and use it to solve eq. (235). We will also assume that all background quantities can be approximated as piecewise-constant functions. The benchmark values we take for each quantity are shown in Table 4, where we introduce the quantity

$$\Sigma \equiv \sqrt{\frac{9\Gamma T}{H^3}} \tag{236}$$

for later convenience.

These parameters have been chosen in such a way that we obtain a primordial spectrum that closely resembles the one derived with the dissipation coefficient (211) for the parameters in eq. (212). We assume the evolution proceeds in four different phases, which we label from 0 to 3. In phases 0 and 3 we have $\Gamma_\phi = 0$ and $\Gamma \ll H$, so that we are in the weak dissipative regime. In phases 1 and 2 we have $\Gamma \gg H$. During phase 1 we have $\Gamma_\phi > 0$, and during phase 2 we have $\Gamma_\phi < 0$. The evolution of the relevant background quantities using this parameterization is compared to their numerical counterpart, obtained using the parameters in eq. (212), in Fig. (22). In addition, we parameterize the time evolution of the root mean

	Phase 0	**Phase 1**	**Phase 2**	**Phase 3**
$\Sigma\ [M_p^{-1/2}]$	2×10^3	3×10^5	3×10^5	2×10^3
$\Gamma\ [M_p]$	10^{-7}	10^{-3}	10^{-3}	10^{-7}
$\frac{\Gamma_\phi}{H}\frac{d\phi}{dN}$	0	3.4×10^2	-3.4×10^2	0

Table 4: *Benchmark parameters for the analytical calculation of the power spectrum. We also take phases 1 and 2 to end at $N_1 = 2.5$ and $N_2 = 5$, respectively (we measure the number of e-folds from the end of phase 0, so that $N_0 = 0$, and we normalize $a(N_0) = 1$, see Fig. 22), as well as $H = 7\times10^{-6}$.*

square of $\delta\rho^{(r)}$ with the following phenomenological expression,[36]

$$\langle|\delta\rho_k^{(r)}|^2\rangle_{\rm S} \simeq \frac{2\pi^2}{15}g_\star T^5\cdot\begin{cases} e^{-3N} & N<N_f,\\ e^{-2N}e^{\Delta N_f}(H/k) & N>N_f,\end{cases} \tag{237}$$

where the time N_f at which the transition occurs is located a couple of e-folds before the horizon crossing time,

$$N_f \equiv \log(k/H) - \Delta N_f, \tag{238}$$

where ΔN_f is an $\mathcal{O}(1)$, k-independent constant. We take $\Delta N_f = 2.1$ for definiteness. Despite its simplicity (which of course cannot capture all the details of the full numerical solution), this parameterization is enough to understand the most important features of the spectrum. Since $\delta\rho^{(r)}$ is a stochastic variable, it is not sufficient to parameterize its root mean square value, but we also need to know its correlation function. To make progress, we will assume that $\delta\rho^{(r)}$ behaves like a Wiener process,

$$\delta\rho_k^{(r)} = \sqrt{\langle|\delta\rho_k^{(r)}|^2\rangle_{\rm S}}\,\xi_N^{\delta\rho}\,, \tag{239}$$

where the correlation function for $\xi_N^{\delta\rho}$ is

$$\langle\xi_N^{\delta\rho}(\boldsymbol{k})\xi_{\hat N}^{\delta\rho}(\boldsymbol{q})\rangle_{\rm S} = (2\pi)^3\delta(N-\hat N)\delta^3(\boldsymbol{k}+\boldsymbol{q}). \tag{240}$$

The homogeneous solution to eq. (235) can be written as

$$\delta\phi_k^{(h)} = e^{-\nu N}\Big[\delta\phi_+ J_\mu\big(ke^{-N}/H\big) + \delta\phi_- J_{-\mu}\big(ke^{-N}/H\big)\Big], \tag{241}$$

where $\delta\phi_\pm$ are constants fixed by the initial conditions, J_μ is the Bessel function of the first

[36] This expression improves over a similar one proposed in [184].

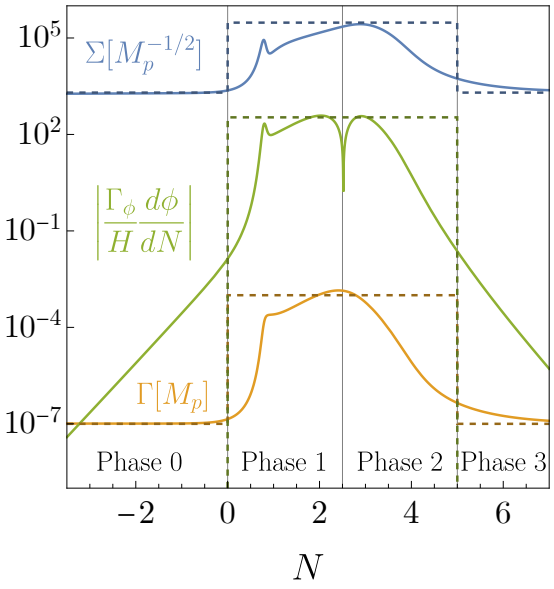

Figure 22: *The dissipative coefficient* Γ*, the quantity* Σ *defined in eq. (236) and the function* $|(\Gamma_\phi/H)d\phi/dN|$ *as functions of the number of e-folds (and in units of* M_p*) for the numerical model with parameters given by eq. (212). The corresponding approximations as piecewise-constant functions from Table 4 are shown with dashed lines.*

kind and

$$\nu = \frac{1}{2}\left(3 + \frac{\Gamma}{H}\right), \quad \mu = \sqrt{\nu^2 - \frac{\Gamma_\phi}{H}\frac{d\phi}{dN}}. \tag{242}$$

This is a straightforward generalization of eq. (61) to the dissipative scenario. This solution and its derivative can also be written in matrix form as

$$\begin{pmatrix} \delta\phi_k^{(h)} \\ \frac{d}{dN}\delta\phi_k^{(h)} \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} \delta\phi_+ \\ \delta\phi_- \end{pmatrix}, \tag{243}$$

where the time-dependent matrix $\boldsymbol{M}$ is given by

$$\boldsymbol{M} = e^{-\nu N} \begin{pmatrix} J_\mu\left(\frac{k}{H}e^{-N}\right) & J_{-\mu}\left(\frac{k}{H}e^{-N}\right) \\ \frac{d}{dN}J_\mu\left(\frac{k}{H}e^{-N}\right) - \nu J_\mu\left(\frac{k}{H}e^{-N}\right) & \frac{d}{dN}J_{-\mu}\left(\frac{k}{H}e^{-N}\right) - \nu J_{-\mu}\left(\frac{k}{H}e^{-N}\right) \end{pmatrix}. \tag{244}$$

The constants μ and ν take different values in each one of the four phases. We denote their values in the j-th phase by μ_j and ν_j. The constants $\delta\phi_\pm$ can be found by imposing continuity of the solution and its derivative at the end of each phase. We denote these constants by $\delta\phi_{\pm j}$ in the j-th phase. We use N_j to refer to the time at which each phase ends. In particular, phase 0 begins at $-\infty$ and ends at $N_0 = 0$, and phase 3 ends at $N_3 = \infty$. To be as general as possible we keep our calculations generic for $n+1$ phases, but we will eventually set $n = 3$.

Following the above procedure we can find the constants in the last phase

$$\begin{pmatrix}\delta\phi_{+n}\\ \delta\phi_{-n}\end{pmatrix} = \left(\prod_{j=1}^{n} \boldsymbol{M}_j^{-1}(N_{j-1})\boldsymbol{M}_{j-1}(N_{j-1})\right)\begin{pmatrix}\delta\phi_{+0}\\ \delta\phi_{-0}\end{pmatrix}. \tag{245}$$

In this equation, terms with smaller j should be placed at the end of the product.[37] The total solution for $\delta\phi$, including both the homogeneous and inhomogeneous solutions is

$$\delta\phi_k = \delta\phi_k^{(h)} + \int_{-\infty}^{N} \frac{\Sigma(\hat{N})}{S(\hat{N})}\frac{G(N,\hat{N})}{a(\hat{N})^{s/2}}\xi_{\hat{N}}^{\delta\rho} d\hat{N}, \tag{246}$$

where G is the Green's function, which we will find below, and

$$s = 3\Theta(N_f - N) + 2\Theta(N - N_f), \tag{247}$$

$$S = \Theta(N_f - N) + \sqrt{\frac{k}{H}}e^{-\Delta N_f/2}\Theta(N - N_f), \tag{248}$$

where Θ is the Heaviside step function. The constants in the homogeneous solution are obtained by imposing Bunch-Davies boundary conditions in the 0-th region,[38]

$$\begin{pmatrix}\delta\phi_{+0}\\ \delta\phi_{-0}\end{pmatrix} = -\frac{1}{2}\sqrt{\frac{\pi}{H}}\begin{pmatrix}1\\ i\end{pmatrix}. \tag{249}$$

The expectation value of the power spectrum at late times is[39]

$$\langle\mathcal{P}_{\delta\phi}(k)\rangle_{\rm S} = \frac{k^3}{2\pi^2}\left[\frac{2^{\mu_n}(k/H)^{-\mu_n}}{\Gamma_{\rm E}(1-\mu_n)}\right]^2|\delta\phi_{-n}|^2 + \frac{k^3}{2\pi^2}\underbrace{\int_{-\infty}^{\infty}\frac{\Sigma(\hat{N})^2}{S(\hat{N})^2}\frac{|G(N\to\infty,\hat{N})|^2}{a(\hat{N})^s}d\hat{N}}_{\mathcal{I}_{\mathcal{P}}}, \tag{250}$$

where $\Gamma_{\rm E}(1-\mu_n)$ denotes Euler's Gamma function evaluated at $(1-\mu_n)$ and Σ was defined in eq. (236).

The Green's function G appearing in the integrand of eq. (250) is

$$G(N,\hat{N}) = \frac{\delta\phi_k^{(1)}(N)\delta\phi_k^{(2)}(\hat{N}) - \delta\phi_k^{(1)}(\hat{N})\delta\phi_k^{(2)}(N)}{\frac{d}{d\hat{N}}\delta\phi_k^{(1)}(\hat{N})\delta\phi_k^{(2)}(\hat{N}) - \frac{d}{d\hat{N}}\delta\phi_k^{(2)}(\hat{N})\delta\phi_k^{(1)}(\hat{N})}, \tag{251}$$

where $\hat{N} < N$ and $\delta\phi_k^{(1,2)}$ are two linearly independent solutions to the homogeneous equation. The calculation of the Green's function is simpler if instead of writing the homogeneous solutions as linear combinations of J_μ and $J_{-\mu}$, we use J_μ and Y_μ (the Bessel function of the

[37] Since these are matrices, the order of the factors is relevant.

[38] We can do this because in this region we are in the weak dissipative regime $\Gamma \ll H$ and thus $\mu \simeq \nu \simeq 3/2$.

[39] See Appendix E for a detailed discussion on the assumptions required to arrive at this result from eq. (246).

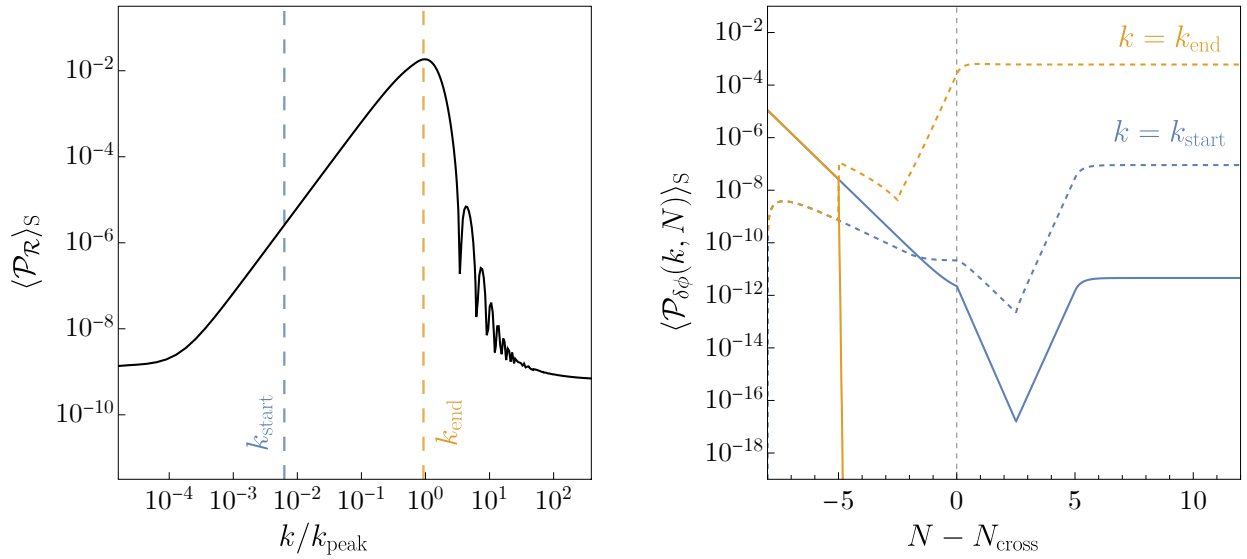

Figure 23: *Left panel: stochastic average of the power spectrum using the analytical approach. Right panel: homogeneous (solid) and inhomogeneous (dashed) solutions –the two terms in eq. (246)– for modes leaving the horizon at the start and the end of the strongly dissipative phase in which* $\Gamma \gg H$*. The inhomogeneous solution, which is independent of initial conditions, always dominates at late times, indicating the presence of an attractor in the equation of motion for the perturbations. The parameters chosen for both panels are shown in Table 4.*

second kind, which is itself a linear combination of J_μ and $J_{-\mu}$). We therefore write

$$\delta\phi_k^{(h)} = e^{-\nu N}\left[\delta\hat{\phi}_+ J_\mu\left(ke^{-N}/H\right) + \delta\hat{\phi}_- Y_\mu\left(ke^{-N}/H\right)\right], \tag{252}$$

which is completely equivalent to eq. (241). Since the Green's function is independent of the boundary conditions chosen for the two linearly independent solutions, we can follow a slightly different procedure from before and arbitrarily choose some linearly independent set of constants in the final region instead of the first. The constants in the previous regions can then be found by matching the solutions and their derivatives at each boundary. We choose $\left(\delta\hat{\phi}_{+n}^{(1)}, \delta\hat{\phi}_{-n}^{(1)}\right) = (0,1)$ and $\left(\delta\hat{\phi}_{+n}^{(2)}, \delta\hat{\phi}_{-n}^{(2)}\right) = (1,0)$ for the two solutions.

The reason for using Y_μ instead of $J_{-\mu}$ and choosing the constants in the final region instead of the first is that we obtain the following simple limits for the two independent solutions at late times,

$$\delta\phi_k^{(1)}(N \to \infty) = -\frac{1}{\pi}\delta\hat{\phi}_{-n}^{(1)}\Gamma(\mu_n)2^{\mu_n}\left(\frac{k}{H}\right)^{-\mu_n}, \tag{253}$$

$$\delta\phi_k^{(2)}(N \to \infty) = 0. \tag{254}$$

If $\hat{N}$ is in the i-th region, the denominator of the Green's function becomes

$$\frac{d}{d\hat{N}}\delta\phi_k^{(1)}(\hat{N})\delta\phi_k^{(2)}(\hat{N}) - \frac{d}{d\hat{N}}\delta\phi_k^{(2)}(\hat{N})\delta\phi_k^{(1)}(\hat{N}) = \frac{2}{\pi}e^{-2\nu_i\hat{N}}\left(\delta\hat{\phi}_{+i}^{(1)}\delta\hat{\phi}_{-i}^{(2)} - \delta\hat{\phi}_{+i}^{(2)}\delta\hat{\phi}_{-i}^{(1)}\right). \quad (255)$$

It is easy to show that this combination of constants is

$$\frac{1}{\mathcal{C}_i} \equiv \left(\delta\hat{\phi}_{+i}^{(1)}\delta\hat{\phi}_{-i}^{(2)} - \delta\hat{\phi}_{+i}^{(2)}\delta\hat{\phi}_{-i}^{(1)}\right) = e^{2(\nu_i-\nu_{i+1})N_i}\frac{1}{\mathcal{C}_{i+1}}. \quad (256)$$

Since in the final region we have $\delta\hat{\phi}_{+n}^{(1)}\delta\hat{\phi}_{-n}^{(2)} - \delta\hat{\phi}_{+n}^{(2)}\delta\hat{\phi}_{-n}^{(1)} = -1$, we obtain the following expression for $\mathcal{C}_i$,

$$\frac{1}{\mathcal{C}_i} = -\prod_{j=i}^{n-1} e^{2(\nu_j-\nu_{j+1})N_j}. \quad (257)$$

Putting everything together, we find that the Green's function at late times is, if $\hat{N}$ is in the i-th region,

$$G(N\to\infty,\hat{N}) = \underbrace{-\Gamma(\mu_n)2^{\mu_n-1}(k/H)^{-\mu_n}}_{\mathcal{B}_n}\delta\phi_k^{(2)}(\hat{N})e^{2\nu_i\hat{N}}\mathcal{C}_i. \quad (258)$$

The integral in eq. (250) then reads

$$\mathcal{I}_{\mathcal{P}} = \sum_{i=0}^{n}\Sigma_i^2\mathcal{B}_n^2\mathcal{C}_i^2\int_{e^{-N_i}}^{e^{-N_{i-1}}}\frac{v^{s-1-2\nu_i}}{S^2}\left[\delta\hat{\phi}_{+i}^{(2)}J_{\mu_i}(kv/H) + \delta\hat{\phi}_{-i}^{(2)}Y_{\mu_i}(kv/H)\right]^2 dv, \quad (259)$$

where we have also switched variables to $v = e^{-N}$. These integrals can be found analytically in terms of hypergeometric functions.

Now that we have all the necessary ingredients we can compute the power spectrum analytically by using eqs. (250, 259) and fixing the parameters as in Table 4. To go from $\delta\phi$ to $\mathcal{R}$ we use eq. (234) in the late time limit, where the ratio between the two is approximately constant. The resulting power spectrum is shown in Fig. 23. The overall size of the peak of the spectrum and the oscillations seen in Fig. 20 are present in the analytical solution. We find that, as with the numerical solution, the peak in the spectrum occurs for modes that leave the horizon around the end of the strongly dissipative phase. This is a consequence of the enhancement being an integrated effect, due to eq. (259), as opposed to a local one.

Having an analytical solution allows us to understand why the initial conditions for the perturbations are irrelevant. All of the information about initial conditions is contained in the homogeneous solution (241) inside the integration constants $\delta\phi_\pm$. However, as shown in the right panel of Fig. 23, this solution is completely negligible at late times. The spectrum is completely dominated by the integral in eq. (259), which is independent of initial conditions. This indicates the presence of an attractor in the equation of motion for the perturbations, as anticipated earlier.

The analytical approximation developed in this section is not enough to reproduce with

accuracy the full averaged primordial spectrum. For instance, the actual slope of the log of the spectrum for $k < k_{\rm peak}$ is about twice the value that the analytical approximation gives. However, it does allow us to understand its features qualitatively, and will be useful in Chapter 5 to estimate the peak value of the gravitational wave signal induced at second order in perturbation theory for this scenario.

CHAPTER 5

Gravitational wave signals

In this chapter we calculate the gravitational wave signals induced at second order in perturbations for each one of the scenarios we have presented throughout the thesis. In Section 5.1 we calculate the solution to the equation of motion of the second-order tensor modes when there is a transition between an early matter-dominated era and a radiation era. In Section 5.2, we examine the gauge-dependence of the result and determine the full gauge-invariant expression for the energy density of gravitational waves through a heuristic argument based on symmetry properties and dimensional analysis. In Section 5.3 we connect the solution for the second-order tensor modes to the energy density of gravitational waves in the aforementioned transition scenario, and examine the bounds on the gravitational wave abundance arising from CMB observations and the abundance of light elements produced during nucleosynthesis. Finally, in Section 5.4 we calculate the gravitational waves induced during and after inflation for the dissipative scenario presented in Chapter 4.

5.1 ■ Induced gravitational waves

The detection of a stochastic background of gravitational waves of primordial origin would have enormous implications for cosmology. At leading order in perturbations and in the absence of anisotropic stress, gravitational waves decouple from other degrees of freedom and propagate freely, carrying information about the mechanism that produced them and the subsequent cosmological history, a fact that makes them a powerful observational tool. Moreover, they can probe processes and energy scales that are difficult to test by other means, such as phase transitions and topological defects in the early Universe [78]. If one goes beyond the leading order description, the stochastic gravitational wave background also becomes a viable tool to test the existence of PBHs. By expanding Einstein's equations to second order, it is possible to show that the tensor degrees of freedom of the metric are sourced by terms quadratic in first-order scalar perturbations [79–81]. This implies that the observable quantity of interest for detectors, the energy density of gravitational waves, is proportional to the square of the power spectrum of curvature perturbations and thus, if PBHs form via gravitational collapse induced by large density fluctuations, we should expect the process to leave an imprint on the stochastic gravitational wave backgrond. Moreover, since the mass of PBHs formed in a radiation-dominated era is related to the peak frequency in the gravitational wave spectrum via eq. (2), if the peak of the mass distribution of PBHs lies in the unconstrained window (1), then future gravitational wave experiments such as

LISA and BBO/DECIGO could potentially detect the signal. Given the fact that, as we have seen, PBH formation during an eMD has significant advantages, it is worth asking whether these facts remain true also in this case.

Throughout this chapter we compute this signal for the different scenarios we have presented in the thesis. These gravitational waves are induced both during and after inflation. In this section we focus on the post-inflationary contribution, which we compute by assuming that a short early matter-dominated era takes place right after inflation. The case in which the Universe enters directly into an radiation-dominated era can be recovered from the results of this section by taking the duration of the eMD era to vanish, as we will see. Thus, the model-independent results derived in this section apply not only for the polynomial model of Section 2.1, where we considered only PBH formation during RD, but also to the monodromy-inspired model of Section 2.3, for which one of the ingredients was the presence of an eMD era. The calculation of the inflationary signal will be performed in Section 5.4 in the context of the warm inflation model presented in Chapter 4. Although gravitational waves induced during an eMD era have been studied before [185], the novel part of this thesis is that we do it in the context of PBH formation by using the results in [92, 93]. We find that due to the fact that the scalar modes of the metric do not decay in time during an eMD era, the signal can be enhanced to the point of violating the bounds on the abundance of gravitational waves today imposed by CMB and BBN observations [98, 99], ruling out part of the parameter space.

The calculation of this signal is not without difficulties, however. It has been pointed out numerous times [84–88] that the transverse-traceless tensor part of the metric at second order is a gauge-dependent variable, and therefore cannot be used to describe gauge-invariant observables such as the energy density of gravitational waves. The issue was addressed in [86, 90], where it was claimed that, despite the above remarks, the transverse-traceless tensor piece of the metric can be used as long as the energy density of gravitational waves is measured late into the radiation era, since in this case the scalars that source the gravitational waves decay quickly in time so that the latter propagate in the same way as the first-order tensor modes do. However, the energy density of gravitational waves should be a gauge-independent quantity at any time, independently of when it is measured. A heuristic derivation of the full expression for the energy density of gravitational waves valid beyond leading order in perturbations will be provided in Section 5.2.

Let us focus, for the time being, on the calculation of the induced gravitational wave signal in the eMD-to-RD transition scenario. We remind the reader that a schematic depiction of the scenario we consider is shown in Fig. 1. Throughout this section we will follow the standard derivation, which can be found, e.g. in [81, 185–187]. The equation of motion for second-order tensor modes obtained by expanding Einstein's equations to second order in

perturbations is, in position space and in the absence of anisotropic stress,

$$\boldsymbol{h}_{ij}^{\mathrm{TT}\prime\prime} + 2\mathcal{H}\boldsymbol{h}_{ij}^{\mathrm{TT}\prime} - \nabla^2\boldsymbol{h}_{ij}^{\mathrm{TT}} = -4\mathcal{T}_{ij}{}^{lm}s_{lm}, \tag{260}$$

where primes denote derivatives with respect to conformal time ($' \equiv d/d\tau$), $\mathcal{H} = a'/a$ denotes the conformal Hubble factor, the bold symbol $\boldsymbol{h}_{ij}^{\mathrm{TT}}$ denotes the second-order transverse, traceless piece of the metric which satisfies $\partial^i\boldsymbol{h}_{ij}^{\mathrm{TT}} = \delta^{ij}\boldsymbol{h}_{ij}^{\mathrm{TT}} = 0$ (see Appendix A for our conventions for the metric perturbations), the source term s_{ij} is given in Newtonian gauge[40] and in the absence of anisotropic stress by (see e.g. [189])

$$s_{ij} = 2\partial_i\psi\partial_j\psi + 4\psi\partial_i\partial_j\psi - \frac{4}{3}\frac{\rho}{\rho+p}\partial_i\left(\frac{\psi'}{\mathcal{H}}+\psi\right)\partial_j\left(\frac{\psi'}{\mathcal{H}}+\psi\right), \tag{261}$$

and $\mathcal{T}_{ij}{}^{lm}$ is the transverse, traceless projector

$$\mathcal{T}_{ij}{}^{lm}s_{lm} = \int\frac{d^3\boldsymbol{k}}{(2\pi)^3}e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\left(\boldsymbol{e}_{ij}^{+}\boldsymbol{e}^{+lm} + \boldsymbol{e}_{ij}^{\times}\boldsymbol{e}^{\times lm}\right)\hat{s}_{lm}(\boldsymbol{k}), \tag{262}$$

where $\hat{s}_{ij}$ denotes the Fourier transform of s_{ij}. The polarization tensors in this expression are

$$\boldsymbol{e}_{ij}^{+} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_i\bar{\boldsymbol{e}}_j - \bar{\boldsymbol{e}}_i\boldsymbol{e}_j), \tag{263}$$

$$\boldsymbol{e}_{ij}^{\times} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_i\bar{\boldsymbol{e}}_j + \bar{\boldsymbol{e}}_i\boldsymbol{e}_j), \tag{264}$$

where $\boldsymbol{e}$ and $\bar{\boldsymbol{e}}$ are two unit vectors satisfying $\boldsymbol{k}\cdot\boldsymbol{e} = \boldsymbol{k}\cdot\bar{\boldsymbol{e}} = \boldsymbol{e}\cdot\bar{\boldsymbol{e}} = 0$, so that $\boldsymbol{h}_{ij}^{\mathrm{TT}}$ is transverse and traceless. In eq. (261) we have also neglected the first-order vector perturbations, which are not typically produced in single-field inflation (in particular, not in the models we consider here), and the first-order tensor modes, which we assume are negligible in comparison to the scalar part of the source.

In what follows we will work in momentum space, where

$$\boldsymbol{h}_{ij}^{\mathrm{TT}}(\boldsymbol{x}) = \int\frac{d^3\boldsymbol{k}}{(2\pi)^3}e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\left(\boldsymbol{h}_k^{+}\boldsymbol{e}_{ij}^{+} + \boldsymbol{h}_k^{\times}\boldsymbol{e}_{ij}^{\times}\right). \tag{265}$$

The equation of motion then becomes

$$\boldsymbol{h}_k^{s\prime\prime} + 2\mathcal{H}\boldsymbol{h}_k^{s\prime} + k^2\boldsymbol{h}_k^{s} = S_k^s, \tag{266}$$

[40] The full expression for the source, valid for all gauges, is shown in eq. (384). For explicit solutions in other gauges, see e.g. [84–89]. See also [188] for solutions in general backgrounds.

where the index s refers to each polarization and

$$S_k^s \equiv -4\boldsymbol{e}_{ij}^s(\boldsymbol{k})\hat{s}_{ij}(\boldsymbol{k}). \tag{267}$$

Writing the Fourier transform of s_{ij} explicitly, we find, assuming the background is dominated by a perfect fluid with $p = w\rho$,

$$S_{k(\text{post})}^s = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})\left[8\psi_p\psi_{k-p} + \frac{16}{3+3w}\left(\psi_p + \frac{1}{\mathcal{H}}\psi_p'\right)\left(\psi_{k-p} + \frac{1}{\mathcal{H}}\psi_{k-p}'\right)\right], \tag{268}$$

where $\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p}) \equiv \boldsymbol{e}_{ij}^s(\boldsymbol{k})p_ip_j$ and we have added the (post) subscript to emphasize the fact that this is the post-inflationary expression for the source.[41] This notation will be used from now on to keep our expressions compact. We also remind the reader that the subindices in the Fourier transforms of the perturbations denote the momenta at which they are evaluated. The solution to eq. (266) is given by

$$\boldsymbol{h}_k^s(\tau) = \int_0^\tau F_{(\text{post})}(\tau,\tau')S_{k(\text{post})}^s(\tau')d\tau', \tag{269}$$

where $F_{(\text{post})}$ is the Green's function for eq. (266) after inflation ends. Since, for the time being, we are only interested in the post-inflationary contribution to the source, in this equation we assume that the amount of gravitational waves induced until the end of inflation is negligible, so we have imposed the initial condition $\boldsymbol{h}_k^s(0) = 0$ where, as is customary, we have fixed the end of inflation at $\tau = 0$. The upper limit τ denotes any posterior time, but we will eventually take it to be today.

If we assume that the transition between the eMD and RD eras is instantaneous, the Hubble factor is given by

$$\mathcal{H} = \begin{cases} \dfrac{2}{\tau} & \text{for } \tau \leq \tau_m, \\ \dfrac{1}{\tau - \tau_m/2} & \text{for } \tau_m \leq \tau, \end{cases} \tag{270}$$

where τ_m is the time at which the transition between eMD and RD occurs. Similarly, the scale factor is

$$a = \begin{cases} a(\tau_m)\left(\dfrac{\tau}{\tau_m}\right)^2 & \text{for } \tau \leq \tau_m, \\ a(\tau_m)\left(\dfrac{\tau - \tau_m/2}{\tau_m/2}\right) & \text{for } \tau_m \leq \tau. \end{cases} \tag{271}$$

The Green's function can be found by computing the homogeneous solutions to eq. (266) in both eras, imposing continuity of the solutions and their derivatives at τ_m to determine the integration constants, and plugging the result into the analogue of eq. (251). The result

[41] As mentioned earlier, the inflationary contribution to the gravitational wave signal will be derived in Section 5.4

is

$$kF_{(\text{post})}(\tau,\tau') = \frac{a(\tau')}{a(\tau)}\frac{\boldsymbol{g}_1(\tau)\boldsymbol{g}_2(\tau') - \boldsymbol{g}_1(\tau')\boldsymbol{g}_2(\tau)}{\boldsymbol{g}_1'(\tau')\boldsymbol{g}_2(\tau') - \boldsymbol{g}_1(\tau')\boldsymbol{g}_2'(\tau')}, \tag{272}$$

where

$$\boldsymbol{g}_i(\tau) \equiv \begin{cases} k\tau[a_i^+ j_1(k\tau) + a_i^- n_1(k\tau)] & \text{for } \tau \leq \tau_m, \\ (k\tau - k\tau_m/2)[b_i^+ j_0(k\tau - k\tau_m/2) + b_i^- n_0(k\tau - k\tau_m/2)] & \text{for } \tau_m \leq \tau. \end{cases} \tag{273}$$

Here, $(a,b)_i^\pm$ are integration constants (not to be confused with the scale factor, which will never have $^\pm$ superscripts) and j_n and n_n are spherical Bessel functions of the first and second kind, respectively. The Green's function formula only requires both solutions to be linearly independent, so we are free to choose the constants in one of the regions (this is analogous to what we did in Section 4.4). The constants in the other region will then be determined by matching the solutions at the boundary, requiring continuity of the functions and their derivatives. We therefore choose $b_1^+ = 1$, $b_1^- = 0$, and $b_2^+ = 0$, $b_2^- = 1$ in the second region, and determine the $a_i^\pm$ to be

$$a_1^+ = \frac{1}{k^2\tau_m^2}\left[\sin\left(\frac{k\tau_m}{2}\right)\left(k^2\tau_m^2 + \cos(k\tau_m)\right) + k\tau_m\cos\left(\frac{k\tau_m}{2}\right)\right], \tag{274}$$

$$a_1^- = \frac{1}{2k^2\tau_m^2}\left[\left(1 - 2k^2\tau_m^2\right)\cos\left(\frac{k\tau_m}{2}\right) + 2k\tau_m\sin\left(\frac{k\tau_m}{2}\right) - \cos\left(\frac{3k\tau_m}{2}\right)\right], \tag{275}$$

$$a_2^+ = -\frac{1}{2k^2\tau_m^2}\left[\left(1 - 2k^2\tau_m^2\right)\cos\left(\frac{k\tau_m}{2}\right) + 2k\tau_m\sin\left(\frac{k\tau_m}{2}\right) + \cos\left(\frac{3k\tau_m}{2}\right)\right], \tag{276}$$

$$a_2^- = \frac{1}{k^2\tau_m^2}\left[\sin\left(\frac{k\tau_m}{2}\right)\left(k^2\tau_m^2 - \cos(k\tau_m) - 1\right) + k\tau_m\cos\left(\frac{k\tau_m}{2}\right)\right]. \tag{277}$$

To perform the integral in eq. (269), we need to relate the value of ψ_k right after inflation ends to the value of the curvature perturbation $\mathcal{R}_k$, which, for the scales of interest, is frozen outside the horizon at this time. The time evolution of ψ_k after the end of inflation is encoded in the transfer function $T_k^\psi(\tau)$, which can be obtained by solving the equation of motion for ψ_k. In the absence of anisotropic stress and in the superhorizon limit, we have the following relation between the time-dependent ψ_k and the frozen curvature perturbation $\mathcal{R}_k$, obtained by straightforward manipulation of Einstein's equations

$$\psi_k(\tau) = \frac{3+3w}{5+3w}T_k^\psi(\tau)\mathcal{R}_k(0). \tag{278}$$

The transfer function is obtained by solving

$$\psi_k'' + 3(1+w)\mathcal{H}\psi_k' + wk^2\psi_k = 0. \tag{279}$$

This equation is obtained, in the absence of anisotropic stress, by manipulating Einstein's

equations and assuming that $\delta p = w\delta\rho$ [189, 190]. The result is, in each epoch,

$$T_k^\psi(\tau) = \begin{cases} a_T^+ + a_T^-(k\tau)^{-5} & \text{for } \tau \leq \tau_m, \\ \dfrac{3\sqrt{3}}{k\tau - k\tau_m/2}\left[b_T^+ j_1\left(\dfrac{k\tau - k\tau_m/2}{\sqrt{3}}\right) + b_T^- n_1\left(\dfrac{k\tau - k\tau_m/2}{\sqrt{3}}\right)\right] & \text{for } \tau_m \leq \tau. \end{cases} \tag{280}$$

The coefficients can be fixed by imposing $T_k^\psi(0) = 1$ and $T_k'^\psi(0) = 0$ and matching the solutions and their derivatives at τ_m. The resulting constants are $a_T^+ = 1$, $a_T^- = 0$, and

$$b_T^+ = \frac{1}{36\sqrt{3}}\left[-\sqrt{3}\left(k^2\tau_m^2 - 36\right)\cos\left(\frac{k\tau_m}{2\sqrt{3}}\right) + 18k\tau_m \sin\left(\frac{k\tau_m}{2\sqrt{3}}\right)\right], \tag{281}$$

$$b_T^- = \frac{1}{36\sqrt{3}}\left[-\sqrt{3}\left(k^2\tau_m^2 - 36\right)\sin\left(\frac{k\tau_m}{2\sqrt{3}}\right) - 18k\tau_m \cos\left(\frac{k\tau_m}{2\sqrt{3}}\right)\right]. \tag{282}$$

Plugging the source (268) into the solution (269) and using eq. (278), we obtain the following expression, valid only in the Newtonian gauge,

$$\boldsymbol{h}_k^s(\tau) = \frac{1}{k^2}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})\mathcal{R}_p(0)\mathcal{R}_{k-p}(0)I(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|), \tag{283}$$

where we have defined the dimensionless quantity

$$I(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|) \equiv \int_0^\tau kF_{(\mathrm{post})}(\tau,\tau')\left[\frac{3+3w(\tau')}{5+3w(\tau')}\right]^2 Q(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|)kd\tau', \tag{284}$$

with

$$Q(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|) \equiv 8T_p^\psi T_{k-p}^\psi + 4\left(T_p^\psi + \frac{1}{\mathcal{H}}T_p^{\psi\prime}\right)\left(T_{k-p}^\psi + \frac{1}{\mathcal{H}}T_{k-p}^{\psi\prime}\right). \tag{285}$$

The integral in eq. (284) can be performed analytically. We write the result as

$$I = \frac{1}{a(\tau)}\left[\cos\left(k\tau - \frac{k\tau_m}{2}\right)I_1 + \sin\left(k\tau - \frac{k\tau_m}{2}\right)I_2\right], \tag{286}$$

where we assume $\tau > \tau_m$ (since we will eventually be interested in the gravitational wave energy density today) and I_1 and I_2 are defined as

$$I_i \equiv 2\frac{a(\tau_m)}{k\tau_m}\left[I_i^{\mathrm{eMD}} + I_i^{\mathrm{RD}}(\tau\to\infty)\right]. \tag{287}$$

The four quantities $I_{(1,2)}^{\mathrm{eMD}}$ and $I_{(1,2)}^{\mathrm{RD}}$ are complicated functions of k and τ_m, computed for the first time in [185] and given in full detail with our conventions in Appendix F. In principle I_i^{RD} should be a function of τ, but since we are only interested in this quantity at late times (because we want the abundance of gravitational waves today) and the transfer functions T_k^ψ decay quickly in time we can simply evaluate it in the $\tau\to\infty$ limit to simplify calculations,

as is standard in the literature, see e.g. [186]. In other words, because of the fast decrease of T_k^ψ, it the integral in I quickly converges and therefore it makes no difference whether the upper limit is set to today (which, strictly speaking, would be the correct choice) or infinity.

5.2 ■ Gauge dependence of the signal

In this section we will explicitly show that the transverse-traceless, second-order tensor mode $\boldsymbol{h}_{ij}^{\rm TT}$ is not a gauge-invariant variable. We will then provide a heuristic derivation of the full gauge-invariant expression for $\Omega_{\rm GW}$ beyond leading order in perturbations. A brief review of the relevant aspects of perturbation theory can be found in Appendix A.

Under an infinitesimal coordinate transformation $x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu$ with $\xi^\mu = (\alpha, \delta^{ij}\partial_j\beta)$ (we set the first-order vector piece β_i of the transformation to zero, since we are neglecting vector modes), the transverse-traceless second-order tensor mode $\boldsymbol{h}_{ij}^{\rm TT}$ transforms as

$$\boldsymbol{h}_{ij}^{\rm TT} \to \tilde{\boldsymbol{h}}_{ij}^{\rm TT} = \boldsymbol{h}_{ij}^{\rm TT} + \mathcal{T}_{ij}{}^{lm}\Sigma_{lm}, \tag{288}$$

with [189]

$$\begin{aligned}\Sigma_{ij} = {}& 2(2\mathcal{H}^2 + \mathcal{H}')\alpha^2\delta_{ij} + 2\mathcal{H}(\alpha\alpha' + \partial_k\alpha\partial^k\beta)\delta_{ij} + 4\alpha(\partial_i\partial_j E' - \psi'\delta_{ij}) \\ & + 8\alpha\mathcal{H}(\partial_i\partial_j E - \psi\delta_{ij}) + 4\partial_k(\partial_i\partial_j E - \psi\delta_{ij})\partial^k\beta + 4(\partial_i\partial_k E - \psi\delta_{ik})\partial^k\partial_j\beta \\ & + 4(\partial_k\partial_j E - \psi\delta_{kj})\partial_i\partial^k\beta + 2(\partial_i B\partial_j\alpha + \partial_j B\partial_i\alpha) + 8\mathcal{H}\alpha\partial_i\partial_j\beta \\ & - 2\partial_i\alpha\partial_j\alpha + 2\partial_k\partial_i\beta\partial^k\partial_j\beta + 2\alpha\partial_i\partial_j\beta' + 2\partial_i\partial_j\partial_k\beta\partial^k\beta + 2\partial_i\partial_k\beta\partial^k\partial_j\beta \\ & + \partial_i\beta'\partial_j\alpha + \partial_j\beta'\partial_i\alpha,\end{aligned} \tag{289}$$

where the transformation goes from one arbitrary gauge to another and the perturbations E, B and ψ are defined in the starting gauge. The perturbations E and B, defined in Appendix A, are the scalar spatial-spatial and spatial-temporal components of the metric perturbation, which are set to zero in the Newtonian gauge. In momentum space, we have

$$\mathcal{T}_{ij}{}^{lm}\Sigma_{lm} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\left(\Sigma_k^+ \boldsymbol{e}_{ij}^+ + \Sigma_k^\times \boldsymbol{e}_{ij}^\times\right), \tag{290}$$

where

$$\begin{aligned}\Sigma_k^s = -\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\boldsymbol{e}_{ij}^s(\boldsymbol{k})p_i p_j \Big\{ & 4\alpha_p\sigma_{k-p} + 8\mathcal{H}\alpha_p\Big[E_{k-p} + \beta_{k-p}\Big] + \\ & \boldsymbol{p}\cdot(\boldsymbol{k}-\boldsymbol{p})\beta_p\Big[4E_{k-p} + 2\beta_{k-p}\Big] - 8\psi_p\beta_{k-p} + 2\alpha_p\alpha_{k-p}\Big\},\end{aligned} \tag{291}$$

with $\sigma \equiv E' - B$ being the so-called shear potential, and where we have used the fact that $\boldsymbol{e}_{ij}^s(\boldsymbol{k})\delta^{ij} = \boldsymbol{e}_{ij}^s(\boldsymbol{k})k^i = 0$. Thus, since the solution to the equation of motion is given (in the

Newtonian gauge) by eq. (283), then under a gauge transformation $(\alpha, \delta^{ij}\partial_j\beta)$ we have

$$\boldsymbol{h}_k^s(\tau) = \frac{1}{k^2}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})\mathcal{R}_p\mathcal{R}_{k-p}\Big[I(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|)+I_\Sigma(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|)\Big], \tag{292}$$

where

$$\begin{aligned} I_\Sigma(x,y,z) = &-\frac{2}{yz}\left(\frac{3+3w}{5+3w}\right)^2\Big\{T_\alpha(xy)\Big[T_\alpha(xz)+2T_\sigma(xz)\Big]-4T_\psi(xy)T_\beta(xz) \\ &+4\frac{\mathcal{H}(x)}{k}T_\alpha(xy)\Big[T_E(xz)+T_\beta(xz)\Big]+yzT_\beta(xy)\Big[2T_E(xz)+T_\beta(xz)\Big]\Big\}. \end{aligned} \tag{293}$$

By combining this equation with eq. (265) we can obtain the solution for $h_{ij}^{\rm TT}$ in position space in any arbitrary gauge. The transfer functions T in this equation are defined in an analogous way to (278), with additional powers of k included appropriately in each function in such a way that they are rendered dimensionless [85, 89]. For instance, for T_α,

$$\alpha(\tau) \equiv \frac{3+3w}{5+3w}\frac{1}{k}T_\alpha(\tau)\mathcal{R}_k(0). \tag{294}$$

We also use the notation $T_\psi(xy) \equiv T_p^\psi(\tau)$, and analogously for the rest of the transfer functions. In the above equation we have also defined the convenient set of dimensionless variables

$$x = k\tau, \qquad y = \frac{p}{k}, \qquad z = \frac{|\boldsymbol{k}-\boldsymbol{p}|}{k}. \tag{295}$$

In Appendix G, we use the above formula to calculate the induced tensor modes $\boldsymbol{h}_{ij}^{\rm TT}$ in three illustrative gauges (uniform curvature, uniform expansion and comoving), and find that the results in the first two gauges coincide with the Newtonian calculation of the previous section only at late times (that is, deep into the RD era).

As shown in eq. (288) and exemplified by the calculations in Appendix G, $\boldsymbol{h}_{ij}^{\rm TT}$ is not a gauge-invariant variable, and thus it cannot describe any physical observables. This gauge dependence has been noted numerous times in the literature [84–88]. To determine the correct, physically relevant gauge-invariant variable we first need to specify which observable we want to compute. In our case, the observable is the energy density of gravitational waves.

The energy density of gravitational waves, which can be obtained by coarse-graining Einstein's equations, depends on the tensor power spectrum [100, 101],

$$\Omega_{\rm GW}(\tau,k) = \frac{1}{6}\left(\frac{k}{\mathcal{H}}\right)^2\langle\mathcal{P}_h(\tau,k)\rangle_{\rm W}, \tag{296}$$

where the brackets $\langle\cdots\rangle_{\rm W}$ denote a spacetime average, which in Fourier space can be thought of as an average over many wavelengths, and $\mathcal{P}_h$ denotes the power spectrum of the first-order transverse, traceless tensor modes $h_{ij}^{\rm TT}$. The reason for this average is that gravitational waves can only be defined when there is a clear separation of scales between the rate of vari-

ation of the background and that of the perturbations, and taking the average allows us to extract the long-wavelength part of the energy density of the gravitational waves that acts as a source for the background metric, see [100, 101]. This expression is only valid *at leading order in perturbations*, since it has been derived by neglecting cubic and higher-order terms in Einstein's equations for the perturbations (see e.g. [100, 101, 191, 192]). In previous studies on induced gravitational waves (e.g. [86, 90]), their energy density has been computed beyond leading order in perturbations by simply replacing $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ in the above expression (assuming that the contribution from the first order tensor modes is negligible). This replacement is not justified, since one would also need to include in the calculation other terms arising from the coarse-graining of Einstein's equations at fourth order in perturbations which would contain, in addition to a piece involving the power spectrum of $\boldsymbol{h}_{ij}^{\rm TT}$, pieces quadratic in first order scalars that would cancel the gauge-dependence. The fact that these terms are not included leads to a gauge-dependent answer for $\Omega_{\rm GW}$ because $\boldsymbol{h}_{ij}^{\rm TT}$ is gauge-dependent, which is a manifestation of the fact that the physical degrees of freedom corresponding to gravitational waves are completely contained in the tensor modes only at leading order in perturbations, whereas at subleading orders they are spread across the tensor and scalar modes.[42]

The energy density of gravitational waves should be a gauge-invariant observable and thus it should be possible to derive an expression for it by appropriately coarse-graining Einstein's equations. This, however, is easier said than done, since many subtleties and difficulties arise when one attempts to derive the full expression for $\Omega_{\rm GW}$ from first principles.[43] Instead of following this path, we will provide a heuristic derivation of the full expression for this quantity by using an argument based on symmetry properties and dimensional analysis.

The energy density of gravitational waves is obtained by splitting the metric into a long-wavelength background and a short-wavelength perturbation

$$g_{\mu\nu} = \hat{g}_{\mu\nu} + h_{\mu\nu}, \tag{297}$$

where the background metric varies over a typical length scale L, and the perturbation over a typical scale λ such that $L \gg \lambda$. This perturbation can be identified with the one that arises in standard cosmological perturbation theory as long as we restrict our attention to

[42] We remark that the first-order vector and tensor modes, which we have assumed are negligible, also contribute.

[43] The derivation of eq. (296) has been performed in a way that is not manifestly gauge-invariant [100, 101, 191, 192], so the gauge-invariance of the result at higher orders is not guaranteed. This is due to the fact that defining a unique, covariant averaging procedure in a generic manifold is a non-trivial problem, and it is unclear whether the existing definitions [191–196] in the context of backreaction are equivalent to each other. In addition, eq. (296) is usually derived in vacuum (so scalar perturbations are usually neglected), whereas we assume that the Universe is filled with a perfect fluid. We also expect scalar perturbations to yield additional backreaction terms if included in the calculation (see e.g. [193]), and it is not clear that a clean separation of the scalar and tensor contributions to the backreaction stress-energy tensor is generically possible.

modes with $\mathcal{H} \ll k$.[44] We are interested in finding an effective equation of motion for the background metric that includes the effect of the short-wavelength perturbations. In order to extract the long-wavelength part of the equations, these must be coarse-grained by taking a spacetime average. By expanding Einstein's equations in powers of $h_{\mu\nu}$ and coarse-graining them, we find

$$\begin{aligned}\langle G_{\mu\nu} - \kappa T_{\mu\nu}\rangle_{\rm W} &= \hat{G}_{\mu\nu} - \kappa\hat{T}_{\mu\nu} + \langle G_{\mu\nu}(h) - \kappa T_{\mu\nu}(h)\rangle_{\rm W} + \langle G_{\mu\nu}(h^2) - \kappa T_{\mu\nu}(h^2)\rangle_{\rm W} + \cdots \\ &\equiv \hat{G}_{\mu\nu} - \kappa\hat{T}_{\mu\nu} - \kappa T^{(\rm GW)}_{\mu\nu} \\ &= 0,\end{aligned} \tag{298}$$

where $\kappa \equiv 1/M_p^2$, hatted quantities depend only on the long-wavelength background $\hat{g}_{\mu\nu}$ and the brackets $\langle\cdots\rangle_{\rm W}$ denotes the aforementioned spacetime average. In Fourier space, the brackets can be thought of as an average over several wavelengths, as we mentioned earlier (hence the W subscript) [100, 101, 191, 192]. The average of $\langle G_{\mu\nu}(h) - \kappa T_{\mu\nu}(h)\rangle_{\rm W}$ is then assumed to vanish, since the quantity inside is linear in $h_{\mu\nu}$, which is a short-wavelength quantity. By expanding the Einstein tensor $G_{\mu\nu}(h^2)$ and manipulating the resulting expression [100, 101], one can then show that at leading (quadratic) order in perturbations, the stress-energy tensor of gravitational waves is given by

$$\kappa T^{\rm GW}_{\mu\nu} = \frac{1}{4}\langle\hat{\nabla}_\mu h_{\alpha\beta}\hat{\nabla}_\nu h^{\alpha\beta}\rangle_{\rm W}, \tag{299}$$

which leads to the expression for the energy density in eq. (296). In this expression $\hat{\nabla}_\mu$ denotes the covariant derivative with respect to the background metric. The cubic order term vanishes for the same reason that the linear term does, and the next-to-leading order term, which is the one we are interested in, turns out to be fourth-order in perturbations.

A natural extension of eq. (296) beyond leading order in perturbations is (assuming that the first order tensor modes vanish, $h^{\rm TT}_{ij} = 0$, so that there are no terms mixing $h^{\rm TT}_{ij}$ and $\boldsymbol{h}^{\rm TT}_{ij}$)

$$\Omega_{\rm GW}(\tau, k) = \frac{1}{24}\left(\frac{k}{\mathcal{H}}\right)^2\langle\mathcal{P}_\Theta(\tau, k)\rangle_{\rm W}, \tag{300}$$

for some gauge-invariant second-order variable Θ_{ij} of the form

$$\Theta_{ij} \equiv \boldsymbol{h}^{\rm TT}_{ij} + \mathcal{T}_{ij}{}^{lm}\theta_{lm}, \tag{301}$$

where θ_{ij} is quadratic in first-order scalars. In what follows we will assume that after the coarse-graining the energy density of gravitational waves has this form, examine the possible terms that could appear in θ_{ij}, and argue that they vanish in the Newtonian gauge.

[44] See also [193], where short-wavelength perturbations are shown to act as a source for the long-wavelength ones.

We assume that $p = w\rho$ and the matter stress-energy tensor is of the form

$$\kappa T_{\mu\nu} = (1+w)\kappa\rho\, u_\mu u_\nu + w\kappa\rho g_{\mu\nu}. \tag{302}$$

If we work in the Newtonian gauge, this implies that the only perturbations available for θ_{ij} are ψ, $\kappa\delta\rho$ and δv. The most general θ_{ij} is then given by

$$\theta_{ij} = c_{\rho\rho}\kappa^2\partial_i\delta\rho\partial_j\delta\rho + c_{vv}\partial_i\delta v\partial_j\delta v + c_{\rho v}\kappa\partial_i\delta\rho\partial_j\delta v + c_{\psi v}\partial_i\psi\partial_j\delta v + c_{\rho\psi}\kappa\partial_i\delta\rho\partial_j\psi + c_{\psi\psi}\partial_i\psi\partial_j\psi, \tag{303}$$

where the coefficients c_{ab} depend only on background quantities. We can also have two derivatives acting on a single perturbation, or, in addition to the two spatial derivatives with free indices $\partial_i\partial_j$, time derivatives and contracted spatial derivatives $\partial_k\partial_k$ acting on the perturbations. These terms will have equal or greater dimensions than the ones above and will therefore be irrelevant in what follows.

Via dimensional analysis, we can check that most of the terms in eq. (303) are forbidden. Since the projector $\mathcal{T}$ is dimensionless, so is θ_{ij}, and thus we must have

$$[c_{\rho\rho}] = -6, \quad [c_{vv}] = 0, \quad [c_{\rho v}] = -3, \quad [c_{\psi v}] = -1, \quad [c_{\rho\psi}] = -4, \quad [c_{\psi\psi}] = -2. \tag{304}$$

The quantity in eq. (300) comes from averaging the fourth-order piece of Einstein's equations (since we assume that the first-order tensor modes are negligible),

$$\kappa T^{(\mathrm{GW})}_{\mu\nu} = \langle G_{\mu\nu}(h^4) - \kappa T_{\mu\nu}(h^4)\rangle_{\mathrm{W}}. \tag{305}$$

For eq. (300) to hold, the fourth-order piece of Einstein's equations must contain a term of the form

$$G_{\mu\nu}(h^4) - \kappa T_{\mu\nu}(h^4) \supset \partial_\mu\Theta_{ij}\partial_\nu\Theta^{ij}. \tag{306}$$

The only dimensionful background quantities that can appear in Einstein's equations once we have assumed that the matter stress-energy is that of eq. (302) are $\mathcal{H}$ and its derivatives (and the background energy density $\hat{\rho}$, which can be related to $\mathcal{H}$ through the background equations), and it is easy to check that they always appear in the numerator. Thus, every coefficient c_{ab} with negative dimensions is forbidden and must therefore vanish. The only coefficient that could in principle be nonzero is c_{vv}, but if the square of the c_{vv} term were present (we consider the square because the term in eq. (306) is quadratic in Θ), it would necessarily come from perturbing the matter stress-energy tensor of eq. (302), as no factors of δv arise from perturbing the Einstein tensor. However, it is easy to see from this expression that there is no way to obtain four copies of δv by perturbing the stress-energy tensor, showing that c_{vv} must also vanish. As anticipated below eq. (303), the coefficients of terms with additional derivatives will have even lower negative dimensions, so they are also forbidden.

The above argument shows that in the Newtonian gauge[45] $\Theta_{ij} = \boldsymbol{h}_{ij}^{\rm TT}$. In principle, in a generic gauge there are many possible candidates for Θ_{ij}, but we will now show that it can in fact be uniquely determined. It is easy to check that the following combination is gauge-invariant,

$$X_{ij} \equiv \boldsymbol{h}_{ij}^{\rm TT} + \mathcal{T}_{ij}{}^{lm}\chi_{lm}, \tag{307}$$

where, in momentum space, we have

$$\mathcal{T}_{ij}{}^{lm}\chi_{lm} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \Big(X_k^+ \boldsymbol{e}_{ij}^+ + X_k^\times \boldsymbol{e}_{ij}^\times \Big), \tag{308}$$

with[46]

$$X_k^s = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \boldsymbol{e}_{ij}^s p_i p_j \Big[2\boldsymbol{p}\cdot(\boldsymbol{k}-\boldsymbol{p}) E_p E_{k-p} - 8 E_p \psi_{k-p} + 2\sigma_p \sigma_{k-p} \Big]. \tag{309}$$

It can also be seen directly from the above expression that in the Newtonian gauge $X_{ij} = \boldsymbol{h}_{ij}^{\rm TT}$. Thus, $X_{ij} - \Theta_{ij} = 0$ in the Newtonian gauge. But since both X_{ij} and Θ_{ij} are gauge-invariant, their difference must be gauge-invariant as well, and since the difference vanishes in the Newtonian gauge, it must vanish in every gauge.[47] We have therefore shown that, as long as our initial assumption holds –namely, that the energy density of gravitational waves beyond leading order in perturbations can be obtained by averaging the power spectrum of a gauge-invariant variable of the form shown in eq. (301)– this energy density must be given by eq. (300), where, in momentum space –compare with eq. (292)–

$$\Theta_k^s(\tau) = \frac{1}{k^2} \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p}) \mathcal{R}_p \mathcal{R}_{k-p} \Big[I(\tau, p, |\boldsymbol{k}-\boldsymbol{p}|) + I_\Theta(\tau, p, |\boldsymbol{k}-\boldsymbol{p}|) \Big], \tag{310}$$

with

$$I_\Theta = \frac{2}{yz} \left(\frac{3+3w}{5+3w} \right)^2 \Big\{ T_E(xy) \Big[yz T_E(xz) - 4 T_\psi(xz) \Big] + T_\sigma(xy) T_\sigma(xz) \Big\}, \tag{311}$$

where we have traded the perturbations in eq. (309) for their transfer functions, see eq. (294). This is a heuristic argument because we have not shown that eq. (300) actually arises from averaging Einstein's equations. As stated earlier, proving this would require more machinery (in particular, a careful study of backreaction beyond leading order in perturbations using a formalism such as the one presented in [193], which is beyond the scope of this thesis).

It has been argued in earlier works [86, 90] that calculating $\Omega_{\rm GW}$ by replacing $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ in eq. (296) is justified, as long as we are only interested in the value of $\Omega_{\rm GW}$ today, because

[45] Had we written θ_{ij} in the most general form, without choosing any specific gauge, we would have simply found that some terms involving E and B are not forbidden by dimensional analysis alone. Since these terms vanish in the Newtonian gauge, we arrive at the same conclusion.

[46] This quantity is obtained simply by guessing which terms must be added to eq. (291) to cancel the gauge transformation. An alternative way to derive it is presented in Appendix A.

[47] In other words, if we suppose there exists a gauge invariant Y_{ij} such that $Y_{ij} - X_{ij} \neq 0$, then since this difference is gauge-invariant, it must hold in particular in the Newtonian gauge. Thus, $Y_{ij} \neq \boldsymbol{h}_{ij}^{\rm TT}$ in this gauge and therefore $Y_{ij} \neq \Theta_{ij}$, proving that Y_{ij} cannot be the quantity that enters in eq. (300).

scalar modes decay quickly during the RD era and therefore second-order tensor modes propagate linearly at late times. In other words, because at late times the source in eq. (260) vanishes, the induced second-order tensor modes obey the same equation of motion as the first-order piece, implying, according to [86, 90], that they effectively become first-order quantities and therefore can be treated as such for the purpose of calculating the gravitational wave energy density at late times. The advantage of our calculation with respect to these arguments is that it is completely independent of the matter content of the Universe, and therefore of the time at which $\Omega_{\rm GW}$ is evaluated. Our explicitly gauge-invariant expression for $\Omega_{\rm GW}$ is valid at all times, provided our initial assumption holds.

5.3 ■ Energy density bounds

In this section we will estimate the abundance of gravitational waves in the eMD $\to$ RD transition scenario by using the monochromatic power spectrum of eq. (45), which will allow us to easily perform the integral in eq. (283). If gravitational waves stop being generated at some late time τ near the end of the RD era and their energy density evolves as that of radiation afterwards, we can relate the abundance of GWs at τ to their abundance today by using entropy conservation and eq. (296) [186],

$$\Omega_{\rm GW}(T_0,k)=\frac{\Omega_\gamma(T_0)}{24}\frac{g_\star(T)}{g_\star(T_0)}\left(\frac{g_{\star s}(T_0)}{g_{\star s}(T)}\right)^{4/3}\left(\frac{k}{\mathcal{H}}\right)^2\langle\mathcal{P}_{\boldsymbol{h}}(T,k)\rangle_{\rm W}, \tag{312}$$

where we have switched to temperature as the time variable and replaced $h_{ij}^{\rm TT}\to\frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ in eq. (296), assuming that $\boldsymbol{h}_{ij}^{\rm TT}$ is computed in the Newtonian gauge, as per the discussion in the previous section. To compute the power spectrum, we begin by noting that the two-point function for $\boldsymbol{h}_{ij}^{\rm TT}$ is, after some manipulation [185, 186],

$$\langle\boldsymbol{h}_k^r\boldsymbol{h}_q^s\rangle_{\rm Q}=64\pi^7\delta^3(\boldsymbol{k}+\boldsymbol{q})\frac{1}{k^4}\int\frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{\boldsymbol{e}^r(\boldsymbol{k},\boldsymbol{p})\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})}{p^3|k-p|^3}\mathcal{P}_{\mathcal{R}}(p)\mathcal{P}_{\mathcal{R}}(|\boldsymbol{k}-\boldsymbol{p}|)I^2(\tau,p,|\boldsymbol{k}-\boldsymbol{p}|), \tag{313}$$

where, on the left-hand side, we have taken the quantum expectation value, see Appendix E. The next step consists on switching to spherical coordinates, performing one of the angular integrals, and switching to the variables $x=k\tau$, $y=p/k$, and $z=|\boldsymbol{k}-\boldsymbol{p}|/k$ defined in the previous section. The result is the following net change in the integrand

$$\int d^3\boldsymbol{p}\,\frac{\boldsymbol{e}^r(\boldsymbol{k},\boldsymbol{p})\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})}{p^3|\boldsymbol{k}-\boldsymbol{p}|^3}\longrightarrow\frac{\pi}{2}\delta^{rs}k\int_0^\infty dy\int_{|1-y|}^{1+y}dz\left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2. \tag{314}$$

The power spectrum is defined by

$$\langle\boldsymbol{h}_k^r\boldsymbol{h}_q^s\rangle_{\rm Q}\equiv(2\pi)^3\delta^3(\boldsymbol{k}+\boldsymbol{q})\delta^{rs}\frac{2\pi^2}{k^3}\mathcal{P}_{\boldsymbol{h}}(\tau,k). \tag{315}$$

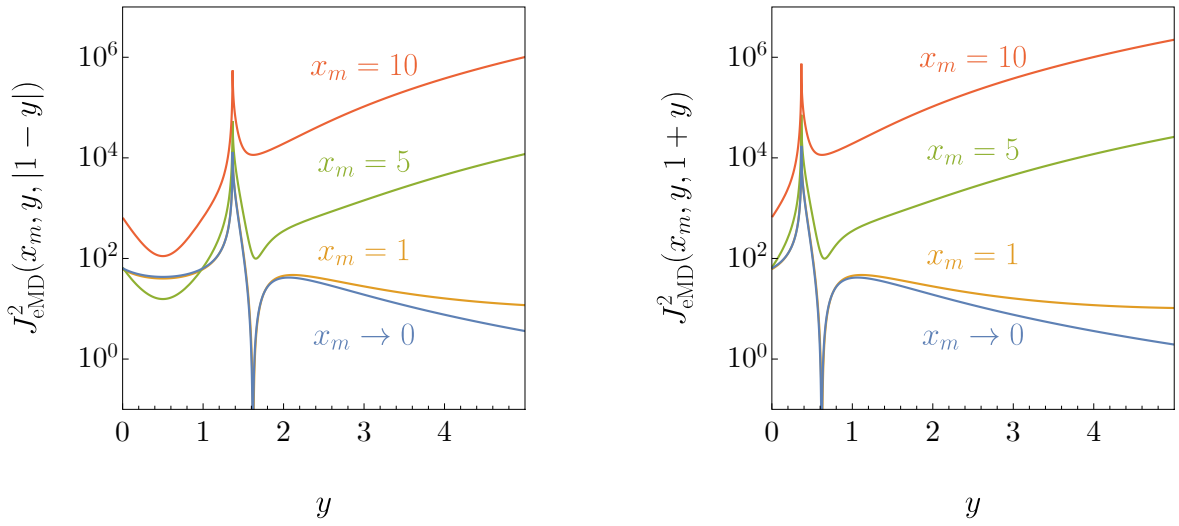

Figure 24: *Integral kernel $J^2_{\rm eMD}(x_m, y, z)$ in the Newtonian gauge for the two limits of integration in eq. (316) and three different values of x_m (and in the limit $x_m \to 0$, corresponding to the case in which the Universe never goes through the eMD phase). For $x_m \lesssim 1$ the kernel is virtually indistinguishable from the $x_m \to 0$ result, so that the eMD era only has a noticeable effect in the spectrum for modes satisfying $x_m \gtrsim 1$.*

Thus, we have the following formula for the dimensionless power spectrum,

$$\mathcal{P}_{\boldsymbol{h}}(\tau, k) = \int_0^\infty dy \int_{|1-y|}^{1+y} dz \left[\frac{4y^2 - (1+y^2-z^2)^2}{8yz}\right]^2 \mathcal{P}_{\mathcal{R}}(ky)\mathcal{P}_{\mathcal{R}}(kz) I^2(x, y, z). \tag{316}$$

The function I is given by eq. (286). We show the full expression in Appendix F. The result for the case in which the Universe never goes through an eMD phase, which is also given in Appendix F, can be obtained by taking the limit $x_m \to 0$ in I.

To obtain $\Omega_{\rm GW}$, we need to square I in eq. (286) and average the result by using

$$\left\langle \cos\left(x - \frac{x_m}{2}\right) \sin\left(x - \frac{x_m}{2}\right) \right\rangle_{\rm W} = 0,$$
$$\left\langle \cos^2\left(x - \frac{x_m}{2}\right) \right\rangle_{\rm W} = \left\langle \sin^2\left(x - \frac{x_m}{2}\right) \right\rangle_{\rm W} = \frac{1}{2}.$$

We find

$$2a(\tau)^2 \langle I^2 \rangle_{\rm W} = 4\frac{a(\tau_m)^2}{x_m^2}\left[\left(I_1^{\rm eMD} + I_1^{\rm RD}\right)^2 + \left(I_2^{\rm eMD} + I_2^{\rm RD}\right)^2\right]. \tag{317}$$

By using eqs. (270, 271) we obtain

$$\frac{k^2}{\mathcal{H}^2}\langle I^2 \rangle_{\rm W} = \frac{1}{2}\left[\left(I_1^{\rm eMD} + I_1^{\rm RD}\right)^2 + \left(I_2^{\rm eMD} + I_2^{\rm RD}\right)^2\right], \tag{318}$$

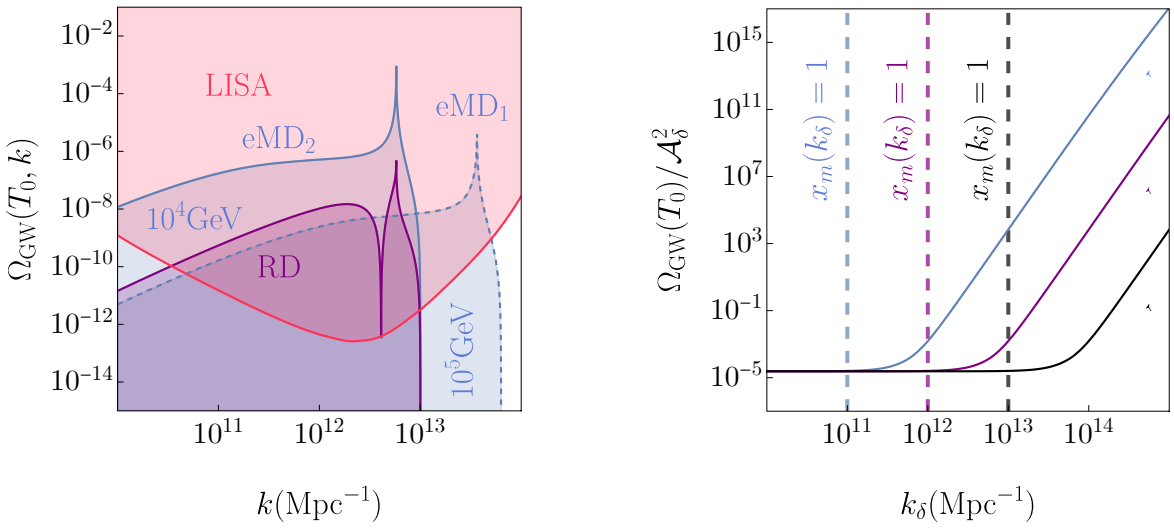

Figure 25: *Left panel: induced abundance $\Omega_{\rm GW}(T_0,k)$ for the three examples of Table 1, together with the LISA sensitivity curve [77]. We remind the reader that the two eMD examples have different associated transition temperatures T_m, shown in the plot on top of the corresponding curves. We remark that the sharp peaks in this figure (and the dip in the RD curve) are a consequence of using the Dirac delta spectrum in eq. (45). These features do not show up for smooth spectra, as exemplified in the next section. Right panel: dependence of the integrated abundance defined by eq.(327) on k_δ for three different transition temperatures T_m, normalized by A_δ^2. Growth in each case begins roughly after $x_m(k_\delta)=1$.*

so that the final result can be written as [185]

$$\Omega_{\rm GW}(T_0,k)=\frac{\Omega_\gamma(T_0)}{48}\frac{g_\star(T)}{g_\star(T_0)}\left(\frac{g_{\star s}(T_0)}{g_{\star s}(T)}\right)^{4/3}\int_0^\infty dy\int_{|1-y|}^{1+y}dz\left[\frac{4y^2-(1+y^2-z^2)^2}{8yz}\right]^2\cdot \\ \cdot\,\mathcal{P}_{\mathcal{R}}(ky)\mathcal{P}_{\mathcal{R}}(kz)J^2_{\rm eMD}(x_m,y,z), \tag{319}$$

where we have defined the shorthand

$$J^2_{\rm eMD}(x_m,y,z)\equiv\left[I_1^{\rm eMD}(x_m,y,z)+I_1^{\rm RD}(x_m,y,z)\right]^2+\left[I_2^{\rm eMD}(x_m,y,z)+I_2^{\rm RD}(x_m,y,z)\right]^2, \tag{320}$$

which is plotted in Fig. 24. If the Universe never goes through an eMD phase, the result is obtained in a straightforward manner by taking the limit $x_m\to 0$,

$$J^2_{\rm eMD}(x_m\to 0,y,z)=\left[I_1^{\rm pRD}(y,z)\right]^2+\left[I_2^{\rm pRD}(y,z)\right]^2, \tag{321}$$

with $I_i^{\rm pRD}$ as defined in Appendix F.

For the Dirac delta spectrum of eq. (45), the integral in eq. (319) can be performed ana-

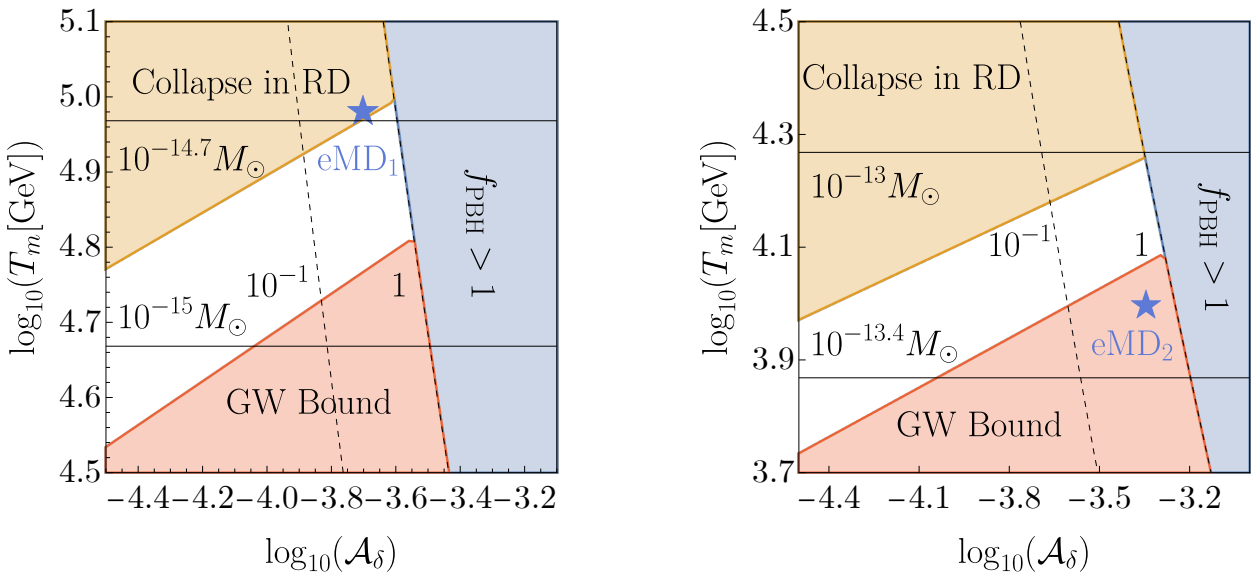

Figure 26: *Available parameter space with k_δ fixed as in the two eMD examples of Table 1 after taking the constraint of eq. (328) into account (compare with Fig. 3), as well as the constraint in eq. (47). The eMD$_2$ example is ruled out by this bound. Although the solid and dashed black lines are affected by uncertainties in the formulas for the PBH mass and abundance, the numbers quoted should be correct as an order of magnitude estimate as per the discussion at the end of Section 2.1.*

lytically, and the result is

$$\Omega_{\rm GW}(T_0,k) = A_\delta^2\,\frac{\Omega_\gamma(T_0)}{3072}\,\frac{g_\star(T)}{g_\star(T_0)}\left(\frac{g_{\star s}(T_0)}{g_{\star s}(T)}\right)^{4/3}\frac{k^2}{k_\delta^2}\left(4\,\frac{k_\delta^2}{k^2}-1\right)^2 J^2_{\rm eMD}\left(x_m,\frac{k_\delta}{k},\frac{k_\delta}{k}\right)\Theta(2k_\delta-k), \tag{322}$$

where the Θ at the end denotes the Heaviside step function. To evaluate this quantity we can use the fact that (taking T as some late time in the RD era before the top quark decouples from the primordial plasma, so that gravitational waves at the scales relevant here have already been induced)

$$x_m = \left(\frac{g_{\star s}(T_m)}{106.75}\right)^{1/3}\left(\frac{106.75}{g_\star(T_m)}\right)^{1/2}\left(\frac{10^7\,\rm GeV}{T_m}\right)\left(\frac{k}{10^{14}\,\rm Mpc^{-1}}\right). \tag{323}$$

The quantity $\Omega_{\rm GW}(T_0,k)$ is plotted in Fig. 25. The left panel of this figure depicts the resulting $\Omega_{\rm GW}(T_0,k)$ for the three examples of Table 1. Interestingly, all of the examples yield spectra with very different amplitudes despite all of them having the same abundance (see Fig. 2). The reason for this is that the height of the induced tensor spectrum depends on the ratio between T_m and k because of eq. (323). As shown in Fig. 24, for $x_m \lesssim 1$ the

integral kernel is virtually indistinguishable from the $x_m \to 0$ case in which the Universe does not go through an eMD phase. For $x_m \gtrsim 1$, however, the kernel grows very quickly, due to the fact that the transfer function T_k^ψ is constant during eMD –see eq. (280)– and thus the corresponding modes are enhanced.

For collapse to occur during the eMD era, the relevant modes must obviously re-enter the horizon before the phase ends. That is, collapse during eMD can only happen for modes with $k > k_m$, where

$$k_m = \frac{\pi}{3}\frac{T_0}{M_p}\left(\frac{g_\star(T_m)}{10}\right)^{1/2}\left(\frac{g_{\star s}(T_0)}{g_{\star s}(T_m)}\right)^{1/3} T_m. \tag{324}$$

It is illustrative to compare the condition $x_m \gtrsim 1$ to this condition. If we take $g_{\star s}(T_m) = g_\star(T_m) = 106.75$, then the two conditions are, respectively,

$$k \gtrsim 10^{12}\text{Mpc}^{-1}\left(\frac{T_m}{10^5\text{GeV}}\right), \qquad \text{(kernel enhancement)} \tag{325}$$

$$k > 4.3\times 10^{10}\text{Mpc}^{-1}\left(\frac{T_m}{10^5\text{GeV}}\right), \qquad \text{(horizon re-entry during eMD)} \tag{326}$$

so we see that all of the enhanced modes re-enter the horizon during the eMD phase. We therefore find that if PBHs form during RD, the presence of an eMD era does not affect the induced GW spectrum at the peak of the distribution, as we would expect.

The net contribution of GWs to the energy budget of the Universe is obtained by integrating

$$\Omega_{\text{GW}}(T_0) = \int \Omega_{\text{GW}}(T_0, k) d\log(k). \tag{327}$$

There exists a bound on this quantity arising from both CMB experiments [98, 99] and from the abundance of light elements produced during Big-Bang Nucleosynthesis[48] [99]. In our figures we show only the latter bound,

$$\Omega_{\text{GW}}h^2 < 1.8\times 10^{-6}, \tag{328}$$

but we remark that the CMB bounds can be slightly stronger, by an $\mathcal{O}(1)$ factor, depending on the choice of initial conditions [98]. Since the enhancement of the tensor spectrum is so large for PBHs formed in the eMD phase, it is important to check whether this bound is satisfied. Indeed, we find that it severely limits the available region of parameter space. The result is plotted in Fig. 26 for the two eMD examples of Table 1 (compare with Fig. 3). The eMD_2 example is ruled out by this new constraint. For PBHs formed during the eMD era, this bound turns out to be stronger than the one coming from galactic γ-ray constraints [31, 32] at relatively low transition temperatures $T_m \lesssim 10^{5.5}\,\text{GeV}$, as shown in Fig. 27. We

[48] We remark that this bound only depends on the assumption that gravitational waves behave as radiation at the time of BBN (or photon decoupling for the CMB), which is indeed the case in our scenario, as can be explicitly checked from eq. (319) [185].

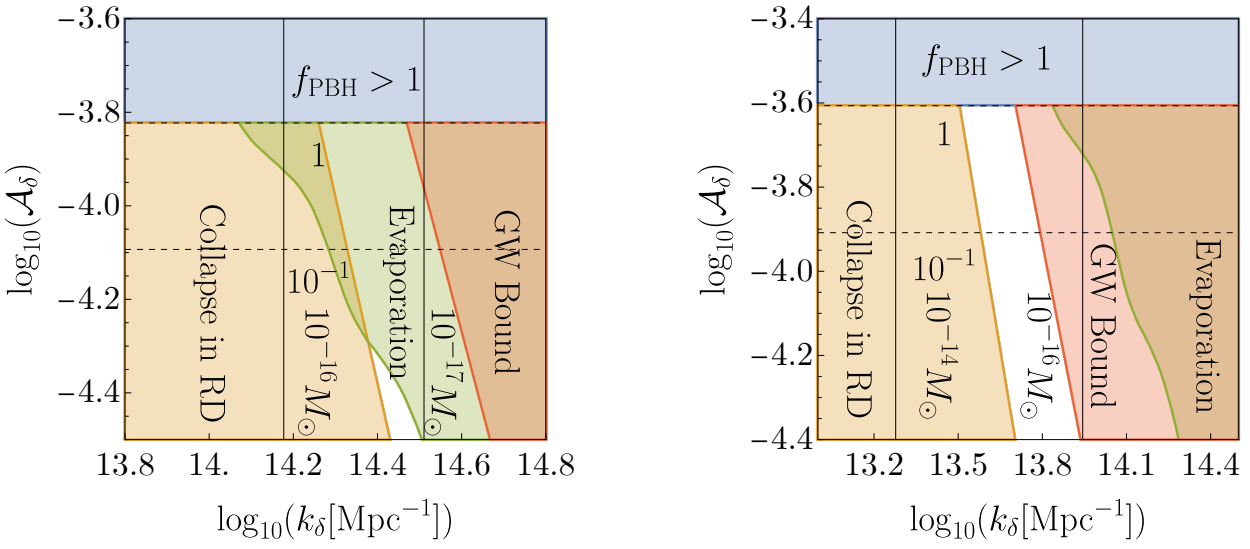

Figure 27: *Available parameter space for $T_m = 10^{5.7}\,\mathrm{GeV}$ (left panel) and $T_m = 10^{5}\,\mathrm{GeV}$ (right panel), together with the bound in eq. (47). For transition temperatures $T_m \lesssim 10^{5.5}\,\mathrm{GeV}$, the constraint of eq. (328) is stronger than the bound arising from galactic γ-ray constraints [32]. The contour lines represent the PBH abundance and masses at the peak of the distribution, which should be correct as an order of magnitude estimate as per the discussion at the end of Section 2.1.*

remark that the reason it is not possible to represent both eMD examples in the same figure is that the bounds depend on the position of the peak in the power spectrum for each one –the parameter k_δ in eq. (45)– which takes different values in each example (see Table 1 and Fig. 3). Finally, we remark that the formula in eq. (319) is of general validity for any scalar power spectra. As an example, we show the induced gravitational wave signals for the polynomial model of Section 2.1 with and without higher-dimensional operators in Fig. 28. The examples coincide with the ones shown in Fig. 8. The signals in this case are much smoother than for the Dirac delta spectrum, although the order of magnitude of the amplitude is the same, as we anticipated in the caption of Fig. 25.

Although we have focused our discussion on the LISA and BBO/DECIGO detectors, there are many more proposed experiments that can probe the gravitational wave signals induced by PBHs. On the right panel of Fig. 28 we show the sensitivity curves for LISA [77], BBO/DECIGO [197], Magis-space and Magis-100 [198], the Einstein Telescope [199] and advanced LIGO [200]. The first four of these experiments, in particular, could probe the range of PBH masses of interest for dark matter (1). Other proposed experiments not shown in the figure with sensitivities in the same range are AION [201] and AEDGE [202]. The advanced LIGO and Einstein Telescope experiments cannot probe the masses in (1),

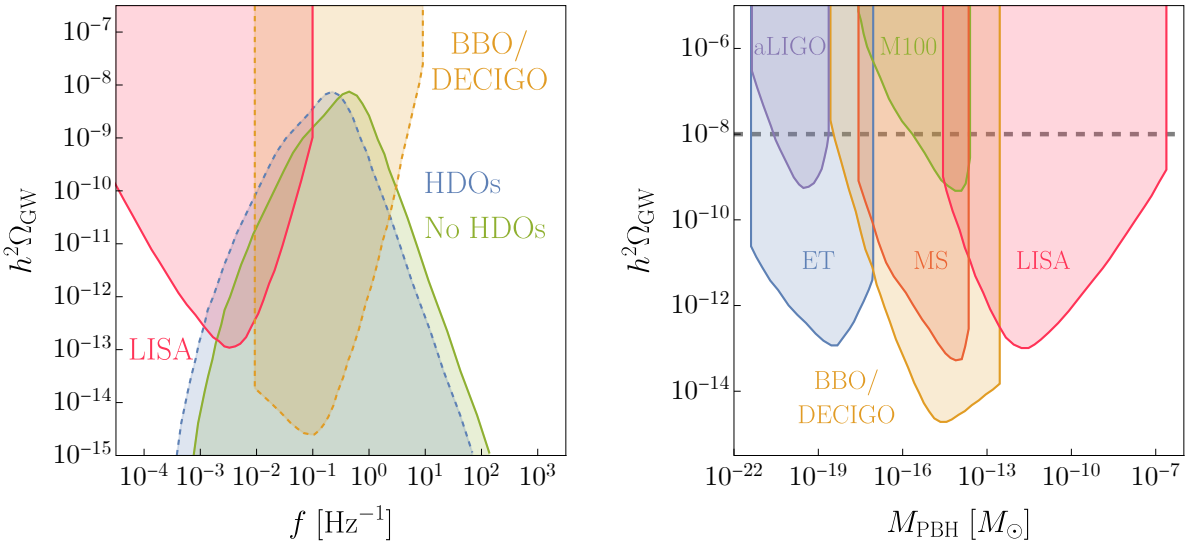

Figure 28: *Left panel: gravitational waves induced in the polynomial model of Section 2.1, with and without higher-dimensional operators, if the black holes form during a RD era, together with the LISA and BBO/DECIGO sensitivity curves. Right panel: projected sensitivities of future gravitational wave experiments as a function of the PBH masses if they form during a RD era. The dashed line represents the approximate expected size of the gravitational wave signal in this case (see left panel).*

but would still be able to detect the gravitational wave signal induced by a population of sufficiently light black holes. PBHs with these masses can evade the evaporation bounds arising from extra-galactic γ-rays [29, 30], but are still constrained by the abundance of light elements produced during nucleosynthesis [203]. In the left panel of Fig. 29 we show a solution of the polynomial model in Section 2.1 that leads to an abundance of order $f_{\rm PBH} \simeq 10^{-4}$ of light PBHs. On the right panel of the same figure we show the induced gravitational wave signal, which falls above the design sensitivity curve of the advanced LIGO experiment.

To close this section, let us point out that although gravitational waves induced during an eMD era have been studied before [185], the novel part of our analysis is in doing so in the context of PBH formation, and including the bound in eq. (328) to determine the available region of parameter space.

5.4 ■ Gravitational waves in warm inflation

In this section we compute the gravitational waves induced in the warm inflation scenario of Chapter 4. The calculation is slightly different to the one presented in the previous sections because both the tensor and scalar power spectra are now stochastic quantities. We also

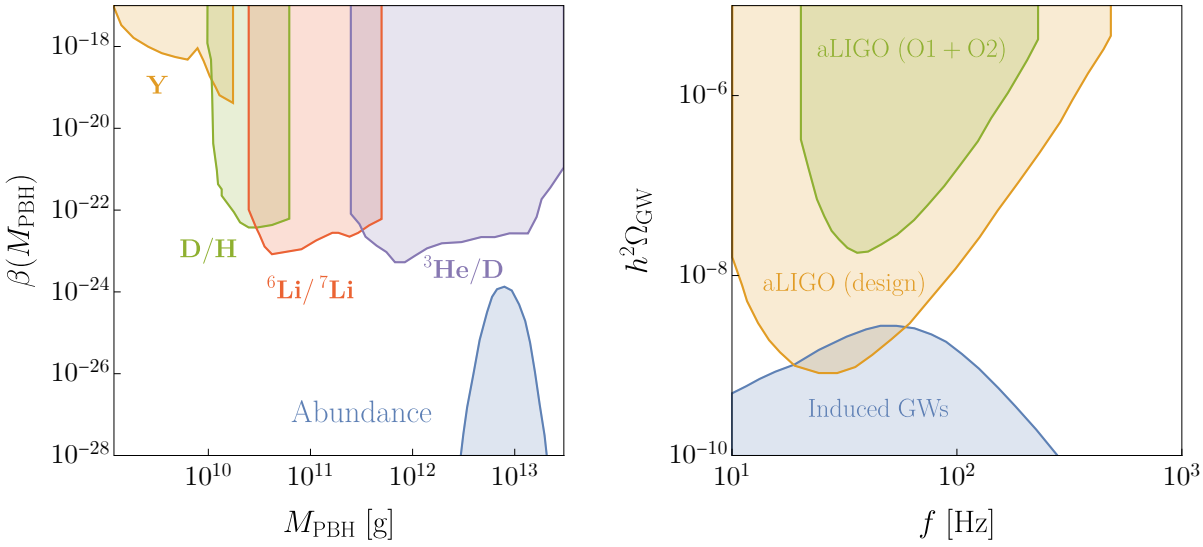

Figure 29: *Left panel: fraction of the Universe's mass in PBHs (41) at their formation time as a function of the PBH mass if the PBHs form during a RD era, together with the BBN bounds in [203]. Right panel: fraction of the energy density in gravitational waves relative to the critical energy density of the Universe as a function of the frequency. The PBH abundance shown in these plots is obtained using the polynomial model of Section 2.1.*

compute, in this case, the gravitational waves induced during inflation, which are suppressed in the inflection-point models.

Let us begin by noting that the source term in eq. (266) can be written, during inflation, in terms of the total momentum perturbation[49] δq (where by total we mean including both the scalar and radiation components) using eq. (201)

$$S^s_{k(\rm pre)} = \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \boldsymbol{e}^s_{ij}(\boldsymbol{k}) p_i p_j \left[8\psi_p \psi_{k-p} + \frac{4M_p^{-2}}{(\rho+p)} \left(\delta q^{(r)}_p - \frac{\phi'}{a}\delta\phi_p \right) \left(\delta q^{(r)}_{k-p} - \frac{\phi'}{a}\delta\phi_{k-p} \right) \right], \quad (329)$$

where the (pre) subscript denotes the fact that this is the source term during inflation. We now make the simplifying assumption that the terms in this expression involving ψ and $\delta q^{(r)}$ can be neglected. This assumption is motivated by the analytical model of Section 4.4 and will be justified at the end of this section. Thus,

$$S^s_{k(\rm pre)} = \left(\frac{4}{\rho+p} \right) \left(\frac{\phi'^2}{a^2 M_p^2} \right) \int \frac{d^3\boldsymbol{p}}{(2\pi)^3} \boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p}) \delta\phi_p \delta\phi_{k-p}. \quad (330)$$

The value of ψ at the end of inflation ($\tau = 0$), which will be the initial condition for the post-inflationary source, is, on superhorizon scales and assuming the Universe enters an RD

[49] The momentum perturbations δq_i are additive, so the total momentum perturbation can be defined as the sum of the individual components, $\delta q \equiv \sum_i \delta q_i$.

era after inflation ends,

$$\psi_k(0) = \frac{2}{3}\mathcal{R}_k(0) = -\frac{2}{3}\frac{H\phi'}{a(\rho+p)}\bigg|_0 \delta\phi_k(0), \tag{331}$$

where we have used eqs. (234, 278). The post-inflationary source can then be written as

$$S^s_{k(\text{post})} = \frac{4}{9}\frac{H^2\phi'^2}{a^2(\rho+p)^2}\bigg|_0 \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{p})\delta\phi_p(0)\delta\phi_{k-p}(0)Q(p,|\boldsymbol{k}-\boldsymbol{p}|,\tau), \tag{332}$$

where the function Q was defined in eq. (285).

The Green's functions for eq. (266) during inflation is

$$kF_{\text{pre}}(\tau,\tau') = -\frac{1}{(k\tau')^2}\Big[k(\tau-\tau')\cos(k\tau-k\tau') - (1+k^2\tau\tau')\sin(k\tau-k\tau')\Big]. \tag{333}$$

where we have used $\mathcal{H} = -1/\tau$ during inflation. The solution is therefore

$$\boldsymbol{h}^s_k(\tau) = T^h_k(\tau)\boldsymbol{h}^s_k(0) + \underbrace{\int_0^\tau F_{\text{post}}(\tau,\tau')S^s_{k(\text{post})}(\tau')d\tau'}_{\boldsymbol{g}^s_k(\tau)}, \tag{334}$$

where T^h_k is the (linear) transfer function of h^s_k in the radiation era,

$$T^h_k \equiv \frac{\sin(k\tau)}{k\tau}, \tag{335}$$

and

$$\boldsymbol{h}^s_k(0) = \int_{\tau_h}^0 F_{\text{pre}}(0,\tau')S^s_{k(\text{pre})}(\tau')d\tau'. \tag{336}$$

The lower integration limit τ_h is some early time at which we assume $\boldsymbol{h}^s_k(\tau_h) = 0$ (that is, no gravitational waves have been induced at this time).

The expectation value of the tensor power spectrum late in the radiation era contains three terms,

$$\langle\mathcal{P}_{\boldsymbol{h}}(k,\tau)\rangle_{\text{S}} = \langle\mathcal{P}_{\text{pre}}(k,\tau)\rangle_{\text{S}} + 2\langle\mathcal{P}_{\text{mix}}(k,\tau)\rangle_{\text{S}} + \langle\mathcal{P}_{\text{post}}(k,\tau)\rangle_{\text{S}}. \tag{337}$$

These three terms will in turn lead to three different contributions to the gravitational wave energy density. The first term corresponds to the gravitational waves induced during the inflationary epoch, and the third term corresponds to the gravitational waves induced during the subsequent radiation-dominated era. The middle term mixes both contributions and its value typically lies between the other two.

To perform the rest of the calculation we will use the analytical results of Section 4.4. The reason for this is that to calculate the tensor power spectrum we need to take the

quantum expectation value in addition to the stochastic one (in order to make the tensor power spectrum a deterministic quantity, as we did with $\mathcal{P}_{\mathcal{R}}$), and, as discussed in Appendix E, when eq. (246) is used to split $\delta\phi$ into the homogeneous and inhomogeneous solutions to its equation of motion, finding the double expectation value is straightforward, since only the homogeneous solution is quantized, and only the inhomogeneous piece is stochastic. This splitting can be done only because the equation of motion for $\delta\phi$ has been decoupled from the rest (up to the source term $\delta\rho^{(r)}$), since the full system of differential equations cannot be solved by Green's function methods. This approach should give a reasonable estimate of the full numerical result. The first term in eq. (337) is defined by[50]

$$T_k^h(\tau)^2\langle \boldsymbol{h}_k^r(0)\boldsymbol{h}_p^s(0)\rangle = (2\pi)^3\frac{2\pi^2}{k^3}\langle \mathcal{P}_{\rm pre}(k,\tau)\rangle_{\rm S}\delta^{rs}\delta^3(\boldsymbol{k}+\boldsymbol{p}). \tag{338}$$

The quantum expectation value is already included inside $\mathcal{P}_{\rm pre}$ on the right-hand side – see eq. (444) for the analogous scalar definition– so we only write the stochastic average. The two-point function on the left-hand side can be computed quantizing the inflaton field perturbation. The result is, using eqs. (334, 336),

$$\begin{aligned}\langle \boldsymbol{h}_k^r(0)\boldsymbol{h}_p^s(0)\rangle = \frac{16}{M_p^4}\int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'} F_{\rm pre}(0,\tau')d\tau' \int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau''} F_{\rm pre}(0,\tau'')d\tau''\cdot \\ \cdot \int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\boldsymbol{e}^r(\boldsymbol{k},\boldsymbol{q})\int \frac{d^3\boldsymbol{l}}{(2\pi)^3}\boldsymbol{e}^s(\boldsymbol{p},\boldsymbol{l})\langle \delta\hat{\phi}_q(\tau')\delta\hat{\phi}_{k-q}(\tau')\delta\hat{\phi}_l(\tau'')\delta\hat{\phi}_{p-l}(\tau'')\rangle.\end{aligned} \tag{339}$$

The four-point function of $\delta\phi$ appearing in this equation can be computed using eq. (246), together with Wick's theorem. This calculation is done in Appendix E. The resulting expression can be plugged back into the above equation and, using one of the Dirac deltas to perform the integral over $\boldsymbol{l}$, we find

$$\begin{aligned}\langle \boldsymbol{h}_k^r(0)\boldsymbol{h}_p^s(0)\rangle = \frac{\delta^3(\boldsymbol{k}+\boldsymbol{p})}{k^{10}}\int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'} kF_{\rm pre}(0,\tau')kd\tau' \int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau''} kF_{\rm pre}(0,\tau'')kd\tau''\cdot \\ \cdot \frac{128\pi^4}{M_p^4}\int d^3\boldsymbol{q}\,\boldsymbol{e}^r(\boldsymbol{k},\boldsymbol{q})\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{q})\mathcal{Q}_{\delta\phi}(q,\tau',\tau'')\mathcal{Q}_{\delta\phi}(|\boldsymbol{k}-\boldsymbol{q}|,\tau',\tau'')\frac{k^3}{q^3}\frac{k^3}{|\boldsymbol{k}-\boldsymbol{q}|^3},\end{aligned} \tag{340}$$

where we have defined[51]

$$\mathcal{Q}_{\delta\phi}(q,\tau',\tau'') \equiv \frac{q^3}{2\pi^2}\left(\delta\phi_q^{(h)}(\tau')^\star\delta\phi_q^{(h)}(\tau'') + \int_{-\infty}^{\min(\tau',\tau'')} G_q(\tau',\hat{\tau})G_q(\tau'',\hat{\tau})\frac{9a^2}{H^2\phi'^2}|\delta\rho_q^{(r)}(\hat{\tau})|^2 d\hat{\tau}\right). \tag{341}$$

[50] As discussed in Appendix E, the brackets without subscripts denote a double expectation value, quantum and stochastic.

[51] $G_k(\tau,\tau')$ denotes the Green's function defined in eq. (251) evaluated at $\tau = -e^{-N}/H$.

This quantity has the following properties

$$\mathcal{Q}_{\delta\phi}(q,\tau,\tau) = \langle \mathcal{P}_{\delta\phi}(q,\tau)\rangle_{\rm S}, \qquad \mathcal{Q}_{\delta\phi}(q,\tau',\tau'') = \mathcal{Q}^{\star}_{\delta\phi}(q,\tau'',\tau'). \tag{342}$$

It is useful to make the integrand in eq. (340) manifestly real and symmetric under $\tau' \leftrightarrow \tau''$. To do so, we take eq. (339), rename the dummy variables $\tau' \leftrightarrow \tau''$ and sum the result with eq. (339) itself. After using the second identity in eq. (342), we find

$$\begin{aligned}\langle \boldsymbol{h}^r_k(0)\boldsymbol{h}^s_p(0)\rangle = \frac{\delta^3(\boldsymbol{k}+\boldsymbol{p})}{k^{10}}\int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'} kF_{\rm pre}(0,\tau')kd\tau' \int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau''} kF_{\rm pre}(0,\tau'')kd\tau''\cdot \\ \cdot \frac{128\pi^4}{M_p^4}\int d^3\boldsymbol{q}\, \boldsymbol{e}^r(\boldsymbol{k},\boldsymbol{q})\boldsymbol{e}^s(\boldsymbol{k},\boldsymbol{q}){\rm Re}\bigg[\mathcal{Q}_{\delta\phi}(q,\tau',\tau'')\mathcal{Q}_{\delta\phi}(|\boldsymbol{k}-\boldsymbol{q}|,\tau',\tau'')\bigg]\frac{k^3}{q^3}\frac{k^3}{|\boldsymbol{k}-\boldsymbol{q}|^3}.\end{aligned} \tag{343}$$

We can now perform the same steps we did below eq. (313) to obtain, for the first term in eq. (337),

$$\begin{aligned}\langle \mathcal{P}_{\rm pre}(k,\tau)\rangle_{\rm S} = T^h_k(\tau)^2\frac{4}{M_p^4}\int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'} kF_{\rm pre}(0,\tau')kd\tau' \int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau''} kF_{\rm pre}(0,\tau'')kd\tau''\cdot \\ \cdot \int_0^\infty dy\int_{|1-y|}^{1+y} dz\left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2 {\rm Re}\bigg[\mathcal{Q}_{\delta\phi}(ky,\tau',\tau'')\mathcal{Q}_{\delta\phi}(kz,\tau',\tau'')\bigg].\end{aligned} \tag{344}$$

To compute the term $2\langle \mathcal{P}_{\rm mix}(k,\tau)\rangle_{\rm S}$, we use

$$T^h_k(\tau)\langle \boldsymbol{h}^r_k(0)\boldsymbol{g}^s_p(\tau)\rangle = (2\pi)^3\frac{2\pi^2}{k^3}\langle \mathcal{P}_{\rm mix}(k,\tau)\rangle_{\rm S}\delta^{rs}\delta^3(\boldsymbol{k}+\boldsymbol{p}), \tag{345}$$

where $g_k(\tau)$ is defined in eq. (334), and similarly for the $\mathcal{P}_{\rm post}$ term. Following a completely analogous procedure to the one we just applied to compute $\langle \mathcal{P}_{\rm pre}(k,\tau)\rangle_{\rm S}$, we obtain

$$\begin{aligned}\langle \mathcal{P}_{\rm mix}(k,\tau)\rangle_{\rm S} = T^h_k(\tau)\frac{4}{9M_p^2}\frac{H^2\phi'^2}{a^2(\rho+p)^2}\bigg|_0 \int_{\tau_h}^0 \frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'} kF_{\rm pre}(0,\tau')kd\tau' \int_0^\tau kF_{\rm post}(\tau,\tau'')kd\tau'' \\ \int_0^\infty dy\int_{|1-y|}^{1+y} dz\left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2 {\rm Re}\bigg[\mathcal{Q}_{\delta\phi}(ky,\tau',0)\mathcal{Q}_{\delta\phi}(kz,\tau',0)\bigg]Q(ky,kz,\tau'').\end{aligned} \tag{346}$$

Similarly,

$$\begin{aligned}\langle \mathcal{P}_{\rm post}(k,\tau)\rangle_{\rm S} = \frac{4}{81}\int_0^\tau kF_{\rm post}(\tau,\tau')kd\tau' \int_0^\tau kF_{\rm post}(\tau,\tau'')kd\tau'' \int_0^\infty dy\int_{|1-y|}^{1+y} dz\cdot \\ \cdot\left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2 \langle \mathcal{P}_{\mathcal{R}}(ky)\rangle_{\rm S}\langle \mathcal{P}_{\mathcal{R}}(kz)\rangle_{\rm S} Q(ky,kz,\tau')Q(ky,kz,\tau''),\end{aligned} \tag{347}$$

where we have used eqs. (331, 342) to relate $\mathcal{Q}_{\delta\phi}$ to $\mathcal{P}_{\mathcal{R}}$. This completes the calculation of the tensor power spectrum late into the radiation era.

To find the energy density (312) we use

$$\langle T_k^h(\tau)^2\rangle_{\rm W} = \frac{1}{2(k\tau)^2}, \tag{348}$$

$$k\langle T_k^h(\tau)F_{\rm post}(\tau,\tau')\rangle_{\rm W} = \frac{1}{2k\tau}\left(\frac{\tau'}{\tau}\right)\cos(k\tau'), \tag{349}$$

$$k^2\langle F_{\rm post}(\tau,\tau')F_{\rm post}(\tau,\tau'')\rangle_{\rm W} = \frac{\tau'\tau''}{2\tau^2}\Big[\cos(k\tau')\cos(k\tau'')+\sin(k\tau')\sin(k\tau'')\Big]. \tag{350}$$

The $\mathcal{H}^2$ factor in the denominator of $\Omega_{\rm GW}$ will cancel out the $1/\tau^2$ in these averages, yielding a finite result in the limit $\tau\to\infty$. The stochastic average of the energy density of gravitational waves today is therefore

$$\langle\Omega_{\rm GW}(k)\rangle_{\rm S} = \frac{\Omega_\gamma(T_0)}{24}\frac{g_\star(T)}{g_\star(T_0)}\left(\frac{g_{\star s}(T_0)}{g_{\star s}(T)}\right)^{4/3}\int_0^\infty dy\int_{|1-y|}^{1+y}dz\left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2 K(ky,kz), \tag{351}$$

where the dimensionless integration kernel $K(ky,kz)$ is

$$K(ky,kz) = K_{\rm pre}(ky,kz)+K_{\rm mix}(ky,kz)+K_{\rm post}(ky,kz), \tag{352}$$

with

$$K_{\rm pre} = \frac{2}{M_p^4}\int_{\tau_h}^0\frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'}kF_{\rm pre}(0,\tau')kd\tau'\int_{\tau_h}^0\frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau''}kF_{\rm pre}(0,\tau'')kd\tau''\cdot$$
$$\cdot\,{\rm Re}\Big[\mathcal{Q}_{\delta\phi}(ky,\tau',\tau'')\mathcal{Q}_{\delta\phi}(kz,\tau',\tau'')\Big], \tag{353}$$

$$K_{\rm mix} = I_s(y,z)\frac{4}{9M_p^2}\frac{H^2\phi'^2}{a^2(\rho+p)^2}\bigg|_0\int_{\tau_h}^0\frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau'}kF_{\rm pre}(0,\tau')kd\tau'\cdot$$
$$\cdot\,{\rm Re}\Big[\mathcal{Q}_{\delta\phi}(ky,\tau',0)\mathcal{Q}_{\delta\phi}(kz,\tau',0)\Big], \tag{354}$$

$$K_{\rm post} = \frac{1}{8}\langle\mathcal{P}_\mathcal{R}(ky)\rangle_{\rm S}\langle\mathcal{P}_\mathcal{R}(kz)\rangle_{\rm S}\,J^2_{\rm eMD}(x_m\to 0,y,z), \tag{355}$$

where $J_{\rm RD}$ was defined in eq. (321) and

$$I_s(y,z) = \int_0^\infty(k\tau')\cos(k\tau')Q(ky,kz,\tau')kd\tau'. \tag{356}$$

This integral can also be performed analytically, see [186].

For completeness, we also present here the expression for the tensor power spectrum valid in the standard cold inflation case; that is, in the absence of the second term in the

parentheses of eq. (341). In this case the spectrum can be written as the square of a sum,

$$\mathcal{P}_{\boldsymbol{h}}(k,\tau) = \int_0^\infty y \int_{|1-y|}^{1+y} z \left[\frac{4y^2-(1+y^2-z^2)^2}{4yz}\right]^2 \mathcal{P}_{\mathcal{R}}(ky)\mathcal{P}_{\mathcal{R}}(kz)$$
$$\left|6T_k^h(\tau)\int_{\eta_h}^0 \frac{\rho+p}{\rho}\right|_{\tau'} kF_{\rm pre}(0,\tau')S(ky,kz,\tau')k\tau' + \frac{2}{9}\int_0^\tau kF_{\rm post}(\tau,\tau')Q(ky,kz,\tau')k\tau'\Bigg|^2, \quad (357)$$

where

$$S(ky,kz,\tau) \equiv \frac{\mathcal{R}_{ky}(\tau)}{\mathcal{R}_{ky}(0)}\frac{\mathcal{R}_{kz}(\tau)}{\mathcal{R}_{kz}(0)}. \quad (358)$$

We now proceed to estimate each one of the terms in eq. (352). Let us focus first on the $K_{\rm pre}$ term. This contribution depends on the lower integration limit τ_h (and as noted in [204] is formally divergent in the limit $\tau_h = -\infty$ due to the behaviour of scalar perturbations in the Bunch-Davies vacuum). We deal with this problem by integrating from a finite value of τ_h that we identify with the time at which the strongly dissipative phase begins. The assumption here is that the contribution from the source prior to this time is neglegible. This assumption is reasonable since up to that time inflation proceeds as in the standard slow roll scenario (up to the presence of a weak dissipative term that does not alter the dynamics significantly), and we do not expect the corresponding gravitational wave signal to be peaked at any particular scale or exhibit any special features, in contrast to the piece arising due to the strongly dissipative phase. In addition, we notice that the inflationary contribution to the energy density of gravitational waves diverges in the $y = \infty$ limit. In principle, this divergence should be renormalized away by properly computing the induced gravitational wave signal using the in-in formalism. However, we just impose a cut-off which renders the result finite. We have verified that our results do not depend on the cutoffs unless unreasonably large values are chosen. Only the $K_{\rm mix}$ and $K_{\rm pre}$ kernels suffer from this issue, and the post-inflationary contribution is finite. We reiterate that the results of this section should be regarded as an accurate order of magnitude estimate of the overall size of the signal.

We can choose the time cutoff around the time at which the dissipative coefficient Γ begins to increase. In terms of the analytical calculation of Section 4.4, this corresponds to the beginning of phase 1. For the momentum cutoff we can choose $k_{\rm cutoff} \sim \mathcal{O}(10-100)\times k_{\rm peak}$. The four-dimensional integral in $\langle\Omega_{\rm pre}\rangle_{\rm S}$ is quite difficult to perform. However, $K_{\rm pre}$ is heavily peaked around a specific time, so the strategy we adopt is to approximate the time integrals by evaluating the integrand at this time and multiplying it by an appropriately chosen integration area.

To determine the point in parameter space at which the integrand is peaked, we use that the integrand is symmetric under $\tau' \leftrightarrow \tau''$ and $y \leftrightarrow z$, so the set of maxima of the function must be symmetric under this transformation. If the function has a unique global maximum in some region (we do not prove that this is the case, but we have checked it numerically),

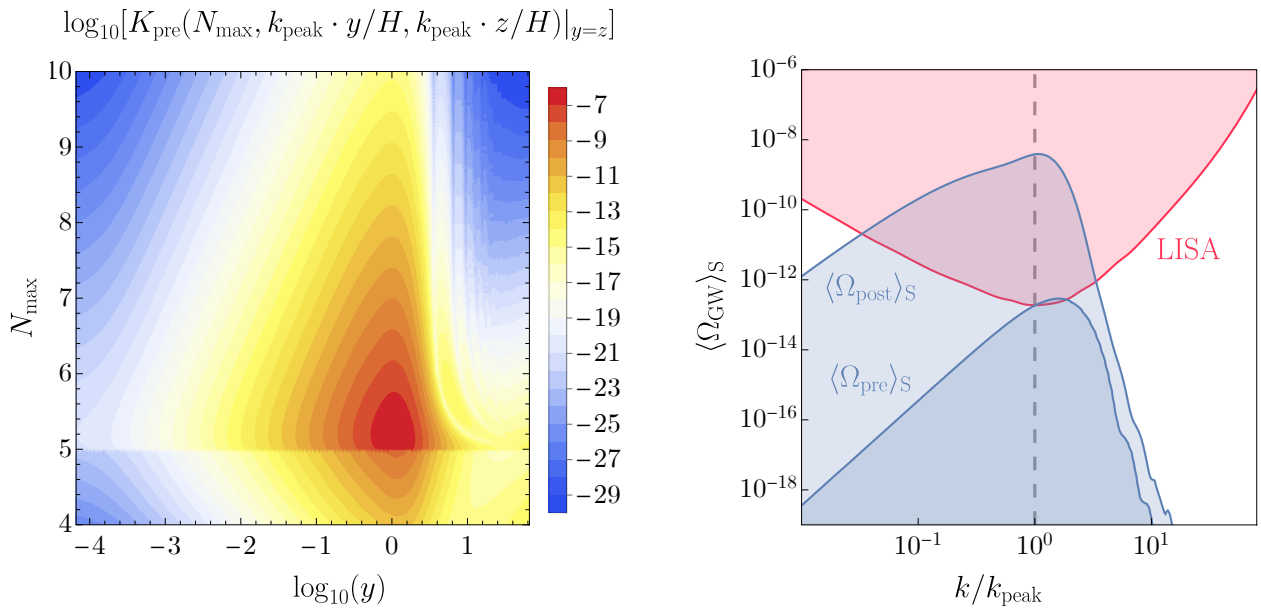

Figure 30: *Left panel: maximum value of the integration kernel in eq. (362) on the surface $y = z$ for the parameters in Table 4. Right panel: gravitational wave signals induced during and after inflation compared with the LISA sensitivity curve [77].*

then it follows that this maximum must be located along the surface with $\tau' = \tau''$ and $y = z$. On this surface,

$$K_{\rm pre} \simeq \frac{1}{2M_p^2}\frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau_{\rm max}} (k\Delta\tau_{\rm max})^2\Big[kF_{\rm pre}(0,\tau_{\rm max})\Big]^2\mathcal{P}_{\delta\phi}(ky,\tau_{\rm max})\mathcal{P}_{\delta\phi}(kz,\tau_{\rm max}), \tag{359}$$

where $\tau_{\rm max}$ is the value of τ at which the local maximum occurs, and $\Delta\tau_{\rm max}$ is the integration area, which must be appropriately chosen as a small square around $\tau_{\rm max}$ requiring, for instance, that the integrand does not decrease by more than an order of magnitude or so, in such a way that the approximation holds. The integration area is, in terms of the number of e-folds,

$$k\Delta\tau_{\rm max} = \frac{k}{H}(e^{-N_a} - e^{-N_b}), \tag{360}$$

where $\Delta N = N_b - N_a$ is the range over which the integrand is large, which spans a couple of e-folds at most. Let us write $N_a = N_{\rm max} - \Delta N_{\rm max}$ and $N_b = N_{\rm max} + \Delta N_{\rm max}$, with $\Delta N_{\rm max} \sim \mathcal{O}(1)$; and where $N_{\rm max}$ is the time in e-folds corresponding to $\tau_{\rm max}$. Then

$$k\Delta\tau_{\rm max} = 2\frac{k}{H}e^{-N_{\rm max}}\sinh(\Delta N_{\rm max}). \tag{361}$$

The function we need to maximize is therefore[52]

$$K_{\rm pre} = \frac{2k^2}{M_p^2 H^2}\frac{\phi'^2}{a^2(\rho+p)}\bigg|_{\tau_{\rm max}} \frac{\sinh^2(\Delta N_{\rm max})}{e^{2N_{\rm max}}}\Big[kF_{\rm pre}(0,\tau_{\rm max})\Big]^2 \mathcal{P}_{\delta\phi}(ky,\tau_{\rm max})\mathcal{P}_{\delta\phi}(kz,\tau_{\rm max}). \quad (362)$$

The quantity in eq. (362) is shown in the left panel of Fig. 30 for the parameters in Table 4. The discontinuity around $N = 5$ e-folds is due to the fact that, as mentioned in Section 4.4, we take the background parameters as piecewise-constant functions for this calculation. Specifically, we take $\phi'^2/a^2(\rho+p) = 1$ in phases 0 and 3, and $\phi'^2/a^2(\rho+p) = 0.02$ in phases 1 and 2. In this figure we also take $\Delta N_{\rm max} = 2$, which is clearly enough to account for the region in which the integrand is large. Changing this number by a factor of $\mathcal{O}(1)$ does not change our results.

To obtain the induced gravitational wave signal, we find the time τ at which $K_{\rm pre}$ is peaked for each k and perform the momentum integrals numerically. The time-dependent power spectrum $\mathcal{P}_{\delta\phi}(k,\tau)$ is calculated using the analytical formalism of Section 4.4, as we anticipated earlier. Specifically, it can be found by keeping the full expression for the Green's function in eq. (251) instead of taking the $N \to \infty$ limit. The resulting signal is shown in the right panel of Fig. 30. We find that the energy density of gravitational waves induced during inflation is much smaller than that of the gravitational waves induced during the radiation era. We do not show in this figure the mixed term from eq. (352), but we find that it is well approximated by $\langle\Omega_{\rm mix}\rangle_{\rm S} \sim \sqrt{\langle\Omega_{\rm post}\rangle_{\rm S}\langle\Omega_{\rm pre}\rangle_{\rm S}}$, and is therefore also suppressed with respect to the post-inflationary contribution.

We stress that approximating the integrand by its peak value is what allows us to neglect the subdominant terms involving ψ and $\delta q^{(r)}$ in eq. (329). Let us show that this is a good approximation by estimating the relative contribution of each term in this equation during the inflationary epoch. Since the integrand in eq. (329) is symmetric under $\boldsymbol{p} \leftrightarrow \boldsymbol{k}-\boldsymbol{p}$, for the purpose of finding out which terms contribute the most at the point at which this integrand reaches its largest value, we can simply evaluate it at $|\boldsymbol{p}| = |\boldsymbol{k}-\boldsymbol{p}|$ as we did for $K_{\rm pre}$. We can then define

$$S_{\rm GW}(N,k) \equiv \underbrace{8\langle\mathcal{P}_\psi(N,k)\rangle}_{S_\psi} + \underbrace{\frac{4}{(\rho+p)M_p^2}\langle\mathcal{P}_{\delta q^{(r)}}(N,k)\rangle}_{S_{\delta q^{(r)}}} + \underbrace{\frac{4\phi'^2}{a^2(\rho+p)M_p^2}\langle\mathcal{P}_{\delta\phi}(N,k)\rangle}_{S_{\delta\phi}}, \quad (363)$$

where we have ignored the mixed terms in eq. (329), since they are always subdominant. This quantity is shown in Fig. 31 for two different modes, both of which become super-Hubble near the end of the strongly dissipative phase. The figure illustrates that the time integral of

[52] The value $N_{\rm max}$ at which $K_{\rm pre}$ peaks is really a function of y, as can be seen in the left panel of Fig. 30. However, since the largest contribution to the integral comes from the region around $ky = k_{\rm peak}$, to make the calculation numerically less demanding we can simply take $N_{\rm max}$ as the value at which the integrand, evaluated at $ky = k_{\rm peak}$, is peaked, and use the same value $N_{\rm max}$ for all y. We have explicitly checked that the peak of the signal remains unchanged if the y-dependence of $N_{\rm max}$ is taken into account.

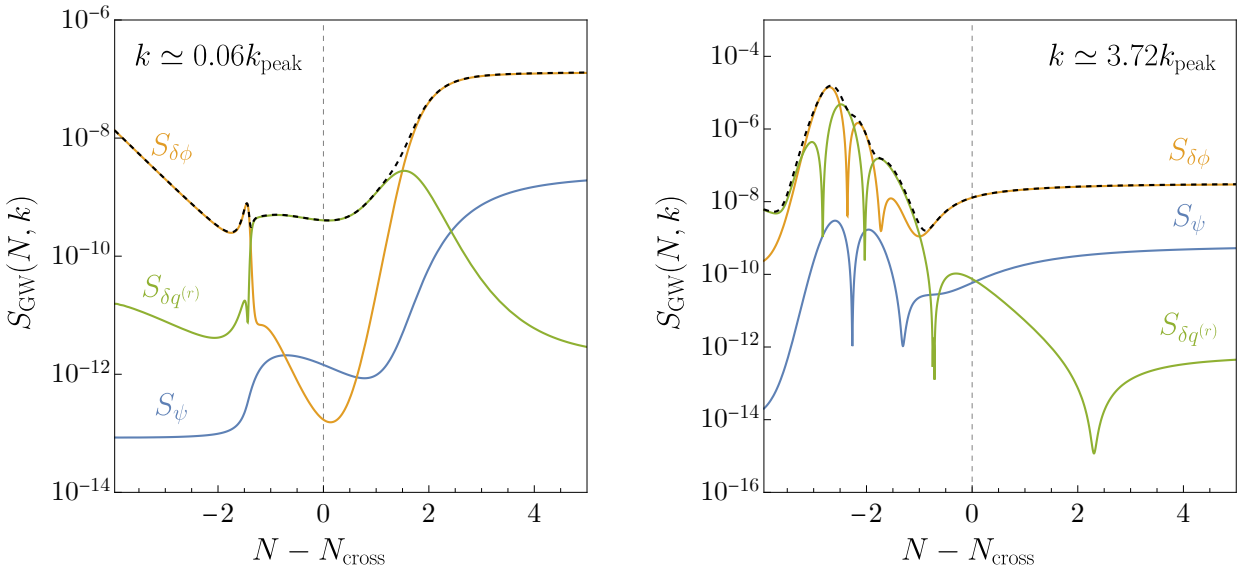

Figure 31: *The different components entering into the quantity $S_{\rm GW}$ defined in eq. (363) are shown as a function of time for two different values of k. We use the numerical results of Section 4.2.*

the scalar source of noise is dominated by the $\delta\phi$ term. Since the other contributions are at most of the same order and will therefore not change the size of the peak in the gravitational wave signal, we can neglect them.

Let us briefly summarize the results of this section. We have made three different approximations in this calculation. The first is that we approximated the time integrals as their peak value times an appropriately chosen area. The second is that we have neglected the contribution of the radiation and metric terms to the gravitational wave source during the inflationary epoch, and we have checked explicitly that the contribution of these terms is indeed subdominant. These two approximations are very good and should not change the order of magnitude of the result. If the time integrals are performed numerically and the metric and radiation perturbations are included in the source term, we expect that the size of the signal will change at most by a factor of $\mathcal{O}(1)$. The final approximation is more delicate, and has to do with the fact that the integrals diverge in the far past (and for large momenta) due to the behaviour of scalar modes in the Bunch-Davies vacuum. We have assumed that this effect can be taken into account correctly by imposing reasonable cutoffs in the integrals. We expect this to be the case because, up to the cutoff scale, the evolution of all quantities proceeds as in standard slow-roll (warm) inflation, so we do not expect the corresponding gravitational wave signal arising from the renormalization of the divergence to be peaked at any particular scale or possess any other features. We have checked that indeed our results are independent of the cutoffs unless unreasonably large choices are made. The uncertainty due to this approximation is difficult to quantify, but we remark that due to the aforementioned argument, our results should be correct as an order of magnitude estimate.

Conclusions

Primordial black holes are intriguing astrophysical objects. As we have seen, their formation does not require extensions to the Standard Model other than inflation, so they are relatively economical dark matter candidates. They also exhibit many astrophysical signatures and, in particular, in the formation mechanisms we have examined here they would generate a large signal in the stochastic background of gravitational waves that could be observed in future experiments. In this thesis we have studied several aspects related to the formation and observable signatures of these black holes. We have presented three specific scenarios in which a large population of black holes that could account for the entirety of the observed dark matter can be generated from single-field inflation. Two of these scenarios rely on the presence of an approximate inflection point in the inflationary potential, which in turn generates large primordial fluctuations on scales much smaller than the ones probed by CMB experiments. In the third scenario, large primordial fluctuations are obtained instead from a transient dissipative phase during inflation in which a source of thermal noise enhances the perturbations. Regions where the density fluctuations are large enough collapse into PBHs when these perturbations become sub-horizon during the radiation epoch.

The first of the models we have presented is characterized by a quartic polynomial inflationary potential, together with a non-minimal coupling between the inflaton and gravity. If the parameters of the potential are chosen appropriately, this model can generate a large population of PBHs in the window (1) able to to account for the entirety of dark matter. If we restrict the analysis to the renormalizable case, we find that the predicted tilt of the power spectrum on CMB scales is in a 3σ tension with the latest Planck analysis for the base ΛCDM model once we require $f_{\rm PBH} \sim \mathcal{O}(1)$. This tension can be alleviated either by considering simple, well-motivated extensions to the base ΛCDM model, such as the addition of neutrino masses, or by considering higher-dimensional operators in the potential, whose presence is expected in the context of suitable UV completions of the theory. In the former case, the spectral index turns out to be smaller than the one reported by Planck, and compatible with our numerical examples, whereas in the latter case the higher-dimensional operators change the potential at large field values, allowing for a better fit of the spectrum at CMB scales.

The second model we have studied is aimed at ameliorating some of the issues present in inflection-point models of PBH formation; in particular, the fact that the presence of the inflection point is usually engineered for the sole purpose of PBH production, and that the

abundance of the PBHs that are generated during the radiation era is exponentially sensitive to the amplitude of the primordial fluctuations at those scales, so that, in order to account for a large fraction of the dark matter, the parameter in the potential that controls the shape of the inflection point must be carefully adjusted. The scenario we present, based on a monodromy-inspired inflationary potential with oscillations superimposed on a term that behaves as a monomial at large field values, together with a phase of early matter domination after inflation, alleviates these issues. In our potential, the existence of the inflection points does not arise from carefully crafting the parameters, but it is instead a rather generic property that emerges whenever the axionic oscillations have a relatively large amplitude. The second ingredient of our scenario can also be easily realized: preheating can be avoided and a long epoch of matter domination after inflation can take place due to perturbative reheating, since the minimum of the potential is approximately quadratic. During this period the likeliness of gravitational collapse is augmented due to the absence of radiation pressure, so that the size of the fluctuations required for PBHs to be able to account for the entirety of dark matter is reduced with respect to the usual radiation scenario. Moreover, the PBH abundance is much less sensitive to small changes in the size of the power spectrum, leading to a reduction in the tuning of the parameters in the potential. The agreement with CMB constraints (except for the tensor-to-scalar ratio r) turns out to be excellent in this model. We have also shown in a model-independent way that PBH formation during an eMD era is most efficient for reheating temperatures of order $\lesssim 10^5$ GeV.

We have shown that a large enhancement of the power spectrum can also be obtained from a transient dissipative phase during inflation. This mechanism is quite different from the other models we present, since the enhancement is not due to a sharp decrease of the inflaton velocity, but rather due to a stochastic source of thermal noise in the equation of motion for the fluctuation that arises from the interaction between the inflaton and an underlying thermalized radiation bath. We have solved the stochastic differential equations for the fluctuations numerically to find the probability distribution of the power spectrum, which turns out to be well-approximated by a skew-normal distribution. Since the quantity of interest is not necessarily the full probability distribution, but its expectation value, we have developed a way to calculate the latter by reducing the full system of stochastic differential equations for the fluctuations to a single, deterministic matrix differential equation for their statistical moments, which is computationally much less demanding to solve. We have also shown that it is possible to simplify the system of equations in such a way that it can be solved analytically. This procedure allowed us to understand qualitatively the most important features of the spectrum.

In addition to the concrete models we have presented, we have also analyzed the role of quantum diffusion during inflation in scenarios featuring an ultra-slow-roll phase, which is particularly relevant for the inflection-point models of PBH production discussed above. The backreaction of quantum fluctuations on the inflationary dynamics is described in the context

of stochastic inflation. In this approach the long-wavelength perturbations of the inflaton are sourced by the short-wavelength ones, which become semi-classical after crossing a suitably-chosen coarse-graining scale. After this stage, the latter behave as a classical source of stochastic noise. We have analyzed the role of the quantum-to-classical transition in this case by using the occupation number density as a test for classicalization, and determined that the choice of the coarse-graining scale is critical: spurious results are obtained for the power spectrum if the modes are not given enough time to classicalize. If the coarse-graining scale is chosen appropriately, the resulting power spectrum coincides, at the linear level, with the result obtained using the standard methods of perturbation theory. This outcome was already known in the case of slow-roll inflation (see e.g. [72]) and we have shown that it holds also in presence of an USR phase. The classicalization of the modes is also shown to be equivalent to the condition that the modes freeze outside the Hubble horizon, much like in the standard slow-roll scenario, so that classicalization only occurs after the USR phase has ended. We have solved the stochastic equations by expanding at first order the coarse-grained inflaton field and its conjugate momentum around their classical trajectories and treating their perturbations as classical random variables. The evolution of the perturbations is then governed by a noise correlation matrix which, in standard slow-roll, is constant in time and vanishes in the field velocity direction of phase space. We find instead that during the USR phase its entries decrease exponentially in all of the phase space directions. The intuitive reason for this is that, as explained before, in this case classicalization is delayed and the inflow of modes sourcing the coarse-grained part of the field stops. We have computed the power spectrum of curvature perturbations using an analytical model that, despite its simplicity, is remarkably powerful and allows us to understand qualitatively the most important features of the spectrum. We have confirmed our results by repeating the analysis in a fully numerical manner making use of the polynomial model presented earlier.

The final part of the thesis was focused on the calculation of the stochastic gravitational wave signals induced at second order in perturbation theory for all of the scenarios we describe here. In particular, we have discussed the gauge-dependence of the signal, which has already been noted several times in the literature for matter- and radiation-dominated universes. We have shown that this gauge-dependence extends to the case in which the Universe goes through an early matter-radiation transition after inflation. The abundance of gravitational waves $\Omega_{\rm GW}$ beyond leading order in perturbations has been obtained in previous studies by simply taking the leading-order expression, which depends on the power spectrum $\mathcal{P}_h$ of $h_{ij}^{\rm TT}$, and making the substitution $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ (where the bold symbol $\boldsymbol{h}_{ij}^{\rm TT}$ denotes the second-order transverse, traceless tensor modes). We have used a heuristic argument to compute the full expression for $\Omega_{\rm GW}$, valid in any gauge and independent of the energy content of the Universe. Our approach shows that the previous substitution yields the correct result only if $\boldsymbol{h}_{ij}^{\rm TT}$ is computed in the Newtonian gauge.

We have examined the tensor spectrum induced by large scalar perturbations in the

scenario in which primordial black holes form during an early matter-dominated era and find that PBH formation during such a phase is fairly constrained. This is due to the fact that the scalar perturbations that source the tensor modes do not decay during an early matter-dominated era, thereby increasing the gravitational wave signal up to the point that the bounds on $\Omega_{\rm GW}$ arising from Big Bang nucleosynthesis and CMB experiments are violated. However, not all the parameter space is ruled out, and points in the remaining available region can be probed by the LISA experiment. We have also computed the gravitational wave signal induced in the warm inflation scenario both during inflation and during the subsequent radiation-dominated era by taking into account the stochastic nature of the fluctuations. We find that the result is dominated by the post-inflationary contribution, which has a size comparable to that of the inflection-point models discussed earlier and is therefore also observable by LISA. To estimate the inflationary contribution, we have used our analytical model and introduced a reasonable cutoff (which in any case does not strongly affect the results) in the time integrals. We have also written an explicit expression for the term that mixes both contributions, and find that it is also subdominant.

There are several possible directions in which our results could be extended. The formulas for the mass and abundance of PBHs formed during an eMD era rely on the results of [92, 93], which are purely analytical and involve many approximations. Although we expect the calculations in Section 1.2 to hold as order of magnitude estimates, more work is needed in terms of numerical simulations in order to precisely determine the effect of non-sphericity and angular momentum on PBH formation during an eMD era. In the stochastic inflation chapter we focused purely on the calculation of the power spectrum, but we remark that it would be interesting to extend the analysis, using our analytical model, to higher order correlators, in order to determine the non-Gaussian corrections. One possible way of accomplishing this would be to follow the approach adopted in [73] in the context of slow-roll inflation, which consists on computing the full probability distribution function of the comoving curvature perturbation using the stochastic δN formalism. In our warm inflation scenario we used a phenomenological parameterization to model the transfer of energy from the inflaton to the radiation background. Peaked dissipative coefficients such as the one we have proposed are, however, not common in standard warm inflation scenarios, and although we have provided a possible route to describe the microphysical origin of this dissipation by starting from a particular Lagrangian (in Appendix D), more work is needed to find concrete, well-motivated models. Finally, our argument to derive the full expression for the gravitational wave energy density is purely heuristic, and it would be interesting to attempt to formalize it by properly defining a covariant coarse-graining procedure and deriving the expression for the energy density of gravitational waves from first principles. We hope that all of these directions can be explored in future work.

Conclusiones

Los agujeros negros primordiales son objetos astrofísicos intrigantes. Como hemos vito, su formación no requiere de extensiones al Modelo Estándar además de la inflación, así que son candidatos a la materia oscura cuyo costo es relativamente bajo. Exhiben además una serie de señales astrofísicas y, en particular, en los mecanismos de formación que hemos examinado en esta tesis se generaría una señal en el fondo estocástico de ondas gravitacionales que podría ser observada en experimentos futuros. En esta tesis hemos estudiado distintos aspectos relacionados con la formación y señales observables de estos agujeros negros. Hemos presentado tres escenarios específicos en los que una gran población de agujeros negros capaz de explicar toda la materia oscura observada puede ser generada a partir de modelos de inflación con un solo campo. Dos de estos escenarios dependen de la presencia de un punto de inflexión aproximado en el potencial inflacionario, que a su vez genera grandes fluctuaciones de energía primordiales a escalas mucho más pequeñas que las observadas por experimentos del fondo cósmico de microondas. En el tercer escenario, las fluctuaciones primordiales son obtenidas de una fase disipativa de corta duración durante la inflación en la cual una fuente de ruido térmico hace crecer las perturbaciones. Las regiones en las que las fluctuaciones de densidad son lo suficientemente grandes colapsan en PBHs cuando estas perturbaciones cruzan de nuevo el horizonte durante la época de radiación.

El primero de los modelos que hemos presentado está caracterizado por un potencial inflacionario que consiste en un polinomio de cuarto orden, así como un acoplamiento no mínimo entre el inflatón y la gravedad. Si los parámetros del potencial son elegidos apropiadamente, este modelo puede generar una población de PBHs en la ventana (1) capaz de explicar la totalidad de la materia oscura. Si restringimos nuestro análisis al caso renormalizable, encontramos que la inclinación del espectro de potencias a escalas del CMB se encuentra en una tensión de 3σ con los últimos análisis de Planck para el modelo base ΛCDM si requerimos $f_{\rm PBH} \sim \mathcal{O}(1)$. Esta tensión puede ser aliviada considerando extensiones simples y bien motivadas al modelo base ΛCDM, tales como la adición de las masas de los neutrinos, o bien considerando operadores de dimensiones altas en el potencial, cuya presencia se espera en el contexto de extensiones ultravioleta de la teoría. En el primer caso, el índice espectral resulta ser más pequeño que el reportado por Planck y compatible con nuestros ejemplos numéricos, mientras que en el segundo caso los operadores cambian el potencial a valores grandes del campo, permitiendo un mejor ajuste del espectro a escalas del CMB.

El segundo modelo que hemos estudiado intenta mejorar algunos de los problemas pre-

sentes en los modelos de formación de PBHs con puntos de inflexión; en particular, el hecho de que el punto de inflexión usualmente esté presente con el único propósito de producir agujeros negros y el que la abundancia de los PBHs generados durante la era de radiación sea exponencialmente sensible a la amplitud de las fluctuaciones primordiales a esas escalas de manera que, para poder explicar una porción significativa de la materia oscura, el parámetro del potencial que controla la forma del punto de inflexión debe de ser ajustado cuidadosamente. El escenario que presentamos, basado en un potencial inspirado por la monodromía axiónica con oscilaciones superpuestas sobre un término que se comporta como un monomio a valores grandes del campo, junto con una fase temprana en la cual el universo estaría dominado por materia no relativista, alivia estos problemas. En nuestro potencial, la existencia del punto de inflexión no surge de ajustar los parámetros cuidadosamente, sino que es una propiedad genérica que emerge cuando las oscilaciones axiónicas tienen una amplitud relativamente grande. El segundo ingrediente de nuestro escenario puede ser realizado fácilmente también: el precalentamiento puede ser evitado y una época larga dominada por materia no relativista puede tener lugar debido al recalentamiento perturbativo, ya que el mínimo del potencial es aproximadamente cuadrático. Durante este periodo, la probabilidad del colapso gravitacional aumenta debido a la ausencia de la presión de radiación, de manera que el tamaño de las fluctuaciones requerido para explicar toda la materia oscura con PBHs se ve reducido con respecto al escenario usual de radiación. Más aún, la abundancia de los PBHs es mucho menos sensible a pequeños cambios en el tamaño del espectro de potencias, lo cual conlleva una reducción en el ajuste de los parámetros del potencial. El ajuste de los parámetros del CMB resulta ser excelente en este modelo (exceptuando el cociente entre escalares y tensores r). Hemos mostrado también de una forma independiente de modelos que la formación de PBHs durante una época dominada por materia no relativista resulta ser más eficiente para temperaturas de recalentamiento de orden $\lesssim 10^5$ GeV.

Hemos mostrado que un incremento del espectro de potencias puede ser obtenido también de una fase disipativa de corta duración durante la inflación. Este mecanismo es muy distinto de los otros modelos que presentamos, ya que el incremento no se debe a un decrecimiento súbito en la velocidad del inflatón, sino a una fuente estocástica de ruido térmico en la ecuación de movimiento para las fluctuaciones que surge de la interacción entre el inflatón y un baño de radiación en equilibrio térmico. Hemos resuelto las ecuaciones diferenciales estocásticas para las fluctuaciones numéricamente para encontrar la distribución de probabilidad del espectro, que resulta ser correctamente aproximada por una distribución normal sesgada. Ya que la cantidad de interés no es necesariamente la distribución de probabilidad total del espectro, sino el valor esperado del mismo, hemos desarrollado un método para calcular esta cantidad reduciendo el sistema completo de ecuaciones diferenciales estocásticas a una única ecuación matricial determinista para sus momentos estadísticos, la cual es computacionalmente mucho más fácil de resolver. Hemos mostrado además que es posible simplificar el sistema de ecuaciones de manera que pueda ser resuelto analíticamente.

Este proceso nos permitió entender cualitativamente las propiedades más importantes del espectro.

Además de los modelos concretos que hemos presentado, hemos analizado también el rol de la difusión cuántica durante la inflación en escenarios que presentan una fase de rodamiento muy lento, lo cual es particularmente relevante para los modelos de formación de PBHs con puntos de inflexión discutidos antes. La reacción de las fluctuaciones cuánticas sobre la dinámica inflacionaria es descrita en el contexto de la inflación estocástica. En este enfoque las perturbaciones del inflación con longitudes de onda cortas actúan como fuente para las de longitudes de onda grandes y se vuelven semiclásicas después de cruzar una escala elegida apropiadamente. Después de este punto, se comportan como una fuente clásica de ruido estocástico. Hemos analizado el rol de la transición del régimen cuántico al clásico en este caso utilizando el número de ocupación como una prueba de clasicalización y hemos determinado que la elección de la escala que separa a la parte clásica del campo de la cuántica es crucial: se obtienen resultados incorrectos para el espectro de potencias si no se le da a los modos suficiente tiempo para volverse clásicos. Si la escala se elige apropiadamente, el espectro coincide, a nivel lineal, con el resultado obtenido utilizando los métodos estándar de la teoría de perturbaciones. Este resultado era conocido ya en el caso inflacionario usual (véase por ejemplo [72]) y hemos demostrado que resulta ser cierto también en presencia de una fase de rodamiento muy lento. El tiempo de clasicalización de los modos resulta ser equivalente al tiempo que tardan los modos en congelarse fuera del horizonte de Hubble, tal como en el caso inflacionario usual, de modo que la clasicalización solo ocurre una vez que ha concluido la fase de rodamiento muy lento. Hemos resuelto las ecuaciones estocásticas expandiendo a primer orden el inflatón y su momento conjugado alrededor de sus trayectorias clásicas y tratando las perturbaciones como variables estocásticas clásicas. La evolución de las perturbaciones está gobernada por una matriz de correlación que, en el caso inflacionario usual, es constante en el tiempo y se anula solo en la dirección de la velocidad en el espacio de fase. En contraste, encontramos que durante una fase de rodamiento muy lento sus entradas decrecen exponencialmente en todas las direcciones. La razón intuitiva de este hecho es que, como hemos explicado antes, en este caso la clasicalización tarda más y la entrada de modos que actúan como fuente para la parte clásica del campo se detiene. Hemos calculado el espectro de potencias utilizando un modelo analítico que, a pesar de su simplicidad, es bastante poderoso y nos permite entender cualitativamente las propiedades más importantes del espectro. Hemos confirmado nuestros resultados repitiendo el análisis numéricamente utilizando el modelo polinómico antes mencionado.

La última parte de la tesis estuvo enfocada en el cálculo del fondo estocástico de ondas gravitacionales inducido a segundo orden en teoría de perturbaciones para todos los escenarios que describimos aquí. En particular, hemos discutido la dependencia de la señal del sistema de coordenadas, que ha sido descrita ya numerosas veces en la literatura para universos dominados por radiación y materia no relativista. Hemos mostrado que esta dependencia

se extiende al caso en que el universo pasa a través de una fase temprana dominada por materia no relativista después de la inflación. La abundancia de las ondas gravitacionales $\Omega_{\rm GW}$ a ordenes subdominantes en perturbaciones ha sido obtenida en estudios previos tomando la expresión a orden dominante, que depende del espectro de potencias $\mathcal{P}_h$ de $h_{ij}^{\rm TT}$ y haciendo la sustitución $h_{ij}^{\rm TT} \to \frac{1}{2}\boldsymbol{h}_{ij}^{\rm TT}$ (donde el símbolo $\boldsymbol{h}_{ij}^{\rm TT}$ denota los modos tensoriales transversos y sin traza de segundo orden). Hemos utilizado un argumento heurístico para calcular la expresión completa para $\Omega_{\rm GW}$, válida en cualquier sistema de coordenadas e independiente del contenido energético del universo. Nuestro enfoque muestra que la sustitución antes mencionada produce el resultado correcto solo si la cantidad $\boldsymbol{h}_{ij}^{\rm TT}$ es calculada en el sistema de coordenadas de Newton.

Hemos examinado el espectro de potencias tensorial inducido por perturbaciones escalares grandes en el escenario en que los agujeros negros primordiales se forman durante una época temprana dominada por materia no relativista y encontramos que la formación de PBHs durante esta fase está bastante constreñida. Esto se debe al hecho de que las perturbaciones escalares que actúan como fuente para las tensoriales no decaen durante esta fase, incrementando la señal de ondas gravitacionales hasta el punto en que se violan los límites sobre $\Omega_{\rm GW}$ que surgen de experimentos de la nucleosíntesis del Big Bang y el CMB. Sin embargo, no todo el espacio de parámetros es eliminado y los puntos que quedan disponibles pueden ser probados por el experimento LISA. Hemos también calculado la señal de ondas gravitacionales inducida en el escenario de inflación caliente tanto durante la inflación como durante la era de radiación subsiguiente, tomando en cuenta la naturaleza estocástica de las fluctuaciones. Encontramos que el resultado está dominado por la contribución post-inflacionaria, que tiene un tamaño comparable al de los modelos con puntos de inflexión antes discutidos y es por lo tanto observable también por el experimento LISA. Para estimar la contribución inflacionaria utilizamos nuestro modelo analítico e introducimos una escala de corte razonable (que en todo caso no afecta fuertemente los resultados) en las integrales temporales. Hemos escrito también expresiones explícitas para el término que mezcla ambas contribuciones y encontramos que es también subdominante.

Existen una serie de direcciones posibles en las cuales se pueden extender los resultados presentados aquí. Las fórmulas para la masa y abundancia de los PBHs formados durante una época temprana dominada por materia no relativista están basadas en los resultados de [92, 93], que son totalmente analíticos e involucran una serie de aproximaciones. Los cálculos de la Sección 1.2 son válidos como estimaciones de órdenes de magnitud, pero es necesario más trabajo en términos de simulaciones numéricas para determinar de forma precisa el efecto de la no-esfericidad y el momento angular sobre la formación de PBHs durante una fase temprana dominada por materia no relativista. En el capítulo de inflación estocástica nos enfocamos únicamente en el cálculo del espectro de potencias, pero sería interesante también extender el análisis, utilizando nuestro modelo analítico, a correladores de orden mayor para determinar las correcciones no gaussianas. Una posible manera de

lograr esto sería seguir el enfoque adoptado en [73] en el contexto del régimen de rodamiento lento, que consiste en calcular la distribución de probabilidades de las perturbaciones utilizando el formalismo δN estocástico. En nuestro escenario de inflación caliente utilizamos una parametrización fenomenológica para modelar la transferencia de energía del inflatón al fondo de radiación. Sin embargo, los coeficientes disipativos localizados como el presentado aquí no son comunes en escenarios estándar de inflación caliente y aunque hemos desarrollado una posible ruta para describir el origen microfísico de esta disipación empezando de un Lagrangiano particular (en el Apéndice D), se necesita más trabajo para encontrar modelos concretos y bien motivados. Finalmente, nuestro argumento para derivar la expresión completa para la densidad de energía de las ondas gravitacionales es puramente heurístico y sería interesante intentar formalizarlo definiendo apropiadamente un procedimiento para promediar las ecuaciones de Einstein y derivar así la expresión para la densidad de energía de las ondas gravitacionales a partir de principios fundamentales. Esperamos que todas estas direcciones puedan ser exploradas en trabajos futuros.

Appendices

A ■ Cosmological perturbation theory

In this appendix we review the aspects of cosmological perturbation theory most relevant for the thesis. A review of second-order perturbation theory can be found in [189]. The starting point for perturbation theory is the assumption that the metric can be split as

$$g_{\mu\nu} = \hat{g}_{\mu\nu} + \epsilon h_{\mu\nu} + \frac{1}{2}\epsilon^2 \boldsymbol{h}_{\mu\nu} + \cdots \tag{364}$$

for some small parameter ϵ, where $\hat{g}_{\mu\nu}$ is some fixed background metric and we use bold symbols for second-order quantities. Indices are raised and lowered with $\hat{g}_{\mu\nu}$, and covariant derivatives taken with respect to it as well. This decomposition is not unique, but any frame in which the metric can be written in this form is related to this one by infinitesimal diffeomorphisms, or gauge transformations. These gauge transformations are generated by some vector v^μ,

$$v^\mu = \epsilon\xi^\mu + \frac{1}{2}\epsilon^2 \boldsymbol{\xi}^\mu + \cdots . \tag{365}$$

A generic tensor then transforms as

$$T^{\mu_1\cdots\mu_n}{}_{\nu_1\cdots\nu_m} \to \tilde{T}^{\mu_1\cdots\mu_n}{}_{\nu_1\cdots\nu_m} = e^{\mathcal{L}_v} T^{\mu_1\cdots\mu_n}{}_{\nu_1\cdots\nu_m} , \tag{366}$$

where $\mathcal{L}_v$ is the Lie derivative with respect to v, and the exponential map is defined by

$$e^{\mathcal{L}_v} = 1 + \epsilon\mathcal{L}_\xi + \frac{1}{2}\epsilon^2 \left(\mathcal{L}_\xi^2 + \mathcal{L}_{\boldsymbol{\xi}}\right) + \cdots . \tag{367}$$

Throughout this thesis we assume the following form for the scalar-vector-tensor decomposition of the metric,

$$ds^2 = -a^2 (1+2\varphi)\, d\tau^2 + 2a^2 \partial_i B dx^i d\tau + a^2 \left[(1-2\psi)\,\delta_{ij} + 2\partial_i\partial_j E + \frac{1}{2}\boldsymbol{h}_{ij}^{\mathrm{TT}} \right] dx^i dx^j . \tag{368}$$

Each one of these quantities can be expanded in powers of ϵ. Their corresponding equations of motion can be found by expanding Einstein's equations in powers of ϵ and equating coefficients. The quantity $\boldsymbol{h}_{ij}^{\mathrm{TT}}$ denotes the transverse, traceless second order piece of the metric, which satisfies $\partial^i \boldsymbol{h}_{ij}^{\mathrm{TT}} = \delta^{ij} \boldsymbol{h}_{ij}^{\mathrm{TT}} = 0$. Throughout this thesis we assume that the first-order vectors and tensor modes vanish, $E_i = B_i = h_{ij}^{\mathrm{TT}} = 0$. Second-order vectors are nonzero, but they are irrelevant to our discussion and thus we do not write them.

Our conventions for the stress-energy perturbations are

$$\delta T^{\tau}{}_{\tau} = -\delta\rho, \tag{369}$$

$$\delta T^{\tau}{}_{i} = (\rho + p)\partial_i(\delta v + B) \equiv \frac{1}{a}\partial_i \delta q, \tag{370}$$

$$\delta T^{i}{}_{j} = \delta p\, \delta^{i}_{j}. \tag{371}$$

If we denote the components of the vector that generates the gauge transformation by $\xi^\mu = (\alpha, \delta^{ij}\partial_j\beta)$, these quantities transform as

$$\delta\rho \to \delta\tilde{\rho} = \delta\rho + \rho'\alpha, \tag{372}$$

$$\delta p \to \delta\tilde{p} = \delta p + p'\alpha, \tag{373}$$

$$\delta q \to \delta\tilde{q} = \delta q - a(\rho + p)\alpha. \tag{374}$$

By using eqs. (366) and (16) we find the following transformation laws for first-order scalars,

$$\varphi \to \tilde{\varphi} = \varphi + \mathcal{H}\alpha + \alpha', \tag{375}$$

$$\psi \to \tilde{\psi} = \psi - \mathcal{H}\alpha, \tag{376}$$

$$E \to \tilde{E} = E + \beta, \tag{377}$$

$$B \to \tilde{B} = B - \alpha + \beta', \tag{378}$$

The gauge-transformation law for the second-order tensor mode is given in eq. (289).

The evolution of first-order scalars can be found from the linearized Einstein equations. In momentum space and conformal time, these are

$$3\mathcal{H}\big(\psi_k' + \mathcal{H}\varphi_k\big) + k^2\big(\psi_k + \mathcal{H}\sigma_k\big) = -a^2\frac{\delta\rho_k}{2M_p^2}, \tag{379}$$

$$\psi_k'' + \mathcal{H}\big(2\psi_k' + \varphi_k'\big) + \big(\mathcal{H}^2 + 2\mathcal{H}'\big)\varphi_k = a^2\frac{\delta p_k}{2M_p^2}, \tag{380}$$

$$\psi_k' + \mathcal{H}\varphi_k = -a\frac{\delta q_k}{2M_p^2}, \tag{381}$$

$$\sigma_k' + 2\mathcal{H}\sigma_k + \psi_k - \varphi_k = 0, \tag{382}$$

where we have neglected anisotropic stress and defined the shear potential $\sigma \equiv E' - B$. In what follows we also assume $\delta p = w\delta\rho$. Thus, we have five equations for seven variables, two of which can be set to zero by fixing the gauge, and so the system is completely determined. The other equation of motion we are concerned with is the one for the second-order tensor modes, which in position space and in the absence of anisotropic stress is given by

$$\boldsymbol{h}_{ij}^{\mathrm{TT}\prime\prime} + 2\mathcal{H}\boldsymbol{h}_{ij}^{\mathrm{TT}\prime} - \nabla^2\boldsymbol{h}_{ij}^{\mathrm{TT}} = -4\mathcal{T}_{ij}{}^{lm}s_{lm}, \tag{383}$$

where $\mathcal{T}_{ij}{}^{lm}$ is the transverse-traceless projector defined in eq. (262) and the source term s_{ij} is, in a generic gauge [85, 89]

$$
\begin{aligned}
s_{ij} = {} & \partial_i\psi\partial_j\psi + \partial_i\varphi\partial_j\varphi - \Big(\varphi' + \psi' - \nabla^2\sigma\Big)\partial_i\partial_j\sigma + \Big(\partial_i\varphi'\partial_j\sigma + \partial_j\varphi'\partial_i\sigma\Big) + 2\Big(\varphi + \psi\Big)\partial_i\partial_j\psi \\
& - \frac{4}{3}\frac{\rho}{\rho + p}\partial_i\left(\frac{\psi'}{\mathcal{H}} + \varphi\right)\partial_j\left(\frac{\psi'}{\mathcal{H}} + \varphi\right) - 2\nabla^2 E\partial_i\partial_j\psi + 2\Big(\psi'' + 2\mathcal{H}\psi' - \nabla^2\psi\Big)\partial_i\partial_j E \\
& - \partial_i\partial_k E'\partial_j\partial^k E' + \partial_i\partial_k\partial_l E\partial_j\partial^k\partial^l E + 2\Big(\partial_j\partial_k\psi\partial_i\partial^k E + \partial_i\partial_k\psi\partial_j\partial^k E\Big) + 2\psi'\partial_i\partial_j E' \\
& - 2\mathcal{H}\Big(\partial_i\psi\partial_j E' + \partial_j\psi\partial_i E'\Big) - \Big(\partial_i\psi'\partial_j E' + \partial_j\psi'\partial_i E'\Big) - \Big(\partial_i\psi\partial_j E'' + \partial_j\psi\partial_i E''\Big) \\
& + \partial_i\partial_j\partial_k\partial^k\Big(E'' + 2\mathcal{H}E' - \nabla^2 E\Big) - \partial_i\partial_k\sigma\partial_j\partial^k\sigma.
\end{aligned}
\tag{384}
$$

Let us review the procedure presented in [189] to construct gauge-invariant quantities by starting from an arbitrary gauge and adding a term to cancel the gauge transformation. The change of the second-order, transverse, traceless tensor mode $\boldsymbol{h}_{ij}^{\mathrm{TT}}$ under a gauge transformation is

$$
\boldsymbol{h}_{ij}^{\mathrm{TT}} \to \tilde{\boldsymbol{h}}_{ij}^{\mathrm{TT}} = \boldsymbol{h}_{ij}^{\mathrm{TT}} + \mathcal{T}_{ij}^{lm}\Sigma_{lm},
\tag{385}
$$

where Σ_{ij} is given in eq. (289) and depends on the parameters α and β of the coordinate transformation as well as the perturbations in the starting gauge, and where the projection tensor $\mathcal{T}_{ij}{}^{lm}$, defined in momentum space in eq. (262), extracts the transverse, traceless piece of Σ_{ij}, and can be written in position space as

$$
\begin{aligned}
\mathcal{T}_{ij}{}^{lm}\Sigma_{lm} \equiv {} & \Sigma_{ij} + \frac{1}{2}\big(\nabla^{-2}\partial_k\partial_\ell\Sigma^{k\ell} - \Sigma^k{}_k\big)\delta_{ij} + \frac{1}{2}\nabla^{-2}\nabla^{-2}\partial_i\partial_j\partial_k\partial_\ell\Sigma^{k\ell} \\
& + \frac{1}{2}\nabla^{-2}\partial_i\partial_j\Sigma^k{}_k - \nabla^{-2}\big(\partial_j\partial^k\Sigma_{ik} + \partial^k\partial_i\Sigma_{jk}\big).
\end{aligned}
\tag{386}
$$

A gauge-invariant quantity X_{ij} can then be easily obtained by simply adding the transformation to $\boldsymbol{h}_{ij}^{\mathrm{TT}}$ (computed in the arbitrary starting gauge) [189],

$$
X_{ij} \equiv \boldsymbol{h}_{ij}^{\mathrm{TT}} + \mathcal{T}_{ij}^{lm}\Sigma_{lm}.
\tag{387}
$$

Suppose we start from a generic gauge and perform a transformation to the Newtonian gauge, so that $E \to \tilde{E} = 0$ and $B \to \tilde{B} = 0$. This fixes the transformation parameters α and β as

$$
\alpha = B - E', \qquad \beta = -E
\tag{388}
$$

according to eqs. (377, 378). We reiterate that these are the perturbations in the starting gauge. By substituting α and β into eq. (289) and writing the resulting expression in mo-

mentum space, we obtain

$$\Sigma_k^s = \int \frac{d^3p}{(2\pi)^3} \boldsymbol{e}_{ij}^s \big[2(E'_{k-p} - B_{k-p}) p_i p_j E'_p + (k-p)_i B_{k-p} p_j E'_p + (k-p)_j B_{k-p} p_i E'_p \\ - 2(k-p)_i B_{k-p} p_j B_p - 8\psi_{k-p} p_i p_j E_p - 4 p_j p_k E_p (k-p)_i (k-p)_k E_{k-p} \\ - 2 p_i p_j p_k E_p (k-p)_k E_{k-p} + \delta_{ij}(\cdots) \big]. \tag{389}$$

where the ellipses stand for all terms proportional to δ_{ij}, and Σ_k^s is the Fourier mode defined in eq. (290). The terms proportional to k_i and δ_{ij} vanish upon contracting with $\boldsymbol{e}_{ij}^s$. Thus, after cleaning up,

$$\Sigma_k^s = \int \frac{d^3p}{(2\pi)^3} \boldsymbol{e}_{ij}^s p_i p_j \big[2\boldsymbol{p} \cdot (\boldsymbol{k} - \boldsymbol{p}) E_p E_{k-p} - 8\psi_{k-p} E_p + 2\sigma_{k-p}\sigma_p \big], \tag{390}$$

which coincides precisely with the quantity defined in eq. (309). There is nothing special about the Newtonian gauge in this procedure. Other gauge invariant quantities can be obtained by performing a transformation to a different gauge, which will not necessarily be equal to X_{ij}. However, the uniqueness argument in Section 5.2 guarantees that X_{ij} is the only one of these gauge-invariant quantities that can enter into the calculation of the gravitational wave energy density. See the corresponding discussion for details.

B ■ Power counting and effective theories

In this appendix we review the basics of power counting as a method to organize HDOs. In order to keep track of the appropriate powers of the couplings in the terms of a Lagrangian, it is useful to restore $\hbar$. An equivalent way to account for the presence of $\hbar$ is to introduce units of energy E and length L, so that $[\hbar] = \mathrm{EL}$ (and natural units are recovered if $\mathrm{E} = \mathrm{L}^{-1}$). A Lagrangian density then has dimension $[\mathcal{L}] = \mathrm{EL}^{-3}$ as we would expect, which implies that a canonically normalized scalar field has $[\phi] = \mathrm{E}^{1/2}\mathrm{L}^{-1/2}$. A coupling constant g (a gauge coupling, for instance) has dimensions $[g] = \mathrm{E}^{-1/2}\mathrm{L}^{-1/2}$, and the quartic scalar coupling in a term such as $\lambda\phi^4$ has dimensions $[\lambda] = \mathrm{E}^{-1}\mathrm{L}^{-1}$, so that $[g^2] = [\lambda]$. Any scale, such as M_p or the scale Λ introduced in eq. (100), has dimensions $[\Lambda] = \mathrm{E}^{1/2}\mathrm{L}^{-1/2}$. Instead of using E and L, it is convenient to switch to units of mass $\mathrm{M} = \mathrm{L}^{-1}$ and coupling $\mathrm{C} = \mathrm{E}^{-1/2}\mathrm{L}^{-1/2}$, so that $[\mathcal{L}] = \mathrm{M}^4\mathrm{C}^{-2}$, $[\phi] = \mathrm{MC}^{-1}$ and $[\hbar] = \mathrm{C}^{-2}$. Thus, for the operator in eq. (99) we have

$$\mathrm{M}^4\mathrm{C}^{-2} = [\mathcal{O}_n] = \left[\frac{1}{\Lambda_n}\right]^{n-4} \mathrm{M}^n\mathrm{C}^{-n}, \tag{391}$$

where the first equality comes from the fact that this is a term in the Lagrangian. This implies

$$\Lambda_n = \mathrm{MC}^{-\frac{n-2}{n-4}}. \tag{392}$$

Tree-level operators have dimensions

$$[\mathcal{O}_n^{\text{tree}}] = \left[\frac{g^{n-2}}{M^{n-4}}\phi^n\right] = \frac{\mathrm{C}^{n-2}}{\mathrm{M}^{n-4}}[\phi^n], \tag{393}$$

so that we can use g to keep tracks of the powers of coupling C. As per our earlier comment, loop operators satisfy $[\mathcal{O}_n^{\text{loop}}] = \hbar g^2 [\mathcal{O}_n^{\text{tree}}]$, and we expect $\mathcal{O}_n^{\text{loop}} \ll \mathcal{O}_n^{\text{tree}}$ in a weakly coupled theory (that is, when $g \ll 1$). We can therefore restrict our attention to operators with tree-level scaling in eq. (100).

It is illustrative to use an explicit toy example to clarify the discussion. Consider the following two-field Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi)(\partial^\mu \Phi) - V(\Phi, \phi), \qquad V(\Phi, \phi) = \frac{1}{2}M^2\Phi^2 + \frac{c_1}{2} g M \phi\, \Phi^2 + \frac{c_2}{2} g M \phi^2 \Phi, \tag{394}$$

where we have used powers of g and M to keep track of C and M, as discussed above. The coefficients $c_{(1,2)}$ are genuine dimensionless numbers that cannot be fixed by dimensional arguments alone. If we assume that the Φ field is heavy (that is, $M \gg g\phi$), then it can be integrated out, producing a Lagrangian that only depends on ϕ and where the effect of Φ is included in HDOs. This can be understood by solving the equation of motion for Φ and plugging the result back into the Lagrangian. Said solution can be formally written as

$$\Phi_{\text{EoM}} \simeq \left[\frac{1}{M^2} - \frac{(\Box + c_1 g M \phi)}{M^4} + \frac{(\Box + c_1 g M \phi)^2}{M^6} + \dots\right]\left(-\frac{1}{2} c_2 g M \phi^2\right). \tag{395}$$

The effective Lagrangian therefore contains, restricting our attention to the last term in eq. (394) for concreteness,

$$\begin{aligned}
\mathcal{L}_{\text{eff}} &= \frac{c_2^2}{4} g^2 M^2 \phi^2 \left[\frac{1}{M^2} - \frac{(\Box + c_1 g M \phi)}{M^4} + \frac{(\Box + c_1 g M \phi)^2}{M^6} + \dots\right] \phi^2 \\
&= \frac{c_2^2}{4} g^2 \phi^4 - \frac{c_1 c_2^2}{4} \frac{g^3}{M} \phi^5 + \frac{c_1^2 c_2^2}{4} \frac{g^4}{M^2} \phi^6 + \dots + \text{operators with derivatives} \\
&= \sum_{n=4}^{\infty} \frac{(-1)^n c_1^{n-4} c_2^2}{4} \frac{g^{n-2}}{M^{n-4}} \phi^n + \text{operators with derivatives},
\end{aligned} \tag{396}$$

so that the tree-level operators without derivatives have the following scaling

$$\mathcal{O}_{n\geqslant 4}^{\text{tree}} = \frac{(-1)^{n+1} c_1^{n-4} c_2^2}{4} \frac{g^{n-2}}{M^{n-4}} \phi^n, \tag{397}$$

where the extra -1 factor comes from the sign of the potential in eq. (394). Notice that this example motivates the choice $c_n = (-1)^n$ that we made in one of the examples of Section 2.2. We see that a quartic coupling ϕ^4 is generated by the UV theory, despite not being present in the original Lagrangian, justifying the equality $g^2 = a_4$ that we assumed for simplicity

in our analysis of Section 2.2. We also find that integrating out Φ generates HDOs with derivatives, which should in general be included. The reason for neglecting them in our analysis of Section 2.2 is that, in the slow-roll regime, where the effect of HDOs is relevant, these operators can be neglected (since the velocity of the field and higher-order derivatives are small). The USR phase occurs instead at small field values, where HDOs are suppressed. Finally, this explicit construction shows that the expansion in terms of effective operators is valid as long as $M \gg g\phi$. This is precisely the condition that led to eq. (102) where we introduced the scale $\Lambda = M/g$.

To close this appendix, let us remark that although we generically expect the scaling in eq. (393) to hold based on dimensional analysis, important differences can arise in the series of HDOs depending on the symmetry properties of the UV theory. For instance, if the heavy field Φ is even under some $\mathbb{Z}_2$ symmetry, the only renormalizable terms involving Φ that can appear in the potential are

$$V(\Phi,\phi) = \frac{1}{2}M^2\Phi^2 + \frac{c_{12}}{2}gM\phi\Phi^2 + \frac{c_{22}}{4}g^2\phi^2\Phi^2 + \frac{c_{04}}{4!}g^2\Phi^4. \tag{398}$$

In this case, only loop-suppressed operators are generated. One can then perform an analysis similar to the one presented in Section 2.2, where the most relevant difference is that the condition $g^2 = a_4$ would now be replaced by $\hbar g^4 = a_4$, which implies $g \sim \mathcal{O}(10^{-2})$.

C ■ Stochastic differential equations

In this appendix we review the basic facts about stochastic differential equations, including the definition of a Wiener process and the derivation of the Fokker-Planck equation. Part of the discussion follows the presentation in [205].

A stochastic differential equation (SDE) driven by a Gaussian noise source for a single time-dependent variable $y(t)$ can be written in differential form as

$$dy = f(y,t)dt + g(y,t)dW_t, \tag{399}$$

where the differentials should be thought of as small increments. The increment dW_t is randomly drawn from a Gaussian distribution at each time step dt,

$$P(dW_t) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-dW_t^2/2\sigma^2}. \tag{400}$$

The dW_t increments at every time step are independent from each other. The quantity dW_t is known as a Wiener increment, and is often written alternatively as $dW_t = \xi_t dt$. Throughout this appendix we focus only on stochastic differential equations of the form (399). The different solutions to eq. (399) obtained via some finite sequence of increments dW_t are referred to as realizations.

An important fact is that the only way for the distribution in eq. (400) to be well-defined is if $\sigma^2 = dt$. Let us illustrate this with a heuristic argument. If we consider the equation $dy = dW_t$, a particular realization can be obtained by summing each random increment, $y = \sum_i^N dW_{t(i)}$. The variance of y, denoted by $V(y)$, is simply the sum of the variance of each increment $dW_{t(i)}$ in the sum, since they are assumed to be independent. Suppose σ^2 is equal to some power of dt,

$$\sigma^2 = dt^n. \tag{401}$$

If $T \equiv Ndt$ denotes the range of time over which we solve the equation, the variance of y is

$$V(y) = \sum_i^N V(dW_{t(i)}) = N\sigma^2 = Ndt^n = N\left(\frac{T}{N}\right)^n = N^{1-n}T^n. \tag{402}$$

In the continuum limit $N \to \infty$, this vanishes for $n > 1$, and diverges if $n < 1$. Thus, the only value that makes sense is $n = 1$.

Let us return to the original equation (399) and compute the expectation value of dW_t^2,

$$\langle dW_t^2 \rangle = \langle dW_t^2 \rangle - \langle dW_t \rangle^2 \equiv V(dW_t) = dt, \tag{403}$$

where the expectation values $\langle \cdots \rangle$ are taken with respect to the distribution in eq. (400). By using eq. (400) we can also show that the variance of the sum of the $dW_{t(i)}^2$ is

$$V\left(\sum_i^N dW_{t(i)}^2\right) = \sum_i^N 2dt^2 = 2\frac{T^2}{N}. \tag{404}$$

Since this quantity vanishes as $N \to \infty$, we see that the sum of all the $dW_{t(i)}^2$ in the continuum limit is not actually random, but deterministic, and must therefore be equal to its mean,

$$\int_0^T dW_t^2 = \left\langle \int_0^T dW_t^2 \right\rangle = \int_0^T \langle dW_t^2 \rangle = \int_0^T dt. \tag{405}$$

Thus, we find that $dW_t^2 = dt$, a result known as Ito's rule. We can use this result to change the time variable in a stochastic differential equation via

$$dW_t = \sqrt{\frac{dt}{ds}} dW_s. \tag{406}$$

Ito's rule also leads to a very important lemma. Consider the Taylor expansion to second order of any function $z(y,t)$ of the stochastic variable y and time t,

$$dz = \frac{\partial z}{\partial t}dt + \frac{\partial z}{\partial y}dy + \frac{1}{2}\frac{\partial^2 z}{\partial y^2}dy^2 + \cdots. \tag{407}$$

We would like to write eq. (399) in terms of z instead of y. Using eq. (399) itself to substitute

dy and dy^2 into eq. (407) and keeping only terms linear in dt and dW_t so that the equation retains the same form of (399), we find, keeping in mind that $dW_t^2 = dt$,

$$dz = \left(\frac{\partial z}{\partial t} + f\frac{\partial z}{\partial y} + \frac{g^2}{2}\frac{\partial^2 z}{\partial y^2}\right)dt + g\frac{\partial z}{\partial y}dW_t. \tag{408}$$

The third term inside the parentheses is the result of this being a stochastic differential equation. This equation, valid for any $z(y,t)$, is known as Ito's lemma, and is the stochastic equivalent of the chain rule. This is the main difference between manipulating regular differential equations and stochastic ones.

The main quantity of interest when solving a stochastic differential equation is the probability distribution $P(y,t)$ for the stochastic variable to take the value y at time t. This distribution can be found via the Fokker-Planck equation, which we now derive. Let us take the expectation value of both sides of eq. (408) with respect to the probability distribution $P(y,t)$, assuming that z is not explicitly time-dependent,

$$\begin{aligned}\frac{d\langle z\rangle_{\mathrm{S}}}{dt} = \left\langle f\frac{\partial z}{\partial y}\right\rangle_{\mathrm{S}} + \left\langle \frac{g^2}{2}\frac{\partial^2 z}{\partial y^2}\right\rangle_{\mathrm{S}} &= \int\left(f\frac{\partial z}{\partial y} + \frac{g^2}{2}\frac{\partial^2 z}{\partial y^2}\right)P(y,t)dy \\ &= \int z\left[-\frac{\partial}{\partial y}(fP) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(g^2P)\right]dy,\end{aligned} \tag{409}$$

where we use $\langle\cdots\rangle_{\mathrm{S}}$ to denote the expectation value with respect to $P(y,t)$. Here we have used the fact that $\langle dW_t\rangle_{\mathrm{S}} = 0$ in the first step and integrated by parts in the last step assuming that the probability distribution vanishes at the boundaries. On the other hand, the time derivative of the mean of z is also given by

$$\frac{d\langle z\rangle_{\mathrm{S}}}{dt} = \frac{d}{dt}\int z(y)P(y,t)dy = \int z(y)\frac{\partial}{\partial t}P(y,t)dy. \tag{410}$$

Since this must be true for any $z(y)$, we can equate both expressions to find the Fokker-Planck equation for the probability density,

$$\frac{\partial}{\partial t}P(y,t) = -\frac{\partial}{\partial y}\Big[f(y,t)P(y,t)\Big] + \frac{1}{2}\frac{\partial^2}{\partial y^2}\Big[g^2(y,t)P(y,t)\Big]. \tag{411}$$

This is a deterministic equation and thus can be solved by standard PDE methods.

Let us assume that the stochastic variable y depends not only on the time t but also on the spatial coordinates $\boldsymbol{x}$. As mentioned earlier, Wiener processes are often written as $dW_t = \xi_t dt$. The quantity ξ_t can be thought of as a distribution, with correlation function given by

$$\langle \xi_t(\boldsymbol{x})\xi_{t'}(\boldsymbol{x}')\rangle_{\mathrm{S}} = \delta(t-t')\delta^3(\boldsymbol{x}-\boldsymbol{x}'). \tag{412}$$

The noise correlators in Fourier space are then

$$\langle \xi_t(\boldsymbol{k})\xi^\star_{t'}(\boldsymbol{k}')\rangle_{\rm S} = \int d^3\boldsymbol{x}\int d^3\boldsymbol{x}'\langle \xi_t(\boldsymbol{x})\xi_{t'}(\boldsymbol{x}')\rangle_{\rm S}\, e^{i\boldsymbol{k}\boldsymbol{x}}e^{-i\boldsymbol{k}'\boldsymbol{x}'} = (2\pi)^3\delta(t-t')\delta^3_-, \tag{413}$$

where we have defined the shorthand $\delta^3_\pm = \delta^3(\boldsymbol{k}\pm\boldsymbol{k}')$. Following a similar procedure for the other entries, the entire correlation matrix can be computed

$$\left\langle \begin{pmatrix} \xi_t(\boldsymbol{k}) \\ \xi^\star_t(\boldsymbol{k}) \end{pmatrix} \begin{pmatrix} \xi_{t'}(\boldsymbol{k}') & \xi^\star_{t'}(\boldsymbol{k}') \end{pmatrix} \right\rangle_{\rm S} = (2\pi)^3 \begin{pmatrix} \delta^3_+ & \delta^3_- \\ \delta^3_- & \delta^3_+ \end{pmatrix} \delta(t-t'). \tag{414}$$

For the real and imaginary parts of the noise $\mathrm{Re}(\xi_t)\equiv\xi^r_t$ and $\mathrm{Im}(\xi_t)\equiv\xi^i_t$, we find

$$\left\langle \begin{pmatrix} \xi^r_t(\boldsymbol{k}) \\ \xi^i_t(\boldsymbol{k}) \end{pmatrix} \begin{pmatrix} \xi^r_{t'}(\boldsymbol{k}') & \xi^i_{t'}(\boldsymbol{k}') \end{pmatrix} \right\rangle_{\rm S} = \frac{1}{2}(2\pi)^3 \begin{pmatrix} \delta^3_- + \delta^3_+ & 0 \\ 0 & \delta^3_- - \delta^3_+ \end{pmatrix} \delta(t-t'). \tag{415}$$

In computing these entries it is necessary to use the parity of the δ-function, $\delta^3(\boldsymbol{k}) = \delta^3(-\boldsymbol{k})$. We therefore find that the real and imaginary parts of the noise are uncorrelated, but ξ_t and $\xi^\star_t$ are not.

Throughout the rest of this appendix we will focus on a specific multivariate version of eq. (399) which is close to the type of equations we deal with in Chapter 4. Let us consider[53]

$$\frac{d\Phi}{dt} + \boldsymbol{A}\Phi = \frac{1}{\sqrt{2}}\boldsymbol{\sigma}\boldsymbol{\xi}_t, \tag{416}$$

where the stochastic time-dependent variable $\Phi(t)$ is an n-dimensional column complex vector, $\boldsymbol{A}$ is an $n\times n$ real matrix, $\boldsymbol{\sigma}$ is an $n\times m$ complex matrix, and $\boldsymbol{\xi}_t$ is an m-dimensional column complex noise vector whose real and imaginary parts have correlation functions given by the multivariate analogue of eq. (415). We have absorbed the overall $1/2$ factor from eq. (415) into the definition of $\boldsymbol{\xi}_t$, leading to the $1/\sqrt{2}$ factor in eq. (416). We want to find the probability distribution $P(\Phi,\Phi^\star,t)$, where $\Phi^\star$ obeys the equation of motion

$$\frac{d\Phi^\star}{dt} + \boldsymbol{A}\Phi^\star = \frac{1}{\sqrt{2}}\boldsymbol{\sigma}\boldsymbol{\xi}^\star_t. \tag{417}$$

The noise vectors $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}^\star_t$ are not independent as per eq. (414). In order to use a generalized multivariate version of the Fokker-Planck equation (411) for the probability distribution $P(\Phi,\Phi^\star,t)$, we need to rewrite eq. (416) in terms of uncorrelated noise sources. To this end, it is convenient to define the vectors $\Psi\equiv(\Phi^{\rm T},\Phi^\dagger)^{\rm T}$ and $\boldsymbol{\chi}_c\equiv(\boldsymbol{\xi}^{\rm T}_t,\boldsymbol{\xi}^\dagger_t)^{\rm T}$. The equation of

[53] Stochastic equations of this form are useful at linear order in perturbation theory and in Fourier space, where spatial derivatives are effectively decoupled. In particular, the system in eqs. (206 – 208) is of this form.

motion for Ψ is then

$$\frac{d\Psi}{dt} + \boldsymbol{\alpha}\Psi = \frac{1}{\sqrt{2}}\boldsymbol{\Sigma}\boldsymbol{\chi}_c, \tag{418}$$

where

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{A} \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \tag{419}$$

This notation might make $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ look like they have the same shapes, but $\boldsymbol{\alpha}$ is a $2n \times 2n$ matrix, whereas $\boldsymbol{\Sigma}$ is a $2n \times 2m$ matrix (the 0 matrices inside have different shapes). The final step is to write this equation with correlated noises $\boldsymbol{\chi}_c$ (hence the subscript c) in terms of the uncorrelated noises $\boldsymbol{\chi}_u \equiv (\boldsymbol{\xi}_t^{r\mathrm{T}}, \boldsymbol{\xi}_t^{i\mathrm{T}})^{\mathrm{T}}$. We have

$$\frac{d\Psi}{dt} + \boldsymbol{\alpha}\Psi = \frac{1}{\sqrt{2}}\boldsymbol{\Sigma}\boldsymbol{M}_\chi\boldsymbol{\chi}_u, \tag{420}$$

with

$$\boldsymbol{M}_\chi = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \quad \longrightarrow \quad \boldsymbol{\Sigma}\boldsymbol{M}_\chi = \begin{pmatrix} \boldsymbol{\sigma} & i\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & -i\boldsymbol{\sigma} \end{pmatrix}. \tag{421}$$

Since the noise sources in eq. (420) are uncorrelated, eq. (411) can be generalized directly,

$$\frac{\partial P}{\partial t} = \sum_{k\ell}\left\{\boldsymbol{\alpha}_{k\ell}\frac{\partial}{\partial\Psi_k}(\Psi_\ell P) + \frac{1}{2}\left[\frac{1}{\sqrt{2}}\boldsymbol{\Sigma}\boldsymbol{M}_\chi\frac{1}{\sqrt{2}}\boldsymbol{M}_\chi^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}}\right]_{k\ell}\frac{\partial^2 P}{\partial\Psi_k\partial\Psi_\ell}\right\}, \tag{422}$$

with

$$\boldsymbol{\Sigma}\boldsymbol{M}_\chi\boldsymbol{M}_\chi^{\mathrm{T}}\boldsymbol{\Sigma}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{\sigma} & i\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & -i\boldsymbol{\sigma} \end{pmatrix}\begin{pmatrix} \boldsymbol{\sigma}^{\mathrm{T}} & \boldsymbol{\sigma}^{\mathrm{T}} \\ i\boldsymbol{\sigma}^{\mathrm{T}} & -i\boldsymbol{\sigma}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} 0 & 2\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}} \\ 2\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}} & 0 \end{pmatrix}. \tag{423}$$

The sums in eq. (422) can be expanded in terms of Φ and $\Phi^\star$ instead of Ψ. We find

$$\frac{\partial P}{\partial t} = \sum_{k\ell}\left[\boldsymbol{A}_{k\ell}\frac{\partial}{\partial\Phi_k}(\Phi_\ell P) + \boldsymbol{A}_{k\ell}\frac{\partial}{\partial\Phi_k^\star}(\Phi_\ell^\star P) + (\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}})_{k\ell}\frac{\partial^2 P}{\partial\Phi_k\partial\Phi_\ell^\star}\right]. \tag{424}$$

The equation of motion for the two-point statistical moments $\boldsymbol{Q} \equiv \langle\Phi\Phi^\dagger\rangle_{\mathrm{S}}$, defined via

$$\langle\Phi\Phi^\dagger\rangle_{\mathrm{S}}(N) \equiv \int\prod_i d\Phi_i \int\prod_j d\Phi_j^\star\; P(\Phi,\Phi^\star,N)\;\Phi\Phi^\dagger. \tag{425}$$

can be found by acting on this equation with a time derivative and using the Fokker-Planck equation (424). It is also necessary to integrate by parts and assume that the probability distribution vanishes at the boundaries. The result is

$$\frac{d\boldsymbol{Q}}{dt} = -\boldsymbol{A}\boldsymbol{Q} - \boldsymbol{Q}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}. \tag{426}$$

D ■ Microphysics of the dissipative coefficient

In this appendix we discuss a particular microphysical realization of a localized dissipative coefficient Γ during inflation. The purpose of the present discussion, however, is not to propose a definitive model but simply to show that obtaining a peaked dissipative coefficient that satisfies all the necessary constraints is, in principle, possible. Specifically, we introduce a Lagrangian which reproduces the form of the dissipative coefficient (211), assuming that the fields that couple to the inflaton are part of a thermalized bath. In order to keep the field content to a minimum we will consider a scenario where, besides the inflaton ϕ, only two more degrees of freedom participate of the dynamics. The first one, denoted by σ and assumed to be a scalar, corresponds to the light radiation field in equilibrium. The second one, also a scalar and denoted by χ, corresponds to a heavy catalyst field. The large effective mass of χ arises via its coupling to the slowly rolling inflaton field. Through the coupling of the (unstable) χ with σ, the inflaton energy density can be efficiently dissipated into radiation. Indirect decay scenarios like this one are among the preferred mechanisms for realistic warm inflation models [177, 206, 207]. One of the advantages of introducing such heavy catalyst fields is to prevent the inflaton potential from receiving strong temperature corrections [207, 208].[54]

Let us consider the following Lagrangian

$$\mathcal{L} = \frac{1}{2}\frac{\partial^\mu\varphi\partial_\mu\varphi}{\mathcal{J}(\varphi)} + \frac{1}{2}\partial^\mu\chi\partial_\mu\chi + \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma - \frac{1}{2}g_\chi^2\varphi^2\chi^2 - \frac{1}{2}g_\sigma^2\varphi\chi\sigma^2 - U(\varphi) + \cdots . \tag{427}$$

Here, φ denotes the inflaton, which we assume to have a non-canonical kinetic term. In this frame, the inflaton φ and the mediator χ interact via a four-legged contact term with coupling g_χ. The term with coupling g_σ connects the three fields. After expanding the inflaton around its classical background value, this term induces the decay of χ into σ. The dots correspond to other interactions which should be present for the light sector σ to be thermalized. If this sector is indeed in equilibrium, the presence of the $\varphi \to \chi \to \sigma$ channel modifies the equation of motion of the inflaton. The canonically normalized inflaton ϕ is related to φ via

$$\frac{d\phi}{d\varphi} = \frac{1}{\sqrt{\mathcal{J}(\varphi)}}, \tag{428}$$

and the equation of motion for ϕ becomes

$$\ddot{\phi} + (3H + \Gamma)\dot{\phi} + V_\phi(\phi) = 0, \tag{429}$$

where $V(\phi) \equiv U[\varphi(\phi)]$. We neglect the temperature dependence of the potential in accor-

[54] In addition, in a simpler construction in which the inflaton directly couples to the radiation field σ, the dissipation rate is determined by the strength of the inflaton self-coupling. This effectively suppresses the value of Γ, due to the requirement of the normalization of such self-coupling by the amplitude of the primordial curvature power spectrum [177].

dance with our earlier comment. The dissipative coefficient Γ encodes the production of σ quanta. The appearance of a local dissipative term relies on the assumption that the microphysical processes which determine Γ operate at time-scales much smaller than those characteristic of the macroscopic slow-roll of the inflaton and the expansion of the Universe (the so-called adiabatic Markovian approximation [177, 209]). Additionally, we assume that the typical interaction time-scale between the constituents of the thermal bath is much shorter than the time-scales associated to the variation of the background quantities. In this approximation, the dissipative coefficient is given by [210–212]

$$\Gamma(\phi, T) = 2\frac{g_\chi^4}{T}\left(\varphi\frac{d\varphi}{d\phi}\right)^2 \int \frac{d^4p}{(2\pi)^4} n(\omega)\big[n(\omega)+1\big]\rho_\chi^2(\omega), \tag{430}$$

where $p = (\omega, \boldsymbol{p})$, $n_\chi(\omega) = 1/(e^{\omega/T}-1)$ is the Bose-Einstein distribution, and

$$\rho_\chi = \frac{4\omega_p\Gamma_\chi}{(\omega^2-\omega_p^2)^2 + 4\omega_p^2\Gamma_\chi^2} \tag{431}$$

is the spectral density of the catalyst field. In this expression $\omega_p^2 = |\boldsymbol{p}|^2 + m_\chi^2$ is the on-shell frequency, with $m_\chi = g_\chi\varphi$ denoting the effective mass of χ in accordance with (427), and Γ_χ is the decay width of χ. Our expression for Γ originates from the g_χ interaction term in (427). In principle, the g_σ term would also contribute to the dissipation rate, but this contribution turns out to be loop-suppressed [212]. The aforementioned adiabatic Markovian approximation translates into the three conditions

$$\Gamma_\chi \gg H \quad \text{(thermalization)}, \qquad \Gamma_\chi \gg \frac{\dot\phi}{\phi}, \frac{\dot T}{T} \quad \text{(adiabaticity)}, \qquad T \gg H. \tag{432}$$

These conditions also ensure that thermal equilibrium is maintained between the light σ and the heavy χ throughout the strongly dissipative phase ($\Gamma \gg H$).

To evaluate the dissipative coefficient Γ we need the decay width Γ_χ, which should be calculated using non-zero temperature QFT. A general expression can be found in e.g. [212]. For simplicity, we will restrict our discussion to the so-called low-temperature limit in what follows, which corresponds to assuming that $T \ll m_\chi$. We remark however that in principle none of our assumptions forbid a peaked dissipative coefficient at higher temperatures. The decay rate for the process $\chi \to \sigma\sigma$ can then be written as

$$\Gamma_\chi(\omega, \boldsymbol{p}) \simeq \frac{g_\sigma^4\varphi^2}{8\pi\omega_p(\boldsymbol{p})} \tag{433}$$

in a frame of reference that is boosted with respect to the χ rest frame. In terms of this rate, the adiabaticity condition corresponds to $\Gamma_\chi \gg |\dot\phi/\phi|$. The low-temperature limit further suppresses the already small temperature corrections to the potential. Let us introduce the

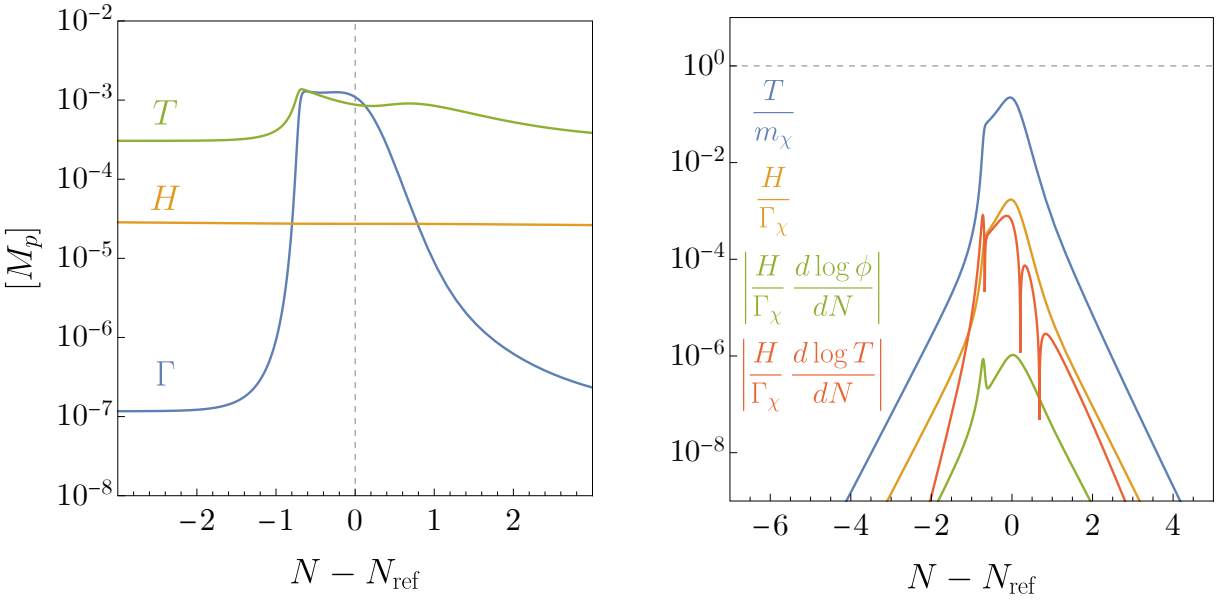

Figure 32: *Left panel: time evolution of the background quantities T, H and Γ for the dissipative coefficient in eq. (439) using the parameters in (440). Right panel: constraints from eq. (432), together with the low-temperature limit $T \ll m_\chi$, for the same parameters as the left panel. The reference time $N_{\rm ref}$ is chosen as the time at which T/m_χ reaches its maximum value.*

dimensionless quantities

$$\alpha \equiv \frac{\omega}{T}, \qquad \beta \equiv \frac{|\boldsymbol{p}|}{m_\chi}, \qquad \lambda \equiv \frac{T}{m_\chi}, \qquad \gamma_\chi \equiv \frac{\Gamma_\chi^{(0)}}{m_\chi}, \tag{434}$$

where $\Gamma_\chi^{(0)} \equiv \Gamma_\chi(m_\chi, \mathbf{0})$ is the decay rate in the χ rest frame. The dissipative rate is then

$$\Gamma = \frac{8g_\chi^4}{m_\chi \pi^3}\left(\varphi\frac{d\varphi}{d\phi}\right)^2 \int_0^\infty d\beta \int_{\beta/\lambda}^\infty d\alpha \left[\frac{\beta\gamma_\chi}{(\alpha^2\lambda^2 - \beta^2 - 1)^2 + 4\gamma_\chi^2}\right]^2 \left(\frac{1}{e^\alpha - 1}\right)\left(\frac{1}{e^\alpha - 1} + 1\right). \tag{435}$$

In the low-temperature limit $\lambda \ll 1$, we can discard the α and γ_χ pieces in the denominator of the term in brackets, which is equivalent to approximating the spectral function as $\rho_\chi \simeq 4\Gamma_\chi/m_\chi^3$ [212]. The dissipative rate then simplifies to

$$\Gamma \simeq \frac{8g_\chi^4\gamma_\chi^2}{m_\chi\pi^3}\left(\varphi\frac{d\varphi}{d\phi}\right)^2 \int_0^\infty d\beta \left(\frac{\beta^2}{1+\beta^2}\right)\left(\frac{1}{e^{\beta/\lambda} - 1}\right) \simeq \frac{20 g_\chi^4 \left(\Gamma_\chi^{(0)}\right)^2}{m_\chi^6 \pi^3}\left(\varphi\frac{d\varphi}{d\phi}\right)^2 T^3, \tag{436}$$

where we have used the fact that the β integral can be well-approximated by $5\lambda^3/2$ for $\lambda \ll 1$.

All of the expressions we have presented so far are independent of the particular shape of $\mathcal{J}(\varphi)$ in the kinetic term of φ in (427). Successful warm inflation can be obtained by setting $\mathcal{J}(\varphi) = 1$, with a dissipative coefficient $\Gamma \propto T^3/\varphi^2$. To recover the phenomenological peaked

dissipative coefficient we presented in Section 4.1, however, let us consider

$$\mathcal{J}(\varphi) = 4\frac{\varphi^2}{\Lambda^2}\left[1 + \frac{4M^2\varphi}{m(\varphi+m)^2}\right], \tag{437}$$

where m, M and Λ are arbitrary dimensionful parameters. The corresponding solution to eq. (428) can then be written as

$$\varphi(\phi) = m\left(\frac{1-f(\phi)}{1+f(\phi)}\right), \qquad f(\phi) \equiv \left\{\frac{M^2+m^2}{M^2+m^2\coth^2\left[(\phi-\phi_\star)/\Lambda\right]}\right\}^{1/2}, \tag{438}$$

where $\phi_\star$ is a free parameter, which we assume is necessary to fix $V(\phi=0)=0$. By plugging this transformation into eq. (436), we find

$$\Gamma = \frac{5g_\sigma^8}{4\pi^5 g_\chi^4\Lambda^2}\left\{\frac{(m^2+M^2)T^3}{m^2+M^2\tanh^2\left[(\phi-\phi_\star)/\Lambda\right]}\right\}. \tag{439}$$

Assuming that $m \ll M$, we recover eq. (211). In the left panel of Fig. 32 we show the time evolution of the dissipative rate Γ, the Hubble parameter and the temperature for the dissipative rate in eq. (439) with parameters

$$\Lambda = 0.1M_p, \qquad M = M_p, \qquad m = 0.05M_p, \qquad g_\chi = 0.08, \qquad g_\sigma = 0.28, \tag{440}$$

together with $\lambda = 4.6\times10^{-14}$ and $g_\star = 10$. The right panel of the same figure shows the conditions in eq. (432), as well as the low-temperature limit $T/m_\chi \ll 1$. All of the constraints are satisfied and our setup is therefore self-consistent. We remark that if we want the potential $V(\phi) = U[\varphi(\phi)]$ to be able to support slow roll inflation, the shape of $U(\varphi)$ must be quite unconventional, due to the complicated field redefinition in eq. (438). As we mentioned earlier, however, the purpose of the present discussion is only to show that obtaining a dissipative coefficient that satisfies all the necessary constraints is possible in principle. We hope the discussion in this appendix encourages further efforts to search for well-motivated models that could produce peaked dissipative coefficients.

E ■ Stochastic and quantum expectation values

As shown in Section 4.4, the solution to the simplified equation of motion for $\delta\phi$ (235) is of the form

$$\delta\phi_k(N) = \delta\phi_k^{(h)}(N) + \int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})\xi_{\hat{N}}(\boldsymbol{k})d\hat{N}, \tag{441}$$

where ξ_N is a Wiener process satisfying (414) and $\mathcal{S}_k$ is a function of time, the form of which is not relevant in what follows. We can quantize the field by writing the homogeneous

solution in terms of creation and annihilation operators

$$\delta\hat{\phi}_k^{(h)} = \delta\phi_k^{(h)}\hat{a}_k + \delta\phi_k^{(h)\star}\hat{a}_k^\dagger. \tag{442}$$

Since we do not quantize the noise, we make the second term in eq. (441) proportional to the identity operator. The quantum expectation value is defined via $\langle\cdots\rangle_{\rm Q} \equiv \langle 0|\cdots|0\rangle$, so we have

$$\langle\delta\hat{\phi}_k\delta\hat{\phi}_q\rangle_{\rm Q} = \langle\delta\hat{\phi}_k^{(h)}\delta\hat{\phi}_q^{(h)}\rangle_{\rm Q} + \int_{-\infty}^{N}\int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})\mathcal{S}_q(N,\tilde{N})\xi_{\hat{N}}(\boldsymbol{k})\xi_{\tilde{N}}(\boldsymbol{q})d\hat{N}d\tilde{N}. \tag{443}$$

The power spectrum is defined via

$$\langle\delta\hat{\phi}_k\delta\hat{\phi}_q\rangle_{\rm Q} \equiv (2\pi)^3\frac{2\pi^2}{k^3}\mathcal{P}_{\delta\phi}(k)\delta^3(\boldsymbol{k}+\boldsymbol{q}). \tag{444}$$

We can show from eq. (442) that

$$\langle\delta\hat{\phi}_k^{(h)}\delta\hat{\phi}_q^{(h)}\rangle_{\rm Q} = (2\pi)^3|\delta\phi_k^{(h)}|^2\delta^3(\boldsymbol{k}+\boldsymbol{q}), \tag{445}$$

and since $\delta\phi_k^{(h)}$ is the solution to the homogeneous equation of motion for $\delta\phi$, which does not involve the stochastic noise, this is a deterministic quantity. Note, however, that the second term in eq. (443) is still stochastic, as it should be, since every realization of the system leads to a different power spectrum. To obtain a deterministic quantity we can take the stochastic expectation value $\langle\cdots\rangle_{\rm S}$, computed by averaging over many realizations. We denote the double expectation value by brackets without subindices, $\langle\cdots\rangle \equiv \langle\langle\cdots\rangle_{\rm Q}\rangle_{\rm S}$. Thus,

$$\begin{aligned}
\langle\delta\hat{\phi}_k\delta\hat{\phi}_q\rangle &= (2\pi)^3|\delta\phi_k^{(h)}|^2\delta^3(\boldsymbol{k}+\boldsymbol{q}) + \int_{-\infty}^{N}\int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})\mathcal{S}_q(N,\tilde{N})\langle\xi_{\hat{N}}(\boldsymbol{k})\xi_{\tilde{N}}(\boldsymbol{q})\rangle_{\rm S}d\hat{N}d\tilde{N} \\
&= (2\pi)^3\delta^3(\boldsymbol{k}+\boldsymbol{q})\left(|\delta\phi_k^{(h)}|^2 + \int_{-\infty}^{N}\int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})\mathcal{S}_q(N,\tilde{N})\delta(\hat{N}-\tilde{N})d\hat{N}d\tilde{N}\right) \\
&= (2\pi)^3\delta^3(\boldsymbol{k}+\boldsymbol{q})\left(|\delta\phi_k^{(h)}|^2 + \int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})^2d\hat{N}\right).
\end{aligned} \tag{446}$$

By combining this with eq. (444), we find

$$\langle\mathcal{P}_{\delta\phi}(k)\rangle_{\rm S} = \frac{k^3}{2\pi^2}\left(|\delta\phi_k^{(h)}|^2 + \int_{-\infty}^{N}\mathcal{S}_k(N,\hat{N})^2d\hat{N}\right). \tag{447}$$

This formula holds as long as we assume that the noise term is not quantized and the stochastic and quantum expectation values are independent from each other.

The four-point function for $\delta\phi$ can be found in a similar manner. Since these results bear direct relation to the calculation of Section 5.4, we present them in conformal time. If we assume that $\delta\phi_k^{(h)}$ is Gaussian with respect to the quantum expectation value, and ξ_τ is

Gaussian with respect to the stochastic one, then the following identities follow from Wick's theorem

$$\langle \delta\hat{\phi}_q^{(h)}(\tau)\delta\hat{\phi}_{k-q}^{(h)}(\tau)\delta\hat{\phi}_l^{(h)}(\tau')\delta\hat{\phi}_{p-l}^{(h)}(\tau')\rangle = (2\pi)^6\delta_{k+p}^3\Big(\delta_{q+l}^3+\delta_{q+p-l}^3\Big) \delta\phi_q^{(h)}(\tau)^\star\delta\phi_l^{(h)}(\tau')\delta\phi_{k-q}^{(h)}(\tau)^\star\delta\phi_{p-l}^{(h)}(\tau'), \quad (448)$$

$$\langle \xi_\tau(\boldsymbol{q})\xi_{\tau'}(\boldsymbol{k}-\boldsymbol{q})\xi_{\tau''}(\boldsymbol{l})\xi_{\tau'''}(\boldsymbol{p}-\boldsymbol{l})\rangle = (2\pi)^6\delta_{k+p}^3\delta_{\tau+\tau'-\tau''-\tau'''} \Big(\delta_{q+p-l}^3\delta_{\tau-\tau'''}+\delta_{q+l}^3\delta_{\tau-\tau''}\Big), \quad (449)$$

$$\langle \xi_\tau(\boldsymbol{q})\delta\hat{\phi}_{k-q}^{(h)}(\tau')\xi_{\tau''}(\boldsymbol{l})\delta\hat{\phi}_{p-l}^{(h)}(\tau''')\rangle = (2\pi)^6\delta_{k+p}^3\delta_{q+l}^3\delta_{\tau-\tau''}\delta\phi_{k-q}^{(h)}(\tau')^\star\delta\phi_{p-l}^{(h)}(\tau'''), \quad (450)$$

where we have introduced the shorthand notation $\delta^3(\boldsymbol{k})=\delta_k^3$ for readability, and we have once again assumed that the quantum and stochastic expectation values are independent from each other. The four-point function is given by the sum of four terms,

$$\langle \delta\hat{\phi}_q(\tau')\delta\hat{\phi}_{k-q}(\tau')\delta\hat{\phi}_l(\tau'')\delta\hat{\phi}_{p-l}(\tau'')\rangle = (2\pi)^6\Bigg(\sum_{i=1}^4 F_{\delta\phi}^{(i)}\Bigg)\delta^3(\boldsymbol{k}+\boldsymbol{p}), \quad (451)$$

where

$$F_{\delta\phi}^{(1)} = \Big(\delta_{q+l}^3+\delta_{q+p-l}^3\Big)\delta\phi_q^{(h)}(\tau')^\star\delta\phi_q^{(h)}(\tau'')\delta\phi_{k-q}^{(h)}(\tau')^\star\delta\phi_{k-q}^{(h)}(\tau''), \quad (452)$$

$$F_{\delta\phi}^{(2)} = \Big(\delta_{q+l}^3+\delta_{q+p-l}^3\Big)\delta\phi_q^{(h)}(\tau')^\star\delta\phi_q^{(h)}(\tau'')\int_{-\infty}^{\min(\tau',\tau'')}\mathcal{S}_{k-q}(\tau'',\hat{\tau})\mathcal{S}_{k-q}(\tau',\hat{\tau})d\hat{\tau}, \quad (453)$$

$$F_{\delta\phi}^{(3)} = \Big(\delta_{q+l}^3+\delta_{q+p-l}^3\Big)\delta\phi_{k-q}^{(h)}(\tau')^\star\delta\phi_{k-q}^{(h)}(\tau'')\int_{-\infty}^{\min(\tau',\tau'')}\mathcal{S}_q(\tau'',\hat{\tau})\mathcal{S}_q(\tau',\hat{\tau})d\hat{\tau}, \quad (454)$$

$$F_{\delta\phi}^{(4)} = \Big(\delta_{q+l}^3+\delta_{q+p-l}^3\Big)\int_{-\infty}^{\min(\tau',\tau'')}\mathcal{S}_q(\tau',\hat{\tau})\mathcal{S}_q(\tau'',\hat{\tau})d\hat{\tau}\int_{-\infty}^{\min(\tau',\tau'')}\mathcal{S}_{k-q}(\tau',\tilde{\tau})\mathcal{S}_{k-q}(\tau'',\tilde{\tau})d\tilde{\tau}. \quad (455)$$

We have ignored some contact interaction terms that are not proportional to $\delta^3(\boldsymbol{k}+\boldsymbol{p})$. The upper limit of the integrals arises because, if $a<c<b$, then

$$\int_a^b\int_a^c f(\tau)\delta(\tau-\tau')d\tau d\tau' = \int_a^c\int_a^c f(\tau)\delta(\tau-\tau')d\tau d\tau' + \underbrace{\int_c^b\int_a^c f(\tau)\delta(\tau-\tau')d\tau d\tau'}_{0} = \int_a^c f(\tau)d\tau, \quad (456)$$

since $a<\tau<c$ but $c<\tau'<b$, so that the regions in the second double integral after the first equality do not overlap, making it impossible to satisfy the constraint $\tau=\tau'$ imposed by the δ function and forcing the integral to vanish. An analogous result holds for $a<b<c$.

F ■ Analytical gravitational wave integrals

The analytical results for the quantities in (287) are, with $x = k\tau$, $y = p/k$ and $z = |\boldsymbol{p}-\boldsymbol{k}|/k$ [185]

$$I_i^{\rm eMD} = \frac{1}{x_m}\frac{12}{5} a_T^+|_{yx_m} a_T^+|_{zx_m} \Big\{ a_i^+ \Big[(3-x^2)\sin x - 3x\cos x\Big] + a_i^- \Big[(x^2-3)\cos x - 3x\sin x\Big]\Big\}\Big|_0^{x_m}, \tag{457}$$

$$\begin{aligned} I_i^{\rm RD} = \frac{3}{y^3 z^3 x^4}\Big\{ &\Big[F^{--}\cos y^{--} + F^{-+}\cos y^{-+} + F^{+-}\cos y^{+-} + F^{++}\cos y^{++} \\ &+ G^{--}\sin y^{--} + G^{-+}\sin y^{-+} + G^{+-}\sin y^{+-} + G^{++}\sin y^{++}\Big] \\ &+ x^4(y^2+z^2-3)^2\Big[H_i^{--}\mathrm{Ci}(y^{--}) + H_i^{-+}\mathrm{Ci}(|y^{-+}|) + H_i^{+-}\mathrm{Ci}(y^{+-}) + H_i^{++}\mathrm{Ci}(y^{++}) \\ &+ I_i^{--}\mathrm{Si}(y^{--}) + I_i^{-+}\mathrm{Si}(y^{-+}) + I_i^{+-}\mathrm{Si}(y^{+-}) + I_i^{++}\mathrm{Si}(y^{++})\Big]\Big\}\Big|_{x_m/2}^{\infty}, \end{aligned} \tag{458}$$

where we have defined

$$y^{\pm\pm} \equiv \left(1 \pm \frac{y\pm z}{\sqrt{3}}\right) x. \tag{459}$$

The functions $H_i^{\pm\pm}$ and $I_i^{\pm\pm}$ are defined by

$$H_i^{++} = \Big(b_T^+|_{zx_m} b_T^+|_{yx_m} - b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- + \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} + b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^+, \tag{460}$$

$$H_i^{+-} = -\Big(b_T^+|_{zx_m} b_T^+|_{yx_m} + b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- + \Big(b_T^+|_{yx_m} b_T^-|_{zx_m} - b_T^+|_{zx_m} b_T^-|_{yx_m}\Big) b_i^+, \tag{461}$$

$$H_i^{-+} = \Big(b_T^+|_{zx_m} b_T^+|_{yx_m} - b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- - \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} + b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^+, \tag{462}$$

$$H_i^{--} = -\Big(b_T^+|_{zx_m} b_T^+|_{yx_m} + b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- + \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} - b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^+, \tag{463}$$

$$I_i^{++} = -\Big(b_T^+|_{zx_m} b_T^+|_{yx_m} - b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- + \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} + b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^-, \tag{464}$$

$$I_i^{+-} = \Big(b_T^+|_{zx_m} b_T^+|_{yx_m} + b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^+ - \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} - b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^-, \tag{465}$$

$$I_i^{-+} = -\Big(b_T^+|_{zx_m} b_T^+|_{yx_m} - b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^- - \Big(b_T^+|_{zx_m} b_T^-|_{yx_m} + b_T^+|_{yx_m} b_T^-|_{zx_m}\Big) b_i^-, \tag{466}$$

$$I_i^{--} = \Big(b_T^+|_{zx_m} b_T^+|_{yx_m} + b_T^-|_{zx_m} b_T^-|_{yx_m}\Big) b_i^+ - \Big(b_T^+|_{yx_m} b_T^-|_{zx_m} - b_T^+|_{zx_m} b_T^-|_{yx_m}\Big) b_i^-. \tag{467}$$

The functions $F^{\pm\pm}$ are, in terms of $H_i^{\pm\pm}$ and $I_i^{\pm\pm}$,

$$\begin{aligned} F^{++} = &\, I^{++}\Big\{18x\big[\sqrt{3}(z+y)-1\big] + x^3\big[\sqrt{3}(z+y)-3\big]\big[(z-y)^2-3\big]\Big\} \\ &- H^{++}\Big\{54 - 3x^2\big[3+z^2+y^2+6yz-2\sqrt{3}(z+y)\big]\Big\}, \end{aligned} \tag{468}$$

$$\begin{aligned} F^{+-} = &- I^{+-}\Big\{18x\big[\sqrt{3}(z-y)+1\big] + x^3\big[\sqrt{3}(z-y)+3\big]\big[(z+y)^2-3\big]\Big\} \\ &- H^{+-}\Big\{54 - 3x^2\big[3+z^2+y^2-6yz+2\sqrt{3}(z-y)\big]\Big\}, \end{aligned} \tag{469}$$

$$F^{-+} = - I^{-+}\Big\{18x\big[\sqrt{3}(z+y)+1\big] + x^3\big[\sqrt{3}(z+y)+3\big]\big[(z-y)^2-3\big]\Big\}$$

$$- H^{-+}\left\{54-3x^2\left[3+z^2+y^2+6yz+2\sqrt{3}(z+y)\right]\right\}, \tag{470}$$

$$F^{--} = I^{--}\left\{18x\left[\sqrt{3}(z-y)-1\right]+x^3\left[\sqrt{3}(z-y)-3\right]\left[(z+y)^2-3\right]\right\}$$
$$- H^{--}\left\{54-3x^2\left[3+z^2+y^2-6yz+2\sqrt{3}(y-z)\right]\right\}. \tag{471}$$

Similarly, for $G^{\pm\pm}$ we have

$$G^{++} = - H^{++}\left\{18x\left[\sqrt{3}(z+y)-1\right]+x^3\left[\sqrt{3}(z+y)-3\right]\left[(z-y)^2-3\right]\right\}$$
$$- I^{++}\left\{54-3x^2\left[3+z^2+y^2+6yz-2\sqrt{3}(z+y)\right]\right\}, \tag{472}$$

$$G^{+-} = H^{+-}\left\{18x\left[\sqrt{3}(z-y)+1\right]+x^3\left[\sqrt{3}(z-y)+3\right]\left[(z+y)^2-3\right]\right\}$$
$$- I^{+-}\left\{54-3x^2\left[3+z^2+y^2-6yz+2\sqrt{3}(z-y)\right]\right\}, \tag{473}$$

$$G^{-+} = H^{-+}\left\{18x\left[\sqrt{3}(z+y)+1\right]+x^3\left[\sqrt{3}(z+y)+3\right]\left[(z-y)^2-3\right]\right\}$$
$$- I^{-+}\left\{54-3x^2\left[3+z^2+y^2+6yz+2\sqrt{3}(z+y)\right]\right\}, \tag{474}$$

$$G^{--} = - H^{--}\left\{18x\left[\sqrt{3}(z-y)-1\right]+x^3\left[\sqrt{3}(z-y)-3\right]\left[(z+y)^2-3\right]\right\}$$
$$- I^{--}\left\{54-3x^2\left[3+z^2+y^2-6yz+2\sqrt{3}(y-z)\right]\right\}. \tag{475}$$

The case in which the Universe never goes through an eMD phase, which we denote as pRD (for pure RD) can be recovered by taking the $x_m \to 0$ limit. The result is

$$I_i^{\rm pRD} \equiv \frac{3(y^2+z^2-3)}{y^3z^3}\Big\{4yzb_i^- - b_i^+\pi(y^2+z^2-3)\Big[\Theta\Big(1-\frac{y+z}{\sqrt{3}}\Big)-1\Big]$$
$$+ b_i^-(y^2+z^2-3)\Big[\log\Big(1-\frac{y-z}{\sqrt{3}}\Big)+\log\Big(1+\frac{y-z}{\sqrt{3}}\Big)$$
$$- \log\Big(1+\frac{y+z}{\sqrt{3}}\Big)-\log\Big(\Big|1-\frac{y+z}{\sqrt{3}}\Big|\Big)\Big]\Big\}. \tag{476}$$

In deriving these results we have used the identities $\lim_{x\to 0}[\mathrm{Ci}(ax)-\mathrm{Ci}(bx)] = \log(a)-\log(b)$, $\lim_{x\to 0}\mathrm{Si}(x) = 0$, $\lim_{x\to\pm\infty}\mathrm{Si}(x) = \pm\pi/2$, and $\lim_{x\to+\infty}\mathrm{Ci}(x) = 0$.

G ■ Induced gravitational waves in different gauges

In this appendix we calculate the induced second-order tensor modes in different gauges using eq. (292) in the eMD-to-RD transition scenario[55] schematically represented in Fig. 1. Since we calculated the second-order tensor modes in the Newtonian gauge in Section 5.1, we can set $T_\psi = T_\varphi$ and $T_E = T_B = T_\sigma = 0$ in eq. (293). Thus, the only transfer function we need to compute is T_ψ –which is given by eq. (280)– and we simply need to determine

[55]The calculation of I_Σ in MD and RD eras was performed in [84, 89] and [85], respectively. The calculation for an eMD → RD transition is presented here for the first time.

T_α and T_β in terms of T_ψ from the gauge conditions. Since the transfer function in this scenario is piecewise-defined, so is I_Σ. Even though the Hubble factor and transfer functions in eq. (293) are continuous by construction, w is not (we assume the eMD $\to$ RD transition is instantaneous), and thus there is a discontinuity at the transition between both eras coming from this numerical prefactor.

We will perform the calculation in three illustrative gauges: the comoving gauge, the uniform curvature gauge, and the uniform expansion gauge. The spatial part β of the gauge transformation vanishes in all three cases if we start from the Newtonian gauge solution computed in Section 5.1, so we find the following compact expression for I_Σ,

$$I_\Sigma = -\frac{2}{yz}\left(\frac{3+3w}{5+3w}\right)^2 T_\alpha(xy)T_\alpha(xz). \tag{477}$$

Comoving gauge

This gauge is defined by $\tilde{E} = \delta\tilde{q} = 0$. The gauge transformation in this case is $\beta = 0$ and

$$\alpha = -\frac{2}{3+3w}\frac{\psi' + \mathcal{H}\psi}{\mathcal{H}^2}, \qquad T_\alpha(x) = -\frac{2}{3+3w}\frac{k}{\mathcal{H}^2}\left[T'_\psi(x) + \mathcal{H}T_\psi(x)\right]. \tag{478}$$

The resulting I_Σ is, for $x < x_m$,

$$I_\Sigma^{\rm eMD} = -\frac{2}{25}x^2. \tag{479}$$

For $x > x_m$, on the other hand, we have

$$\begin{aligned} I_\Sigma^{\rm RD} = &-\frac{2}{y^3 z^3 (x_m - 2x)^4}\cdot \\ &\left\{\sin\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[12b_T^-|_{yx_m} y(x_m - 2x) + \sqrt{3}b_T^+|_{yx_m}\left(-4x^2y^2 + 4xx_m y^2 - x_m^2 y^2 + 24\right)\right]\right. \\ &\left.+\cos\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[\sqrt{3}b_T^-|_{yx_m}\left(4x^2y^2 - 4xx_m y^2 + x_m^2 y^2 - 24\right) + 12b_T^+|_{yx_m} y(x_m - 2x)\right]\right\}\cdot \\ &\left\{\sin\left[\frac{z}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[12b_T^-|_{zx_m} z(x_m - 2x) + \sqrt{3}b_T^+|_{zx_m}\left(-4x^2z^2 + 4xx_m z^2 - x_m^2 z^2 + 24\right)\right]\right. \\ &\left.+\cos\left[\frac{z}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[\sqrt{3}b_T^-|_{zx_m}\left(4x^2z^2 - 4xx_m z^2 + x_m^2 z^2 - 24\right) + 12b_T^+|_{zx_m} z(x_m - 2x)\right]\right\}. \end{aligned} \tag{480}$$

The constants $b_T^\pm$ are given in eqs. (281, 282). The behaviour of I_Σ, depicted in Fig. 33, is problematic in this gauge; during the eMD era it grows as x^2, whereas in the RD era it oscillates as $x \to \infty$. The resulting $\boldsymbol{h}_{ij}^{\rm TT}$ is therefore completely different from the solution in the Newtonian gauge at late times. We remark that this difference propagates also to the calculation of $\Omega_{\rm GW}$, which diverges at late times in this gauge, see e.g. [84]. ■

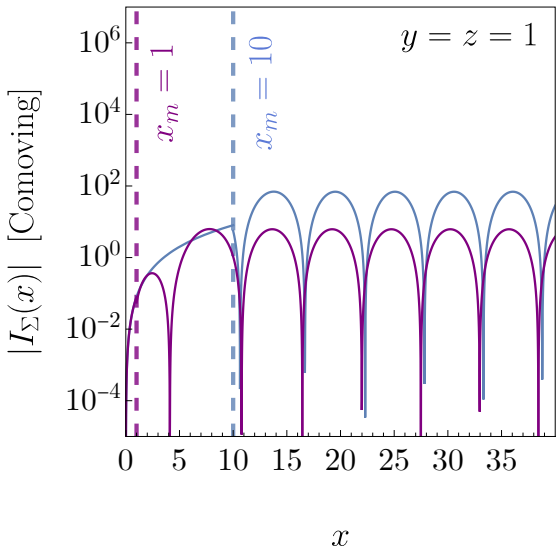

Figure 33: *Integral kernel I_Σ for a gauge transformation from the Newtonian to the comoving gauge for two different values of x_m, setting $y = z = 1$. The result oscillates as $x \to \infty$, and thus the power spectrum of $\boldsymbol{h}_{ij}^{\rm TT}$ diverges at late times in this gauge.*

Uniform curvature gauge

This gauge is defined by $\tilde{E} = \tilde{\psi} = 0$. The gauge transformation in this case is $\beta = 0$ and

$$\alpha = \frac{\psi}{\mathcal{H}}, \qquad T_\alpha(x) = \frac{k}{\mathcal{H}} T_\psi(x). \tag{481}$$

The resulting I_Σ is, for $x < x_m$,

$$I_\Sigma^{\rm eMD} = -\frac{9}{50} x^2. \tag{482}$$

For $x > x_m$ we obtain

$$\begin{aligned} I_\Sigma^{\rm RD} = &- \frac{96}{y^3 z^3 (x_m - 2x)^4} \Big\{ \cos\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right] \left[6 b_T^-|_{y x_m} + \sqrt{3} b_T^+|_{y x_m} y (2x - x_m)\right] \\ &- \sin\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right] \left[\sqrt{3} b_T^-|_{y x_m} y (x_m - 2x) + 6 b_T^+|_{y x_m}\right] \Big\} \cdot \\ &\Big\{ \cos\left[\frac{z}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right] \left[6 b_T^-|_{z x_m} + \sqrt{3} b_T^+|_{z x_m} z (2x - x_m)\right] \\ &- \sin\left[\frac{z}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right] \left[\sqrt{3} b_T^-|_{z x_m} z (x_m - 2x) + 6 b_T^+|_{z x_m}\right] \Big\}. \end{aligned} \tag{483}$$

The result for I_Σ in this gauge is depicted in Fig. 34. During the eMD era the solution still grows as x^2, but with a different prefactor to the one in eq. (479). During the RD era, on the other hand, I_Σ decays as x^{-2} when $x \to \infty$. Thus, as we pointed out earlier, this means that the result in this gauge coincides with the Newtonian one at late times. ■

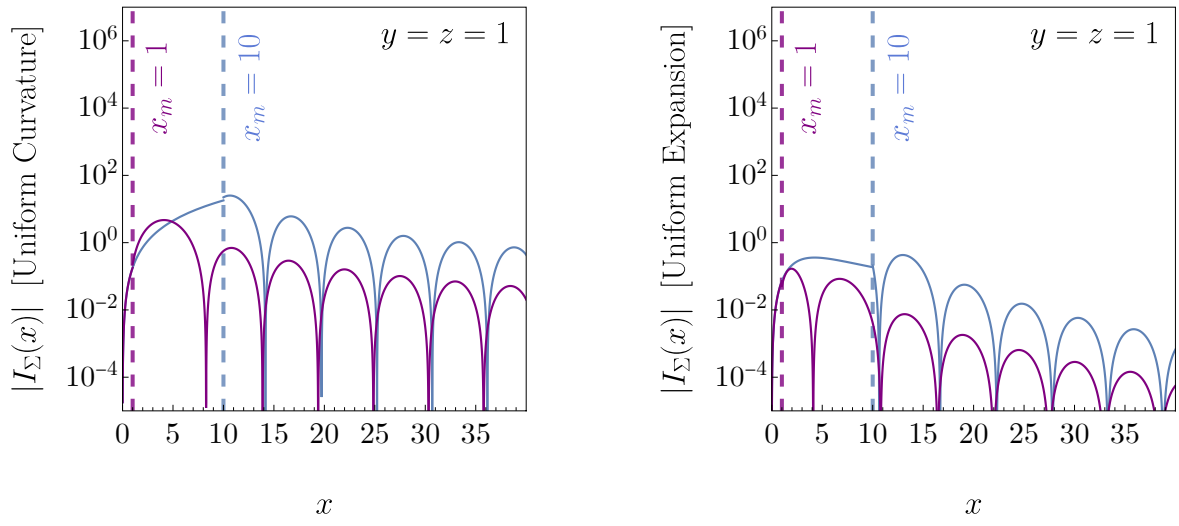

Figure 34: *Integral kernels I_Σ for gauge transformations from the Newtonian gauge to the uniform curvature gauge (left panel) and the uniform expansion gauge (right panel), for two different values of x_m and setting $y = z = 1$. In both cases the result decays faster than x^{-1} and thus the solution coincides with the Newtonian one only at late times.*

Uniform expansion gauge

This gauge is defined by $\tilde{E} = 3(\mathcal{H}\tilde{\phi} + \tilde{\psi}') + k^2\tilde{\sigma} = 0$. The gauge transformation in this case is $\beta = 0$ and

$$\alpha = -\frac{\psi' + \mathcal{H}\psi}{\mathcal{H}^2 - \mathcal{H}' + k^2/3}, \qquad T_\alpha(x) = -\frac{6}{9 + 9w + 2(k^2/\mathcal{H}^2)}\frac{k}{\mathcal{H}^2}\left[T'_\psi(x) + \mathcal{H}T_\psi(x)\right]. \tag{484}$$

The resulting I_Σ in the eMD era is

$$I_\Sigma^{\rm eMD} = -\frac{648}{25}\frac{x^2}{(18 + x^2y^2)(18 + x^2z^2)}. \tag{485}$$

After the transition ($x > x_m$), we obtain

$$\begin{aligned} I_\Sigma^{\rm RD} = &-\frac{1152}{y^3z^3(x_m - 2x)^4\left(4x^2y^2 - 4xx_my^2 + x_m^2y^2 + 24\right)\left(4x^2z^2 - 4xx_mz^2 + x_m^2z^2 + 24\right)}\cdot \\ &\left\{\sin\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[12b_T^-|_{yx_m}y(x_m - 2x) + \sqrt{3}b_T^+|_{yx_m}\left(-4x^2y^2 + 4xx_my^2 - x_m^2y^2 + 24\right)\right]\right. \\ &\left. + \cos\left[\frac{y}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[\sqrt{3}b_T^-|_{yx_m}\left(4x^2y^2 - 4xx_my^2 + x_m^2y^2 - 24\right) + 12b_T^+|_{yx_m}y(x_m - 2x)\right]\right\}\cdot \\ &\left\{\sin\left[\frac{z}{\sqrt{3}}\left(x - \frac{x_m}{2}\right)\right]\left[12b_T^-|_{zx_m}z(x_m - 2x) + \sqrt{3}b_T^+|_{zx_m}\left(-4x^2z^2 + 4xx_mz^2 - x_m^2z^2 + 24\right)\right]\right. \end{aligned}$$

$$+\cos\left[\frac{z}{\sqrt{3}}\left(x-\frac{x_m}{2}\right)\right]\left[\sqrt{3}b_T^-|_{zx_m}\left(4x^2z^2-4xx_mz^2+x_m^2z^2-24\right)+12b_T^+|_{zx_m}z(x_m-2x)\right]\Big\}. \tag{486}$$

Once again, the results in this gauge are completely different from the previous ones. In this case, I_Σ decays faster than x^{-1} as $x\to\infty$, and so the solution coincides with the one obtained in the Newtonian gauge only at late times. The result is depicted in Fig. 34. ■

References

[1] Guillermo Ballesteros, Julián Rey, and Fabrizio Rompineve. **Detuning primordial black hole dark matter with early matter domination and axion monodromy**. *JCAP* 06 (2020). arXiv: astro-ph.CO/1912.01638.

[2] Guillermo Ballesteros, Julián Rey, Marco Taoso, and Alfredo Urbano. **Primordial black holes as dark matter and gravitational waves from single-field polynomial inflation**. *JCAP* 07 (2020). arXiv: astro-ph.CO/2001.08220.

[3] Guillermo Ballesteros, Julián Rey, Marco Taoso, and Alfredo Urbano. **Stochastic inflationary dynamics beyond slow-roll and consequences for primordial black hole formation**. *JCAP* 08 (2020). arXiv: astro-ph.CO/2006.14597.

[4] Guillermo Ballesteros, Marcos A. G. García, Alejandro Pérez Rodríguez, Mathias Pierre, and Julián Rey. **Primordial black holes and gravitational waves from dissipation during inflation** (Aug. 2022). arXiv: astro-ph.CO/2208.14978.

[5] Gianfranco Bertone and Dan Hooper. **History of dark matter**. *Rev. Mod. Phys.* 90.4 (2018). arXiv: astro-ph.CO/1605.04909.

[6] F. Zwicky. **On the Masses of Nebulae and of Clusters of Nebulae**. *Astrophys. J.* 86 (1937).

[7] Vera C. Rubin and W. Kent Ford Jr. **Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions**. *Astrophys. J.* 159 (1970).

[8] C. Alcock et al. **The MACHO project: Microlensing results from 5.7 years of LMC observations**. *Astrophys. J.* 542 (2000). arXiv: astro-ph/0001272.

[9] D. N. Spergel et al. **First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters**. *Astrophys. J. Suppl.* 148 (2003). arXiv: astro-ph/0302209.

[10] S. Perlmutter et al. **Measurements of Ω and Λ from 42 high redshift supernovae**. *Astrophys. J.* 517 (1999). arXiv: astro-ph/9812133.

[11] Adam G. Riess et al. **Observational evidence from supernovae for an accelerating universe and a cosmological constant**. *Astron. J.* 116 (1998). arXiv: astro-ph/9805201.

[12] Jonathan L. Feng. **Dark Matter Candidates from Particle Physics and Methods of Detection**. *Ann. Rev. Astron. Astrophys.* 48 (2010). arXiv: astro-ph.CO/1003.0904.

[13] Ya. B. Zel'dovich and I. D. Novikov. **The Hypothesis of Cores Retarded during Expansion and the Hot Cosmological Model**. *Sov. Ast.* 10 (1967).

[14] Bernard J. Carr and S. W. Hawking. **Black holes in the early Universe**. *Mon. Not. Roy. Astron. Soc.* 168 (1974).

[15] I. D. Novikov, A. G. Polnarev, A. A. Starobinskii, and Ia. B. Zeldovich. **Primordial black holes**. *Astronomy and Astrophysics* 80.1 (1979).

[16] George F. Chapline. **Cosmological effects of primordial black holes**. *Nature* 253.5489 (1975).

[17] Arno A. Penzias and Robert Woodrow Wilson. **A Measurement of excess antenna temperature at 4080-Mc/s**. *Astrophys. J.* 142 (1965).

[18] Alexei A. Starobinsky. **Spectrum of relict gravitational radiation and the early state of the universe**. *JETP Lett.* 30 (1979). Ed. by I. M. Khalatnikov and V. P. Mineev.

[19] Alan H. Guth. **The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems**. *Phys. Rev. D* 23 (1981). Ed. by Li-Zhi Fang and R. Ruffini.

[20] 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 (1982). Ed. by Li-Zhi Fang and R. Ruffini.

[21] Ruth Durrer. **The cosmic microwave background: the history of its experimental investigation and its significance for cosmology**. *Class. Quant. Grav.* 32.12 (2015). arXiv: astro-ph.CO/1506.01907.

[22] George F. Smoot et al. **Structure in the COBE differential microwave radiometer first year maps**. *Astrophys. J. Lett.* 396 (1992).

[23] Y. Akrami et al. **Planck 2018 results. X. Constraints on inflation**. *Astron. Astrophys.* 641 (2020). arXiv: astro-ph.CO/1807.06211.

[24] P. A. R. Ade et al. **Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season**. *Phys. Rev. Lett.* 127.15 (2021). arXiv: astro-ph.CO/2110.00483.

[25] Kevork N. Abazajian et al. **CMB-S4 Science Book, First Edition** (Oct. 2016). arXiv: astro-ph.CO/1610.02743.

[26] T. Matsumura et al. **Mission design of LiteBIRD**. *J. Low Temp. Phys.* 176 (2014). arXiv: astro-ph.IM/1311.2847.

[27] Misao Sasaki, Teruaki Suyama, Takahiro Tanaka, and Shuichiro Yokoyama. **Primordial black holes—perspectives in gravitational wave astronomy**. *Class. Quant. Grav.* 35.6 (2018). arXiv: astro-ph.CO/1801.05235.

[28] Hiroko Niikura et al. **Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations**. *Nature Astron.* 3.6 (2019). arXiv: astro-ph.CO/1701.02151.

[29] B.J. Carr, Kazunori Kohri, Yuuiti Sendouda, and Jun'ichi Yokoyama. **Constraints on primordial black holes from the Galactic gamma-ray background**. *Phys. Rev. D* 94.4 (2016). arXiv: astro-ph.CO/1604.05349.

[30] Alexandre Arbey, Jérémy Auffinger, and Joseph Silk. **Constraining primordial black hole masses with the isotropic gamma ray background**. *Phys. Rev. D* 101.2 (2020). arXiv: astro-ph.CO/1906.04750.

[31] Ranjan Laha, Julian B. Muñoz, and Tracy R. Slatyer. **INTEGRAL constraints on primordial black holes and particle dark matter**. *Phys. Rev. D* 101.12 (2020). arXiv: astro-ph.CO/2004.00627.

[32] J. Berteaud, F. Calore, J. Iguaz, P. D. Serpico, and T. Siegert. **Strong constraints on primordial black hole dark matter from 16 years of INTEGRAL/SPI observations**. *Phys. Rev. D* 106.2 (2022). arXiv: astro-ph.HE/2202.07483.

[33] Vivian Poulin, Julien Lesgourgues, and Pasquale D. Serpico. **Cosmological constraints on exotic injection of electromagnetic energy**. *JCAP* 03 (2017). arXiv: astro-ph.CO/1610.10051.

[34] Patrick Stöcker, Michael Krämer, Julien Lesgourgues, and Vivian Poulin. **Exotic energy injection with ExoCLASS: Application to the Higgs portal model and evaporating black holes**. *JCAP* 03 (2018). arXiv: astro-ph.CO/1801.01871.

[35] Mathieu Boudaud and Marco Cirelli. **Voyager 1 $e^{\pm}$ Further Constrain Primordial Black Holes as Dark Matter**. *Phys. Rev. Lett.* 122.4 (2019). arXiv: astro-ph.HE/1807.03075.

[36] William DeRocco and Peter W. Graham. **Constraining Primordial Black Hole Abundance with the Galactic 511 keV Line**. *Phys. Rev. Lett.* 123.25 (2019). arXiv: astro-ph.CO/1906.07740.

[37] Ranjan Laha. **Primordial Black Holes as a Dark Matter Candidate Are Severely Constrained by the Galactic Center 511 keV γ -Ray Line**. *Phys. Rev. Lett.* 123.25 (2019). arXiv: astro-ph.HE/1906.09994.

[38] Basudeb Dasgupta, Ranjan Laha, and Anupam Ray. **Neutrino and positron constraints on spinning primordial black hole dark matter**. *Phys. Rev. Lett.* 125.10 (2020). arXiv: hep-ph/1912.01014.

[39] P. Tisserand et al. **Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds**. *Astron. Astrophys.* 469 (2007). arXiv: astro-ph/0607207.

[40] Hiroko Niikura, Masahiro Takada, Shuichiro Yokoyama, Takahiro Sumi, and Shogo Masaki. **Constraints on Earth-mass primordial black holes from OGLE 5-year microlensing events**. *Phys. Rev. D* 99.8 (2019). arXiv: astro-ph.CO/1901.07120.

[41] Yacine Ali-Haïmoud and Marc Kamionkowski. **Cosmic microwave background limits on accreting primordial black holes**. *Phys. Rev. D* 95.4 (2017). arXiv: astro-ph.CO/1612.05644.

[42] Pasquale D. Serpico, Vivian Poulin, Derek Inman, and Kazunori Kohri. **Cosmic microwave background bounds on primordial black holes including dark matter halo accretion**. *Phys. Rev. Res.* 2.2 (2020). arXiv: astro-ph.CO/2002.10771.

[43] Martti Raidal, Christian Spethmann, Ville Vaskonen, and Hardi Veermäe. **Formation and Evolution of Primordial Black Hole Binaries in the Early Universe**. *JCAP* 02 (2019). arXiv: astro-ph.CO/1812.01930.

[44] A. Barnacka, J. F. Glicenstein, and R. Moderski. **New constraints on primordial black holes abundance from femtolensing of gamma-ray bursts**. *Phys. Rev. D* 86 (2012). arXiv: astro-ph.CO/1204.2056.

[45] Andrey Katz, Joachim Kopp, Sergey Sibiryakov, and Wei Xue. **Femtolensing by Dark Matter Revisited**. *JCAP* 12 (2018). arXiv: astro-ph.CO/1807.11495.

[46] Peter W. Graham, Surjeet Rajendran, and Jaime Varela. **Dark Matter Triggers of Supernovae**. *Phys. Rev.* D92.6 (2015). arXiv: hep-ph/1505.04444.

[47] Fabio Capela, Maxim Pshirkov, and Peter Tinyakov. **Constraints on primordial black holes as dark matter candidates from capture by neutron stars**. *Phys. Rev. D* 87.12 (2013). arXiv: astro-ph.CO/1301.4984.

[48] Paulo Montero-Camacho, Xiao Fang, Gabriel Vasquez, Makana Silva, and Christopher M. Hirata. **Revisiting constraints on asteroid-mass primordial black holes as dark matter candidates**. *JCAP* 08 (2019). arXiv: astro-ph.CO/1906.05950.

[49] Yoann Génolini, Pasquale Serpico, and Peter Tinyakov. **Revisiting primordial black hole capture into neutron stars**. *Phys. Rev. D* 102.8 (2020). arXiv: astro-ph.HE/2006.16975.

[50] Bernard J. Carr. **The Primordial black hole mass spectrum**. *Astrophys. J.* 201 (1975).

[51] William H. Kinney. **Horizon crossing and inflation with large eta**. *Phys. Rev. D* 72 (2005). arXiv: gr-qc/0503017.

[52] Juan Garcia-Bellido, Andrei D. Linde, and David Wands. **Density perturbations and black hole formation in hybrid inflation**. *Phys. Rev. D* 54 (1996). arXiv: astro-ph/9605094.

[53] A. Vilenkin. **Cosmological Density Fluctuations Produced by Vacuum Strings**. *Phys. Rev. Lett.* 46 (1981). [Erratum: Phys. Rev. Lett.46,1496(1981)].

[54] Joaquim Fort and Tanmay Vachaspati. **Do global string loops collapse to form black holes?** *Phys. Lett.* B311 (1993). arXiv: hep-th/hep-th/9305081.

[55] Maxim. Yu. Khlopov, Sergei G. Rubin, and Alexander S. Sakharov. **Primordial structure of massive black hole clusters**. *Astropart. Phys.* 23 (2005). arXiv: astro-ph/astro-ph/0401532.

[56] Heling Deng, Jaume Garriga, and Alexander Vilenkin. **Primordial black hole and wormhole formation by domain walls**. *JCAP* 1704.04 (2017). arXiv: gr-qc/1612.03753.

[57] Francesc Ferrer, Eduard Masso, Giuliano Panico, Oriol Pujolas, and Fabrizio Rompineve. **Primordial Black Holes from the QCD axion**. *Phys. Rev. Lett.* 122.10 (2019). arXiv: hep-ph/1807.01707.

[58] Hideo Kodama, Misao Sasaki, and Katsuhiko Sato. **Abundance of Primordial Holes Produced by Cosmological First Order Phase Transition**. *Prog. Theor. Phys.* 68 (1982).

[59] S. W. Hawking, I. G. Moss, and J. M. Stewart. **Bubble Collisions in the Very Early Universe**. *Phys. Rev. D* 26 (1982).

[60] Kiyoharu Kawana and Ke-Pan Xie. **Primordial black holes from a cosmic phase transition: The collapse of Fermi-balls**. *Phys. Lett. B* 824 (2022). arXiv: astro-ph.CO/2106.00111.

[61] Bernard Carr, Sebastien Clesse, Juan García-Bellido, and Florian Kühnel. **Cosmic conundra explained by thermal history and primordial black holes**. *Phys. Dark Univ.* 31 (2021). arXiv: astro-ph.CO/1906.08217.

[62] Michael J. Baker, Moritz Breitbach, Joachim Kopp, and Lukas Mittnacht. **Detailed Calculation of Primordial Black Hole Formation During First-Order Cosmological Phase Transitions** (Sept. 2021). arXiv: astro-ph.CO/2110.00005.

[63] P. Ivanov, P. Naselsky, and I. Novikov. **Inflation and primordial black holes as dark matter**. *Phys. Rev. D* 50 (1994).

[64] Alexei A. Starobinsky. **Spectrum of adiabatic perturbations in the universe when there are singularities in the inflation potential**. *JETP Lett.* 55 (1992). [Pisma Zh. Eksp. Teor. Fiz.55,477(1992)].

[65] B. P. Abbott et al. **Observation of Gravitational Waves from a Binary Black Hole Merger**. *Phys. Rev. Lett.* 116.6 (2016). arXiv: gr-qc/1602.03837.

[66] Juan García-Bellido and Ester Ruiz Morales. **Primordial black holes from single field models of inflation**. *Phys. Dark Univ.* 18 (2017). arXiv: astro-ph.CO/1702.03901.

[67] Jose Maria Ezquiaga, Juan Garcia-Bellido, and Ester Ruiz Morales. **Primordial Black Hole production in Critical Higgs Inflation**. *Phys. Lett. B* 776 (2018). arXiv: astro-ph.CO/1705.04861.

[68] Guillermo Ballesteros and Marco Taoso. **Primordial black hole dark matter from single field inflation**. *Phys. Rev. D* 97.2 (2018). arXiv: hep-ph/1709.05565.

[69] Ogan Ozsoy, Susha Parameswaran, Gianmassimo Tasinato, and Ivonne Zavala. **Mechanisms for Primordial Black Hole Production in String Theory**. *JCAP* 1807 (2018). arXiv: hep-th/1803.07626.

[70] Ioannis Dalianis, Alex Kehagias, and George Tringas. **Primordial black holes from α-attractors**. *JCAP* 1901 (2019). arXiv: astro-ph.CO/1805.09483.

[71] Alexei A. Starobinsky. **STOCHASTIC DE SITTER (INFLATIONARY) STAGE IN THE EARLY UNIVERSE**. *Lect. Notes Phys.* 246 (1986).

[72] Vincent Vennin and Alexei A. Starobinsky. **Correlation Functions in Stochastic Inflation**. *Eur. Phys. J. C* 75 (2015). arXiv: hep-th/1506.04732.

[73] José María Ezquiaga, Juan García-Bellido, and Vincent Vennin. **The exponential tail of inflationary fluctuations: consequences for primordial black holes**. *JCAP* 03 (2020). arXiv: astro-ph.CO/1912.05399.

[74] Chris Pattison, Vincent Vennin, David Wands, and Hooshyar Assadullahi. **Ultra-slow-roll inflation with quantum diffusion**. *JCAP* 04 (2021). arXiv: astro-ph.CO/2101.05741.

[75] Daniel G. Figueroa, Sami Raatikainen, Syksy Rasanen, and Eemeli Tomberg. **Non-Gaussian Tail of the Curvature Perturbation in Stochastic Ultraslow-Roll Inflation: Implications for Primordial Black Hole Production**. *Phys. Rev. Lett.* 127.10 (2021). arXiv: astro-ph.CO/2012.06551.

[76] Daniel G. Figueroa, Sami Raatikainen, Syksy Rasanen, and Eemeli Tomberg. **Implications of stochastic effects for primordial black hole production in ultra-slow-roll inflation**. *JCAP* 05.05 (2022). arXiv: astro-ph.CO/2111.07437.

[77] Pau Amaro-Seoane et al. **Laser Interferometer Space Antenna** (Feb. 2017). arXiv: astro-ph.IM/1702.00786.

[78] Chiara Caprini, Mark Hindmarsh, Stephan Huber, Thomas Konstandin, Jonathan Kozaczuk, Germano Nardini, Jose Miguel No, Antoine Petiteau, Pedro Schwaller, Géraldine Servant, and David J. Weir. **Science with the space-based interferometer eLISA. II: Gravitational waves from cosmological phase transitions**. *JCAP* 04 (2016). arXiv: astro-ph.CO/1512.06239.

[79] Kenji Tomita. **Non-Linear Theory of Gravitational Instability in the Expanding Universe**. *Progress of Theoretical Physics* 37.5 (May 1967).

[80] Kishore N. Ananda, Chris Clarkson, and David Wands. **The Cosmological gravitational wave background from primordial density perturbations**. *Phys. Rev. D* 75 (2007). arXiv: gr-qc/0612013.

[81] Daniel Baumann, Paul J. Steinhardt, Keitaro Takahashi, and Kiyotomo Ichiki. **Gravitational Wave Spectrum Induced by Primordial Scalar Perturbations**. *Phys. Rev. D* 76 (2007). arXiv: hep-th/0703290.

[82] N. Bartolo, V. De Luca, G. Franciolini, A. Lewis, M. Peloso, and A. Riotto. **Primordial Black Hole Dark Matter: LISA Serendipity**. *Phys. Rev. Lett.* 122.21 (2019). arXiv: astro-ph.CO/1810.12218.

[83] N. Bartolo, V. De Luca, G. Franciolini, M. Peloso, D. Racco, and A. Riotto. **Testing primordial black holes as dark matter with LISA**. *Phys. Rev. D* 99.10 (2019). arXiv: astro-ph.CO/1810.12224.

[84] Jai-Chan Hwang, Donghui Jeong, and Hyerim Noh. **Gauge dependence of gravitational waves generated from scalar perturbations**. *Astrophys. J.* 842.1 (2017). arXiv: astro-ph.CO/1704.03500.

[85] Yizhou Lu, Arshad Ali, Yungui Gong, Jiong Lin, and Fengge Zhang. **Gauge transformation of scalar induced gravitational waves**. *Phys. Rev. D* 102.8 (2020). arXiv: gr-qc/2006.03450.

[86] V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto. **On the Gauge Invariance of Cosmological Gravitational Waves**. *JCAP* 03 (2020). arXiv: gr-qc/1911.09689.

[87] Keisuke Inomata and Takahiro Terada. **Gauge Independence of Induced Gravitational Waves**. *Phys. Rev. D* 101.2 (2020). arXiv: gr-qc/1912.00785.

[88] Chen Yuan, Zu-Cheng Chen, and Qing-Guo Huang. **Scalar induced gravitational waves in different gauges**. *Phys. Rev. D* 101.6 (2020). arXiv: astro-ph.CO/1912.00885.

[89] Arshad Ali, Yungui Gong, and Yizhou Lu. **Gauge transformation of scalar induced tensor perturbation during matter domination**. *Phys. Rev. D* 103.4 (2021). arXiv: gr-qc/2009.11081.

[90] Guillem Domènech and Misao Sasaki. **Approximate gauge independence of the induced gravitational wave spectrum**. *Phys. Rev. D* 103.6 (2021). arXiv: gr-qc/2012.14016.

[91] M. Yu. Khlopov and A. G. Polnarev. **Primordial Black Holes as a Cosmological Test of Grand Unification**. *Phys. Lett.* 97B (1980).

[92] Tomohiro Harada, Chul-Moon Yoo, Kazunori Kohri, Ken-ichi Nakao, and Sanjay Jhingan. **Primordial black hole formation in the matter-dominated phase of the Universe**. *Astrophys. J.* 833.1 (2016). arXiv: astro-ph.CO/1609.01588.

[93] Tomohiro Harada, Chul-Moon Yoo, Kazunori Kohri, and Ken-Ichi Nakao. **Spins of primordial black holes formed in the matter-dominated phase of the Universe**. *Phys. Rev. D* 96.8 (2017). [Erratum: Phys.Rev.D 99, 069904 (2019)]. arXiv: gr-qc/1707.03595.

[94] Takafumi Kokubu, Koutarou Kyutoku, Kazunori Kohri, and Tomohiro Harada. **Effect of Inhomogeneity on Primordial Black Hole Formation in the Matter Dominated Era**. *Phys. Rev.* D98.12 (2018). arXiv: astro-ph.CO/1810.03490.

[95] Arjun Berera and Li-Zhi Fang. **Thermally induced density perturbations in the inflation era**. *Phys. Rev. Lett.* 74 (1995). arXiv: astro-ph/9501024.

[96] Arjun Berera. **Warm inflation**. *Phys. Rev. Lett.* 75 (1995). arXiv: astro-ph/9509049.

[97] Matteo Biagetti, Valerio De Luca, Gabriele Franciolini, Alex Kehagias, and Antonio Riotto. **The formation probability of primordial black holes**. *Phys. Lett. B* 820 (2021). arXiv: astro-ph.CO/2105.07810.

[98] Thomas J. Clarke, Edmund J. Copeland, and Adam Moss. **Constraints on primordial gravitational waves from the Cosmic Microwave Background**. *JCAP* 10 (2020). arXiv: astro-ph.CO/2004.11396.

[99] Luca Pagano, Laura Salvati, and Alessandro Melchiorri. **New constraints on primordial gravitational waves from Planck 2015**. *Phys. Lett. B* 760 (2016). arXiv: astro-ph.CO/1508.02393.

[100] Charles W. Misner, K.S. Thorne, and J.A. Wheeler. **Gravitation**. San Francisco: W. H. Freeman, 1973.

[101] Michele Maggiore. **Gravitational Waves. Vol. 1: Theory and Experiments**. Oxford Master Series in Physics. Oxford University Press, 2007.

[102] Daniel Baumann. **Inflation**. 2011. arXiv: hep-th/0907.5424.

[103] N. Aghanim et al. **Planck 2018 results. VI. Cosmological parameters**. *Astron. Astrophys.* 641 (2020). [Erratum: Astron.Astrophys. 652, C4 (2021)]. arXiv: astro-ph.CO/1807.06209.

[104] Michael S. Turner. **Coherent Scalar Field Oscillations in an Expanding Universe**. *Phys. Rev.* D28 (1983).

[105] Alba Kalaja, Nicola Bellomo, Nicola Bartolo, Daniele Bertacca, Sabino Matarrese, Ilia Musco, Alvise Raccanelli, and Licia Verde. **From Primordial Black Holes Abundance to Primordial Curvature Power Spectrum (and back)**. *JCAP* 10 (2019). arXiv: astro-ph.CO/1908.03596.

[106] Jens C. Niemeyer and K. Jedamzik. **Dynamics of primordial black hole formation**. *Phys. Rev.* D59 (1999). arXiv: astro-ph/astro-ph/9901292.

[107] D. J. Fixsen. **The Temperature of the Cosmic Microwave Background**. *Astrophysical Journal* 707 (Dec. 2009). arXiv: 0911.1955.

[108] Nora Elisa Chisari and Matias Zaldarriaga. **Connection between Newtonian simulations and general relativity**. *Phys. Rev. D* 83 (2011). [Erratum: Phys.Rev.D 84, 089901 (2011)]. arXiv: astro-ph.CO/1101.3555.

[109] Guillermo Ballesteros, Jose Beltran Jimenez, and Mauro Pieroni. **Black hole formation from a general quadratic action for inflationary primordial fluctuations**. *JCAP* 1906 (2019). arXiv: astro-ph.CO/1811.03065.

[110] William H. Press and Paul Schechter. **Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation**. *Astrophys. J.* 187 (1974).

[111] Ilia Musco, John C. Miller, and Luciano Rezzolla. **Computations of primordial black hole formation**. *Class. Quant. Grav.* 22 (2005). arXiv: gr-qc/0412063.

[112] Ilia Musco, John C. Miller, and Alexander G. Polnarev. **Primordial black hole formation in the radiative era: Investigation of the critical nature of the collapse**. *Class. Quant. Grav.* 26 (2009). arXiv: gr-qc/0811.1452.

[113] Ilia Musco and John C. Miller. **Primordial black hole formation in the early universe: critical behaviour and self-similarity**. *Class. Quant. Grav.* 30 (2013). arXiv: gr-qc/1201.2379.

[114] Ya. B. Zeldovich. **Gravitational instability: An Approximate theory for large density perturbations**. *Astron. Astrophys.* 5 (1970).

[115] A. G. Doroshkevich. **Spatial structure of perturbations and origin of galactic rotation in fluctuation theory**. *Astrophysics* 6.4 (Oct. 1970).

[116] K. S. Thorne. **Nonspherical Gravitational Collapse–A Short Review**. 1972.

[117] P.J.E. Peebles. **Origin of the Angular Momentum of Galaxies**. *Astrophys. J.* 155 (1969).

[118] Hiroko Niikura, Masahiro Takada, Naoki Yasuda, Robert H. Lupton, Takahiro Sumi, Surhud More, Toshiki Kurita, Sunao Sugiyama, Anupreeta More, Masamune Oguri, and et al. **Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations**. *Nature Astron.* 3.6 (2019). arXiv: astro-ph.CO/1701.02151.

[119] Lev Kofman, Andrei D. Linde, and Alexei A. Starobinsky. **Towards the theory of reheating after inflation**. *Phys. Rev.* D56 (1997). arXiv: hep-ph/hep-ph/9704452.

[120] Dmitry I. Podolsky, Gary N. Felder, Lev Kofman, and Marco Peloso. **Equation of state and beginning of thermalization after preheating**. *Phys. Rev.* D73 (2006). arXiv: hep-ph/hep-ph/0507096.

[121] Michele Cicoli, Joseph P. Conlon, and Fernando Quevedo. **Dark radiation in LARGE volume models**. *Phys. Rev.* D87.4 (2013). arXiv: hep-ph/1208.3562.

[122] Vijay Balasubramanian, Per Berglund, Joseph P. Conlon, and Fernando Quevedo. **Systematics of moduli stabilisation in Calabi-Yau flux compactifications**. *JHEP* 03 (2005). arXiv: hep-th/hep-th/0502058.

[123] Joseph P. Conlon, Fernando Quevedo, and Kerim Suruliz. **Large-volume flux compactifications: Moduli spectrum and D3/D7 soft supersymmetry breaking**. *JHEP* 08 (2005). arXiv: hep-th/hep-th/0505076.

[124] Andrew R Liddle and Samuel M Leach. **How long before the end of inflation were observable perturbations produced?** *Phys. Rev.* D68 (2003). arXiv: astro-ph/astro-ph/0305263.

[125] Marco Taoso and Alfredo Urbano. **Non-gaussianities for primordial black hole formation**. *JCAP* 08 (2021). arXiv: astro-ph.CO/2102.03610.

[126] Christian T. Byrnes, Philippa S. Cole, and Subodh P. Patil. **Steepest growth of the power spectrum and primordial black holes**. *JCAP* 06 (2019). arXiv: astro-ph.CO/1811.11158.

[127] Samuel M. Leach and Andrew R. Liddle. **Inflationary perturbations near horizon crossing**. *Phys. Rev.* D63 (2001). arXiv: astro-ph/astro-ph/0010082.

[128] Samuel M Leach, Misao Sasaki, David Wands, and Andrew R Liddle. **Enhancement of superhorizon scale inflationary curvature perturbations**. *Phys. Rev.* D64 (2001). arXiv: astro-ph/astro-ph/0101406.

[129] Kristjan Kannike, Luca Marzola, Martti Raidal, and Hardi Veermäe. **Single Field Double Inflation and Primordial Black Holes**. *JCAP* 09 (2017). arXiv: astro-ph.CO/1705.06225.

[130] M. Tristram et al. **Improved limits on the tensor-to-scalar ratio using BICEP and Planck data**. *Phys. Rev. D* 105.8 (2022). arXiv: astro-ph.CO/2112.07961.

[131] Shouvik Roy Choudhury, Steen Hannestad, and Thomas Tram. **Massive neutrino self-interactions and Inflation** (July 2022). arXiv: astro-ph.CO/2207.07142.

[132] Martina Gerbino, Katherine Freese, Sunny Vagnozzi, Massimiliano Lattanzi, Olga Mena, Elena Giusarma, and Shirley Ho. **Impact of neutrino properties on the estimation of inflationary parameters from current and future observations**. *Phys. Rev. D* 95.4 (2017). arXiv: astro-ph.CO/1610.08830.

[133] Elcio Abdalla et al. **Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies**. *JHEAp* 34 (2022). arXiv: astro-ph.CO/2203.06142.

[134] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. **Cosmological constraints in extended parameter space from the Planck 2018 Legacy release**. *JCAP* 01 (2020). arXiv: astro-ph.CO/1908.01391.

[135] Cristiano Germani and Ilia Musco. **Abundance of Primordial Black Holes Depends on the Shape of the Inflationary Power Spectrum**. *Phys. Rev. Lett.* 122.14 (2019). arXiv: astro-ph.CO/1805.04087.

[136] Albert Escrivà, Cristiano Germani, and Ravi K. Sheth. **Universal threshold for primordial black hole formation**. *Phys. Rev. D* 101.4 (2020). arXiv: gr-qc/1907.13311.

[137] Stanley J. Brodsky and Paul Hoyer. **The $\bar{h}$ Expansion in Quantum Field Theory**. *Phys. Rev. D* 83 (2011). arXiv: hep-ph/1009.2313.

[138] C. P. Burgess, Hyun Min Lee, and Michael Trott. **Power-counting and the Validity of the Classical Approximation During Inflation**. *JHEP* 09 (2009). arXiv: hep-ph/0902.4465.

[139] J. L. F. Barbon and J. R. Espinosa. **On the Naturalness of Higgs Inflation**. *Phys. Rev. D* 79 (2009). arXiv: hep-ph/0903.0355.

[140] Eva Silverstein and Alexander Westphal. **Monodromy in the CMB: Gravity Waves and String Inflation**. *Phys. Rev.* D78 (2008). arXiv: hep-th/0803.3085.

[141] Liam McAllister, Eva Silverstein, and Alexander Westphal. **Gravity Waves and Linear Inflation from Axion Monodromy**. *Phys. Rev.* D82 (2010). arXiv: hep-th/0808.0706.

[142] Xi Dong, Bart Horn, Eva Silverstein, and Alexander Westphal. **Simple exercises to flatten your potential**. *Phys. Rev.* D84 (2011). arXiv: hep-th/1011.4521.

[143] Liam McAllister, Eva Silverstein, Alexander Westphal, and Timm Wrase. **The Powers of Monodromy**. *JHEP* 09 (2014). arXiv: hep-th/1405.3652.

[144] Raphael Flauger, Liam McAllister, Eva Silverstein, and Alexander Westphal. **Drifting Oscillations in Axion Monodromy**. *JCAP* 1710.10 (2017). arXiv: hep-th/1412.1814.

[145] Nima Arkani-Hamed, Lubos Motl, Alberto Nicolis, and Cumrun Vafa. **The String landscape, black holes and gravity as the weakest force**. *JHEP* 06 (2007). arXiv: hep-th/hep-th/0601001.

[146] Tom Rudelius. **Constraints on Axion Inflation from the Weak Gravity Conjecture**. *JCAP* 1509.09 (2015). arXiv: hep-th/1503.00795.

[147] Jon Brown, William Cottrell, Gary Shiu, and Pablo Soler. **Fencing in the Swampland: Quantum Gravity Constraints on Large Field Inflation**. *JHEP* 10 (2015). arXiv: hep-th/1503.04783.

[148] Daniel Klaewer and Eran Palti. **Super-Planckian Spatial Field Variations and Quantum Gravity**. *JHEP* 01 (2017). arXiv: hep-th/1610.00010.

[149] Sumit K. Garg and Chethan Krishnan. **Bounds on Slow Roll and the de Sitter Swampland**. *JHEP* 11 (2019). arXiv: hep-th/1807.05193.

[150] Hirosi Ooguri, Eran Palti, Gary Shiu, and Cumrun Vafa. **Distance and de Sitter Conjectures on the Swampland**. *Phys. Lett.* B788 (2019). arXiv: hep-th/1810.05506.

[151] Aitor Landete, Fernando Marchesano, Gary Shiu, and Gianluca Zoccarato. **Flux Flattening in Axion Monodromy Inflation**. *JHEP* 06 (2017). arXiv: hep-th/1703.09729.

[152] Peter W. Graham, David E. Kaplan, and Surjeet Rajendran. **Cosmological Relaxation of the Electroweak Scale**. *Phys. Rev. Lett.* 115.22 (2015). arXiv: hep-ph/1504.07551.

[153] Sidney R. Coleman. **The Fate of the False Vacuum. 1. Semiclassical Theory**. *Phys. Rev.* D15 (1977). [Erratum: Phys. Rev.D16,1248(1977)].

[154] Arthur Hebecker, Joerg Jaeckel, Fabrizio Rompineve, and Lukas T. Witkowski. **Gravitational Waves from Axion Monodromy**. *JCAP* 1611.11 (2016). arXiv: hep-ph/1606.07812.

[155] Mark P. Hertzberg and Masaki Yamada. **Primordial Black Holes from Polynomial Potentials in Single Field Inflation**. *Phys. Rev.* D97.8 (2018). arXiv: astro-ph.CO/1712.09750.

[156] David Polarski and Alexei A. Starobinsky. **Semiclassicality and decoherence of cosmological perturbations**. *Class. Quant. Grav.* 13 (1996). arXiv: gr-qc/9504030.

[157] Claus Kiefer and David Polarski. **Emergence of classicality for primordial fluctuations: Concepts and analogies**. *Annalen Phys.* 7 (1998). arXiv: gr-qc/9805014.

[158] Matteo Biagetti, Gabriele Franciolini, Alex Kehagias, and Antonio Riotto. **Primordial Black Holes from Inflation and Quantum Diffusion**. *JCAP* 07 (2018). arXiv: astro-ph.CO/1804.07124.

[159] José María Ezquiaga and Juan García-Bellido. **Quantum diffusion beyond slow-roll: implications for primordial black-hole production**. *JCAP* 08 (2018). arXiv: astro-ph.CO/1805.06731.

[160] Diego Cruces, Cristiano Germani, and Tomislav Prokopec. **Failure of the stochastic approach to inflation beyond slow-roll**. *JCAP* 03 (2019). arXiv: gr-qc/1807.09057.

[161] Hassan Firouzjahi, Amin Nassiri-Rad, and Mahdiyar Noorbala. **Stochastic Ultra Slow Roll Inflation**. *JCAP* 01 (2019). arXiv: hep-th/1811.02175.

[162] Chris Pattison, Vincent Vennin, Hooshyar Assadullahi, and David Wands. **Stochastic inflation beyond slow roll**. *JCAP* 07 (2019). arXiv: astro-ph.CO/1905.06300.

[163] S. Winitzki and A. Vilenkin. **Effective noise in stochastic description of inflation**. *Phys. Rev. D* 61 (2000). arXiv: gr-qc/9911029.

[164] M. Liguori, S. Matarrese, M. Musso, and A. Riotto. **Stochastic inflation and the lower multipoles in the CMB anisotropies**. *JCAP* 08 (2004). arXiv: astro-ph/0405544.

[165] H. P. Breuer and K. E. Kunze. **Stochastic inflation with coloured noise**. *AIP Conf. Proc.* 841.1 (2006). Ed. by Lysiane Mornas and Joaquin Diaz Alonso.

[166] Rafid Mahbub and Aritra De. **Smooth coarse-graining and colored noise dynamics in stochastic inflation** (Apr. 2022). arXiv: astro-ph.CO/2204.03859.

[167] Laurence Perreault Levasseur, Vincent Vennin, and Robert Brandenberger. **Recursive Stochastic Effects in Valley Hybrid Inflation**. *Phys. Rev. D* 88 (2013). arXiv: hep-th/1307.2575.

[168] Laurence Perreault Levasseur. **Lagrangian formulation of stochastic inflation: Langevin equations, one-loop corrections and a proposed recursive approach**. *Phys. Rev. D* 88.8 (2013). arXiv: hep-th/1304.6408.

[169] F. Finelli, G. Marozzi, A. A. Starobinsky, G. P. Vacca, and G. Venturi. **Stochastic growth of quantum fluctuations during slow-roll inflation**. *Phys. Rev. D* 82 (2010). arXiv: hep-th/1003.1327.

[170] F. Finelli, G. Marozzi, A. A. Starobinsky, G. P. Vacca, and G. Venturi. **Generation of fluctuations during inflation: Comparison of stochastic and field-theoretic approaches**. *Phys. Rev. D* 79 (2009). arXiv: hep-th/0808.1786.

[171] Julien Grain and Vincent Vennin. **Stochastic inflation in phase space: Is slow roll a stochastic attractor?** *JCAP* 05 (2017). arXiv: gr-qc/1703.00447.

[172] Alexei A. Starobinsky and Junichi Yokoyama. **Equilibrium state of a selfinteracting scalar field in the De Sitter background**. *Phys. Rev. D* 50 (1994). arXiv: astro-ph/9407016.

[173] Andrew J. Tolley and Mark Wyman. **Stochastic Inflation Revisited: Non-Slow Roll Statistics and DBI Inflation**. *JCAP* 04 (2008). arXiv: hep-th/0801.1854.

[174] Chris Pattison, Vincent Vennin, Hooshyar Assadullahi, and David Wands. **Quantum diffusion during inflation and primordial black holes**. *JCAP* 10 (2017). arXiv: hep-th/1707.00537.

[175] Jun'ichi Yokoyama and Andrei D. Linde. **Is warm inflation possible?** *Phys. Rev. D* 60 (1999). arXiv: hep-ph/9809409.

[176] Arjun Berera, Marcelo Gleiser, and Rudnei O. Ramos. **A First principles warm inflation model that solves the cosmological horizon / flatness problems**. *Phys. Rev. Lett.* 83 (1999). arXiv: hep-ph/9809583.

[177] Arjun Berera and Rudnei O. Ramos. **The Affinity for scalar fields to dissipate**. *Phys. Rev. D* 63 (2001). arXiv: hep-ph/0101049.

[178] Richa Arya. **Formation of Primordial Black Holes from Warm Inflation**. *JCAP* 09 (2020). arXiv: astro-ph.CO/1910.05238.

[179] Mar Bastero-Gil and Marta Subías Díaz-Blanco. **Gravity waves and primordial black holes in scalar warm little inflation**. *JCAP* 12.12 (2021). arXiv: hep-ph/2105.08045.

[180] Chung-Pei Ma and Edmund Bertschinger. **Cosmological perturbation theory in the synchronous and conformal Newtonian gauges**. *Astrophys. J.* 455 (1995). arXiv: astro-ph/9506072.

[181] Hideo Kodama and Misao Sasaki. **Cosmological Perturbation Theory**. *Prog. Theor. Phys. Suppl.* 78 (1984).

[182] Marcelo Gleiser and Rudnei O. Ramos. **Microphysical approach to nonequilibrium dynamics of quantum fields**. *Phys. Rev. D* 50 (1994). arXiv: hep-ph/9311278.

[183] Mar Bastero-Gil, Arjun Berera, Ian G. Moss, and Rudnei O. Ramos. **Cosmological fluctuations of a random field and radiation fluid**. *JCAP* 05 (2014). arXiv: astro-ph.CO/1401.1149.

[184] Lisa M. H. Hall, Ian G. Moss, and Arjun Berera. **Scalar perturbation spectra from warm inflation**. *Phys. Rev. D* 69 (2004). arXiv: astro-ph/0305015.

[185] Kazunori Kohri and Takahiro Terada. **Semianalytic calculation of gravitational wave spectrum nonlinearly induced from primordial curvature perturbations**. *Phys. Rev. D* 97.12 (2018). arXiv: gr-qc/1804.08577.

[186] José Ramón Espinosa, Davide Racco, and Antonio Riotto. **A Cosmological Signature of the SM Higgs Instability: Gravitational Waves**. *JCAP* 09 (2018). arXiv: hep-ph/1804.07732.

[187] Keisuke Inomata, Kazunori Kohri, Tomohiro Nakama, and Takahiro Terada. **Enhancement of Gravitational Waves Induced by Scalar Perturbations due to a Sudden Transition from an Early Matter Era to the Radiation Era**. *Phys. Rev. D* 100.4 (2019). arXiv: astro-ph.CO/1904.12879.

[188] Guillem Domènech. **Induced gravitational waves in a general cosmological background**. *Int. J. Mod. Phys. D* 29.03 (2020). arXiv: gr-qc/1912.05583.

[189] Karim A. Malik and David Wands. **Cosmological perturbations**. *Phys. Rept.* 475 (2009). arXiv: astro-ph/0809.4944.

[190] 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 (1992).

[191] Richard A. Isaacson. **Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics**. *Phys. Rev.* 166 (5 Feb. 1968).

[192] Richard A. Isaacson. **Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Effective Stress Tensor**. *Phys. Rev.* 166 (5 Feb. 1968).

[193] Daniel Baumann, Alberto Nicolis, Leonardo Senatore, and Matias Zaldarriaga. **Cosmological Non-Linearities as an Effective Fluid**. *JCAP* 07 (2012). arXiv: astro-ph.CO/1004.2488.

[194] Stephen R. Green and Robert M. Wald. **A new framework for analyzing the effects of small scale inhomogeneities in cosmology**. *Phys. Rev. D* 83 (2011). arXiv: gr-qc/1011.4920.

[195] Stephen R. Green and Robert M. Wald. **A simple, heuristic derivation of our 'no backreaction' results**. *Class. Quant. Grav.* 33.12 (2016). arXiv: gr-qc/1601.06789.

[196] L. Raul W. Abramo, Robert H. Brandenberger, and Viatcheslav F. Mukhanov. **The Energy - momentum tensor for cosmological perturbations**. *Phys. Rev. D* 56 (1997). arXiv: gr-qc/9704037.

[197] Kent Yagi and Naoki Seto. **Detector configuration of DECIGO/BBO and identification of cosmological neutron-star binaries**. *Phys. Rev. D* 83 (2011). [Erratum: Phys.Rev.D 95, 109901 (2017)]. arXiv: astro-ph.CO/1101.3940.

[198] Jon Coleman. **Matter-wave Atomic Gradiometer InterferometricSensor (MAGIS-100) at Fermilab**. *PoS* ICHEP2018 (2019). arXiv: physics.ins-det/1812.00482.

[199] Michele Maggiore et al. **Science Case for the Einstein Telescope**. *JCAP* 03 (2020). arXiv: astro-ph.CO/1912.02622.

[200] Benjamin P. Abbott et al. **Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy**. *Phys. Rev. D* 93.11 (2016). [Addendum: Phys.Rev.D 97, 059901 (2018)]. arXiv: astro-ph.IM/1604.00439.

[201] L. Badurina et al. **AION: An Atom Interferometer Observatory and Network**. *JCAP* 05 (2020). arXiv: astro-ph.CO/1911.11755.

[202] Yousef Abou El-Neaj et al. **AEDGE: Atomic Experiment for Dark Matter and Gravity Exploration in Space**. *EPJ Quant. Technol.* 7 (2020). arXiv: gr-qc/1908.00802.

[203] B. J. Carr, Kazunori Kohri, Yuuiti Sendouda, and Jun'ichi Yokoyama. **New cosmological constraints on primordial black holes**. *Phys. Rev.* D81 (2010). arXiv: astro-ph.CO/0912.5297.

[204] Jacopo Fumagalli, Gonzalo A. Palma, Sébastien Renaux-Petel, Spyros Sypsas, Lukas T. Witkowski, and Cristobal Zenteno. **Primordial gravitational waves from excited states**. *JHEP* 03 (2022). arXiv: astro-ph.CO/2111.14664.

[205] Kurt Jacobs. **Stochastic Processes for Physicists: Understanding Noisy Systems**. Cambridge University Press, 2010.

[206] Arjun Berera and Rudnei O. Ramos. **Construction of a robust warm inflation mechanism**. *Phys. Lett. B* 567 (2003). arXiv: hep-ph/0210301.

[207] Mar Bastero-Gil and Arjun Berera. **Warm inflation model building**. *Int. J. Mod. Phys. A* 24 (2009). arXiv: hep-ph/0902.0521.

[208] Lisa M H Hall and Ian G Moss. **Thermal effects on pure and hybrid inflation**. *Phys. Rev. D* 71 (2005). arXiv: hep-ph/0408323.

[209] Arjun Berera, Marcelo Gleiser, and Rudnei O. Ramos. **Strong dissipative behavior in quantum field theory**. *Phys. Rev. D* 58 (1998). arXiv: hep-ph/9803394.

[210] Akio Hosoya and Masa-aki Sakagami. **Time Development of Higgs Field at Finite Temperature**. *Phys. Rev. D* 29 (1984).

[211] Arjun Berera, Ian G. Moss, and Rudnei O. Ramos. **Warm Inflation and its Microphysical Basis**. *Rept. Prog. Phys.* 72 (2009). arXiv: hep-ph/0808.1855.

[212] Mar Bastero-Gil, Arjun Berera, and Rudnei O. Ramos. **Dissipation coefficients from scalar and fermion quantum field interactions**. *JCAP* 09 (2011). arXiv: hep-ph/1008.1929.