MITP-25-036

Full-Spectrum Analysis of Gravitational Wave Production from Inflation to Reheating

Xun-Jie Xu[a], **Yong Xu**[b], **Qiqin Yin**[c], **and Junyu Zhu**[a,d]

[a] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

[b] *PRISMA$^+$ Cluster of Excellence and Mainz Institute for Theoretical Physics Johannes Gutenberg University, 55099 Mainz, Germany*

[c] *School of Physics, Nanjing University, Nanjing 210093, China*

[d] *School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

E-mail: xuxj@ihep.ac.cn, yonxu@uni-mainz.de, 211870080@smail.nju.edu.cn, zhujunyu@ihep.ac.cn

Abstract: In this work, we systematically study gravitational wave (GW) production during both the inflationary and post-inflationary epochs. While inflationary GWs can be readily derived from tensor perturbations during inflation, post-inflationary GWs arise from a variety of processes during reheating and require detailed treatment for quantitative analysis. We consider four distinct production channels: (*i*) pure inflaton annihilation, (*ii*) graviton bremsstrahlung from inflaton decay, (*iii*) radiation-catalyzed inflaton-graviton conversion, and (*iv*) scattering among fully thermalized radiation particles. For each channel, we solve the corresponding Boltzmann equation to obtain the GW spectrum and derive a simple yet accurate analytical expression for it. By employing a consistent treatment of all production channels, our analysis yields for the first time the full spectrum of GWs produced during the inflationary and post-inflationary epochs. We find that, while inflationary GWs dominate at low frequencies, post-inflationary processes generally produce high-frequency GWs with considerably high energy densities that may significantly exceed that of inflationary GWs.

Contents

1 Introduction

Cosmic inflation offers an elegant framework for addressing key problems in standard cosmology [1–4], such as the horizon and flatness problems, while also providing a natural source for generating a stochastic background of gravitational waves (GWs)—see [5, 6] for reviews. In the simplest realization, a single inflaton field undergoes slow-roll evolution along a nearly flat potential, driving a period of exponential expansion. During this phase, quantum fluctuations of the inflaton field generate metric perturbations, sourcing tensor modes. After inflation, as these modes reenter the horizon, they become dynamical and form a primordial GW background [7–10]. The exponential expansion stretches these primordial GWs to long wavelengths and low frequencies, with their amplitude directly linked to the inflationary energy scale. Consequently, detecting primordial GWs could offer critical information of inflation; see Refs. [6, 11] for recent reviews for inflationary primordial GWs.

After inflation ends, the vacuum energy of the inflaton field must be transferred to radiations, ultimately forming a thermal bath of Standard Model (SM) particles. This process, known as reheating [12–15] (see also Refs. [16–19] for reviews), may also play an important role in the production of GWs. During reheating, the inflaton oscillates around the minimum of its potential, producing particles coupled to it, including gravitons. Once produced, these gravitons propagate freely through the early universe and, after undergoing cosmological red-shift, eventually form a stochastic background of GWs in the present universe. This is analogous to how photons from the hot Big Bang eventually gave rise to the cosmic microwave background (CMB).

Several specific mechanisms for graviton production during reheating have been investigated in the literature. These include (*i*) pair production of gravitons via inflaton annihilation [20–24], (*ii*) graviton bremsstrahlung via inflaton decay [25–36], (*iii*) inflaton scattering with the daughter particles produced from inflaton decay [37, 38], and (*iv*) scattering of thermal species [39–42]. After reheating, the universe enters a radiation-dominated era, likely composed primarily of the SM plasma, in which the thermal production of gravitons has been computed in Refs. [39–41].

The primary objective of this study is to provide a systematic and consistent analysis of the GW production via the aforementioned mechanisms through the period from inflation to reheating, offering a comprehensive treatment that is currently lacking in the literature. We consider a rather generic framework spanning over vacuum-energy-dominated (VD), matter-dominated (MD), and radiation-dominated (RD) epochs, and analyze the dominant GW production channels in each epoch. For each production channel, we derive a simple yet accurate analytical expression for the GW spectrum. Compared to existing calculations in the literature, a large part of our analytical results are new or contain significant improvements. For instance, by solving the Boltzmann equation of the graviton phase space distribution, we are able to significantly improve the analytical description of the GW energy

spectra generated by (ii) and (iii), which in the previous studies [27, 37] were calculated by neglecting terms that can be potentially important at high frequencies. While previous analytical results for (ii) and (iii) are only valid in the low-frequency regime, our analytical results are accurate over the entire range. Moreover, for (iv), we demonstrate that the production of gravitons from the SM thermal plasma can be well approximated by a toy model with only three parameters, with the major characteristics such as the logarithmic dependence on the graviton energy arising from t-channel scattering fully included. This not only justifies our model-agnostic approach but also offers a simplified yet effective method for calculating graviton production in a Beyond-the-Standard-Model (BSM) thermal bath.

The remainder of this article is organized as follows. In Sec. 2, we introduce the framework and sketch out GW production mechanisms considered in this framework. In Sec. 3, we calculate various background quantities (e.g., the temperature, the inflaton and radiation energy densities) which will be used in the subsequent GW calculations. A comprehensive analysis of GW production is presented in Sec. 4. Our main results are discussed in Sec. 5. Finally, we conclude with a summary of our findings in Sec. 6 and delegate various detailed calculations to appendices.

2 The framework

In this work, we consider a quite general framework that contains an inflaton field (ϕ) and generic radiation species (R), both minimally coupled to gravity via the action

$$S = \int d^4x\sqrt{-g}\left[\frac{M_P^2}{2}\mathcal{R} + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi) + \mathcal{L}_R\right] . \tag{2.1}$$

Here, g is the determinant of the metric $g_{\mu\nu}$ tensor; $M_P \equiv 1/\sqrt{8\pi G_N} \simeq 2.4\times10^{18}$ GeV is the reduced Planck mass; $\mathcal{R}$ denotes the Ricci scalar; $V(\phi)$ is the potential of the inflaton; and $\mathcal{L}_R$ denotes the Lagrangian of particles that constitute the thermal plasma of the radiation-dominated universe. It can be the Lagrangian of the Standard Model (SM) or beyond, such as grand unified theories (GUTs) or supersymmetric theories. In order to let the inflaton eventually release its energy into radiation (i.e., to allow for the reheating process), ϕ should be coupled to at least one of the particles in the R sector. We include such coupling terms into $\mathcal{L}_R$. The model-dependent content of $\mathcal{L}_R$ will be discussed later.

Eq. (2.1) gives rise to all gravitational interactions responsible for GW production in this work. More specifically, we expand the metric $g_{\mu\nu}$ around the Minkowski metric $\eta_{\mu\nu} = (+,-,-,-)$ as follows [43, 44]:

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{2}{M_P}\,h_{\mu\nu} . \tag{2.2}$$

In linearized classical General Relativity, the second term $\frac{2}{M_P}\,h_{\mu\nu}$ is regarded as the metric perturbation, reflecting its geometric interpretation. We consider an energy

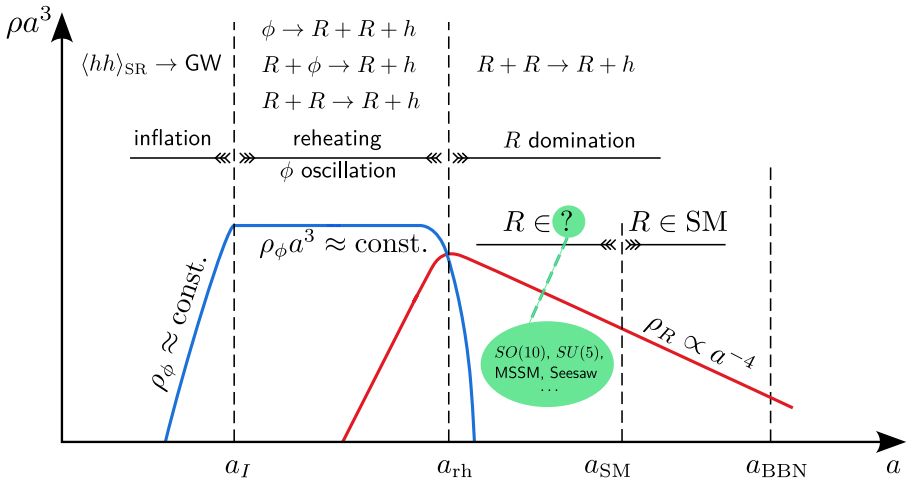

Figure 1. A schematic representation of the three phases from inflation, through reheating, and eventually to radiation (R) domination. Each phase is dominated by a different cosmological ingredient (the scalar potential energy, matter, and radiation) and features GW production with distinct spectral shapes.

scale well below the Planck scale, where the Einstein-Hilbert term can be treated within the framework of quantum effective field theory [44]. In this context, $h_{\mu\nu}$ can be interpreted as a dynamical field—namely, the graviton—propagating in a Minkowski background, similar to other quantum fields. Given $g_{\mu\nu}$, one can further determine the contravariant metric $g^{\mu\nu}$, which appears in the Lagrangian densities. Substituting the metric expansion into the action leads to the effective interaction between the graviton and the energy-momentum tensor [43, 44]:

$$\sqrt{-g}\mathcal{L} \supset \frac{1}{M_P} h_{\mu\nu} \sum_i T_i^{\mu\nu} \,, \tag{2.3}$$

where $T_i^{\mu\nu}$ represents the energy-momentum tensor of a given particle species i.

During different phases of the very early universe, different cosmological ingredients play the dominant role in Eq. (2.3), accounting for GW production at different frequency bands with distinct spectral shapes. This work aims at a full-spectrum analysis of the GW production from inflation, through reheating, and eventually to radiation domination.

Figure 1 illustrates the scope of our framework, where a is the scale factor, and ρ denotes the energy density. We start from the inflationary phase, in which the inflaton ϕ undergoes slow-roll (SR) and GWs are produced from the quantum fluctuations of $h_{\mu\nu}$ during SR, as indicated by the label "$\langle hh \rangle_{\rm SR} \to$ GW" in Fig. 1. Although this is a well-studied subject of cosmological inflation, we will revisit the SR production of GWs in this work for completeness and extend the analysis to include horizon crossing during reheating.

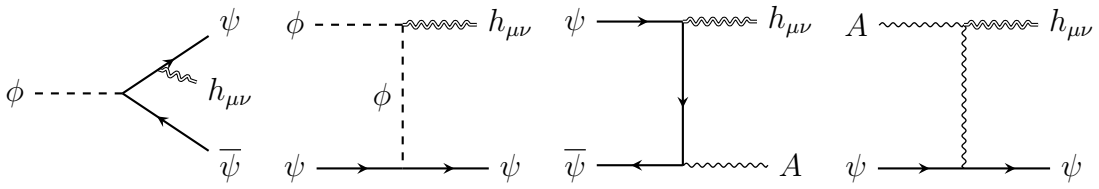

Figure 2. Representative Feynman diagrams for graviton production to the order of $1/M_P$. For a more complete set, see Fig. 6 in Appendix B.

After SR, one of the most likely scenarios is that the inflaton oscillates at the bottom of the potential and it is a natural assumption that the bottom of the potential is quadratic:

$$V(\phi) \approx \frac{1}{2} m_\phi^2 \, \phi^2 \,, \quad \text{(for small } \phi) \,. \tag{2.4}$$

This is the case for various inflationary models, including Starobinsky inflation [1], certain classes of α-attractor T-models [45, 46], and both small- and large-field polynomial inflation scenarios [47–51]. In this phase, the universe is dominated by an oscillating ϕ field, which is physically equivalent to a condensate of ϕ particles at rest with mass m_ϕ. These ϕ particles should eventually be converted to radiation in order to accommodate the success of Big Bang Nucleosynthesis (BBN). We assume that such conversion is accomplished by ϕ directly decaying to some generic species R. In the presence of a decay channel such as $\phi \to R + R$, gravitons (h) can be produced from bremsstrahlung of the process ($\phi \to R + R + h$) [27]. Moreover, as the yield of R increases, h can also be produced via $R + \phi \to h + R$ and $R + R \to h + R$. A few of representative Feynman diagrams of these processes are illustrated in Fig. 2, where we consider $R \in \{\psi,\ A\}$ with ψ and A a generic fermion and a generic gauge boson.

After ϕ decay, the universe becomes radiation-dominated. In this phase, h is produced mainly via scattering of particles in the thermal plasma. In the SM plasma, the thermal production of h has been comprehensively calculated in Ref. [39, 40] (see also [41, 52] for related phenomenological studies). In Ref. [39, 40], scattering between any two SM species, including fermions, gauge bosons, and the Higgs boson, has been taken into account. Given a variety of potentially more fundamental theories beyond the SM at very high energy scales (e.g., $SO(10)$, $SU(5)$, MSSM, Seesaw), we adopt a model-agnostic approach to study the thermal production of gravitons during the R-domination phase. In order to simply the analysis, we assume that the plasma mainly consists of fermions (ψ) and gauge bosons (A), with their gauge interactions ($\bar{\psi} A_\mu \gamma^\mu \psi$) responsible for maintaining thermal equilibrium and the thermal production of h. Such a toy model, albeit missing the contribution of scalar particles (e.g., the Higgs), is capable to capture the major characteristics of graviton production in the SM plasma. Indeed, as we will show, the thermal production rate calculated in

this toy model agrees well with the results obtained in the complete SM calculation, up to an overall normalization factor that can be accounted for by the multiplicity of ψ and A. For more unified gauge theories such as $SO(10)$ with only one universal gauge coupling for all fermions, we expect that the toy model may offer a good and simple approximation to the full calculation.

3 The post-inflationary background evolution

In this section, we briefly discuss the post-inflationary evolution of the universe and revisit formulae of background quantities including the energy densities of inflaton (ρ_ϕ) and radiation (ρ_R), the temperature T, and other relevant variables.

3.1 From a_I to $a_{\rm rh}$

In many inflationary models, it is quite common that after the slow-roll phase the inflaton field rolls toward the minimum of its potential and undergoes oscillations. The oscillating ϕ field with the quadratic potential in Eq. (2.4) can be treated as a condensate of ϕ particles at rest. Hence its energy density, defined as

$$\rho_\phi(\phi) \equiv \frac{\dot{\phi}^2}{2} + V(\phi), \tag{3.1}$$

scales as a^{-3} and behaves like matter. Given a certain decay rate of ϕ to radiation, the evolution of ρ_ϕ and ρ_R is governed by the following Boltzmann equations:

$$\frac{d\rho_\phi}{dt} + 3H\rho_\phi = -\Gamma_\phi\,\rho_\phi\,, \tag{3.2}$$

$$\frac{d\rho_R}{dt} + 4H\rho_R = +\Gamma_\phi\,\rho_\phi\,, \tag{3.3}$$

where Γ_ϕ denotes the energy conversion rate of ρ_ϕ to ρ_R. If the thermal effect can be neglected, Γ_ϕ is identical to the decay rate of ϕ in vacuum and depends only on the relevant Lagrangian parameters. If the thermal effect is significant, Γ_ϕ can also be temperature-dependent—see e.g. Eq. (5) in Ref. [53]. Here we only view Γ_ϕ as an effective parameter to quantify the energy conversion rate of ρ_ϕ to ρ_R. More details regarding Γ_ϕ will be discussed in the next section. The Hubble parameter is defined as $H \equiv \dot{a}/a$, which is determined by the Friedmann equation:

$$H^2 = \frac{\rho_\phi + \rho_R}{3M_P^2}\,. \tag{3.4}$$

When the universe is dominated by the condensate of ϕ particles (i.e., dominated by matter), the Hubble parameter scales as

$$H \approx H_I \left(\frac{a_I}{a}\right)^{3/2}\,, \tag{3.5}$$

where the subscript "I" denotes the moment at the end of inflation. When using Eq. (3.5), the underlying assumption is that

$$\Gamma_\phi \ll H_I \,, \tag{3.6}$$

such that there is a significant matter-dominated period. We refer to this as the slow-decaying regime and will use it extensively in our analytical calculations below.

Note that H_I can be determined by inflationary observables:

$$\frac{H_I}{M_P} = 1.0 \times 10^{-5} \left(\frac{\mathcal{P}_\mathcal{R}}{2.1 \times 10^{-9}}\right)^{1/2} \left(\frac{r}{0.01}\right)^{1/2} , \tag{3.7}$$

where r is the tensor-to-scalar ratio and $\mathcal{P}_\mathcal{R}$ represents the amplitude of the scalar power spectrum. In this work, we remain agnostic about the specific inflationary model and use experimental inputs to set H_I. Adopting the central value $\mathcal{P}_\mathcal{R} = 2.1 \times 10^{-9}$ from Planck 2018 [54] and impose the latest constraint $r < 0.035$ from BICEP/Keck 2018 [55], we obtain

$$H_I < 1.9 \times 10^{-5} M_P \,. \tag{3.8}$$

Substituting $H = \dot{a}/a$ into Eq. (3.5), one can solve it as a differential equation of $a(t)$ and obtain

$$t \approx \frac{2}{3H_I} \left[\left(\frac{a}{a_I}\right)^{3/2} - 1\right] , \tag{3.9}$$

$$a \approx a_I \left(1 + \frac{3}{2} H_I t\right)^{2/3} . \tag{3.10}$$

Here we have set $t = 0$ at the moment "I".

Using $\frac{d\rho_\phi}{dt} + 3H\rho_\phi = \frac{d\left(\rho_\phi a^3\right)}{a^3 dt}$, we can write Eq. (3.2) as $\dot{Y} = -\Gamma_\phi Y$ with $Y \equiv \rho_\phi a^3$, which implies

$$\rho_\phi(a) = \rho_\phi(a_I) \left(\frac{a_I}{a}\right)^3 e^{-\Gamma_\phi t} \tag{3.11}$$

$$\approx 3H_I^2 M_P^2 \left(\frac{a_I}{a}\right)^3 , \quad (\text{for } t \lesssim \Gamma_\phi^{-1}) \,. \tag{3.12}$$

Despite that the a-t relation in Eq. (3.9) is approximate, Eq. (3.11) is exact. Its physical meaning is clear: the energy of ϕ in a comoving volume (a^3) decreases exponentially because the number of ϕ particles in this volume decays exponentially. In Eq. (3.12), we have approximated $e^{-\Gamma_\phi t} \approx 1$. This is valid for $t \lesssim t_{\rm rh}$ where $t_{\rm rh}$ is defined as

$$t_{\rm rh} \equiv \Gamma_\phi^{-1} \,. \tag{3.13}$$

Correspondingly, the scale factor at this point is

$$a_{\rm rh} \approx a_I \left(1 + \frac{3}{2} H_I t_{\rm rh}\right)^{2/3} \approx a_I \left(\frac{3H_I}{2\Gamma_\phi}\right)^{\frac{2}{3}} , \tag{3.14}$$

where in the second step, we have assumed $\frac{3H_I}{2\Gamma_\phi} \gg 1$.

Using Eq. (3.12), the solution to Eq. (3.3) for the radiation energy density is approximately given by

$$\rho_R(a) \approx \frac{6}{5} M_P^2 \Gamma_\phi H_I \left(\frac{a_I}{a}\right)^{3/2} \left[1 - \left(\frac{a_I}{a}\right)^{5/2}\right], \quad (\text{for } t \lesssim t_{\rm rh})\,. \tag{3.15}$$

We assume that the radiation, which is likely to contain multiple species as already discussed in the previous section, always maintains thermal equilibrium among the multiple species. Therefore, it has a well-defined temperature T, determined by

$$\rho_R(T) \equiv \frac{g_\star\, \pi^2}{30} T^4\,, \tag{3.16}$$

where $g_\star$ denotes the number of relativistic degrees of freedom in the thermal bath.

Here we shall clarify potential ambiguities of the reheating temperature $T_{\rm rh}$. If we define it as the temperature of the thermal bath at $t = t_{\rm rh}$, then by substituting Eq. (3.14) into Eq. (3.15) and using Eq. (3.16), we get

$$T_{\rm rh} \approx \frac{2}{\sqrt{\pi}} \left(\frac{3}{2g_\star}\right)^{1/4} \sqrt{M_P \Gamma_\phi}\,. \tag{3.17}$$

Alternatively, one can also extrapolate the radiation domination backwards until ρ_R reaches $3M_P^2 H_{\rm rh}^2$ with $H_{\rm rh} \equiv H(t_{\rm rh}) \approx \frac{2}{3t_{\rm rh}}$, and define $T_{\rm rh}$ as the temperature at this point. This gives

$$T_{\rm rh} \approx \sqrt{\frac{2}{\pi}} \left(\frac{10}{g_\star}\right)^{1/4} \sqrt{M_P \Gamma_\phi}\,. \tag{3.18}$$

The two reheating temperatures only differ by a factor of 0.88, implying that power-law extrapolation from either side is a good approximation. In this work, we choose Eq. (3.18) as the definition of $T_{\rm rh}$.

According to Eq. (3.15), the radiation energy density $\rho_R(a)$ reaches its maximum at $a = a_{\rm max} = (8/3)^{2/5} a_I$, corresponding to a maximum temperature

$$T_{\rm max}^4 = \frac{60}{\pi^2 g_\star} \left(\frac{3}{8}\right)^{8/5} M_P^2 \Gamma_\phi H_I\,. \tag{3.19}$$

Comparing it to $T_{\rm rh}$ in Eq. (3.18), we obtain

$$\frac{T_{\rm max}}{T_{\rm rh}} \approx 0.75 \left(\frac{H_I}{\Gamma_\phi}\right)^{1/4}. \tag{3.20}$$

Since we are mainly concerned with the slow-decaying regime [see Eq. (3.6)], Eq. (3.20) implies that $T_{\rm max}$ in general exceeds $T_{\rm rh}$ [56].

3.2 After $a_{\rm rh}$

After inflatons decay to radiations, the universe becomes radiation-dominated, corresponding to the epoch of $a \gtrsim a_{\rm rh}$. In this epoch, inflatons due to their exponentially suppressed abundance, play a negligible role in the production of GWs or gravitons. Instead, radiation-radiation scattering becomes the dominant production channel, which has been investigated as the thermal production of gravitons in the literature [39–42, 57–62]. The thermal production crucially relies on the temperature, with the production rate decreasing rapidly as the universe expands and cools down. Consequently, one can assume that it ceases producing gravitons effectively after the universe expands by a significantly large factor.

Nevertheless, the subsequent evolution still has an important effect on the GW spectrum caused by $g_\star$ which decreases from $\mathcal{O}(100)$ or higher in the early universe to a value of a few in the present universe. The decreasing $g_\star$ causes the temperature of the thermal bath to be higher than it would be if $g_\star$ were a constant, thereby increasing the ratio of the mean GW frequancy to the CMB frequency. To take this into account, we use entropy conservation to obtain the relation of T with a:

$$T = T_0 \frac{a_0}{a} \left(\frac{g_{\star s,0}}{g_{\star s}} \right)^{1/3} , \tag{3.21}$$

where $g_{\star s}$ denotes the number effective degrees of freedom in entropy, and the subscript "0" denotes values in the present universe. Note that Eq. (3.21) remains valid after neutrino decoupling, provided that T is interpreted as the temperature of photons, not neutrinos. It is also valid in the matter-dominated era with $a > a_{\rm eq}$ where $a_{\rm eq} \approx 1/3400$ is the scale factor at matter-radiation equality, since matter and dark energy do not contribute to the total entropy.

In this work, we take the following present-day values of relevant quantities:

$$a_0 = 1\,, \quad T_0 = 2.73 \text{ K}\,, \qquad g_{\star s,0} \approx 2 + \frac{7}{8} \times \frac{4}{11} \times 6 \approx 3.9\,, \tag{3.22}$$

$$H_0 = 100h \text{ kms}^{-1}\text{Mpc}^{-1} \quad \text{with} \quad h \approx 0.67\,. \tag{3.23}$$

The critical energy density is defined as $\rho_c = 3H_0^2 M_P^2 \simeq 1.05 \times 10^{-5}\ h^2 \text{ GeV/cm}^3$ [63].

4 Gravitational wave production

With the background evolution determined, we are ready to calculate the production of GWs in our framework covering both the inflationary and post-inflationary epochs.

4.1 Inflationary GW

During inflation, GWs can be generated by metric perturbations ($h_{\mu\nu}$) sourced by quantum fluctuations of the inflaton, corresponding to the well-known primordial

tensor power spectrum, $\mathcal{P}_t$. The calculation of the tensor power spectrum in the slow-roll inflationary paradigm has been well established (see, e.g., [11, 64]) and the result is typically formulated as

$$\mathcal{P}_t = r\,\mathcal{P}_{\mathcal{R}}(k_\star)\left(\frac{k}{k_\star}\right)^{n_t}\,, \tag{4.1}$$

where r denotes the tensor-to-scalar ratio, n_t is the tensor spectral index, and $\mathcal{P}_{\mathcal{R}}(k_\star)$ is the amplitude of the scalar power spectrum at the pivot scale $k_\star$. Using the Planck measurement [65], we take $k_\star = 0.05\ \mathrm{Mpc}^{-1}$ and $\mathcal{P}_{\mathcal{R}}(k_\star) \simeq 2.1\times 10^{-9}$ in this work. In single-field slow-roll inflation, the spectral index satisfies $n_t \simeq -r/8$. Given the current constraint $r < 0.036$ from BICEP/Keck [55], the dependence of $\mathcal{P}_t$ on k is rather weak and therefore is neglected in our work, as we aim to minimize the model dependence of our full-spectrum analysis.

The corresponding GW spectrum at the present time is related to the primordial tensor power spectrum $\mathcal{P}_t$ by a gravitational transfer function—see e.g. [64, 66]. The transfer function encapsulates the expansion history from the horizon crossing of a given mode k at $a = a_{\rm hc}$ to a later time $a > a_{\rm hc}$, where $a_{\rm hc}$ is the scale factor at horizon crossing, determined by

$$a_{\rm hc} H(a_{\rm hc}) = k\,. \tag{4.2}$$

Different k modes reenter the horizon at different epochs, either during radiation domination or reheating. For both epochs, it is straightforward to solve Eq. (4.2) and obtain (see Appendix. A for details):

$$a_{\rm hc} = \begin{cases} \sqrt{\frac{\pi^2 g_{\star,\rm hc}}{90}}\left(\frac{g_{\star s,0}}{g_{\star s,\rm hc}}\right)^{2/3}\frac{T_0^2}{kM_P} & (k_{\rm eq} < k < k_{\rm rh}) \\ \frac{\pi^2 g_{\star,\rm rh}}{90}\left(\frac{g_{\star s,0}}{g_{\star s,\rm rh}}\right)\frac{T_0^3 T_{\rm rh}}{k^2 M_P^2} & (k_{\rm rh} < k < k_I) \end{cases}\,, \tag{4.3}$$

where the subscripts "hc", "rh", and "0" in $g_\star$ or $g_{\star s}$ indicate that one should evaluate $g_\star$ or $g_{\star s}$ at $a = a_{\rm hc}$, $a_{\rm rh}$, and a_0, respectively. The first and second cases in Eq. (4.3) correspond to k reentering the horizon during radiation domination and reheating, respectively. Here $k_{\rm rh}$, $k_{\rm eq}$, and k_I are defined as $k_X \equiv a_X H(a_X)$ with $X \in \{\mathrm{rh},\ \mathrm{eq},\ I\}$. By equating the two cases with each other, we can obtain $k_{\rm rh}$ and the corresponding frequency $f_{\rm rh} = k_{\rm rh}/(2\pi)$, which reads

$$f_{\rm rh} = \frac{1}{6}\sqrt{\frac{g_{\star,\rm rh}}{10}}\left(\frac{g_{\star s,0}}{g_{\star s,\rm rh}}\right)^{\frac{1}{3}}\frac{T_0 T_{\rm rh}}{M_P} \simeq 3\times 10^4\left(\frac{T_{\rm rh}}{10^{12}\ \mathrm{GeV}}\right)\ \mathrm{Hz}\,. \tag{4.4}$$

Similarly, for $k_I = a_I H_I$, we find

$$\begin{aligned} f_I \equiv \frac{a_I H_I}{2\,\pi} &= \frac{T_0}{2}\left(\frac{g_{\star,\rm rh}}{90\pi}\,\frac{g_{\star s,0}}{g_{\star s,\rm rh}}\,\frac{H_I T_{\rm rh}}{M_P^2}\right)^{1/3} \\ &\simeq 7\times 10^6\left(\frac{T_{\rm rh}}{10^{12}\ \mathrm{GeV}}\right)^{1/3}\left(\frac{H_I}{10^{13}\ \mathrm{GeV}}\right)^{1/3}\mathrm{Hz}\,. \end{aligned} \tag{4.5}$$

We note that GWs with frequencies $f > f_I$, corresponding to extremely short wavelengths, have never exited the horizon. However, this does not imply that such modes cannot be excited. As we will discuss in the next section, GWs with $f > f_I$ can be generated through inflaton-inflaton annihilation, $\phi\,\phi \to h\,h$. This shall be understood as the high frequency tail of the primordial gravitational waves [22, 67, 68].

By combining the primordial tensor power spectrum and the transfer function, one can obtain the present-day GW spectrum sourced by quantum fluctuations of the inflaton [11, 64, 69, 70]:

$$\Omega_{\rm GW}(k) = \frac{1}{12}\left(\frac{k}{a_0 H_0}\right)^2 \mathcal{P}_t\, \mathcal{T}(a_0, k)\,, \tag{4.6}$$

where $\mathcal{T}(a,k)$ is the transfer function accounting for the evolution of tensor modes at later times. In the WKB approximation, it is given by [64, 66]:

$$\mathcal{T}(a,k) = \frac{1}{2}\left(\frac{a_{\rm hc}}{a}\right)^2, \tag{4.7}$$

where $a_{\rm hc}$ depends on k, with the explicit form given in Eq. (4.3).

Substituting Eq. (4.7) into Eq. (4.6) and using Eq. (4.3), we get

$$\Omega_{\rm GW}(f) = \Omega_\gamma^0 \frac{\mathcal{P}_t}{24}\frac{g_{\star,\rm hc}}{g_\gamma} \times \begin{cases} \left(\frac{g_{\star s,0}}{g_{\star s,\rm hc}}\right)^{\frac{4}{3}} & (f_{\rm eq} < f < f_{\rm rh}) \\ \frac{g_{\star,\rm rh}}{g_{\star,\rm hc}}\left(\frac{g_{\star s,0}}{g_{\star s,\rm rh}}\right)^{\frac{4}{3}}\left(\frac{f}{f_{\rm rh}}\right)^{-2} & (f_{\rm rh} < f < f_I) \end{cases}. \tag{4.8}$$

Here, $\Omega_\gamma^0 \equiv \rho_\gamma^0/\rho_c = 2.47\times 10^{-5}\,h^{-2}$ [54] is the present-day photon energy density fraction, and $g_\gamma = 2$ accounts for the two degrees of freedom of the photon.

4.2 Post-inflationary GW

After inflation, the universe is filled with cold ϕ particles together with subsequently generated thermal radiations. As previously illustrated in Fig. 2, gravitons can be produced via various particle decay or scattering processes in the post-inflationary epoch. These gravitons would constitute a high-frequency GW spectrum, with the typical frequency significantly higher than f_I in Eq. (4.5).

Since this part is produced via particle decay or scattering processes, we adopt the Boltzmann equation to calculate the spectrum. Denoting the phase-space distribution function of gravitons by $f_h(t,\ k)$ where k is the graviton momentum, the Boltzmann equation for f_h reads:

$$\left[\frac{\partial}{\partial t} - H(a)\,k\,\frac{\partial}{\partial k}\right] f_h(t,\ k) = \mathcal{C}_h(a,\ k)\,, \tag{4.9}$$

where $\mathcal{C}_h$ is the collision term for graviton production. In the freeze-in regime[1], one can assume that it is a function of a and k, independent of f_h. Under this assumption, Eq. (4.9) can be solved by (see, e.g., [38, 71, 72] for derivations)

$$f_h = \int_{a_I}^{a} \frac{da'}{a' H(a')} \mathcal{C}_h \left(a', \frac{ak}{a'} \right) . \tag{4.10}$$

We note that Eq. (4.10) takes into account graviton production throughout the reheating phase and the subsequent radiation-dominated phase with $a > a_{\rm rh}$. For channels involving the inflaton, the contribution after reheating becomes exponentially suppressed. For channels that mainly produces gravitons with $\omega \sim T$, the low- and high-frequency parts of the resulting GW spectrum are typically dominated by earlier and later epochs, respectively. This is because T approximately scales as $a^{-3/8}$ during reheating, while the frequency of a graviton once produced scales as a^{-1}. Since the latter decreases faster than the former, later epochs during reheating generally produce gravitons of higher frequencies than earlier epochs.

Given the phase space distribution function f_h, we define the number density n_h, the energy density ρ_h, and its differential energy distribution $\frac{d\rho_h}{dk}$ as follows:

$$n_h = g_h \int \frac{4\pi\, k^2\, dk}{(2\pi)^3} f_h \,, \tag{4.11}$$

$$\rho_h = g_h \int \frac{4\pi\, k^3\, dk}{(2\pi)^3} f_h \,, \tag{4.12}$$

$$\frac{d\rho_h}{dk} \equiv g_h \frac{4\pi\, k^3}{(2\pi)^3} f_h \,, \tag{4.13}$$

where $g_h = 2$ is the number of graviton degrees of freedom. The GW spectrum at present, as a function of the frequency $f = k/(2\pi)$, is then given by

$$\Omega_{\rm GW}(f) \equiv \frac{1}{\rho_c} \frac{d\rho_h}{d\ln k} = 8\pi^2 g_h \frac{f^4}{\rho_c} f_h \,, \tag{4.14}$$

where f_h should be evaluated at the present time.

Finally, we shall briefly introduce the collision term $\mathcal{C}_h$ in the Boltzmann equation. For a general scattering process, $X_1 + X_2 + \cdots + X_n \to X_{n+1} + X_{n+2} + \cdots + X_{n+m} + h$, neglecting quantum statistics factors (e.g., the Pauli blocking factor), the collision term for the production of h is given by the phase space integral

$$\mathcal{C}_h = \frac{g_{n+m}}{2\omega} \int \left(\prod_{i=1}^{n+m} d\Pi_i \right) f_1 \cdots f_n \, \mathcal{S} |\overline{\mathcal{M}}|^2 \, (2\pi)^4 \delta^{(4)} \left(\sum_{i=1}^{n} p_i - \sum_{j=n+1}^{n+m} p_j \right) , \tag{4.15}$$

[1]Due to the weakness of graviton interactions, the yield of gravitons produced via particle processes considered in this work can never reach equilibrium ($f_h \ll 1$), implying that backreactions of these processes can be safely neglected.

with

$$d\Pi_i \equiv \frac{d^3\mathbf{p}_i}{(2\pi)^3 2E_i}\,. \tag{4.16}$$

Here, ω is the energy of the graviton, p_i and E_i denote the momentum and energy of the i^{th} particle, f_i is its phase-space distribution, $|\overline{\mathcal{M}}|^2$ is the squared amplitude of this process (with an average taken over the spins and polarizations of all initial and final state particles), $\mathcal{S}$ is the symmetry factor, and the delta function $(2\pi)^4\delta^{(4)}(\dots)$ ensures energy-momentum conservation. Note that we have factored out the multiplicity factors of all particles (except for h) from the integral in order to make the dependence on these factors explicit and define

$$g_{n+m} \equiv \prod_{i=1}^{n+m} g_i\,, \tag{4.17}$$

where g_i is the multiplicity factor of the i^{th} particle. The multiplicity factor of the graviton, $g_h = 2$, is not included in Eq. (4.17) because f_h in Eq. (4.9) does not include g_h—see also Eq. (4.12) in which g_h is eventually included.

Below we start our analyses for specific processes, using the above prescription to calculate their collision terms and the corresponding f_h and Ω_{GW}.

4.2.1 GWs from inflaton annihilation

Well before the inflaton starts decaying, the post-inflationary universe is exclusively filled with non-relativistic ϕ particles. In this epoch, gravitons are produced via the inflaton annihilation channel: $\phi\,\phi \to h\,h$. This production channel has been previously studied in Refs. [23, 38] and the squared matrix element reads (see Appendix B for the detailed calculation):

$$|\overline{\mathcal{M}}_{\phi\phi\to hh}|^2 = \frac{2\,m_\phi^4}{M_P^4} \times \frac{1}{4}\,, \tag{4.18}$$

where the factor of 1/4 arises from taking the average of the polarizations of the two gravitons[2]. Substituting Eq. (4.18) into Eq. (4.15) and performing the phase space integral (see Appendix C for details), we obtain

$$\mathcal{C}_h = \frac{\pi}{32}\,\frac{n_\phi^2}{M_P^4}\,\delta(\omega - m_\phi)\,. \tag{4.19}$$

Here n_ϕ is the number density of ϕ and we have taken into account the symmetry factor as well as double graviton production. The $\delta(\omega - m_\phi)$ function indicates that

[2]Note that in Eq. (10) of Ref. [23], the squared matrix element is also multiplied by a factor of 1/4. But that one is a symmetry factor, corresponding to our $\mathcal{S}$ in Eq. (4.15). Our factor of 1/4, arising from polarization averaging, will eventually be canceled out by two g_h's when $|\overline{\mathcal{M}}_{\phi\phi\to hh}|^2$ is used to compute the production rate for ρ_h or n_h—see Eq. (4.22) below, which agrees with Eq. (11) of Ref. [23].

the produced graviton is monochromatic. Due to the monochromatic spectrum, the Boltzmann equations of ρ_h and n_h share the same production rate:

$$\dot{\rho}_h + 4H\rho_h = \Gamma_h \rho_\phi \,, \tag{4.20}$$

$$\dot{n}_h + 3Hn_h = \Gamma_h n_\phi \,, \tag{4.21}$$

with

$$\Gamma_h \equiv \frac{g_h}{n_\phi} \int \mathcal{C}_h \frac{d^3k}{(2\pi)^3} = \frac{g_h n_\phi m_\phi^2}{64\pi M_P^4} \,. \tag{4.22}$$

At the end of reheating, the graviton phase-space distribution obtained by calculating the integral in Eq. (4.10) is given by:

$$f_h(a_{\rm rh}, k) \simeq \frac{9\pi}{32} \left(\frac{H_I}{m_\phi}\right)^3 \left(\frac{a_I}{a_{\rm rh}} \frac{m_\phi}{k}\right)^{9/2} \tag{4.23}$$

where we have used $n_\phi \simeq n_{\phi I} a_I^3/a^3 e^{-\Gamma_\phi t}$ during reheating, with $n_{\phi I}$ the initial value of n_ϕ.

Assuming that this channel effectively stops producing h when $a > a_{\rm rh}$, the present-day distribution $f_h(a_0, k)$ is related to $f_h(a_{\rm rh}, k)$ by simple red-shifting:

$$f_h(a_0, k) = f_h(a_{\rm rh}, k a_0/a_{\rm rh}) \,. \tag{4.24}$$

Then using Eq. (4.14), we find

$$\Omega_{\rm GW} h^2(f) \simeq 9 \cdot 10^{-22} \left(\frac{m_\phi}{10^{13}\,{\rm GeV}} \cdot \frac{T_{\rm rh}}{10^{13}\,{\rm GeV}}\right)^{\frac{3}{2}} \left(\frac{\rm GHz}{f}\right)^{\frac{1}{2}} . \tag{4.25}$$

We remind the reader that the graviton energy at production is the same as the inflaton mass, i.e., $\omega(a_p) = m_\phi$, with a_p the scale factor at production. This implies that the graviton energy at the end of reheating lies in the range $m_\phi a_I/a_{\rm rh} \leq \omega \leq m_\phi$. After taking the red-shift from $a_{\rm rh}$ to a_0 into account, the GW energy at present should be red-shifted by a factor is $a_{\rm rh}/a_0$, leading to the following frequency band:

$$f_{\rm min} \leq f \lesssim f_{\rm max} \,, \tag{4.26}$$

where

$$f_{\rm min} = \frac{m_\phi}{2\pi} \frac{a_I}{a_0} \simeq 7 \cdot 10^6 \left(\frac{m_\phi}{10^{13}\ {\rm GeV}}\right) \left(\frac{T_{\rm rh}}{10^{12}\ {\rm GeV}}\right)^{1/3} \left(\frac{10^{13}\ {\rm GeV}}{H_I}\right)^{2/3} {\rm Hz} \,, \tag{4.27}$$

$$f_{\rm max} = \frac{m_\phi}{2\pi} \frac{a_{\rm rh}}{a_0} \simeq 2 \cdot 10^{11} \left(\frac{m_\phi}{10^{13}\ {\rm GeV}}\right) \left(\frac{10^{12}\ {\rm GeV}}{T_{\rm rh}}\right) {\rm Hz} \,. \tag{4.28}$$

Note that gravitons with frequencies above $f_{\rm max}$ can, in principle, be produced after reheating; however, their amplitude is exponentially suppressed due to the rapid depletion of the inflaton number density.

4.2.2 GWs from bremsstrahlung of inflaton decay

In our framework, the inflaton ϕ is unstable and has a decay rate, Γ_ϕ, which is essential to reheating. The specific form of Γ_ϕ depends on how ϕ is coupled to lighter species (radiation). Although this is model-dependent, for any given decay channel of ϕ, one can generally expect that a graviton can be emitted via bremsstrahlung of the decay channel, and the bremsstrahlung decay rate is roughly given by

$$\Gamma_{\phi\to h+\cdots} \sim \frac{1}{16\pi^2}\frac{m_\phi^2}{M_P^2}\Gamma_{\phi\to\cdots} \tag{4.29}$$

where "$\cdots$" denotes arbitrary final states of the given decay channel. The factor of $1/(16\pi^2)$ is a typical phase-space suppression factor of bremsstrahlung, and m_ϕ^2/M_P^2 accounts for the suppression of a gravitational interaction vertex.

Eq. (4.29) offers a crude estimation of the bremsstrahlung decay rate, which may deviate from the exact value by a factor of a few, depending on whether ϕ decays to scalar/vector bosons or fermions [27]. For concreteness, we focus on the scenario that ϕ dominantly decays to fermions via a Yukawa interaction:

$$\mathcal{L} \supset y\overline{\psi}\psi\phi\,, \tag{4.30}$$

where y is the Yukawa coupling and ψ is a light fermion with a negligible mass. We note that the couplings between the inflaton and the daughter fields induce effective mass terms for the latter, proportional to the inflaton field value ϕ, thereby altering the decay kinematics [73–75]. To properly account for this effect, one shall average over the inflaton oscillations, which yields an effective coupling $y_{\rm eff}$ the interaction in Eq. (4.30). In the case of the quadratic inflaton potential considered here, this effect has been shown to be moderate, with $y_{\rm eff} \simeq y$ [73, 75].

Eq. (4.30) gives rise to the dominant decay channel $\phi \to \overline{\psi}\psi$ with the decay rate

$$\Gamma_\phi = \frac{y^2}{8\pi} m_\phi\,. \tag{4.31}$$

Here, we have assumed an inflaton decay rate in vacuum. It has been shown that backreaction effects on the inflaton decay can arise if the daughter field acquires a thermal mass [53, 76–78]. Throughout our numerical analysis, we will restrict ourselves to scenarios in which the reheating temperature remains smaller than the inflaton mass. This is typically the case for a small inflaton decay rate (or equivalently, small Yukawa couplings). In such cases, assuming a vacuum inflaton decay rate and neglecting the thermal mass of the daughter field is a good assumption. Incorporating it into the Einstein-Hilbert action, Eq. (4.30) also implies that gravitons can be produced via the bremsstrahlung decay channel, $\phi \to \overline{\psi}\psi h$, for which the squared matrix element reads (see Appendix B):

$$|\overline{\mathcal{M}}_{\phi\to\bar{\psi}\psi h}|^2 = \frac{1}{8}\times\frac{y^2\,m_\phi^2}{M_P^2}\left(1-\frac{2\omega}{m_\phi}\right)\left[2-\frac{2m_\phi}{\omega}+\left(\frac{m_\phi}{\omega}\right)^2\right]. \tag{4.32}$$

Here the factor of 1/8 arises from averaging the spins and polarizations of the final states.

Note that, for fixed ω, Eq. (4.32) is independent of the energy of ψ. This can be used to greatly simplify the calculation of the corresponding collision term (see Appendix C for details). The resulting collision term is

$$\mathcal{C}_h = \frac{y^2\rho_\phi}{64\pi\omega M_P^2}\left(1-\frac{2\omega}{m_\phi}\right)\left[2-\frac{2m_\phi}{\omega}+\left(\frac{m_\phi}{\omega}\right)^2\right]\Theta\left(\frac{m_\phi}{2}-\omega\right), \tag{4.33}$$

where Θ denotes the Heaviside theta function, reflecting the kinematical constraint that the graviton produced from a three-body decay process must have $\omega < m_\phi/2$.

Solving the Boltzmann equation using Eq. (4.10), we find

$$f_h(a_{\rm rh},k) \simeq \frac{y^2\,H_I\,m_\phi^2}{32\,\pi\,k^3} r_a^{3/2} \times \begin{cases} \mathcal{J}\left(\frac{k}{m_\phi}\right) & \text{for } \frac{r_a}{2} < \frac{k}{m_\phi} < \frac{1}{2} \\ \mathcal{J}\left(\frac{k}{m_\phi}\right) - r_a^{3/2}\mathcal{J}\left(\frac{k}{r_a m_\phi}\right) & \text{for } \frac{k}{m_\phi} < \frac{r_a}{2} \end{cases}, \tag{4.34}$$

with $r_a \equiv a_I/a_{\rm rh}$ and

$$\mathcal{J}(x) \equiv 1 - 12x + 18\sqrt{2}x^{3/2} - 18x^2 + 4x^3\,. \tag{4.35}$$

The piecewise feature of Eq. (4.34) originates from the Θ function in Eq. (4.33). The $\mathcal{J}(x)$ function monotonically decreases from 1 to 0 for $x \in [0, 1/2]$.

Then using Eq. (4.24) and Eq. (4.14), we obtain the GW spectrum:

$$\Omega_{\rm GW}h^2(f) \simeq 6.2\times10^{-17}\cdot\frac{m_\phi}{10^{13}\ {\rm GeV}}\cdot\frac{f}{\rm GHz}\cdot g_{\star 106}^{-\frac{1}{4}}\left(\frac{\Gamma_\phi}{10^{-5}M_P}\right)^{\frac{1}{2}}\mathcal{J}\left(\frac{f}{f_{\rm max}}\right), \tag{4.36}$$

with $g_{\star 106} \equiv g_{\star,{\rm rh}}/106.75$, and $f_{\rm max}$ corresponding to the frequency of a graviton produced at $a = a_{\rm rh}$ with $\omega = m_\phi$ and red-shifted to $a = a_0$; see Eq. (4.28). In deriving Eq. (4.36), we have replaced y^2 by $8\pi m_\phi/\Gamma_\phi$ and expressed $T_{\rm rh}$ in terms of Γ_ϕ according to Eqs. (4.31) and (3.18). For a crude estimate, one can ignore the factor $\mathcal{J}(f/f_{\rm max})$ in Eq. (4.36) because it is typically of $\mathcal{O}(1)$. Eq. (4.36) is valid only for $f \le f_{\rm max}2$ and it vanishes at $f = f_{\rm max}/2$. However, higher frequency GWs with $f > f_{\rm max}/2$ can still be generated via the bremsstrahlung channel, due to a small amount of inflatons decaying in the radiation-dominated epoch, but this contribution is exponentially suppressed, similar to that discussed at the end of Sec. 4.2.1.

4.2.3 Radiation-catalyzed inflaton-graviton conversion

In the presence of the decay channel $\phi \to \bar{\psi}\psi h$ discussed above, the accompanying process $\phi\psi \to \psi h$ is also possible, provided that a significant amount of ψ particles have already been produced from ϕ decay. In this process, ψ can be viewed as a catalyst that facilitates the conversion of ϕ to h, with itself not consumed after the reaction. This radiation-catalyzed inflaton-graviton conversion can be very efficient,

causing an enhanced GW production rate compared to the bremsstrahlung decay under certain circumstances, as we will show later.

Note that this process shares essentially the same Feynman diagrams as $\phi \to \bar{\psi}\psi h$: pulling the $\bar{\psi}$ leg in the diagrams for the latter from final states to initial state yields exactly the diagrams for the latter. Hence the squared matrix element of $\phi\psi \to \psi h$ can be obtained from that of $\phi \to \bar{\psi}\psi h$ using crossing symmetry—see Appendix B. The result is given as follows:

$$|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 = \frac{1}{8} \times \frac{y^2\, m_\phi^2}{M_P^2} \left(\frac{2\omega}{m_\phi} - 1\right) \left[2 - \frac{2m_\phi}{\omega} + \left(\frac{m_\phi}{\omega}\right)^2\right]. \tag{4.37}$$

The subsequent calculation of $\mathcal{C}_h$ and f_h is straightforward and similar to the previous one, leading to

$$\mathcal{C}_h = \frac{|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 n_\phi}{8\pi m_\phi \omega^2} \left(\int_{\omega - m_\phi/2}^{\infty} dp_\psi f_\psi\right) \Theta(\omega - m_\phi/2) \tag{4.38}$$

$$\approx \frac{y^2 \rho_\phi}{64\pi\omega M_P^2} \left(\frac{2\omega}{m_\phi} - 1\right) \left[2 - \frac{2m_\phi}{\omega} + \left(\frac{m_\phi}{\omega}\right)^2\right] \frac{T}{\omega} e^{-\frac{\omega - m_\phi/2}{T}} \Theta\left(\omega - \frac{m_\phi}{2}\right), \tag{4.39}$$

where $f_\psi(p_\psi)$ denotes the phase space distribution function of ψ and in the last step we have used the Boltzmann approximation: $f_2 \approx e^{-p_2/T}$.

Substituting Eq. (4.39) into Eq. (4.10) cannot yield an analytically calculable integral. However, we observe that, in practice, one can take $\frac{2\omega}{m_\phi} \gg 1$ in Eq. (4.39) to obtain a simple yet accurate analytical expression for f_h:

$$f_h(a_{\rm rh}, k) \approx \frac{y^2 H_I T_{\rm rh}}{2\pi m_\phi^2} \left(1 - \frac{r_a m_\phi}{2k}\right)^3 r_a^{\frac{3}{2}} \frac{3m_\phi}{5k} \left(\frac{T_{\rm rh}}{k}\right)^{\frac{7}{5}} \Upsilon_{\frac{7}{5}}\left(\frac{k}{T_{\rm rh}}\right), \tag{4.40}$$

with

$$r_a \equiv \frac{a_I}{a_{\rm rh}} \approx \left(\frac{2\Gamma_\phi}{3H_I}\right)^{\frac{2}{3}}. \tag{4.41}$$

In deriving the above f_h, we have used $\rho_R \propto a^{-3/2}$ [see Eq. (3.15)], which implies $T \approx T_{\rm rh}(a_{\rm rh}/a)^{3/8}$ during reheating. The Υ function used in Eq. (4.40) [also used in Eq. (4.51) below] is defined as follows:

$$\Upsilon_\xi(x) \equiv \Gamma(\xi,\ x) - \Gamma\left(\xi,\ r_a^{-\frac{5}{8}} x\right), \tag{4.42}$$

where Γ is the incomplete gamma function.

Since Eq. (4.40) relies on $\frac{2\omega}{m_\phi} \gg 1$ while the thermal average of ω is typically about $3T$, we expect that Eq. (4.40) is accurate when

$$T_{\rm rh} \gg \frac{1}{6} m_\phi. \tag{4.43}$$

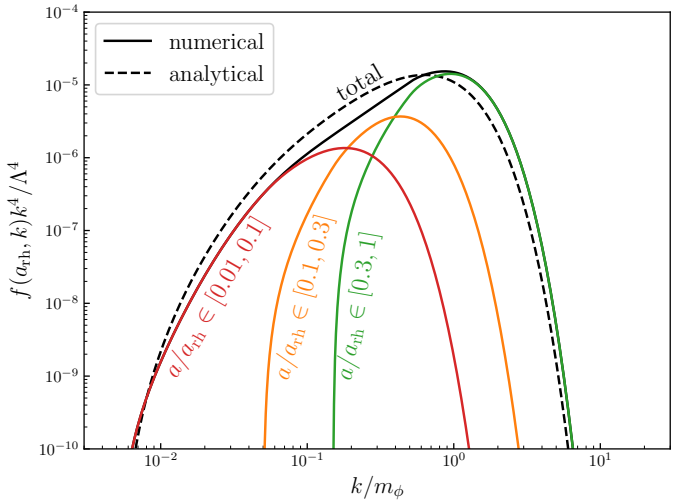

Figure 3. Comparison of Eq. (4.40) with the numerical result obtained by performing the integration in Eq. (4.10) with $\mathcal{C}_h$ in Eq. (4.39). Since $\Omega_{\rm GW} \propto k^4 f_h$, we multiply f_h by a dimensionless factor k^4/Λ^4 with $\Lambda^4 \equiv \frac{y^2}{2\pi} H_I\, T_{\rm rh} m_\phi^2$. In the shown example, we have used $r_a = 10^{-2}$ and $T_{\rm rh}/m_\phi = 1/3$.

Figure 3 shows a comparison of Eq. (4.40) with the numerical result and demonstrates modest accuracy of our analytical calculation. In the shown example, $r_a = 10^{-2}$, implying that the universe expands by two orders of magnitude from the end of inflation to the end of reheating. In order to gain a better understanding of when the spectrum receives the dominant contribution, we decompose it into contributions from three different epochs: $a/a_{\rm rh} \in [0.01, 0.1]$, $a/a_{\rm rh} \in [0.1, 0.3]$, and $a/a_{\rm rh} \in [0.3, 1]$. They are presented in Fig. 3 by colored lines. From the decomposition, one can see that the spectrum at its peak receives the dominant contribution from the last epoch. Earlier epochs mainly contribute to the low-frequency part of the spectrum due to redshift.

Using Eq. (4.40), we obtain the following the GW spectrum

$$\Omega_{\rm GW} h^2(f) \simeq 4 \times 10^{-9} \left(\frac{10^{13}\ {\rm GeV}}{m_\phi} \cdot \frac{\Gamma_\phi}{10^{-5} M_P} \right)^2 \left(\frac{f}{10^9\ {\rm Hz}} \right)^{\frac{8}{5}} \frac{\Upsilon_{\frac{7}{5}}(f/f_c)}{g_{\star 106}^{4/5}}\,, \tag{4.44}$$

where $g_{\star 106} \equiv g_{\star,{\rm rh}}/106.75$ and

$$f_c \equiv \frac{18.9\ {\rm GHz}}{g_{\star 106}^{1/3}}\,. \tag{4.45}$$

The spectrum peaks at $f \sim 3 f_c$, corresponding to $k \sim 3 T_{\rm rh}$ in Fig. 3. To estimate the peak of the spectrum, one can take $\Upsilon_{7/5}(3) \sim 0.1$. The approximations in this section are helpful for understanding the behavior of the radiation-catalyzed GWs. We would like to stress that the GW spectrum presented in the next section is based on full numerical computations.

Comparing Eq. (4.44) to Eq. (4.36), we see that the radiation-catalyzed channel can generate a much higher GW spectrum than the bremsstrahlung channel under certain circumstances. We will also show this more explicitly in the next section.

4.2.4 GWs from radiation-radiation scattering

As has been illustrated in Fig. 1, during reheating and after reheating, gravitons can be produced from radiation-radiation scattering. The exact production rate depends on how thermal species in the thermal bath interact with each other. For the SM thermal bath in the radiation-dominated universe (corresponding to the epoch indicated by $R \in \text{SM}$ in Fig. 1), this production has been calculated in Refs. [39, 40]. For more extensive studies on the production of gravitons from the SM or BSM thermal bath, see [41, 42, 57–62].

Since the early universe at a very high temperature (well above the electroweak scale) could be dominated by very different particle contents (e.g., $SO(10)$ plasma), we remain agnostic regarding the particle physics models and consider a simplified model consisting of N_f fermions (ψ_1, ψ_2, ψ_3, ..., ψ_{N_f}) and N_g gauge bosons (A_1, A_2, A_3, ..., A_{N_g}). Each gauge boson is coupled to every fermion with a universal gauge coupling g, i.e.,

$$\mathcal{L}_R \supset \sum_{i,j} g\overline{\psi_i} A_j^\mu \gamma_\mu \psi_i \,. \tag{4.46}$$

Self-interactions of gauge bosons, which would be present in non-Abelian gauge theories, are neglected for simplicity. Given that most interactions in the SM and many BSM models are gauge interactions, the simplified model is expected to capture the essential characteristics of graviton production in the SM and BSM thermal bath.

In the simplified model, the leading-order production of gravitons arises from the processes $\psi\overline{\psi} \to hA$ and $A\psi \to h\psi$—see the last two diagrams in Fig. 2. The spin/polarization averaged matrix elements for these two processes read

$$|\overline{\mathcal{M}}_{\psi\overline{\psi}\to hA}|^2 = \frac{g^2}{4M_P^2}\frac{t^2+u^2}{s}\,, \tag{4.47}$$

$$|\overline{\mathcal{M}}_{A\psi\to h\psi}|^2 = \frac{g^2}{4M_P^2}\frac{u^2+s^2}{t}\,, \tag{4.48}$$

where s, t, and u are the Mandelstam variables of these two-to-two scattering processes.

Note that $|\overline{\mathcal{M}}_{A\psi\to h\psi}|^2$ would be divergent in the limit of $t \to 0$. This divergence is regulated by Debye-Hückel screening—see Appendix C for details. After the screening effect is taken into account, we obtain the following collision term:

$$\mathcal{C}_h \simeq \frac{N_f N_g g^2}{\pi^3 M_P^2} T^3 e^{-\omega/T}\left[\frac{1}{12} + \frac{1}{8}G\left(\frac{\omega}{\kappa}\right)\right], \tag{4.49}$$

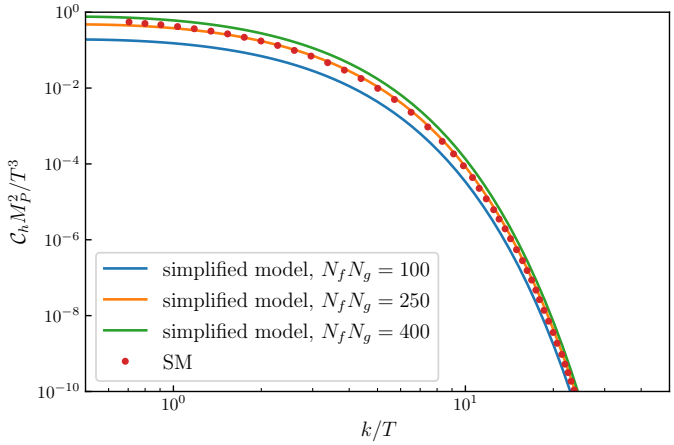

Figure 4. Comparison of our thermal graviton production rate in the simplified model (solid lines) with the SM calculation (red dots). The solid lines are produced using Eq. (4.49), with $g^2 = 0.1$ and $N_f N_g$ specified in the plot. The SM values are taken from Ref. [41].

where $\kappa = g\sqrt{n_\psi/T}$ is the Debye-Hückel screening scale and

$$G(x) \equiv -\frac{3}{2} - \frac{1}{4x^2} + \left(2 + \frac{1}{2x^2} + \frac{1}{16x^4}\right) \ln\left(1 + 4x^2\right) . \tag{4.50}$$

In Eq. (4.49), the first and second terms correspond to the contributions of $\psi\overline{\psi} \to hA$ and $A\psi \to h\psi$, respectively. Note that $G(x)$ in Eq. (4.50) is always positive, despite its first two negative terms. In particular, when x is small, the logarithmic term approaches $3/2 + 1/(4x^2) + 16x^2/3 + \mathcal{O}(x^4)$, implying that $\lim_{x\to 0} G(x) = 16x^2/3$. In general, the G-term in Eq. (4.49) is greater than the other term, as long as $\omega/\kappa \gtrsim 0.4$.

In Fig. 4, we compare the production rate in Eq. (4.49) with the SM production rate, which has been previously calculated in Refs. [39–41]. The SM values shown in this figure are taken from Figure 1 of Ref. [41], where a dimensionless quantity $\hat{\eta}$ was used to present the result. According to Eq. (2.4) in Ref. [41], we use $\hat{\eta} = \frac{M_P^2}{4T^4} k\mathcal{C}_h$ to recast the result. As is shown in Fig. 4, the SM results can be well approximated by our simplified model with $N_f N_g = 250$ and $g^2 = 0.1$.

From the collision term to f_h, the calculation is similar to that in Sec. 4.2.3 and yields

$$f_h(a_{\rm rh}, k) \approx \frac{N_f N_g g^2 T_{\rm rh}^3}{15\pi^3 H_I M_P^2} \frac{2 + 3\overline{G}}{r_a^{3/2}} \left(\frac{k}{T_{\rm rh}}\right)^{\frac{3}{5}} \Upsilon_{-\frac{3}{5}}\left(\frac{k}{T_{\rm rh}}\right) , \tag{4.51}$$

where $\overline{G}$ denotes the thermal average of the G factor in Eq. (4.50). Its specific value depends on the Debye-Hückel screening scale κ and hence on g^2. However, due to the logarithmic function in G, its dependence on g^2 is weak. For g^2 varying from $\sim 10^{-2}$ to ~ 1, $\overline{G}$ only varies from ~ 10 to ~ 20. So in practice, we recommend taking $\overline{G} \approx 15$ for SM-like thermal plasma.

Using Eq. (4.51), we obtain the following the GW spectrum

$$\Omega_{\rm GW}h^2(f) \simeq 6.5\times10^{-17}\left(\frac{N_f N_g g^2}{25}\right)\left(\frac{\Gamma_\phi}{10^{-5}M_P}\right)^{\frac{1}{2}}\left(\frac{f}{\rm GHz}\right)^{\frac{23}{5}}\frac{\Upsilon_{-\frac{3}{5}}(f/f_c)}{g_{\star 106}^{11/20}}\,, \quad (4.52)$$

where the last part containing the Υ function and $g_{\star 106}$ is typically around $\mathcal{O}(1)$ at the peak of the spectrum. Note that the $\Upsilon_{-\frac{3}{5}}$ function asymptotically behaves as $\frac{5}{3}(f/f_c)^{-3/5}$ for $f \ll f_c$ and $r_a \to 0$. Therefore, in the low-f regime, Eq. (4.52) implies a power law of $f^{23/5}\cdot f^{-3/5} = f^4$, provided that r_a is sufficiently small.

5 Results

In this section, we present the complete gravitational wave (GW) spectra generated in our framework spanning over the period from inflation through reheating and eventually to the radiation-dominated era. For the post-inflationary GW spectra, we numerically solve the Boltzmann equation for the graviton phase space distribution function in each scenario, along with the background evolution equations, Eq. (3.2) and Eq. (3.3). The computation begins at the onset of reheating and extends through several e-folds beyond its completion. This approach ensures that our result fully includes the contributions from the epochs before and after reheating.

The results are shown in Fig. 5, along with the projected sensitivity curves of future detectors, including the Big Bang Observer (BBO) [79, 80], DECIGO [81, 82], LISA [83], μAres [84], the Cosmic Explorer (CE) [85], and the Einstein Telescope (ET) [86–89]. These sensitivity curves are shown in yellow color. The energy stored in GWs behaves like dark radiation and contributes to the effective number of neutrino species, $N_{\rm eff}$.[3] The Planck 2018 results provide a 95% confidence level measurement of $N_{\rm eff} = 2.99 \pm 0.34$ [54]. The proposed CMB-S4 experiment is projected to reach a sensitivity of $\Delta N_{\rm eff} \lesssim 0.06$ [90]. Future surveys, such as COrE [91] and Euclid [92], are expected to tighten this constraint to $\Delta N_{\rm eff} \lesssim 0.013$ at the 2σ level. These limits are illustrated by horizontal dashed lines in Fig. 5.

We consider four benchmark scenarios characterized by different values of the coupling y and corresponding reheating temperatures $T_{\rm rh}$: ① $y = 10^{-4}$ with $T_{\rm rh} \simeq 10^{10}\,$GeV (top left), ② $y = 10^{-3}$ with $T_{\rm rh} \simeq 10^{11}\,$GeV (top right), ③ $y = 10^{-2}$ with $T_{\rm rh} \simeq 10^{12}\,$GeV (bottom left), and ④ $y = 10^{-1}$ with $T_{\rm rh} \simeq 10^{13}\,$GeV (bottom right). In all cases, we fix the remaining model parameters to $g^2 = 10^{-1}$, $N_g N_f = 250$, $m_\phi = 10^{13}\,$GeV, $r = 0.02$, and $H_I \simeq 10^{13}\,$GeV (cf. Eq. (3.7)). We emphasize that once the coupling y, the inflaton mass, and the inflationary Hubble scale H_I are specified, the reheating temperature is determined accordingly by solve the background Eq. (3.2) and Eq. (3.3).

Below we list and discuss some features of the GW spectra under consideration.

[3] To constrain this contribution, we use the bound $\Omega_{\rm GW}h^2 \lesssim 5.6\times10^{-6}\Delta N_{\rm eff}$, which applies to the total energy density integrated over logarithmic frequency [11].

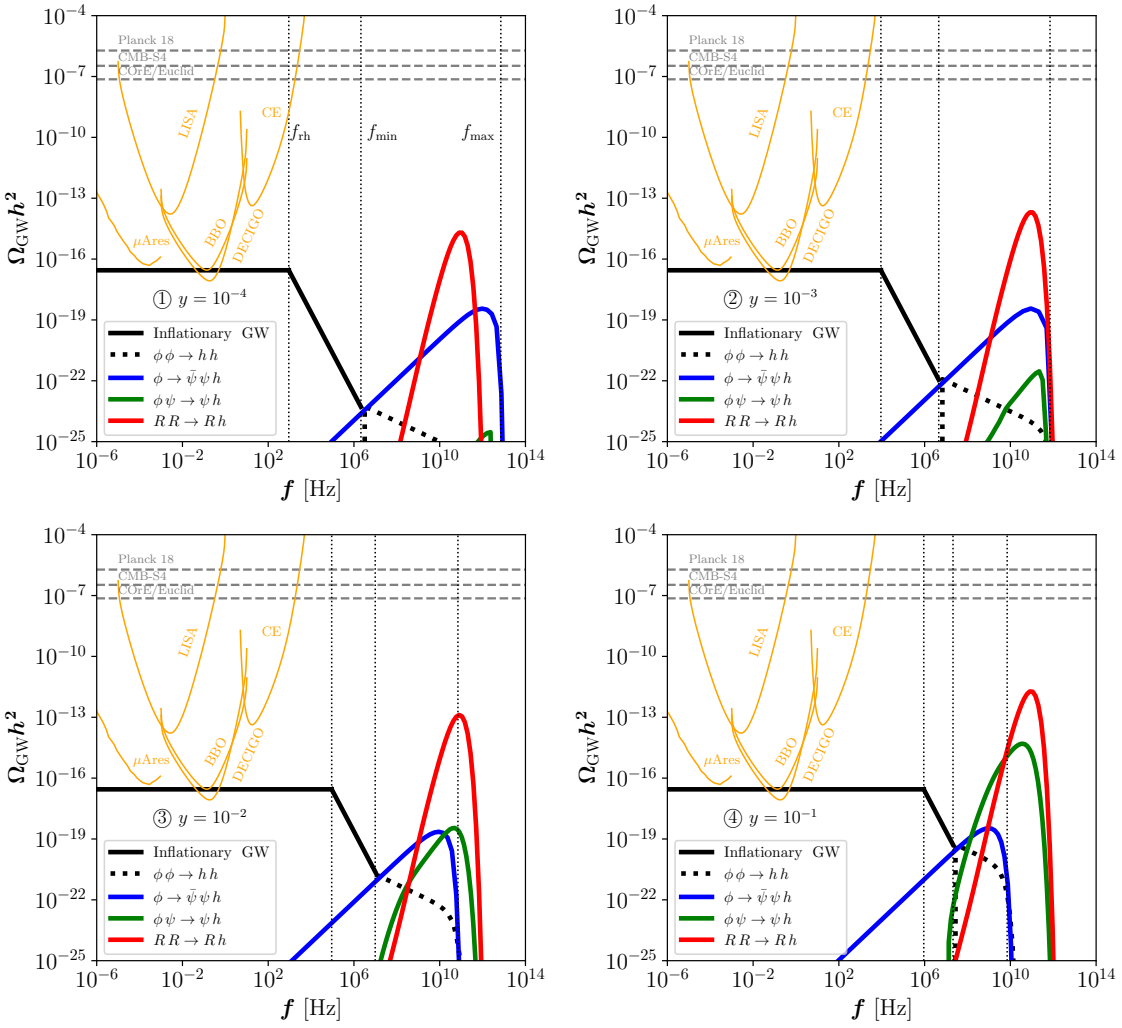

Figure 5. Comparison of the different sources of GWs in three benchmarks with ① $y = 10^{-4}$, ② $y = 10^{-3}$, ③ $y = 10^{-2}$, and ④ $y = 10^{-1}$. Other parameters are commonly set at $g^2 = 10^{-1}$, $m_\phi = 10^{13}$ GeV, $r = 0.02$, $H_I \simeq 10^{13}$ GeV, and $N_f N_g = 250$. The three vertical lines correspond to $f = f_{\rm rh}$ and $f = f_I$, and $f = f_{\rm max}$, given by Eqs. (4.4), (4.5), and (4.28), respectively.

- The solid black line represents the inflationary gravitational wave spectrum generated during inflation. The flat portion corresponds to tensor modes that re-enter the horizon during the radiation-dominated era, while the subsequent f^{-2} decline reflects modes that re-enter during the reheating phase, as is formulated in the second line of Eq. (4.8).

- The black dotted line represents the high-frequency tail of tensor modes, which never exit the horizon but can still be excited during reheating via inflaton annihilation. According to Eq. (4.25), this component of the spectrum decreases as $f^{-1/2}$ until it reaches the cutoff frequency $f_{\rm max}$ given by Eq. (4.28), beyond

which the signal is exponentially suppressed due to inflaton decay.

- The solid blue line represents the GW spectrum generated from bremsstrahlung. The signal peaks when the energy of the emitted graviton reaches its kinematic threshold, corresponding to $\omega = m_\phi/2$ at the end of reheating. This sets the peak frequency at $f_{\rm peak} \simeq f_{\rm max}/2$. In the regime $f < f_{\rm peak}$, the GW spectrum scales as $\Omega_{\rm GW} \propto f$ (cf. Eq. (4.36)), while for $f \gtrsim f_{\rm peak}$, the spectrum is exponentially suppressed due to the rapid depletion of the inflaton number density.

- The solid green curve represents the GW spectrum generated from the radiation-catalyzed scattering process $\phi\psi \to \psi h$. The produced gravitons carry energy $\omega > m_\phi/2$, which implies a lower limit on the frequency, given by $f \simeq f_{\rm min}/2$, where $f_{\rm min}$ is defined in Eq. (4.27). At low temperatures, the Boltzmann suppression can be neglected, and the collision term in Eq. (4.39) peaks at $\omega \simeq m_\phi$, corresponding to a GW frequency $f \simeq f_{\rm max}$. At higher temperatures, gravitons are produced by more energetic radiation, leading to a shift of the peak frequency toward that of the thermal GW spectrum, which will be discussed in the next bullet point.

- The solid red line represents the GW spectrum arising from pure thermal scatterings via the processes $\bar{\psi}\psi \to A\,h$ and $A\psi \to \psi h$, with $N_f N_g = 250$. As discussed in the previous section, this scenario can reproduce the GW signal from Standard Model (SM) thermal scatterings. Similar to the SM case, the spectrum exhibits a peak at $f_{\rm peak} \simeq 100$ GHz [41]. In the low-frequency regime, the spectrum is dominated by the $\bar{\psi}\psi \to A\,h$ process, leading to a scaling of $\Omega_{\rm GW} \propto f^4$—see the discussion below Eq. (4.52). At higher frequencies close to the peak, the spectrum receives the dominant contribution from $A\psi \to \psi h$. Most gravitons are produced with energies on the order of the temperature at the time of production, while excessively high energy gravitons are exponentially suppressed due to the Boltzmann factor $e^{-\omega/T}$ in Eq. (4.49).

By comparing the spectra in Fig. 5 for the four benchmark model parameters, we find that the inflationary GW signal dominates in the regime $f < f_I$, where $f = f_I$ is depicted as the second vertical gray line. For $f > f_I$, the contributions from the other four sources begin to interplay. In general, the bremsstrahlung channel $\phi \to \bar{\psi}\psi h$ dominates over the pure inflaton annihilation process $\phi\phi \to hh$ when $m_\phi \gtrsim T_{\rm rh}$, except in the high-frequency tail and a narrow low-frequency region. Compared to the $1 \to 3$ bremsstrahlung process, the radiation-catalyzed $2 \to 2$ scatterings involving the inflaton and its daughter fields become significant when $T_{\rm rh}$ is not much smaller than m_ϕ, as shown in the lower panels of Fig. 5.

Our main results concern the GW spectrum sourced by thermal scatterings. Since pure thermal scatterings correspond to a UV freeze-in process, the production

rate increases with temperature. This explains why the peak amplitude of the pure thermal GW spectrum grows with larger values of y or $T_{\rm rh}$; see the peaks of the red curves in Fig. 5. We note that the peak can reach $\Omega_{\rm GW}h^2 \sim \mathcal{O}(10^{-12})$ for a reheating temperature of $T_{\rm rh} \simeq 10^{13}\,\text{GeV}$, as shown in the bottom right panel of Fig. 5. More interestingly, we find that the pure thermal scattering channel can dominate GW production even in scenarios with $T_{\rm rh} \ll m_\phi$, and even with a single species of fermions and gauge bosons, i.e., $N_f N_g = 1$. This can be seen in the upper left panel of Fig. 5, where the thermal GW peak reaches approximately $\mathcal{O}(10^{-15})/250 \sim \mathcal{O}(10^{-17})$ for $N_f N_g = 1$. When additional species are included—i.e., for large $N_f N_g$—the amplitude of the thermal GW signal increases further, as shown in Eq. (4.52).

6 Conclusions

In this work, we aim to provide a full-spectrum analysis of GW production within a generic framework that assumes slow-roll inflation followed by a reheating phase, during which cold inflatons gradually decay into radiation, ultimately driving the universe into a radiation-dominated era. By employing the Boltzmann equation of the graviton phase space distribution function, we systematically compute the GW spectra generated by (i) pure inflaton annihilation, (ii) graviton bremsstrahlung from inflaton decay, (iii) radiation-catalyzed inflaton-graviton conversion, and (iv) scattering among fully thermalized radiation particles. For each channel, we obtain the corresponding collision term with the calculation presented in great detail in the appendixes, and solve the Boltzmann equation numerically to get the GW spectrum. Moreover, we derive accurate and simple analytical formulae for these GW spectra—see Eqs. (4.25), (4.36), (4.44), and (4.52)—which we believe would be useful to relevant studies.

Our main results are illustrated in Fig. 5, where we compare the inflationary GW spectrum with contributions from reheating-era sources. We find that inflationary GWs dominate the low-frequency regime, $f < f_I$, where f_I denotes the frequency of modes re-entering the horizon at the onset of reheating. At higher frequencies, the GW spectrum is shaped by the interplay of reheating-era sources. Among them, the dominant contribution arises from (iv)—even when the reheating temperature is much lower than the inflaton mass.

It is worth mentioning that for GWs produced in (iv), we have considered a simplified model consisting solely of N_f fermions and N_g gauge bosons. Remarkably, we find that our result obtained in this simplified model with $N_f N_g \sim 250$ can approximate the SM thermal production rate of gravitons very well—see Fig. 4. We believe that our result in the simplified model can be used to facilitate the calculation of GW production in various BSM thermal plasma.

In summary, we have presented the first full-spectrum analysis of inflationary and post-inflationary GWs. Our work demonstrates that post-inflationary physics can

generate high-frequency GWs with rich structures, which, if observable by any means, could reveal crucial information about inflation and post-inflationary reheating.

Acknowledgments

We thank Nicolás Bernal for his helpful comments on the draft and Marco Drewes for valuable discussions. YX would like to thank Shi Pi for discussions during his visit at the Institute of Theoretical Physics, Chinese Academy of Sciences in 2024. The work of XJX is supported in part by the National Natural Science Foundation of China (NSFC) under grant No. 12141501 and also by the CAS Project for Young Scientists in Basic Research (YSBR-099). YX has received support from the Cluster of Excellence "Precision Physics, Fundamental Interactions, and Structure of Matter" (PRISMA$^+$ EXC 2118/1) funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the German Excellence Strategy (Project No. 390831469).

A Calculation of horizon crossing

Solving the horizon crossing scale factor $a_{\rm hc}$ in Eq. (4.2) requires the explicit form of $H(a)$ during the radiation-dominated and reheating epochs.

Let us first consider the radiation-dominated epoch, in which the T-a relation is given by Eq. (3.21).

Eq. (3.21) implies that the temperature at matter-radiation equality is

$$T_{\rm eq} \simeq T_0 \frac{a_0}{a_{\rm eq}} \approx 0.8 \text{ eV} \,. \tag{A.1}$$

Using Eq. (3.21), one can express ρ_R and $H = \sqrt{\rho_R/(3M_P^2)}$ as functions of a:

$$\rho_R(a) = \left(\frac{a_0}{a}\right)^4 \frac{\pi^2 g_\star}{30} T_0^4 \left(\frac{g_{\star s,0}}{g_{\star s}}\right)^{4/3}, \quad \text{for} \quad a \geq a_{\rm rh} \,, \tag{A.2}$$

$$H(a) = \frac{\pi T_0^2}{3M_P} \left(\frac{a_0}{a}\right)^2 \sqrt{\frac{g_\star}{10}} \left(\frac{g_{\star s,0}}{g_{\star s}}\right)^{2/3}, \quad \text{for} \quad a \in [a_{\rm rh},\ a_{\rm eq}] \,. \tag{A.3}$$

Substituting Eq. (A.3) into Eq. (4.2), we obtain the first case in Eq. (4.3).

Next, we turn to the reheating epoch, in which the expansion is dominantly driven by matter, $\rho_\phi \propto a^{-3}$. Its specific form can be inferred from the end of this epoch, at which the energy density is approximately $\rho_R(a_{\rm rh})$:

$$\rho_\phi \approx \left(\frac{a_{\rm rh}}{a}\right)^3 \rho_R(a_{\rm rh}) \approx \left(\frac{a_{\rm rh}}{a}\right)^3 \frac{\pi^2 g_{\star,\rm rh}}{30} T_{\rm rh}^4 \,, \quad \text{for} \quad a \in [a_I,\ a_{\rm rh}] \,. \tag{A.4}$$

Therefore, the Hubble parameter is given by

$$H(a) = \frac{\pi T_{\rm rh}^2}{3M_P} \left(\frac{a_{\rm rh}}{a}\right)^{\frac{3}{2}} \sqrt{\frac{g_{\star,\rm rh}}{10}}, \quad \text{for} \quad a \in [a_I,\ a_{\rm rh}] \,. \tag{A.5}$$

Substituting Eq. (A.5) into Eq. (4.2), we obtain the second case in Eq. (4.3).

The frequency f_I in Eq. (4.5) is obtained by expressing a_I in terms of $T_{\rm rh}$:

$$a_I \approx a_{\rm rh}\left(\frac{2\Gamma_\phi}{3H_I}\right)^{\frac{2}{3}} \approx a_{\rm rh}\left(\frac{\pi^2 g_{\star,\rm rh}T_{\rm rh}^4}{90H_I^2M_P^2}\right)^{\frac{1}{3}} \approx \left(\frac{\pi^2 g_{\star,\rm rh}T_{\rm rh}T_0^3}{90H_I^2M_P^2}\cdot\frac{g_{\star s,0}}{g_{\star s,\rm rh}}\right)^{\frac{1}{3}}, \tag{A.6}$$

where the first, second, and third steps have used Eqs. (3.14), (3.18), and (3.21), respectively.

B Calculation of matrix elements

In this appendix, we present the detailed calculation of the matrix elements used in this work. We shall mention here that some of the matrix elements have already been calculated in the literature [23, 27, 37, 38]. Only the matrix elements for $\psi\overline{\psi} \to hA$ and $A\psi \to h\psi$ are new. Nevertheless, we believe it is useful to include all the matrix elements in a self-contained manner, with unified conventions and notations.

B.1 Gravitational Polarization Summation (GPS)

For spin-2 gravitons with four-momentum $q^\mu = (\omega,\ \vec{q})$, the polarization tensor $\epsilon^{\mu\nu}$ satisfies the following conditions:

$$\begin{aligned} \epsilon^{i\mu\nu} &= \epsilon^{i\nu\mu} \text{ symmetric} \\ q_\mu\epsilon^{i\mu\nu} &= 0 \text{ transverse} \\ \eta_{\mu\nu}\epsilon^{i\mu\nu} &= 0 \text{ traceless} \\ \epsilon^{i\mu\nu}\epsilon^{*j}_{\mu\nu} &= \delta^{ij} \text{ orthonormal} \end{aligned}, \tag{B.1}$$

where i, j are polarization indices. Introducing the auxiliary vector $\bar{q}^\mu = (\omega,\ -\vec{q})$, the gravitational polarization summation (GPS) reads [27]:

$$\mathrm{GPS}_{\mu\nu,\alpha\beta} \equiv \sum_{\rm pol}\epsilon^*_{\mu\nu}\epsilon_{\alpha\beta} = \frac{1}{2}\left(\hat{\eta}_{\mu\alpha}\hat{\eta}_{\nu\beta} + \hat{\eta}_{\mu\beta}\hat{\eta}_{\nu\alpha} - \hat{\eta}_{\mu\nu}\hat{\eta}_{\alpha\beta}\right), \tag{B.2}$$

where

$$\hat{\eta}_{\mu\nu} \equiv \eta_{\mu\nu} - \frac{q_\mu\bar{q}_\nu + \bar{q}_\mu q_\nu}{q\cdot\bar{q}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{q_1^2}{\omega^2}-1 & \frac{q_1q_2}{\omega^2} & \frac{q_1q_3}{\omega^2} \\ 0 & \frac{q_1q_2}{\omega^2} & \frac{q_2^2}{\omega^2}-1 & \frac{q_2q_3}{\omega^2} \\ 0 & \frac{q_1q_3}{\omega^2} & \frac{q_2q_3}{\omega^2} & \frac{q_3^2}{\omega^2}-1 \end{pmatrix}. \tag{B.3}$$

Due to the vanishing temporal components of $\hat{\eta}_{\mu\nu}$, the GPS satisfies

$$\mathrm{GPS}_{\mu\nu,\alpha\beta}X^\mu = 0 \quad \text{for } \forall\ X^\mu = \left(X^0,\ 0,\ 0,\ 0\right). \tag{B.4}$$

This feature allows us to significantly simplify calculations involving GPS and non-relativistic particles.

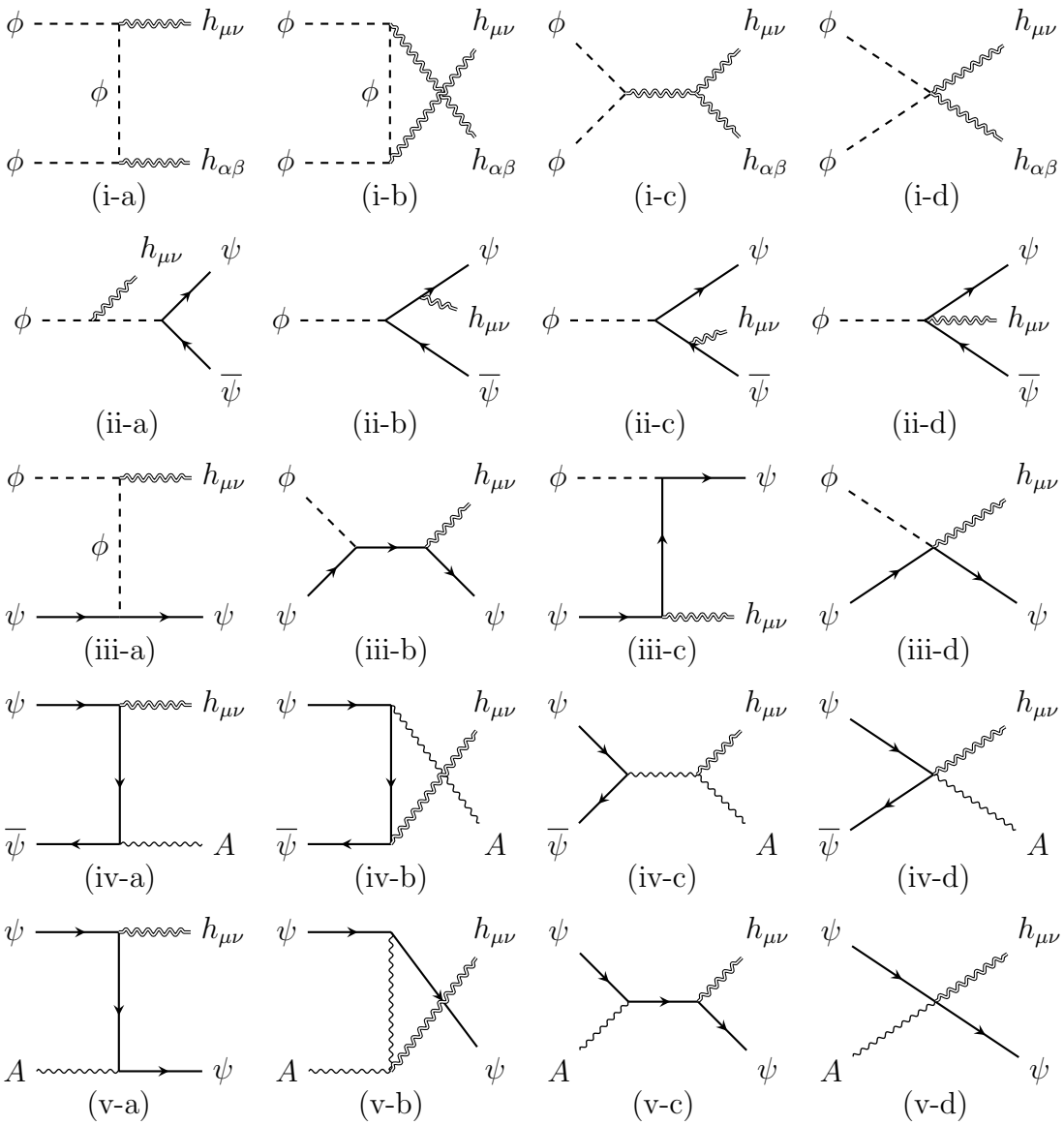

Figure 6. Feynman diagrams of graviton production processes considered in this work.

B.2 $\phi\phi \to hh$

A pair of inflatons can annihilate into a pair of gravitons according to diagrams (i-a)-(i-d) presented in Fig. 6.

In this work, the inflatons are at rest so the kinematics is simple. Denoting the momenta of the two initial states by p_1 and p_2, and the final ones by k_1 and k_2, we have

$$p_1^\mu = p_2^\mu = (m_\phi, \vec{0})\,, \quad k_1^\mu = (m_\phi, \vec{k})\,, \; k_2^\mu = (m_\phi, -\vec{k})\,, \tag{B.5}$$

with $\left|\vec{k}\right| = m_\phi$.

In general, the amplitude of each diagram can be written as

$$\mathcal{M} = \epsilon_{\mu\nu}(k_1)\epsilon_{\alpha\beta}(k_2)\mathcal{A}^{\mu\nu,\alpha\beta}\,,$$

where $\mathcal{A}^{\mu\nu,\alpha\beta}$ denotes the remaining part of amplitude after the two graviton legs are removed. In practice, one can drop many terms in $\mathcal{A}^{\mu\nu,\alpha\beta}$ and use the following replacement:

$$\overline{\mathcal{A}}^{\mu\nu,\alpha\beta} \equiv \mathcal{A}^{\mu\nu,\alpha\beta}\big|_{X\to 0} \quad \text{for } X \in \{\eta_{\mu\nu},\ \eta_{\alpha\beta},\ k_1^\mu,\ k_1^\nu,\ k_2^\alpha,\ k_2^\beta\}\,. \tag{B.6}$$

Here terms proportional to $\eta_{\mu\nu}$ or $\eta_{\alpha\beta}$ can be safely set to zero due to the traceless condition in Eq. (B.1). Terms proportional to those k's vanish due to the transverse condition in Eq. (B.1).

Using Eq. (B.6) can greatly simplify the calculation. For instance, $\mathcal{A}^{\mu\nu,\alpha\beta}$ for diagram (a) reduces to to a single term proportional to $p_1^\mu p_1^\nu p_2^\alpha p_2^\beta$, which in the nonrelativistic limit leads to $|\mathcal{M}_{\text{(i-a)}}|^2 = 0$ according to Eq. (B.4). Similarly, diagram (i-b) also yields a null result.

For diagrams (i-c) and (i-d), a similar reduction of the amplitudes leads to

$$\overline{\mathcal{A}}^{\mu\nu,\alpha\beta}_{\text{(i-c)}} = -\frac{3}{2}\frac{m_\phi^2}{M_P^2}\left(\eta^{\alpha\nu}\eta^{\beta\mu} + \eta^{\alpha\mu}\eta^{\beta\nu}\right)\,, \tag{B.7}$$

$$\overline{\mathcal{A}}^{\mu\nu,\alpha\beta}_{\text{(i-d)}} = 2\frac{m_\phi^2}{M_P^2}\left(\eta^{\alpha\nu}\eta^{\beta\mu} + \eta^{\alpha\mu}\eta^{\beta\nu}\right)\,. \tag{B.8}$$

From Eqs. (B.7) and (B.8), it is straightforward to compute the squared amplitude:

$$\begin{aligned}|\mathcal{M}_{\phi\phi\to hh}|^2 &= \left(\overline{\mathcal{A}}^{\mu\nu,\alpha\beta}_{\text{(i-c)}} + \overline{\mathcal{A}}^{\mu\nu,\alpha\beta}_{\text{(i-d)}}\right)\left(\overline{\mathcal{A}}^{\mu'\nu',\alpha'\beta'}_{\text{(i-c)}} + \overline{\mathcal{A}}^{\mu'\nu',\alpha'\beta'}_{\text{(i-d)}}\right)^* \text{GPS}_{\mu\nu,\mu'\nu'}\text{GPS}_{\alpha\beta,\alpha'\beta'} \\ &= \frac{2m_\phi^4}{M_P^4}\,.\end{aligned} \tag{B.9}$$

Note that in the above calculation, we have used gravitational polarization summation instead of taking the average. For the latter, one should further multiply the result by a factor of 1/4:

$$|\overline{\mathcal{M}}_{\phi\phi\to hh}|^2 = |\mathcal{M}_{\phi\phi\to hh}|^2 \times \frac{1}{4}\,. \tag{B.10}$$

The above result agrees with Eq. (10) in Ref. [23] and Eq. (A.18) in Ref. [38].

B.3 $\phi \to \psi\overline{\psi}h$

Next, we consider the gravitational bremsstrahlung from the inflaton decay — see diagrams (ii-a)-(ii-d) in Fig. 6.

We label the particle momenta as $\phi(l^\mu) \to \psi(p^\mu)\overline{\psi}(q^\mu)h(k^\mu)$ with $l^\mu = (m_\phi,\ \vec{0})$, $p^\mu = (E_p,\ \vec{p})$, $k^\mu = (\omega,\ \vec{k})$, and $q^\mu = (m_\phi - E_p - \omega,\ -\vec{p} - \vec{k})$. For later use, we shall mention that

$$p \cdot q = \frac{1}{2} m_\phi (m_\phi - 2\omega)\,, \tag{B.11}$$

which is independent of E_p for fixed ω. Eq. (B.11) can be obtained from $q^2 = 0$ which implies $\vec{p}\cdot\vec{k} = \frac{m_\phi^2}{2} + E_p\omega - E_p m_\phi - m_\phi\omega$

Using the same reduction technique in Eq. (B.6), it is straightforward to simplify the matrix elements of diagrams (ii-a)-(ii-d) to

$$i\mathcal{M}_{\text{(ii-a)}} = \frac{-i\,y}{2l\cdot k M_P}\ (2l_\mu\, l_\nu)\ \bar{u}(p)v(q)\,\epsilon^{\mu\nu}\,, \tag{B.12}$$

$$i\mathcal{M}_{\text{(ii-b)}} = \frac{i\,y}{2p\cdot k M_P}\,[\bar{u}(p)(p_\mu\gamma_\nu)(l\cdot\gamma)v(q)]\,\epsilon^{\mu\nu}\,, \tag{B.13}$$

$$i\mathcal{M}_{\text{(ii-c)}} = \frac{i\,y}{2q\cdot k M_P}\,[\bar{u}(p)(l\cdot\gamma)(q_\mu\gamma_\nu)v(q)]\,\epsilon^{\mu\nu}\,, \tag{B.14}$$

$$i\mathcal{M}_{\text{(ii-d)}} \propto \eta_{\mu\nu}\epsilon^{\mu\nu} = 0\,. \tag{B.15}$$

where $\mathcal{M}_{\text{(ii-d)}}$ vanishes due to the traceless condition in Eq. (B.1). In addition, according to Eq. (B.4), we have $|\mathcal{M}_{\text{(ii-a)}}|^2 = 0$, which implies that $\mathcal{M}_{\text{(ii-a)}}$ can be neglected. Therefore, the combined matrix element reads

$$\begin{aligned}|\mathcal{M}_{\phi\to\psi\bar{\psi}h}|^2 &\equiv \left|\mathcal{M}_{\text{(ii-b)}} + \mathcal{M}_{\text{(ii-c)}}\right|^2 \\ &= \frac{y^2\, m_\phi^2}{M_P^2}\left(1 - \frac{2\omega}{m_\phi}\right)\left[2 - \frac{2m_\phi}{\omega} + \left(\frac{m_\phi}{\omega}\right)^2\right],\end{aligned} \tag{B.16}$$

where ω is the energy of the graviton. Eq. (B.16) agrees with Appendix B.2 in Ref. [27] in the the massless fermion limit.

Similar to Eq. (B.10), one should divide the above result by a factor of 8 to obtain the spin/polarization-averaged matrix element

$$|\overline{\mathcal{M}}_{\phi\to\psi\bar{\psi}h}|^2 = \frac{1}{8}|\mathcal{M}_{\phi\to\psi\bar{\psi}h}|^2\,. \tag{B.17}$$

B.4 $\phi\psi \to \psi h$

It is also possible to produce gravitons via ϕ-ψ scattering [37]. The relevant Feynman diagrams are shown in Fig. 6. We label the particle momenta as $\phi(l)\psi(q) \to \psi(p)h(k)$, and define the Mandelstam variables as $s = (l+q)^2$, $t = (l-p)^2$, and $u = (l-k)^2$.

The matrix elements of the diagrams (iii-a)-(iii-d) are given by:

$$i\mathcal{M}_{\text{(iii-a)}} = \frac{-i\,y}{2l\cdot kM_P}\,(2l_\mu\,l_\nu)\;\bar{u}(p)u(q)\,\epsilon^{\mu\nu}\,, \tag{B.18}$$

$$i\mathcal{M}_{\text{(iii-b)}} = \frac{i\,y}{2p\cdot kM_P}\,[\bar{u}(p)(p_\mu\gamma_\nu)(l\cdot\gamma)u(q)]\,\epsilon^{\mu\nu}\,, \tag{B.19}$$

$$i\mathcal{M}_{\text{(iii-c)}} = \frac{i\,y}{2q\cdot kM_P}\,[\bar{u}(p)(l\cdot\gamma)(q_\mu\gamma_\nu)u(q)]\,\epsilon^{\mu\nu}\,, \tag{B.20}$$

$$i\mathcal{M}_{\text{(iii-d)}} \propto \eta_{\mu\nu}\epsilon^{\mu\nu} = 0\,. \tag{B.21}$$

We note that crossing symmetry allows us to relate these $2\to 2$ scattering matrix elements to those in the $1\to 3$ process calculated above. Therefore, we expect that the result can be obtained from the previous one via crossing symmetry[4]:

$$\left|\overline{\mathcal{M}}_{\phi\psi\to\psi h}(l,p,q,k)\right|^2 = (-1)^{\#\text{FC}}\left|\overline{\mathcal{M}}_{\phi\to\psi\bar{\psi}h}(l,p,-q,k)\right|^2\,, \tag{B.22}$$

where the bar over $\mathcal{M}$ indicates averaging all spins and polarizations in both initial and final states; the minus sign before q takes effect on both its energy and momentum; and "#FC" denotes the fermion crossing number (i.e., the number of fermions moved from initial to final or final to initial states) [94]. For $\phi\psi\to\psi h$ and $\phi\to\psi\bar{\psi}h$, we have $\#\text{FC}=1$ and hence an overall minus sign on the right-hand side. However, this does not imply that $\left|\overline{\mathcal{M}}_{\phi\psi\to\psi h}(l,p,q,k)\right|^2$ would be negative, since the analytic continuation of $|\overline{\mathcal{M}}_{\phi\to\psi\bar{\psi}h}|^2$ from q to $-q$ also introduces an extra minus sign.

A straightforward calculation of the matrix element leads to

$$|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 = \frac{1}{8}\frac{y^2(-2u)(m_\phi^4+u^2)}{M_P^2(m_\phi^2-u)^2}\,, \tag{B.23}$$

where we have used $s+t+u=m_\phi^2$. The factor of $\frac{1}{8}$ accounts for spin/polarization-averaging of all initial and final states—similar to Eq. (B.17). In the non-relativistic limit of inflaton, $l=(m_\phi,\vec{0})$, $u=m_\phi^2-2\,m_\phi\,\omega$, Eq. (B.23) becomes

$$|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 = \frac{1}{8}\frac{y^2\,m_\phi^2}{M_P^2}\left(\frac{2\omega}{m_\phi}-1\right)\left[2-\frac{2m_\phi}{\omega}+\left(\frac{m_\phi}{\omega}\right)^2\right]\,. \tag{B.24}$$

As one can see, Eq. (B.24) explicitly verifies the identity of crossing symmetry in Eq. (B.22). Compared to Eqs. (B.17) and (B.16), Eq. (B.24) is almost the same, except for an overall minus sign, which has been included in Eq. (B.22).

B.5 $\psi\bar{\psi}\to hA$

We compute this process first in the center-of-mass (CM) frame and then generalize it to a general frame. We label the particle momenta as $\psi(p_1)\bar{\psi}(p_2)\to h(k_1)A(k_2)$

[4] See Eq. (5.67) in Ref. [93] and also Refs. [94, 95].

and write the momenta as follows:

$$p_1 = (E, \vec{p})\,, \quad p_2 = (E, -\vec{p})\,, \tag{B.25}$$

$$k_1 = (E, \vec{k})\,, \quad k_2 = (E, -\vec{k})\,, \tag{B.26}$$

with $|\vec{p}| = |\vec{k}| = E$. The Mandelstam variables therefore read:

$$s = 2p_1 \cdot p_2 = 2k_1 \cdot k_2 = 4E^2\,, \tag{B.27}$$

$$t = -2p_1 \cdot k_1 = -2p_2 \cdot k_2 = -2E^2(1 - \cos\theta)\,, \tag{B.28}$$

$$u = -2p_2 \cdot k_1 = -2p_1 \cdot k_2 = -2E^2(1 + \cos\theta)\,, \tag{B.29}$$

where θ is the angle between $\vec{p}$ and $\vec{k}$. Note that there are two independent kinematic parameters (E and $\cos\theta$). Meanwhile, only two of the three Mandelstam variables are independent ($s + t + u = 0$ for this process). So the generalization from the CM frame to a general one can be done by mapping E and $\cos\theta$ to two of the Mandelstam variables.

The Feynman diagrams responsible for this process are given by diagrams (iv-a)-(iv-d). Applying similar reduction techniques, we find that the amplitudes of these diagrams reduce to

$$i\mathcal{M}_{\text{(iv-a)}} = i\kappa g \frac{\bar{v}(p_2) \cdot \gamma^\rho \cdot [\not{k}_1 \cdot (\gamma^\nu p_1^\mu + \gamma^\mu p_1^\nu) - 4p_1^\mu p_1^\nu] \cdot u(p_1)}{4t} \epsilon_{\mu\nu}\epsilon_\rho\,, \tag{B.30}$$

$$i\mathcal{M}_{\text{(iv-b)}} = -i\kappa g \frac{\bar{v}(p_2) \cdot [(\gamma^\nu p_2^\mu + \gamma^\mu p_2^\nu) \cdot \not{k}_1 - 4p_2^\mu p_2^\nu] \cdot \gamma^\rho \cdot u(p_1)}{4u} \epsilon_{\mu\nu}\epsilon_\rho\,, \tag{B.31}$$

$$\begin{aligned} i\mathcal{M}_{\text{(iv-c)}} = &-\frac{g}{4s} i\kappa\ \bar{v}(p_2) \cdot \{2\ \not{k}_1 \cdot (\eta^{\rho\nu}\,(p_1^\mu + p_2^\mu) + \eta^{\rho\mu}\,(p_1^\nu + p_2^\nu)) \\ &+ \gamma^\nu\,(s\eta^{\rho\mu} + 2k_1^\mu p_1^\rho - 2k_1^\rho p_1^\mu + 2k_1^\mu p_2^\rho - 2k_1^\rho p_2^\mu) \\ &+ \gamma^\mu\,(s\eta^{\rho\nu} + 2k_1^\nu p_1^\rho - 2k_1^\rho p_1^\nu + 2k_1^\nu p_2^\rho - 2k_1^\rho p_2^\nu) \\ &-2\gamma^\rho\,[k_1^\mu\,(p_1^\nu + p_2^\nu) + (p_1^\mu + p_2^\mu)\,(k_1^\nu - 2p_1^\nu - 2p_2^\nu)]\} \cdot u(p_1)\epsilon_{\mu\nu}\epsilon_\rho\,, \end{aligned} \tag{B.32}$$

$$i\mathcal{M}_{\text{(iv-d)}} = \frac{i\kappa g}{4} \bar{v}(p_2) \cdot (\eta^{\rho\mu}\gamma^\nu + \eta^{\rho\nu}\gamma^\mu) \cdot u(p_1)\epsilon_{\mu\nu}\epsilon_\rho\,. \tag{B.33}$$

Averaging all the initial spins and final polarizations yields

$$|\overline{\mathcal{M}}_{\psi\bar{\psi}\to hA}|^2 = \frac{g^2}{4M_P^2} E^2\,(3 + \cos 2\theta) = \frac{g^2}{4M_P^2} \frac{t^2 + u^2}{s}\,, \tag{B.34}$$

where a factor of $1/4$ accounting for spin averaging of initial fermion states and a factor of $1/4$ for polarization averaging of final states have been included. The $\frac{t^2+u^2}{s}$ part in the above result matches Eq. (2.46) of Ref. [40].

B.6 $\psi A \to h\psi$

The Feynman diagrams for this process are given in the last row of Fig. 6. Using crossing symmetry, the squared amplitude for $\psi A \to h\psi$ can be directly obtained

from the $\psi\overline{\psi} \to hA$ result through the Mandelstam variable substitution $s \leftrightarrow u$:

$$|\overline{\mathcal{M}}_{\psi A\to\psi h}|^2 = \frac{g^2}{4M_P^2}\frac{t^2+s^2}{u}\,. \tag{B.35}$$

Note that the ordering of the particles can affect the definition of s, t, and u. We denote the momentum of i-th particles by p_i and define $s = (p_1+p_2)^2$, $t = (p_2-p_4)^2$, $u = (p_1-p_4)^2$. From $\psi\overline{\psi} \to hA$ to $\psi A \to h\psi$, we perform an interchange of the 2nd and the 4th particle so t is not changed while s and u are interchanged.

One can further interchange the two initial states of $\psi A \to h\psi$ to get the result for $A\psi \to h\psi$. Although it is essentially the same process, the t-channel enhancement in the soft-scattering regime (or forward scattering limit) becomes manifest:

$$|\overline{\mathcal{M}}_{A\psi\to h\psi}|^2 = \frac{g^2}{4M_P^2}\frac{u^2+s^2}{t}\,. \tag{B.36}$$

C Calculation of collision terms

This appendix presents a detailed calculation of the collisions terms used in our work. The final results are summarized as follows:

$$\mathcal{C}_h^{\phi\phi\to hh} = \frac{\pi n_\phi^2}{32M_P^4}\delta(\omega - m_\phi)\,, \tag{C.1}$$

$$\mathcal{C}_h^{\phi\to\bar{\psi}\psi h} = \frac{y^2\rho_\phi}{64\pi\omega M_P^2}F\left(\frac{\omega}{m_\phi}\right)\Theta\left(\frac{m_\phi}{2}-\omega\right), \tag{C.2}$$

$$\mathcal{C}_h^{\phi\psi\to\psi h} = \frac{y^2\rho_\phi}{64\pi\omega M_P^2}(-1)F\left(\frac{\omega}{m_\phi}\right)\frac{T}{\omega}e^{-\frac{\omega-m_\phi/2}{T}}\Theta\left(\omega-\frac{m_\phi}{2}\right), \tag{C.3}$$

$$\mathcal{C}_h^{\bar{\psi}\psi\to Ah} = \frac{g^2}{12\pi^3 M_P^2}T^3 e^{-\omega/T}\,, \tag{C.4}$$

$$\mathcal{C}_h^{\psi A\to\psi h} = \frac{g^2}{(2\pi)^3 M_P^2}G\left(\frac{\omega}{\kappa}\right)T^3 e^{-\omega/T}\,, \tag{C.5}$$

where $\kappa = \sqrt{g^2 n_\psi/T}$ is the Debye-Hückel screening scale, and

$$F(x) \equiv (1-2x)\left(2-2x^{-1}+x^{-2}\right), \tag{C.6}$$

$$G(x) \equiv -\frac{3}{2}-\frac{1}{4x^2}+\left(2+\frac{1}{2x^2}+\frac{1}{16x^4}\right)\ln\left(1+4x^2\right). \tag{C.7}$$

In the following calculation, for convenience, we use subscripts "1", "2", "3", and "4" to denote the first, second, third, and fourth particles in a given process.

C.1 $\phi\phi \to hh$

After inflation, the inflaton field forms a cold, non-relativistic condensate, implying that its phase space distribution can be approximated by a Dirac delta function:

$$f_\phi(t,\mathbf{p}) = (2\pi)^3 n_\phi(t)\delta^{(3)}(\mathbf{p})\,. \tag{C.8}$$

The polarization-averaged matrix element of this process is given by Eq. (B.10), which is momentum-independent and can be factored out of the phase space integral in Eq. (4.15). The symmetry factor in Eq. (4.15) for this process is

$$\mathcal{S} = \frac{1}{2!} \times \frac{1}{1!}\,, \tag{C.9}$$

where the factor $\frac{1}{2!}$ comes from two identical ϕ particles in $\phi\phi \to hh$ and the factor $\frac{1}{1!}$ comes from gravitons. Here one might naively take two identical h particles in the final state, which would lead to a factor of $\frac{1}{2!}$. This would be correct if the symmetry factor is used to calculate the collision term for $\left[\frac{\partial}{\partial t} - 3H\right] n_h$ instead of $\left[\frac{\partial}{\partial t} - Hk\frac{\partial}{\partial k}\right] f_h$. For the Boltzmann equation of f_h, we have to select one h from the final states and assign it to the production state in Eq. (4.15). For this h, its phase space is not integrated out in Eq. (4.15). After this selection, the total number of identical particles in the final states reduces to one, hence the contribution is $\frac{1}{1!}$. For further clarifications, we refer to the discussion below Eq. (A21) in Ref. [96].

The overall multiplicity factor in Eq. (4.15) is

$$g_{n+m} = 1 \times 1 \times 2\,. \tag{C.10}$$

Putting the above pieces together, we have

$$g_{n+m}\mathcal{S}|\overline{\mathcal{M}}_{\phi\phi\to hh}|^2 = 2 \times \frac{1}{2!} \times \frac{1}{1!} \times \frac{2m_\phi^4}{M_P^4} \times \frac{1}{4} = \frac{m_\phi^4}{2M_P^4}\,. \tag{C.11}$$

Substituting Eqs. (C.11) and (C.8) into Eq. (4.15), the phase space integral can be computed straightforwardly:

$$\begin{aligned}
\mathcal{C}_h^{\phi\phi\to hh} &= \frac{1}{2\omega} \cdot \frac{m_\phi^4}{2M_P^4} \int d\Pi_1 d\Pi_2 d\Pi_3 f_1 f_2 (2\pi)^4 \delta^{(4)}(p_1^\mu + p_2^\mu - p_3^\mu - p_4^\mu) \\
&= \frac{1}{2\omega} \cdot \frac{m_\phi^4}{2M_P^4} \int \frac{n_\phi}{2E_1}\frac{n_\phi}{2E_2} d^3\mathbf{p}_3 \frac{2\pi}{2E_3} \delta^{(4)}(p_1^\mu + p_2^\mu - p_3^\mu - p_4^\mu) \\
&= \frac{\pi n_\phi^2 m_\phi^2}{16\omega M_P^4} \int d^3\mathbf{p}_3 \frac{1}{E_3} \delta^{(3)}(\mathbf{p}_3 + \mathbf{p}_4)\delta(2m_\phi - p_3 - p_4) \\
&= \frac{\pi n_\phi^2 m_\phi^2}{16\omega M_P^4} \left[\frac{1}{E_3}\delta(2m_\phi - p_3 - p_4)\right]_{\mathbf{p}_3\to -\mathbf{p}_4} \\
&= \frac{\pi n_\phi^2 m_\phi^2}{16\omega M_P^4} \left[\frac{1}{\omega}\delta(2m_\phi - 2\omega)\right] \\
&= \frac{\pi n_\phi^2}{32 M_P^4}\delta(\omega - m_\phi)\,.
\end{aligned} \tag{C.12}$$

C.2 $\phi \to \bar{\psi}\psi h$

The squared amplitude of this process is given by Eq. (B.17). Using the F function in Eq. (C.6), we rewrite it into a more compact form:

$$|\overline{\mathcal{M}}_{\phi\to\bar{\psi}\psi h}|^2 = \frac{y^2 m_\phi^2}{8M_P^2} F\left(\frac{\omega}{m_\phi}\right). \tag{C.13}$$

Since it depends solely on the energy carried by the radiated graviton, we can factor it out from the phase space integral of the collision term. So the collision term reads

$$\begin{aligned}\mathcal{C}_h^{\phi\to\bar{\psi}\psi h} &= \frac{\mathcal{S}g_{n+m}|\overline{\mathcal{M}}_{\phi\to\bar{\psi}\psi h}|^2}{2\omega}\int d\Pi_1 f_1 d\Pi_2 d\Pi_3 (2\pi\delta)^4 \\ &= \frac{\mathcal{S}g_{n+m}|\overline{\mathcal{M}}_{\phi\to\bar{\psi}\psi h}|^2}{2\omega}\frac{n_\phi}{2m_\phi}\int d\Pi_2 d\Pi_3 (2\pi\delta)^4\,, \end{aligned} \tag{C.14}$$

with the symmetry factor and overall multiplicity factor given by

$$\mathcal{S}=1, \quad g_{n+m}=4\,. \tag{C.15}$$

The remaining part of the phase space integral is worked out at follows:

$$\begin{aligned}\int d\Pi_2 d\Pi_3 (2\pi\delta)^4 &= \int \frac{d^3\mathbf{p}_2}{(2\pi)^3 2E_2}\frac{d^3\mathbf{p}_3}{(2\pi)^3 2E_3}(2\pi\delta)^4 \\ &= \frac{1}{16\pi^2}\int \frac{\delta\left(m_\phi - E_2(\mathbf{p}_2) - E_3(-\mathbf{p}_2-\mathbf{p}_4)-\omega\right)}{E_2(\mathbf{p}_2)E_3(-\mathbf{p}_2-\mathbf{p}_4)}d^3\mathbf{p}_2 \\ &= \frac{1}{16\pi^2}\int \frac{d^3\mathbf{p}_2}{p_2\,|\mathbf{p}_2+\mathbf{p}_4|}\delta\left(m_\phi-\omega-p_2-|\mathbf{p}_2+\mathbf{p}_4|\right) \\ &= \frac{2\pi}{16\pi^2}\int_0^\infty p_2^2 dp_2 \int_{-1}^1 \frac{\delta\left(m_\phi-\omega-p_2-|\mathbf{p}_2+\mathbf{p}_4|\right)}{p_2\,|\mathbf{p}_2+\mathbf{p}_4|}d\cos\theta\,, \end{aligned} \tag{C.16}$$

where θ denotes the angle between $\mathbf{p}_2$ and $\mathbf{p}_4$, and $|\mathbf{p}_2+\mathbf{p}_4|$ in the last step should be interpreted as a function of θ and p_2:

$$|\mathbf{p}_2+\mathbf{p}_4| = \sqrt{p_2^2+\omega^2+2p_2\omega\cos\theta}\,. \tag{C.17}$$

The Dirac delta function in Eq. (C.16) is integrated out as follows:

$$\int_{-1}^1 \frac{\delta\left(m_\phi-\omega-p_2-|\mathbf{p}_2+\mathbf{p}_4|\right)}{|\mathbf{p}_2+\mathbf{p}_4|}d\cos\theta = \frac{1}{p_2\omega}, \tag{C.18}$$

provided that the equation $m_\phi-\omega-p_2-|\mathbf{p}_2+\mathbf{p}_4|=0$ with respect to $\cos\theta$ has a solution in the allowed range $[-1,1]$. Solving this equation gives rise to the following solution

$$\cos\theta = \frac{m_\phi^2+2\omega p_2-2m_\phi(\omega+p_2)}{2\omega p_2}\,. \tag{C.19}$$

Imposing the condition $\cos\theta\in[-1,1]$ on Eq. (C.19), we obtain

$$p_2 \in \left[\frac{m_\phi}{2}-\omega, \frac{m_\phi}{2}\right]\,. \tag{C.20}$$

Assembling the above pieces together, Eq. (C.16) becomes

$$\int d\Pi_2 d\Pi_3 (2\pi\delta)^4 = \frac{2\pi}{16\pi^2}\int_{m_\phi/2-\omega}^{m_\phi/2} p_2^2 dp_2 \frac{1}{p_2}\frac{1}{p_2\omega}\Theta\left(\frac{m_\phi}{2}-\omega\right)$$
$$= \frac{1}{8\pi}\Theta\left(\frac{m_\phi}{2}-\omega\right). \tag{C.21}$$

Hence, the collision term is given by

$$\mathcal{C}_h^{\phi\to\bar\psi\psi h} = \frac{\mathcal{S} g_{n+m}|\overline{\mathcal{M}}_{\phi\to\bar\psi\psi h}|^2}{2\omega}\frac{n_\phi}{2m_\phi}\frac{1}{8\pi}\Theta\left(\frac{m_\phi}{2}-\omega\right)$$
$$= \frac{\mathcal{S} g_{n+m}|\overline{\mathcal{M}}_{\phi\to\bar\psi\psi h}|^2}{32\pi\omega m_\phi} n_\phi \Theta\left(\frac{m_\phi}{2}-\omega\right)$$
$$= \frac{y^2}{8\pi\omega}\frac{m_\phi n_\phi}{8M_P^2} F\left(\frac{\omega}{m_\phi}\right)\Theta\left(\frac{m_\phi}{2}-\omega\right). \tag{C.22}$$

C.3 $\phi\psi \to \psi h$

The squared amplitude of this process, given by Eq. (B.24), can also be factored out from the phase space integral of the collision term. This leads to

$$\mathcal{C}_h^{\phi\psi\to\psi h} = \frac{\mathcal{S} g_{n+m}|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2}{2\omega}\int d\Pi_1 f_1 d\Pi_2 f_2 d\Pi_3 (2\pi\delta)^4$$
$$= \frac{\mathcal{S} g_{n+m}|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2}{2\omega}\frac{n_\phi}{2m_\phi}\int d\Pi_2 f_2 d\Pi_3 (2\pi\delta)^4, \tag{C.23}$$

The next step is similar to Eq. (C.16) except that now we have an extra factor of f_2 and different kinematics. So Eq. (C.16) changes to

$$\int d\Pi_2 f_2 d\Pi_3 (2\pi\delta)^4 = \int \frac{d^3\mathbf{p}_2}{(2\pi)^3 2E_2}\frac{d^3\mathbf{p}_3}{(2\pi)^3 2E_3}(2\pi\delta)^4$$
$$= \frac{2\pi}{16\pi^2}\int_0^\infty p_2^2 dp_2 f_2 \int_{-1}^1 \frac{\delta\left(m_\phi + p_2 - \omega - |\mathbf{p}_2-\mathbf{p}_4|\right)}{p_2\,|\mathbf{p}_2-\mathbf{p}_4|} d\cos\theta\,. \tag{C.24}$$

Following a similar analysis to that in the previous subsection, we find that Eq. (C.20) changes to

$$p_2 \in \left[\omega - \frac{M}{2},\ \infty\right], \tag{C.25}$$

while the integration of $\cos\theta$ still leads to $\frac{1}{p_2\omega}$. Therefore, (C.24) becomes

$$\int d\Pi_2 f_2 d\Pi_3 (2\pi\delta)^4 = \frac{2\pi}{16\pi^2}\int_{\omega-\frac{m_\phi}{2}}^\infty p_2^2 dp_2 f_2 \frac{1}{p_2}\frac{1}{p_2\omega}\Theta\left(\omega-\frac{m_\phi}{2}\right)$$
$$= \frac{1}{8\pi\omega}\Theta\left(\omega-\frac{m_\phi}{2}\right)\int_{\omega-\frac{m_\phi}{2}}^\infty dp_2 f_2\,. \tag{C.26}$$

And the collision term reduces to

$$\begin{aligned}\mathcal{C}_h^{\phi\psi\to\psi h} &= \frac{\mathcal{S}g_{n+m}|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2}{2\omega}\frac{n_\phi}{2m_\phi}\frac{1}{8\pi\omega}\Theta\left(\omega-\frac{m_\phi}{2}\right)\int_{\omega-\frac{m_\phi}{2}}^{\infty}dp_2 f_2 \\ &= \frac{\mathcal{S}g_{n+m}|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 n_\phi}{32\pi m_\phi\omega^2}\Theta\left(\omega-\frac{m_\phi}{2}\right)\int_{\omega-\frac{m_\phi}{2}}^{\infty}f_2 dp_2 \\ &\approx \frac{\mathcal{S}g_{n+m}|\overline{\mathcal{M}}_{\phi\psi\to\psi h}|^2 n_\phi}{32\pi m_\phi\omega^2}Te^{-\frac{\omega-m_\phi/2}{T}}\Theta\left(\omega-\frac{m_\phi}{2}\right), \end{aligned} \tag{C.27}$$

where in the last step we have used the Boltzmann approximation: $f_2 \approx e^{-p_2/T}$.

C.4 $\bar{\psi}\psi \to Ah$

The collision term of $\bar{\psi}\psi \to Ah$ where all particles are relativistic can be written as the following form:

$$\mathcal{C}_h^{\bar{\psi}\psi\to Ah} \approx \frac{\mathcal{S}g_{n+m}}{2\omega}\int d\Pi_3 f_1 f_2 \int d\Pi_1 d\Pi_2 (2\pi\delta)^4 |\overline{\mathcal{M}}_{\bar{\psi}\psi\to Ah}|^2, \tag{C.28}$$

where $f_1 f_2$ can be treated as a quantity independent of p_1 and p_2 using the Boltzmann approximation and energy conservation:

$$f_1 f_2 = e^{-(p_1+p_2)/T} = e^{-(p_3+\omega)/T}\,. \tag{C.29}$$

For this process, we have

$$\mathcal{S} = 1\,, \quad g_{n+m} = g_\psi g_{\bar{\psi}} g_A = 8\,. \tag{C.30}$$

Next, we shall work out the integral $\int d\Pi_1 d\Pi_2 (2\pi\delta)^4 |\overline{\mathcal{M}}_{\bar{\psi}\psi\to Ah}|^2$, which is Lorentz invariant. Therefore, one can calculate it in the center-of-mass (CM) frame without loss of generality.

In the CM frame, the four-momentum delta function can be written as

$$\delta^{(4)} = \delta(\sqrt{s} - E_1 - E_2)\delta^{(3)}(\mathbf{p}_1 + \mathbf{p}_2)\,, \tag{C.31}$$

where the Mandelstam variable s is fixed by the final-state kinematics (i.e., $\mathbf{p}_3$ and $\mathbf{p}_4$) instead of $\mathbf{p}_1$ and $\mathbf{p}_2$. So s can be treated as a constant in the phase space integral of $\mathbf{p}_1$ and $\mathbf{p}_2$.

The matrix element in the CM frame can be written as [see Eq. (B.34)]

$$|\overline{\mathcal{M}}_{\bar{\psi}\psi\to Ah}|^2_{\rm CM} = \frac{g^2}{16M_P^2}s\,(3+\cos 2\theta)\,, \tag{C.32}$$

where θ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_4$.

Therefore, in the CM frame, the integral reads:

$$\begin{aligned}\int d\Pi_1 d\Pi_2 (2\pi\delta)^4 |\overline{\mathcal{M}}_{\bar{\psi}\psi\to Ah}|^2 &= \frac{1}{(2\pi)^2}\int \frac{p_1^2 dp_1 d\Omega}{4p_1^2}\delta(\sqrt{s}-2p_1)|\overline{\mathcal{M}}_{\bar{\psi}\psi\to Ah}|^2_{\rm CM}, \\ &= \frac{1}{8(2\pi)^2}\int d\Omega \frac{g^2}{16M_P^2} s\,(3+\cos 2\theta) \\ &= \frac{g^2}{48\pi M_P^2} s\,. \end{aligned} \tag{C.33}$$

Although Eq. (C.33) is derived in the CM frame, due to the Lorentz invariance of the integral, the result remains valid in a general frame.

Then the full integration proceeds as:

$$\begin{aligned}\mathcal{C}_h^{\bar{\psi}\psi\to Ah} &\approx \frac{\mathcal{S} g_{n+m}}{2\omega}\int d\Pi_3 f_1 f_2 \frac{g^2}{48\pi M_P^2} s \\ &= \frac{\mathcal{S} g_{n+m}}{4}\frac{g^2}{24\pi M_P^2}\int \frac{p_3^2 dp_3 d\cos\theta_3}{(2\pi)^3} e^{-(\omega+p_3)/T}(1-\cos\theta_3) \\ &= \frac{\mathcal{S} g_{n+m} g^2}{12(2\pi)^3 M_P^2} T^3 e^{-\omega/T}\,, \end{aligned} \tag{C.34}$$

where θ_3 denotes the angle between $\mathbf{p}_3$ and $\mathbf{p}_4$ and we have used $s = 2\omega p_3(1-\cos\theta_3)$.

C.5 $\psi A \to \psi h$

This process exhibits kinematic similarities to $\bar{\psi}\psi \to Ah$ scattering, but there is a t-channel divergence which requires a careful treatment in the soft scattering limit.

First, following a similar analysis to that in the previous subsection, we obtain

$$\mathcal{C}_h^{\psi A\to\psi h} \approx \frac{\mathcal{S} g_{n+m}}{16(2\pi)^2\omega}\int d\Pi_3 e^{-(\omega+p_3)/T}\int d\Omega |\overline{\mathcal{M}}_{\psi A\to\psi h}|^2_{\rm CM}\,, \tag{C.35}$$

where

$$|\overline{\mathcal{M}}_{\psi A\to\psi h}|^2_{\rm CM} = \frac{g^2 s\,(5+2\cos\theta+\cos^2\theta)}{8M_P^2(1-\cos\theta)}\,, \tag{C.36}$$

with θ the angle between $\mathbf{p}_1$ and $\mathbf{p}_4$.

As is obvious, there is a divergence at $\theta = 0$ when $\cos\theta$ is integrated from -1 to 1. It originates from the divergence of $t \to 0$ in Eq. (B.36). Such divergences are common in thermal production of light particles via diagrams containing light t-channel mediators. One well-studied example is the axion production in thermal plasma via the Primakoff process—see e.g. Refs. [97, 98], where this divergence is regulated by the Debye-Hückel screening.

For an electrically neutral particle scattering off a electrically charged one, the Debye-Hückel screening is included by $|\mathcal{M}|^2 \to |\mathcal{M}|^2\mathcal{F}^2_{\rm Debye}$ with $\mathcal{F}^2_{\rm Debye}$ the Debye-Hückel form factor give as followed [97, 98]:

$$\mathcal{F}^2_{\rm Debye} = \frac{|q^2|}{\kappa^2+|q^2|}\,, \tag{C.37}$$

where q is the momentum transfer and $\kappa = \sqrt{g^2 n_\psi / T}$ is the Debye-Hückel screening scale. We note that a recent calculation of thermal production of gravitons in the Sun also adopted the same Debye-Hückel screening to handle the divergence [99].

The momentum transfer $q = p_1 - p_3$ is a space-like momentum ($q^2 < 0$). In the CM frame, we obtain

$$q^2 = -2p_1 \cdot p_3 = -2(\omega^2 - \mathbf{p}_1 \cdot \mathbf{p}_3) = -2\omega^2(1 - c_\theta)\,, \tag{C.38}$$

where $c_\theta \equiv \cos\theta$.

With the Debye-Hückel screening effect included, the regulated angular integral becomes:

$$\begin{aligned}\int d\Omega |\overline{\mathcal{M}}_{\psi A\to\psi h}|^2_{\rm CM} \mathcal{F}^2_{\rm Debye} &= \frac{2\pi g^2 s}{8M_P^2}\int_{-1}^{1} dc_\theta \frac{5+2c_\theta+c_\theta^2}{1-c_\theta}\frac{2(1-c_\theta)}{(\kappa/\omega)^2+2(1-c_\theta)} \\ &= \frac{\pi g^2 s}{M_P^2}\left[-\frac{3}{2}-\frac{\kappa^2}{4\omega^2}+\left(2+\frac{\kappa^2}{2\omega^2}+\frac{\kappa^4}{16\omega^4}\right)\ln\left(1+\frac{4\omega^2}{\kappa^2}\right)\right] \\ &= \frac{\pi g^2 s}{M_P^2} G\left(\frac{\omega}{\kappa}\right), \end{aligned} \tag{C.39}$$

where G has been defined in Eq. (C.7).

Finally, similar to the calculation in Eq. (C.34), we perform the phase space integral of $\mathbf{p}_3$:

$$\begin{aligned}\mathcal{C}_h^{\psi A\to\psi h} &\simeq \frac{\mathcal{S} g_{n+m} g^2}{32(2\pi)^3 M_P^2} G\left(\frac{\omega}{\kappa}\right)\int p_3^2 e^{-(\omega+p_3)/T} dp_3 \int (1-\cos\theta_3) d\cos\theta_3 \\ &= \frac{g^2}{(2\pi)^3 M_P^2} G\left(\frac{\omega}{\kappa}\right) T^3 e^{-\omega/T}, \end{aligned} \tag{C.40}$$

where θ_3 denotes the angle between $\mathbf{p}_3$ and $\mathbf{p}_4$.

References

[1] A.A. Starobinsky, *A New Type of Isotropic Cosmological Models Without Singularity*, *Phys. Lett. B* **91** (1980) 99.

[2] A.H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, *Phys. Rev. D* **23** (1981) 347.

[3] A.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) 389.

[4] A. Albrecht and P.J. Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, *Phys. Rev. Lett.* **48** (1982) 1220.

[5] M.C. Guzzetti, N. Bartolo, M. Liguori and S. Matarrese, *Gravitational waves from inflation*, *Riv. Nuovo Cim.* **39** (2016) 399 [1605.01615].

[6] M. Giovannini, *Primordial backgrounds of relic gravitons*, *Prog. Part. Nucl. Phys.* **112** (2020) 103774 [1912.07065].

[7] A.A. Starobinsky, *Spectrum of relict gravitational radiation and the early state of the universe*, *JETP Lett.* **30** (1979) 682.

[8] B. Allen, *The Stochastic Gravity Wave Background in Inflationary Universe Models*, *Phys. Rev. D* **37** (1988) 2078.

[9] V. Sahni, *The Energy Density of Relic Gravity Waves From Inflation*, *Phys. Rev. D* **42** (1990) 453.

[10] M.S. Turner, M.J. White and J.E. Lidsey, *Tensor perturbations in inflationary models as a probe of cosmology*, *Phys. Rev. D* **48** (1993) 4613 [astro-ph/9306029].

[11] C. Caprini and D.G. Figueroa, *Cosmological Backgrounds of Gravitational Waves*, *Class. Quant. Grav.* **35** (2018) 163001 [1801.04268].

[12] L.F. Abbott, E. Farhi and M.B. Wise, *Particle Production in the New Inflationary Cosmology*, *Phys. Lett. B* **117** (1982) 29.

[13] A.D. Dolgov and D.P. Kirilova, *ON PARTICLE CREATION BY A TIME DEPENDENT SCALAR FIELD*, *Sov. J. Nucl. Phys.* **51** (1990) 172.

[14] L. Kofman, A.D. Linde and A.A. Starobinsky, *Reheating after inflation*, *Phys. Rev. Lett.* **73** (1994) 3195 [hep-th/9405187].

[15] L. Kofman, A.D. Linde and A.A. Starobinsky, *Towards the theory of reheating after inflation*, *Phys. Rev. D* **56** (1997) 3258 [hep-ph/9704452].

[16] R. Allahverdi, R. Brandenberger, F.-Y. Cyr-Racine and A. Mazumdar, *Reheating in Inflationary Cosmology: Theory and Applications*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 27 [1001.2600].

[17] M.A. Amin, M.P. Hertzberg, D.I. Kaiser and J. Karouby, *Nonperturbative Dynamics Of Reheating After Inflation: A Review*, *Int. J. Mod. Phys. D* **24** (2014) 1530003 [1410.3808].

[18] K.D. Lozanov, *Lectures on Reheating after Inflation*, 1907.04402.

[19] B. Barman, N. Bernal and J. Rubio, *Two or three things particle physicists (mis)understand about (pre)heating*, 3, 2025 [2503.19980].

[20] Y. Ema, R. Jinno, K. Mukaida and K. Nakayama, *Gravitational Effects on Inflaton Decay*, *JCAP* **05** (2015) 038 [1502.02475].

[21] Y. Ema, R. Jinno, K. Mukaida and K. Nakayama, *Gravitational particle production in oscillating backgrounds and its cosmological implications*, *Phys. Rev. D* **94** (2016) 063517 [1604.08898].

[22] Y. Ema, R. Jinno and K. Nakayama, *High-frequency Graviton from Inflaton Oscillation*, *JCAP* **09** (2020) 015 [2006.09972].

[23] G. Choi, W. Ke and K.A. Olive, *Minimal production of prompt gravitational waves during reheating*, *Phys. Rev. D* **109** (2024) 083516 [2402.04310].

[24] M. Gross, Y. Mambrini, E. Kpatcha, M.O. Olea-Romacho and R. Roshan, *Gravitational wave production during reheating: From the inflaton to primordial black holes*, *Phys. Rev. D* **111** (2025) 035020 [2411.04189].

[25] K. Nakayama and Y. Tang, *Stochastic Gravitational Waves from Particle Origin*, *Phys. Lett. B* **788** (2019) 341 [1810.04975].

[26] D. Huang and L. Yin, *Stochastic Gravitational Waves from Inflaton Decays*, *Phys. Rev. D* **100** (2019) 043538 [1905.08510].

[27] B. Barman, N. Bernal, Y. Xu and O. Zapata, *Gravitational wave from graviton Bremsstrahlung during reheating*, *JCAP* **05** (2023) 019 [2301.11345].

[28] B. Barman, N. Bernal, Y. Xu and O. Zapata, *Bremsstrahlung-induced gravitational waves in monomial potentials during reheating*, *Phys. Rev. D* **108** (2023) 083524 [2305.16388].

[29] S. Kanemura and K. Kaneta, *Gravitational waves from particle decays during reheating*, *Phys. Lett. B* **855** (2024) 138807 [2310.12023].

[30] N. Bernal, S. Cléry, Y. Mambrini and Y. Xu, *Probing reheating with graviton bremsstrahlung*, *JCAP* **01** (2024) 065 [2311.12694].

[31] A. Tokareva, *Gravitational waves from inflaton decay and bremsstrahlung*, *Phys. Lett. B* **853** (2024) 138695 [2312.16691].

[32] W. Hu, K. Nakayama, V. Takhistov and Y. Tang, *Gravitational wave probe of Planck-scale physics after inflation*, *Phys. Lett. B* **856** (2024) 138958 [2403.13882].

[33] K.-Y. Choi, E. Lkhagvadorj and S. Mahapatra, *Gravitational wave sourced by decay of massive particle from primordial black hole evaporation*, *JCAP* **07** (2024) 064 [2403.15269].

[34] B. Barman, N. Bernal, S. Cléry, Y. Mambrini, Y. Xu and O. Zapata, *Probing Reheating with Gravitational Waves from Graviton Bremsstrahlung*, in *58th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, 5, 2024 [2405.09620].

[35] R. Inui, Y. Mikura and S. Yokoyama, *Gravitational waves from graviton bremsstrahlung with kination phase*, *Phys. Rev. D* **111** (2025) 043511 [2408.10786].

[36] Y. Jiang and T. Suyama, *Spectrum of high-frequency gravitational waves from graviton bremsstrahlung by the decay of inflaton: case with polynomial potential*, *JCAP* **02** (2025) 041 [2410.11175].

[37] Y. Xu, *Ultra-high frequency gravitational waves from scattering, Bremsstrahlung and decay during reheating*, *JHEP* **10** (2024) 174 [2407.03256].

[38] N. Bernal, Q.-f. Wu, X.-J. Xu and Y. Xu, *Pre-thermalized Gravitational Waves*, 2503.10756.

[39] J. Ghiglieri and M. Laine, *Gravitational wave background from Standard Model physics: Qualitative features*, *JCAP* **07** (2015) 022 [1504.02569].

[40] J. Ghiglieri, G. Jackson, M. Laine and Y. Zhu, *Gravitational wave background from Standard Model physics: Complete leading order*, *JHEP* **07** (2020) 092 [2004.11392].

[41] A. Ringwald, J. Schütte-Engel and C. Tamarit, *Gravitational Waves as a Big Bang Thermometer*, *JCAP* **03** (2021) 054 [2011.04731].

[42] N. Bernal and Y. Xu, *Thermal gravitational waves during reheating*, *JHEP* **01** (2025) 137 [2410.21385].

[43] S.Y. Choi, J.S. Shim and H.S. Song, *Factorization and polarization in linearized gravity*, *Phys. Rev. D* **51** (1995) 2751 [hep-th/9411092].

[44] J.F. Donoghue, *General relativity as an effective field theory: The leading quantum corrections*, *Phys. Rev. D* **50** (1994) 3874 [gr-qc/9405057].

[45] R. Kallosh and A. Linde, *Universality Class in Conformal Inflation*, *JCAP* **07** (2013) 002 [1306.5220].

[46] R. Kallosh and A. Linde, *Non-minimal Inflationary Attractors*, *JCAP* **10** (2013) 033 [1307.7938].

[47] M. Drees and Y. Xu, *Small field polynomial inflation: reheating, radiative stability and lower bound*, *JCAP* **09** (2021) 012 [2104.03977].

[48] N. Bernal and Y. Xu, *Polynomial inflation and dark matter*, *Eur. Phys. J. C* **81** (2021) 877 [2106.03950].

[49] M. Drees and Y. Xu, *Large field polynomial inflation: parameter space, predictions and (double) eternal nature*, *JCAP* **12** (2022) 005 [2209.07545].

[50] M. Drees and Y. Xu, *Parameter space of leptogenesis in polynomial inflation*, *JCAP* **04** (2024) 036 [2401.02485].

[51] N. Bernal, J. Harz, M.A. Mojahed and Y. Xu, *Graviton- and inflaton-mediated dark matter production after large field polynomial inflation*, *Phys. Rev. D* **111** (2025) 043517 [2406.19447].

[52] Y. Xu, *Probing gravitational dark matter with ultra-high frequency gravitational waves*, *Phys. Lett. B* **865** (2025) 139483 [2412.21137].

[53] E.W. Kolb, A. Notari and A. Riotto, *On the Reheating Stage after Inflation*, *Phys. Rev. D* **68** (2003) 123505 [hep-ph/0307241].

[54] Planck collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].

[55] BICEP, Keck collaboration, *Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season*, *Phys. Rev. Lett.* **127** (2021) 151301 [2110.00483].

[56] G.F. Giudice, E.W. Kolb and A. Riotto, *Largest temperature of the radiation era and its cosmological implications*, *Phys. Rev. D* **64** (2001) 023508 [hep-ph/0005123].

[57] P. Klose, M. Laine and S. Procacci, *Gravitational wave background from non-Abelian reheating after axion-like inflation*, *JCAP* **05** (2022) 021 [2201.02317].

[58] A. Ringwald and C. Tamarit, *Revealing the cosmic history with gravitational waves*, *Phys. Rev. D* **106** (2022) 063027 [2203.00621].

[59] P. Klose, M. Laine and S. Procacci, *Gravitational wave background from vacuum and thermal fluctuations during axion-like inflation*, *JCAP* **12** (2022) 020 [2210.11710].

[60] J. Ghiglieri, J. Schütte-Engel and E. Speranza, *Freezing-in gravitational waves*, *Phys. Rev. D* **109** (2024) 023538 [2211.16513].

[61] F. Muia, F. Quevedo, A. Schachner and G. Villa, *Testing BSM physics with gravitational waves*, *JCAP* **09** (2023) 006 [2303.01548].

[62] M. Drewes, Y. Georis, J. Klaric and P. Klose, *Upper bound on thermal gravitational wave backgrounds from hidden sectors*, *JCAP* **06** (2024) 073 [2312.13855].

[63] Particle Data Group collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.

[64] K. Saikawa and S. Shirai, *Primordial gravitational waves, precisely: The role of thermodynamics in the Standard Model*, *JCAP* **05** (2018) 035 [1803.01038].

[65] Planck collaboration, *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10 [1807.06211].

[66] L.A. Boyle and P.J. Steinhardt, *Probing the early universe with inflationary gravitational waves*, *Phys. Rev. D* **77** (2008) 063504 [astro-ph/0512014].

[67] S. Pi, M. Sasaki, A. Wang and J. Wang, *Revisiting the ultraviolet tail of the primordial gravitational wave*, *Phys. Rev. D* **110** (2024) 103529 [2407.06066].

[68] M. Giovannini, *Ultra-high frequency spikes of relic gravitons*, 2503.12175.

[69] M. Maggiore, *Gravitational wave experiments and early universe cosmology*, *Phys. Rept.* **331** (2000) 283 [gr-qc/9909001].

[70] Y. Watanabe and E. Komatsu, *Improved Calculation of the Primordial Gravitational Wave Spectrum in the Standard Model*, *Phys. Rev. D* **73** (2006) 123515 [astro-ph/0604176].

[71] Q.-f. Wu and X.-J. Xu, *High-energy and ultra-high-energy neutrinos from Primordial Black Holes*, *JCAP* **02** (2025) 059 [2409.09468].

[72] S.-P. Li and X.-J. Xu, *Dark matter produced from right-handed neutrinos*, *JCAP* **06** (2023) 047 [2212.09109].

[73] K. Ichikawa, T. Suyama, T. Takahashi and M. Yamaguchi, *Primordial Curvature Fluctuation and Its Non-Gaussianity in Models with Modulated Reheating*, *Phys. Rev. D* **78** (2008) 063545 [0807.3988].

[74] M. Drewes, *Measuring the inflaton coupling in the CMB*, *JCAP* **09** (2022) 069 [1903.09599].

[75] M.A.G. Garcia, K. Kaneta, Y. Mambrini and K.A. Olive, *Inflaton Oscillations and Post-Inflationary Reheating*, *JCAP* **04** (2021) 012 [2012.10756].

[76] J. Yokoyama, *Can oscillating scalar fields decay into particles with a large thermal mass?*, *Phys. Lett. B* **635** (2006) 66 [hep-ph/0510091].

[77] M. Drewes, *On the Role of Quasiparticles and thermal Masses in Nonequilibrium Processes in a Plasma*, 1012.5380.

[78] K. Mukaida and K. Nakayama, *Dissipative Effects on Reheating after Inflation*, *JCAP* **03** (2013) 002 [1212.4985].

[79] J. Crowder and N.J. Cornish, *Beyond LISA: Exploring future gravitational wave missions*, *Phys. Rev. D* **72** (2005) 083005 [gr-qc/0506015].

[80] V. Corbin and N.J. Cornish, *Detecting the cosmic gravitational wave background with the big bang observer*, *Class. Quant. Grav.* **23** (2006) 2435 [gr-qc/0512039].

[81] N. Seto, S. Kawamura and T. Nakamura, *Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space*, *Phys. Rev. Lett.* **87** (2001) 221103 [astro-ph/0108011].

[82] H. Kudoh, A. Taruya, T. Hiramatsu and Y. Himemoto, *Detecting a gravitational-wave background with next-generation space interferometers*, *Phys. Rev. D* **73** (2006) 064006 [gr-qc/0511145].

[83] LISA collaboration, *Laser Interferometer Space Antenna*, 1702.00786.

[84] A. Sesana et al., *Unveiling the gravitational universe at μ-Hz frequencies*, *Exper. Astron.* **51** (2021) 1333 [1908.11391].

[85] D. Reitze et al., *Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO*, *Bull. Am. Astron. Soc.* **51** (2019) 035 [1907.04833].

[86] S. Hild et al., *Sensitivity Studies for Third-Generation Gravitational Wave Observatories*, *Class. Quant. Grav.* **28** (2011) 094013 [1012.0908].

[87] M. Punturo et al., *The Einstein Telescope: A third-generation gravitational wave observatory*, *Class. Quant. Grav.* **27** (2010) 194002.

[88] B. Sathyaprakash et al., *Scientific Objectives of Einstein Telescope*, *Class. Quant. Grav.* **29** (2012) 124013 [1206.0331].

[89] ET collaboration, *Science Case for the Einstein Telescope*, *JCAP* **03** (2020) 050 [1912.02622].

[90] K. Abazajian et al., *CMB-S4 Science Case, Reference Design, and Project Plan*, 1907.04473.

[91] COrE collaboration, *COrE (Cosmic Origins Explorer) A White Paper*, 1102.2181.

[92] EUCLID collaboration, *Euclid Definition Study Report*, 1110.3193.

[93] M.E. Peskin and D.V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA (1995), 10.1201/9780429503559.

[94] V. Kaplunovsky, *Crossing Symmetry*, https://web2.ph.utexas.edu/~vadim/Classes/2022f/crossing.pdf.

[95] B. Bellazzini, *Softness and amplitudes' positivity for spinning particles*, *JHEP* **02** (2017) 034 [1605.06111].

[96] X. Luo, W. Rodejohann and X.-J. Xu, *Dirac neutrinos and $N_{\rm eff}$*, *JCAP* **06** (2020) 058 [2005.01629].

[97] G.G. Raffelt, *Astrophysical axion bounds diminished by screening effects*, *Phys. Rev. D* **33** (1986) 897.

[98] Q.-f. Wu and X.-J. Xu, *A comprehensive calculation of the Primakoff process and the solar axion flux*, *JCAP* **07** (2024) 013 [2402.16083].

[99] C. García-Cely and A. Ringwald, *Complete Gravitational-Wave Spectrum of the Sun*, 2407.18297.