

How well do we know the primordial black hole abundance? The crucial role of nonlinearities when approaching the horizon

Valerio De Luca,^{1,*} Alex Kehagias,^{2,3,†} and Antonio Riotto^{4,‡}

¹*Center for Particle Cosmology, Department of Physics and Astronomy,
University of Pennsylvania 209 S. 33rd St., Philadelphia, PA 19104, USA*

²*Physics Division, National Technical University of Athens, 15780 Zografou Campus, Athens, Greece*

³*CERN, Theoretical Physics Department, Geneva, Switzerland*

⁴*Département de Physique Théorique and Gravitational Wave Science Center (GWSC),
Université de Genève, CH-1211 Geneva, Switzerland CH-1211 Geneva, Switzerland*

We discuss the non-linear corrections entering in the calculation of the primordial black hole abundance from the non-linear radiation transfer function and the determination of the true physical horizon crossing. We show that the current standard techniques to calculate the abundance of primordial black holes suffer from uncertainties and argue that the primordial black hole abundance may be much smaller than what routinely considered. This would imply, among other consequences, that the interpretation of the recent pulsar timing arrays data from scalar-induced gravitational waves may not be ruled out because of an overproduction of primordial black holes.

Introduction. The various detections of gravitational waves (GWs) originated from the mergers of black hole binaries [1–4] have resurrected the interest in the physics of Primordial Black Holes (PBHs) [5–7]. Some of the LIGO/Virgo/KAGRA data may be in fact of primordial origin [8–15] and forthcoming GW experiments might shed light on the possible existence of PBHs [16–20].

In the standard scenario, PBHs are considered to be born in the radiation-dominated phase by the collapse of large overdensities created during inflation on small scales [5]. Upon horizon re-entry, the very same sizeable fluctuations generate GWs at second-order in perturbation theory (see Ref. [21] for a review).

The recent pulsar timing arrays (PTA) data releases by the NANOGrav [22, 23], EPTA [24–26], PPTA [27–29] and CPTA [30] collaborations have shown evidence for the presence of a stochastic background of GWs and have raised the question if it can be ascribed to the scalar-induced GWs that may be sourced along with PBH formation. Whether or not this is possible depends crucially on the PBH abundance. The argument goes as follows. The amount of GWs induced at second-order depends on the square of the amplitude of the dimensionless curvature perturbation power spectrum \mathcal{P}_ζ , $\Omega_{\text{GW}} \sim \mathcal{P}_\zeta^2$. On the other hand, the abundance of PBHs is exponentially sensitive to the same amplitude, $f_{\text{PBH}} \sim \exp(-1/\mathcal{P}_\zeta)$, where f_{PBH} is the PBH abundance with respect to the total dark matter. A large stochastic background of GWs requires large values of \mathcal{P}_ζ , automatically generating a PBH abundance which tends to be too large to be compatible with the 15-year NANOGrav data, (unless some negative non-Gaussianity is introduced), see Refs. [31–38] for recent works along this direction.

Given what is at stake, the compatibility of the first discovery of a stochastic background of GWs with the PBH scenario, but also for more general reasons, a natural and fundamental question to ask is how well do we know the PBH abundance to make any definite conclusion.

The goal of this paper is to argue that there are sources of uncertainties plaguing the standard calculation of the PBH abundance, mostly coming from the details of horizon crossing and the effect of the non-linear radiation transfer function. These effects might change the PBH abundance with respect to what considered in the literature.

A quick summary of the standard PBH abundance calculation. In this section we briefly summarize what it is routinely done, and currently the best way to our knowledge, to compute the PBH abundance in the literature, see for instance Refs. [39, 40].

As we already mentioned, we will focus on the PBH formation from the collapse of sizeable overdensities that are generated during inflation and re-enter the cosmological horizon during the subsequent radiation-dominated era. One key quantity is the curvature perturbation ζ which appears in the metric in the comoving uniform-energy density gauge as

$$ds^2 = -dt^2 + a^2(t)e^{2\zeta}d\mathbf{x}^2, \quad (1)$$

where $a(t)$ is the scale factor in terms of the cosmic time. On superhorizon scales, one applies the gradient expansion [41] to relate the non-linear density contrast $\delta^{\text{com}}(r, t)$ on comoving orthogonal slicings and the time independent curvature perturbation $\zeta(r)$ as [42]

$$\delta^{\text{com}}(t, r) = -\frac{8}{9} \frac{1}{a^2 H^2} e^{-5\zeta(r)/2} \nabla^2 e^{\zeta(r)/2}, \quad (2)$$

where H is the Hubble rate. Whether or not cosmological perturbations may gravitationally collapse to form a PBH depends on the amplitude measured at the peak of the compaction function, defined to be the mass excess compared to the background value in a given radius [42–44]. On superhorizon scales it reads

$$\mathcal{C}(r) = -\frac{2}{3} r \zeta'(r) [2 + r \zeta'(r)], \quad (3)$$

where the prime stands for differentiation with respect to r . The compaction function has a maximum at the comoving length scale r_m satisfying

$$\mathcal{C}'(r_m) = 0 \quad \text{or} \quad \zeta'(r_m) + r_m \zeta''(r_m) = 0. \quad (4)$$

One can therefore define a smoothed perturbation amplitude as the volume average of the energy density contrast within the scale r_m at the cosmological horizon crossing time t_H [43]

$$\delta_m = \frac{3}{(r_m e^{\zeta(r_m)})^3} \int_0^{r_m} dr \delta^{\text{com}}(r, t_H) \left(r e^{\zeta(r)} \right)^2 \left(r e^{\zeta(r)} \right)', \quad (5)$$

where a top-hat window function is adopted to account for the treatment of the threshold [45]. This represents the main quantity determining the abundance of PBHs. When computed at the cosmological horizon crossing time t_H (the perturbations, if large enough, collapse into a PBH very rapidly after horizon crossing)

$$\epsilon(t_H) = \frac{r_H}{r_m e^{\zeta(r_m)}} = \frac{1}{r_m e^{\zeta(r_m)} a H} = 1, \quad (6)$$

it simplifies to

$$\delta_m = \delta_l - \frac{3}{8} \delta_l^2, \quad \delta_l = -\frac{4}{3} r_m \zeta'(r_m). \quad (7)$$

The PBH abundance is then computed by integrating the probability distribution function of the smoothed density contrast from a threshold value δ_c on, as

$$\beta = \int_{\delta_c} \left(\frac{M_{\text{PBH}}}{M_H} \right) P(\delta_m) d\delta_m, \quad (8)$$

in terms of the PBH mass M_{PBH} and the mass enclosed in the cosmological horizon at the horizon crossing time M_H . However, from the relation shown above, one can use the conservation of probability to write

$$P(\delta_l) d\delta_l = P(\delta_m) d\delta_m, \quad (9)$$

such that the linear smoothed density contrast δ_l is the ultimate key parameter which we have to compute the probability of [39, 40, 46–49], with a corresponding threshold given by [39]

$$\delta_{l,c} = \frac{4}{3} \left(1 - \sqrt{1 - \frac{3}{2} \delta_c} \right). \quad (10)$$

For Gaussian curvature perturbations, the probability for the linear density contrast is Gaussian and is exponentially sensitive to the threshold $\delta_{l,c}$ and the variance σ_l^2

$$P(\delta_l) = \frac{\delta_{l,c}}{\sqrt{2\pi}\sigma_l} e^{-\delta_{l,c}^2/2\sigma_l^2},$$

$$\sigma_l^2 = \int \frac{dk}{k} \mathcal{P}_{\delta_l}(k)$$

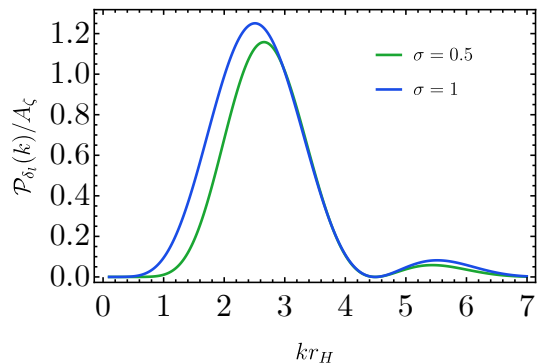


FIG. 1. The integrand of the linear variance in Eq. (11) for a lognormal curvature perturbation power spectrum $\mathcal{P}_\zeta(k) = A_\zeta \exp[-\ln(2k/3k_*)^2/2\sigma^2]$, assuming $k_* r_m = 2.1$ (1.75) for different widths $\sigma = 0.5$ (1).

$$= \frac{16}{81} \int \frac{dk}{k} (kr_H)^4 W^2(k, r_m) T^2(kr_H) \mathcal{P}_\zeta(k), \quad (11)$$

where $W(k, r_m)$ is the Fourier transform of the top-hat window function in real space with radius r_m , and

$$T(x) = 3 \frac{\sin(x/\sqrt{3}) - (x/\sqrt{3}) \cos(x/\sqrt{3})}{(x/\sqrt{3})^3} \quad (12)$$

is the linear radiation transfer function which has to be finally computed at the time of horizon crossing t_H .

A critical look at the standard PBH abundance calculation. At this stage one can outline a few inconsistencies in the standard calculation of the PBH abundance:

i) The initial density perturbation is defined on super-horizon scales, and then evolved forward through horizon re-entry to check if a PBH will form. The starting point, Eq. (2), to calculate the threshold is rigorously valid only on superhorizon scales, much before the large overdensities re-enter the Hubble radius. The threshold is computed by using the time-independent part of the profile at times such that $r_m \gg r_H$, and then extrapolating its value at horizon crossing, for which $\epsilon(t_H) = 1$, with the growth in time $\sim (aH)^{-2} \sim a^2$. This procedure based on the gradient expansion breaks down close to horizon crossing and neglects the effect of the radiation pressure, which would be increasingly more important as we approach horizon crossing. This is clear from the expression (5) where there is no radiation transfer function.

The point is that the variance σ_l^2 , which contains the information of the radiation pressure through the linear radiation transfer function, is obtained by integrating over all momenta and the integrand in its definition, Eq. (11), is peaked at scales beneath the horizon. Fig. 1 shows this point for some representative examples of comoving curvature perturbation power spectra.

This means that the variance gets contributions from scales which at $\epsilon(t_H) = 1$ are already well inside the

horizon. The action of the radiation pressure becomes so important that the density contrast stops growing, thus changing completely the time dependence with respect to the linearly extrapolated case with no radiation pressure. This is because, for the relevant scales, the time dependence in the square root of the variance is of the form (up to oscillations) $\sim (aH)^{-2}T(k/aH) \sim \text{constant}$.

The comparison between the extrapolated threshold with no transfer function and σ_l , which determines the PBH abundance, is therefore not much consistent. Even at the linear level the time dependence between the threshold, which is computed considering only cosmic expansion and without pressure effects at horizon crossing, and the square root of the variance does not cancel in the standard procedure.

One could of course compute both the density contrast and the square root of the variance on superhorizon scales to cancel the time dependence. This would require computing the variance without including the radiation transfer function [45, 50, 51]. However, the integral over the momenta in the variance would be sensitive to the ultraviolet cut-off due to the $\sim k^4$ term and therefore again sensitive to the details around horizon crossing.

Forgetting the issue introduced by the radiation pressure and accounted for by the radiation transfer function does not come without a price already at the linear level. Indeed, calling k_* the typical momentum at which PBHs form, one has typically at the linear level $k_* r_m \simeq k_*/aH \simeq \mathcal{O}(2 \div 3)$ [43, 52]¹, and $T^2(k_* r_m) \simeq (0.76 \div 0.53)$. The effect of the transfer function is not negligible because the length scale $\sim k_*^{-1}$ is already well inside the horizon when $r_m \simeq r_H$. The radiation pressure therefore tends to suppress the variance and makes it more difficult to build up perturbations able to overcome a given barrier. As we shall see, this remains true at second-order.

ii) The quantity δ_m is intrinsically of the second-order. Evaluating the effects of the radiation pressure will therefore require at least a second-order transfer function to be consistent both for the calculation of the threshold and the variance. This is a major source of uncertainty in the calculation of the PBH abundance. In the following we will adopt a perturbative approach to discuss the role of non-linearities in the evolution of the density contrast by expanding each quantity as $\mathcal{O}(t, \mathbf{x}) = \mathcal{O}_1(t, \mathbf{x}) + \mathcal{O}_2(t, \mathbf{x})/2 + \dots$ (see the Appendix for the perturbative expansion of the cosmological perturbations).²

Consider the density contrast in the comoving gauge of Eq. (2) close to the Hubble crossing time t_H and imagine an effective expansion at second-order of the form

$$\begin{aligned} \delta^{\text{com}}(t_H, \mathbf{x}) &= f_\zeta(t_H) \nabla^2 \zeta_1(\mathbf{x}) \\ &+ g_\zeta(t_H) \zeta_1(\mathbf{x}) \nabla^2 \zeta_1(\mathbf{x}) \\ &+ h_\zeta(t_H) \partial_i \zeta_1(\mathbf{x}) \partial^i \zeta_1(\mathbf{x}) + \dots, \end{aligned} \quad (13)$$

in terms of some time-dependent functions $f_\zeta(t_H)$, $g_\zeta(t_H)$, $h_\zeta(t_H)$, \dots . The variance of such a quantity will depend only (assuming $\zeta_1(\mathbf{x})$ is a Gaussian field) upon the square root of the combination

$$f_\zeta^2(t_H) \langle (\nabla^2 \zeta_1)^2 \rangle + g_\zeta^2(t_H) \langle (\zeta_1 \nabla^2 \zeta_1)^2 \rangle + \dots, \quad (14)$$

as $\zeta_1(\mathbf{x})$ is treated as a stochastic quantity. The threshold, on the other hand, is determined from the profile in real space, and will depend on the linear combination of the functions $f_\zeta(t_H)$, $g_\zeta(t_H)$, $h_\zeta(t_H)$, \dots . Since close to the horizon the functions are not expected to have the same time dependence, the cancellation operating at the linear level will be in general no longer operative;

iii) The threshold δ_c (or $\delta_{l,c}$) is routinely calculated extrapolating its value at horizon crossing in an unperturbed universe. However, it has been already noted that the non-linear effects arising at the true horizon crossing in an inhomogeneous universe, where the gradient expansion fails, increase the value of the threshold by about a factor $\mathcal{O}(1.7 \div 2)$ with respect to the one computed on superhorizon scales and linearly extrapolated at $\epsilon(t_H) = 1$ [54] (see their Fig. 8). Part of this extra growth is due to the longer time necessary to reach the non-linear horizon crossing, part is due to higher orders in the gradient expansion when $\epsilon(t_H) \sim 1$ and finally part to non-linear effects due to the radiation pressure, which are accounted for numerically in Ref. [54]. In general, increasing the critical threshold by a factor $\mathcal{O}(1.7 \div 2)$ might drastically reduce the PBH abundance, unless the variance of the density contrast does increase correspondingly.

Overall, non-linear effects become important close to horizon crossing and may give rise to large corrections to the probability of collapse estimated in the literature.

The rest of the paper is dedicated to inspect more closely such sources of uncertainties due to non-linear effects. We start by commenting about the last point we mentioned.

Crossing the real horizon later. There are various effects to consider. First, the physical horizon in a perturbed universe expands more slowly than in an unperturbed universe, making the density profile grow more. Secondly, non-linear effects and the effect of the radiation pressure close to the horizon crossing change the density contrast profile; both lead to an increase of the critical threshold, as discussed numerically in Section 5 of Ref. [54].

¹ For a monochromatic curvature spectrum peaked at a given momentum, k_* coincides with such a momentum and $k_* r_m \simeq 2.7$, while for a broad spectrum it coincides with the maximum momentum scale, as the PBH mass function peaks at that scale [53], and $k_* r_m \simeq 3.5$.

² This perturbative scheme will apply as well to the treatment of $r_m e^{\zeta(r_m)}$ in the definition of $\epsilon(t_H)$, such that at the linear level $r_m \simeq r_H$ at the time of horizon crossing.

We discuss here the first effect, which is already present at horizon crossing at the linear level, and we will come back to the second effect in the next section.

Consider the equation describing the expansion rate $H(t, \mathbf{x})$ as measured by a free-falling observer which is instantaneously at rest with respect to the radiation fluid, i.e. on comoving slicings [55]

$$H^2(t, \mathbf{x}) = H^2(t) \left(1 + \delta_1^{\text{com}}(t, \mathbf{x}) \right) + \frac{2}{3} \frac{\nabla^2}{a^2} \zeta_1(t, \mathbf{x}). \quad (15)$$

Since for large amplitudes the peaks of the density contrast are identifiable with the peaks of the comoving curvature perturbation [56] for which $\nabla^2 \zeta_1 < 0$, we automatically deduce that the physical Hubble rate is smaller than the background value provided that the last term in Eq. (15) is larger than the term $H^2(t) \delta_1^{\text{com}}$. One can easily convince oneself that this is indeed the case.

At the linear level and for any scale (not only super-horizon), one can relate the linear density contrast in the comoving gauge to the linear gauge-invariant Bardeen's potential $\Phi_1(t, \mathbf{x})$ through the following equation [57] (see also the Appendix):

$$\delta_1^{\text{com}}(t, \mathbf{x}) = \frac{2}{3} \frac{\nabla^2 \Phi_1(t, \mathbf{x})}{a^2 H^2}. \quad (16)$$

Approximating at horizon crossing $\Phi_1(t_H, \mathbf{x}) \simeq -2\zeta_1(\mathbf{x})/3$, one obtains

$$H^2(t_H, \mathbf{x}) \simeq H^2(t_H) \left(1 - \frac{1}{2} \delta_1^{\text{com}}(t_H, \mathbf{x}) \right). \quad (17)$$

One can estimate the effect on the threshold by averaging over a sphere of radius r_m and assuming the critical profile for $\delta_1^{\text{com}}(t_H, \mathbf{x})$, obtaining

$$\frac{3}{r_m^3} \int_0^{r_m} dr r^2 \left(\frac{H^2(t_H, \mathbf{x})}{H^2(t_H)} - 1 \right) \simeq -\frac{1}{2} \delta_c, \quad (18)$$

from which one can see already the effect of the curvature term in slowing down the Hubble radius expansion. In the last equation, to be consistent with the linear perturbation theory adopted so far, one needs to take the value of δ_c obtained from the linear Gaussian threshold [54] for the volume average of the first-order density contrast $\delta_1^{\text{com}}(t_H, \mathbf{x})$.

One can do a better estimate including the effect of higher-order gradients, still remaining at the linear level. At horizon crossing, the comoving curvature term starts growing due to the scalar shear and the effect of higher-order gradients starts becoming manifest [57]

$$\frac{\dot{\zeta}_1}{H} \simeq \frac{1}{4} \delta_1^{\text{com}} - \frac{3}{4} \frac{\nabla^2 \delta_1^{\text{com}}}{a^2 H^2}, \quad (19)$$

where the dot denotes differentiation with respect to the cosmic time. The rate $\dot{\zeta}_1$ is positive since the second term grows faster (as a^2) than the first one and around the

peak of the density contrast $\nabla^2 \delta_1^{\text{com}} < 0$. Integrating over time and inserting the result in Eq. (15) gives

$$\left(\frac{H^2(t_H, \mathbf{x})}{H^2(t_H)} - 1 \right) \simeq \delta_1^{\text{com}} + \frac{1}{12} \frac{\nabla^2 \delta_1^{\text{com}}}{a^2 H^2} - \frac{1}{4} \frac{\nabla^4 \delta_1^{\text{com}}}{a^4 H^4}. \quad (20)$$

When the perturbation has just crossed the horizon, the last two terms dominate and averaging again over a sphere of radius r_m we obtain an extra shift in the Hubble rate

$$\begin{aligned} & \frac{3}{r_m^3} \int_0^{r_m} dr r^2 \left(\frac{H^2(t_H, \mathbf{x})}{H^2(t_H)} - 1 \right) \\ & \simeq \frac{7}{4} r_m \delta_1^{\text{com}}(r_m) - \frac{3}{2} r_m^2 \delta_1^{\prime\prime\text{com}}(r_m) - \frac{3}{4} r_m^3 \delta_1^{\prime\prime\prime\text{com}}(r_m) \\ & \sim -\frac{5}{2} \delta_1^{\text{com}}(r_m) \sim -\frac{5}{6} \delta_c, \end{aligned} \quad (21)$$

where we have approximated $r_m \delta_1^{\text{com}}(r_m) \sim -r_m^2 \delta_1^{\prime\prime\text{com}}(r_m) \sim r_m^3 \delta_1^{\prime\prime\prime\text{com}}(r_m) \sim -\delta_1^{\text{com}}(r_m)$ for a peaked perturbation, and used the relation $\delta_c = 3\delta_1^{\text{com}}(t_H, r_m)$ for the critical profile [43], to get an order of magnitude estimate. Taking $\delta_c \simeq 0.51$ for the linear Gaussian threshold [54], we obtain a change in the horizon-crossing time $\epsilon(t_H) \propto 1/H$ of

$$\epsilon(t_H) \simeq \frac{1}{(1 - 5\delta_c/6)^{1/2}} \simeq 1.3, \quad (22)$$

which is already in good agreement with what was numerically found in Ref. [54] (see their Fig. 5, where a monochromatic spectrum corresponds formally to the case $\alpha = 6.33$).

Being the variance dominated by the momentum modes well inside the horizon at $r_m = r_H$, when the linear density contrast has already stopped growing, increasing $\epsilon(t_H)$ will not necessarily change the linear variance by the same factor. For instance, for a broad spectrum of the curvature perturbation, the linear variance does not change shifting r_H as the integral is over the momenta kr_H once r_m is set to be equal to r_H . For a monochromatic power spectrum $\mathcal{P}_\zeta(k) = A_\zeta k_* \delta(k - k_*)$, the change of the variance is set by the square of the radiation transfer function and it does not automatically cancel the change of $\epsilon^2(t_H)$ in the critical threshold.

We do not attempt here to calculate the effect of nonlinearities on the slowing down of horizon crossing, which have been studied numerically in Ref. [54]. We limit ourselves to the observation that one expects corrections of the order of $\mathcal{O}(\delta_c^2)$, see for instance Eq. (A.21) of the Appendix. For instance, the leading (and growing in time) second-order correction to the comoving curvature perturbation at horizon crossing has the form [57–59]

$$\frac{2}{3} \nabla^2 \zeta_2(t, \mathbf{x}) \simeq -\frac{1}{48a^2 H^2} \sum_{i \neq j} \partial_i \partial_j \delta_1^{\text{com}}(t, \mathbf{x}) \partial^i \partial^j \delta_1^{\text{com}}(t, \mathbf{x}), \quad (23)$$

which is again negative and increasing like a^2 . Its effect is to further slow down the expansion rate and increase the PBH threshold, but only for non-spherical peaks.

The non-linear transfer function: the second-order case. As we have stressed in the introduction, one of the main uncertainties in the calculation of the PBH abundance is that the effect of the radiation pressure is standardly either not accounted for (in the threshold, with the exception of Ref. [54]) or it is, but only at the linear level (in the variance). We would like to offer some considerations about the non-linear effects, limiting ourselves to the second-order.

Our starting point is the second-order equation for the density contrast in the comoving orthogonal gauge expressed in terms of the Bardeen potentials constructed from the longitudinal gauge (see the Appendix)

$$\begin{aligned} \frac{3}{2}a^2H^2\delta_2^{\text{com}} &= \nabla^2\Psi_2 + 3\partial_i\Phi_1\partial^i\Phi_1 + 8\Phi_1\nabla^2\Phi_1 \\ &+ 3\dot{\Phi}_1^2 + \frac{6H}{\nabla^2}(\partial^i\dot{\Phi}_1\partial_i\Phi_1 + \dot{\Phi}_1\nabla^2\Phi_1). \end{aligned} \quad (24)$$

This expression is valid at all scales, not only on super-horizon. It shows a fundamental point, that close to horizon crossing the overdensity on comoving orthogonal slices contains many non-linear terms, even non-local, and therefore it is unlikely that the non-linear smoothed density contrast will have a quadratic relation of the local type with the linear one, as in Eq. (7). In other words, Eq. (7), obtained on superhorizon scales, fails to capture the non-linearities in the critical threshold due to the radiation pressure. This implies that neither δ_c nor $\delta_{l,c}$ may be used and that a priori there is no simple relation of the local type between the linear density contrast and the non-linear one.

Eq. (24) needs to be supplemented with the equation for the evolution of the second-order gravitational potential Ψ_2 [59, 60]

$$\begin{aligned} \ddot{\Psi}_2 + 5H\dot{\Psi}_2 - \frac{1}{3}\frac{\nabla^2\Psi_2}{a^2} &= \frac{1}{a^2}\left[\frac{2}{3}\partial^i\Phi_1\partial_i\Phi_1 + \frac{8}{3}\Phi_1\nabla^2\Phi_1\right. \\ &+ 2\dot{\Phi}_1^2 + HN^j{}_i\left(A_j^{(2)i}\right)^k \\ &+ \left.\frac{1}{3}N^j{}_i\left(A_j^{(2)i}\right)^k\right] \quad (25) \\ A_j^{(2)i} &= 6\partial^i\Phi_1\partial_j\Phi_1 + \frac{2}{H}(\partial^i\Phi_1\partial_j\Phi_1) \\ &+ \frac{2}{H^2}\partial^i\dot{\Phi}_1\partial_j\dot{\Phi}_1, \end{aligned}$$

where

$$N^j{}_i = \frac{3}{2}\frac{1}{\nabla^2}\left(\frac{\partial^j\partial_i}{\nabla^2} - \frac{1}{3}\delta^j{}_i\right). \quad (26)$$

Eq. (25) has to be solved with the superhorizon initial condition [60, 61]

$$\Psi_2(\mathbf{x}) = -2\Phi_1^2(\mathbf{x}) + 2N^j{}_i(\partial^i\Phi_1(\mathbf{x})\partial_j\Phi_1(\mathbf{x})). \quad (27)$$

In momentum space the solution is the sum of two pieces

$$\Psi_2(t, \mathbf{k}) = \Psi_2(\mathbf{k})T(k/aH)$$

$$+ \int \frac{d^3p}{(2\pi)^3} uv \left(\frac{2}{3}\right)^2 I_\Psi\left(u, v, \frac{k}{aH}\right) \zeta_1(\mathbf{p})\zeta_1(\mathbf{k}-\mathbf{p}), \quad (28)$$

where $u = |\mathbf{k} - \mathbf{p}|/k$, $v = p/k$. This equation shows the dependence on the linear radiation transfer function in the first term and of the second-order radiation transfer function through the function $I_\Psi(u, v, x)$ [59–61].

To account for the effect of the radiation pressure in the PBH threshold at second-order in perturbation theory we proceed as follows. At first-order the equation of motion of the gravitational potential is

$$\ddot{\Phi}_1 + 5H\dot{\Phi}_1 - \frac{1}{3}\frac{\nabla^2\Phi_1}{a^2} = 0. \quad (29)$$

We can treat the last term as a perturbation, which gives close to horizon crossing

$$\begin{aligned} \Phi_1(t, \mathbf{x}) &= \int^t \frac{dt'}{a^5(t')} \int^{t'} dt'' a^5(t'') \frac{1}{3} \frac{\nabla^2\Phi_1(\mathbf{x})}{a^2(t'')} \\ &= \left(1 + \frac{1}{30} \frac{\nabla^2}{a^2 H^2}\right) \Phi_1(\mathbf{x}). \end{aligned} \quad (30)$$

This is quite a good approximation: going to momentum space and evaluating at $k_* r_m \simeq k_* r_H = 2.74$ for a monochromatic power spectrum, we find $\Phi_1(t, \mathbf{x}) \simeq 0.75\Phi_1(\mathbf{x})$, while the linear radiation transfer function gives $T(2.74)\Phi_1(\mathbf{x}) = 0.77\Phi_1(\mathbf{x})$.

At second-order, proceeding in a similar way to obtain $\Psi_2(t, \mathbf{x})$ by using Eq. (27), then shown in Eq. (A.25), we obtain (see the Appendix for details)

$$\begin{aligned} \delta_2^{\text{com}} &\simeq \frac{1}{a^2 H^2} \left(\frac{2}{3} \nabla^2 \Psi_2 + 2\partial_i \Phi_1 \partial^i \Phi_1 + \frac{16}{3} \Phi_1 \nabla^2 \Phi_1 \right) \\ &= \frac{1}{a^2 H^2} \left(-\frac{4}{3} \nabla^2 \Phi_1^2(\mathbf{x}) + 2 \frac{\partial^j \partial_i}{\nabla^2} (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) \right. \\ &+ \frac{4}{3} \partial^i \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) + \frac{16}{3} \Phi_1 \nabla^2 \Phi_1(\mathbf{x}) \\ &+ \frac{1}{a^4 H^4} \left(\frac{2}{3} \partial^i \nabla^2 \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) \right. \\ &+ \left. \left. \frac{8}{15} \nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \frac{2}{15} \Phi_1(\mathbf{x}) \nabla^4 \Phi_1(\mathbf{x}) \right) \right). \end{aligned} \quad (31)$$

When evaluated at the time of horizon crossing $aH = 1/r_H \simeq 1/r_m$, the first-order curvature profile $\Phi_1(t_H, \mathbf{x}) \simeq -2\zeta_1(\mathbf{x})/3$ in Fourier space is dominated by large momenta, see for instance Fig. 1 for different shapes of a lognormal power spectrum. This allows us to focus only on the last three lines in the previous expression, which are expected to dominate for large momenta.

At horizon crossing the density contrast becomes

$$\delta_2^{\text{com}}(t_H, \mathbf{x}) \simeq \frac{4r_m^4}{9} \left(\frac{2}{3} \partial^i \nabla^2 \zeta_1(\mathbf{x}) \partial_i \zeta_1(\mathbf{x}) \right)$$

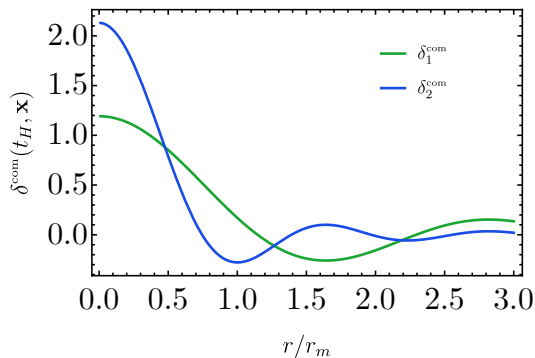


FIG. 2. The density profile corresponding to the critical threshold at first- and second-order in perturbation theory, for a monochromatic curvature spectrum.

$$+ \frac{8}{15} \nabla^2 \zeta_1(\mathbf{x}) \nabla^2 \zeta_1(\mathbf{x}) + \frac{2}{15} \zeta_1(\mathbf{x}) \nabla^4 \zeta_1(\mathbf{x}) \Big). \quad (32)$$

Adopting Gaussian peak theory to a monochromatic curvature perturbation power spectrum, for which $k_* r_m \simeq 2.74$, one can derive the shape for the linear curvature profile [54]

$$\zeta_1(\mathbf{x}) = \zeta_0 \frac{\sin(k_* r)}{k_* r}, \quad (33)$$

from which we obtain the critical density profile at second-order shown in Fig. 2, where the overall amplitude ζ_0 has been fixed at the linear level by requiring that $\delta_c = 0.51 = 3\delta_1^{\text{com}}(t_H, r_m)$ [43]. The impact of the second-order radiation transfer function on the PBH threshold is estimated to be

$$\frac{3}{r_m^3} \int_0^{r_m} dr r^2 \delta_2^{\text{com}}(t_H, r) \simeq 0.13. \quad (34)$$

This result shows that the second-order radiation transfer function increases further the PBH threshold for a monochromatic spectrum, due to the radiation pressure which makes the collapse of the overdensities into a PBH more difficult. To this effect one should further add the one described in the previous section related to the slowing down of the horizon. The corresponding critical threshold will be given roughly by (the factor 1/2 comes from the normalization of the second-order variables)

$$\epsilon^2(t_H) \delta_c + \frac{1}{2} \epsilon^4(t_H) \frac{3}{r_m^3} \int_0^{r_m} dr r^2 \delta_2^{\text{com}} \simeq 1.08, \quad (35)$$

where the corrective factor $\epsilon^4(t_H)$ has been included for the slowing down of the horizon based on the dependence $1/(aH)^4$ in Eq. (32). This result is in good agreement with the numerical fit for the PBH threshold, $\alpha^{0.06} + 0.025 \simeq 1.14$, for a monochromatic curvature power spectrum with $\alpha = 6.33$, provided in Ref. [54] (see also the left panel of their Fig. 8).

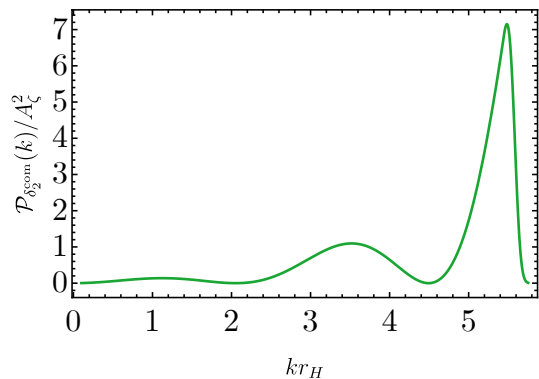


FIG. 3. The integrand of the second-order variance in Eq. (38) for a monochromatic curvature perturbation power spectrum evaluated at the time of horizon crossing.

As for the variance at second-order, we offer the following considerations. When a given scale re-enters the horizon, the linear gravitational potential $\Phi_1(t, \mathbf{x})$ decays like a^{-2} (up to oscillations). By inspecting the time dependence of Eq. (28) it is easy to convince ourselves that the second-order gravitational potential $\Psi_2(t, \mathbf{x})$ decays like a^{-2} as well [59]. As a consequence, one can keep the leading time dependent term in Eq. (31) and approximate

$$\frac{3}{2} a^2 H^2 \delta_2^{\text{com}} \simeq \nabla^2 \Psi_2, \quad (36)$$

because all the other terms decay faster at horizon crossing. From the solution of the second-order scalar perturbation in Eq. (28), one can estimate the variance of the second-order smoothed density contrast in the comoving gauge at horizon crossing by computing its two-point connected correlation function, which gives

$$\begin{aligned} \sigma_2^2 &= \int \frac{dk}{k} \mathcal{P}_{\delta_2^{\text{com}}}(k, r_H), \\ \mathcal{P}_{\delta_2^{\text{com}}}(k, r_H) &= \int_0^\infty dv \int_{|v-1|}^{v+1} du \left(\frac{2}{3}\right)^6 (kr_H)^4 I_\Psi^2(u, v, kr_H) \\ &\quad \cdot W^2(k, r_m) \mathcal{P}_{\zeta_1}(kv) \mathcal{P}_{\zeta_1}(ku). \end{aligned} \quad (37)$$

In Fig. 3 we plot the power spectrum of the second-order smoothed density contrast in the comoving gauge for a monochromatic curvature perturbation power spectrum, for which we get

$$\begin{aligned} \mathcal{P}_{\delta_2^{\text{com}}}(k, r_H) &= \left(\frac{2}{3}\right)^6 (kr_H)^4 \left(\frac{k_*}{k}\right)^2 W^2(k_*, r_m) \\ &\quad \cdot I_\Psi^2\left(\frac{k_*}{k}, \frac{k_*}{k}, k_* r_H\right) \theta\left(2 - \frac{k}{k_*}\right), \end{aligned} \quad (38)$$

evaluated for $k_* r_m \simeq k_* r_H = 2.74$, where θ indicates the Heaviside step function. The figure confirms that, as in the linear case, the variance is dominated by momenta larger than the Hubble rate where the second-order radiation transfer function is non-negligible. At this stage the

time dependence $\sim a^{-2}$ of the gravitational potential Ψ_2 is the same as at the linear level and the second-order variance does not change as well with time (up to oscillations). The values of the linear and second-order variances for $k_* r_m = 2.74$ are (the factor 1/4 in the second-order term accounts for the normalization of the second-order density contrast)

$$\sigma_l^2 \simeq 1.19 (0.41) A_\zeta \quad \text{and} \quad \frac{1}{4} \sigma_2^2 \simeq 0.25 (1.02) A_\zeta^2, \quad (39)$$

where in the parenthesis we have indicated the values when the Hubble rate is shifted by a factor $\epsilon(t_H) \simeq 1.3$ due to the crossing of the physical horizon, as discussed in the previous section. We deduce that the second-order contribution is always negligible compared to the linear one, even for the preferred value of $A_\zeta \simeq 6 \cdot 10^{-2}$ to explain the recent PTA data [31]. We think this conclusion is quite robust and does not depend on the details of the time of horizon crossing and the shape of the comoving curvature power spectrum, as long as $k_* r_m \gtrsim 1$. For instance, for a broad spectrum, even at second-order the variance is insensitive to changes in the physical horizon as it is an integral over all the momenta kr_H once one sets $r_m = r_H$ at the corresponding order in perturbation theory.

As a final remark we notice that, because of the non-Gaussian nature of the second-order density contrast δ_2^{om} , the expression for the PBH probability will depend also on higher-order cumulants of the density field beyond the two-point function. We leave the investigation of the form of such a probability to future work.

Conclusions. In this article we have addressed various issues related to the calculation of the abundance of PBHs and point out the uncertainties the standard procedure suffers of. They come above all from the incomplete treatment of the radiation pressure and of the non-linear effects close to horizon crossing.

Even if our calculations are certainly incomplete, we can offer the following hints: *i)* the critical threshold is larger than what routinely assumed. We have provided some physical explanations for such an increase; *ii)* The variance of the non-linear density contrast is unlikely to be changed with respect to the linear one. This is because it is suppressed by an extra power of the curvature perturbation power spectrum; *iii)* The time dependence between the critical threshold and the square root of the variance does not cancel at horizon crossing as the latter is dominated by momenta larger than the horizon, when the density contrast has already frozen in (up to oscillations). This is because the characteristic length scale is such that $k_*^{-1} \gtrsim r_m \simeq r_H$. The reason why the linear and non-linear variances are suppressed is because of the action of the radiation pressure at the scales close to k_* . Therefore, it is difficult to build up fluctuations which overcome a given barrier; *iv)* The relation between

the non-linear and linear threshold is not of the local type, therefore making the calculation of the formation probability not straightforward; *v)* Given the fact that the non-linear effects increase the critical threshold, but not the variance, it is likely that the PBH abundance is much smaller than what considered in the literature so far.

If confirmed by a more complete all-order (most likely numerical) calculation, such conclusions would make the PTA data compatible with the PBH scenario, since the same amplitude of the stochastic GW background would correspond to much smaller PBH abundances.

Acknowledgments. We thank I. Musco for several discussions about the procedure to calculate the critical threshold and I. Musco and G. Franciolini for comments on the draft. V.DL. is supported by funds provided by the Center for Particle Cosmology at the University of Pennsylvania. A.R. is supported by the Boninchi Foundation for the project ‘‘PBHs in the Era of GW Astronomy’’.

* vdeluca@sas.upenn.edu

† kehagias@central.ntua.gr

‡ antonio.riotto@unige.ch

- [1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. Lett.* **116**, 061102 (2016), [arXiv:1602.03837 \[gr-qc\]](https://arxiv.org/abs/1602.03837).
- [2] B. P. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. X* **9**, 031040 (2019), [arXiv:1811.12907 \[astro-ph.HE\]](https://arxiv.org/abs/1811.12907).
- [3] R. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. X* **11**, 021053 (2021), [arXiv:2010.14527 \[gr-qc\]](https://arxiv.org/abs/2010.14527).
- [4] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), (2021), [arXiv:2111.03606 \[gr-qc\]](https://arxiv.org/abs/2111.03606).
- [5] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, *Class. Quant. Grav.* **35**, 063001 (2018), [arXiv:1801.05235 \[astro-ph.CO\]](https://arxiv.org/abs/1801.05235).
- [6] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, *Rept. Prog. Phys.* **84**, 116902 (2021), [arXiv:2002.12778 \[astro-ph.CO\]](https://arxiv.org/abs/2002.12778).
- [7] A. M. Green and B. J. Kavanagh, *J. Phys. G* **48**, 043001 (2021), [arXiv:2007.10722 \[astro-ph.CO\]](https://arxiv.org/abs/2007.10722).
- [8] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, *Phys. Rev. Lett.* **116**, 201301 (2016), [arXiv:1603.00464 \[astro-ph.CO\]](https://arxiv.org/abs/1603.00464).
- [9] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, *Phys. Rev. Lett.* **117**, 061101 (2016), [Erratum: *Phys.Rev.Lett.* 121, 059901 (2018)], [arXiv:1603.08338 \[astro-ph.CO\]](https://arxiv.org/abs/1603.08338).
- [10] S. Clesse and J. García-Bellido, *Phys. Dark Univ.* **15**, 142 (2017), [arXiv:1603.05234 \[astro-ph.CO\]](https://arxiv.org/abs/1603.05234).
- [11] Y. Ali-Haïmoud, E. D. Kovetz, and M. Kamionkowski, *Phys. Rev. D* **96**, 123523 (2017), [arXiv:1709.06576 \[astro-ph.CO\]](https://arxiv.org/abs/1709.06576).
- [12] G. Hütsi, M. Raidal, V. Vaskonen, and H. Veermäe, *JCAP* **03**, 068 (2021), [arXiv:2012.02786 \[astro-ph.CO\]](https://arxiv.org/abs/2012.02786).
- [13] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *JCAP* **05**, 003 (2021), [arXiv:2102.03809 \[astro-ph.CO\]](https://arxiv.org/abs/2102.03809).

- [14] G. Franciolini, V. Baibhav, V. De Luca, K. K. Y. Ng, K. W. K. Wong, E. Berti, P. Pani, A. Riotto, and S. Vitale, *Phys. Rev. D* **105**, 083526 (2022), arXiv:2105.03349 [gr-qc].
- [15] G. Franciolini, I. Musco, P. Pani, and A. Urbano, *Phys. Rev. D* **106**, 123526 (2022), arXiv:2209.05959 [astro-ph.CO].
- [16] Z.-C. Chen and Q.-G. Huang, *JCAP* **08**, 039 (2020), arXiv:1904.02396 [astro-ph.CO].
- [17] O. Pujolas, V. Vaskonen, and H. Veermäe, *Phys. Rev. D* **104**, 083521 (2021), arXiv:2107.03379 [astro-ph.CO].
- [18] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *JCAP* **11**, 039 (2021), arXiv:2106.13769 [astro-ph.CO].
- [19] S. Barsanti, V. De Luca, A. Maselli, and P. Pani, *Phys. Rev. Lett.* **128**, 111104 (2022), arXiv:2109.02170 [gr-qc].
- [20] S. S. Bavera, G. Franciolini, G. Cusin, A. Riotto, M. Zevin, and T. Fragos, *Astron. Astrophys.* **660**, A26 (2022), arXiv:2109.05836 [astro-ph.CO].
- [21] G. Domènech, *Universe* **7**, 398 (2021), arXiv:2109.01398 [gr-qc].
- [22] G. Agazie *et al.* (NANOGrav), *Astrophys. J. Lett.* **951**, L8 (2023), arXiv:2306.16213 [astro-ph.HE].
- [23] G. Agazie *et al.* (NANOGrav), *Astrophys. J. Lett.* **951**, L9 (2023), arXiv:2306.16217 [astro-ph.HE].
- [24] J. Antoniadis *et al.* (EPTA), (2023), arXiv:2306.16214 [astro-ph.HE].
- [25] J. Antoniadis *et al.* (EPTA), (2023), arXiv:2306.16224 [astro-ph.HE].
- [26] J. Antoniadis *et al.* (EPTA), (2023), arXiv:2306.16227 [astro-ph.HE].
- [27] D. J. Reardon *et al.*, *Astrophys. J. Lett.* **951**, L6 (2023), arXiv:2306.16215 [astro-ph.HE].
- [28] A. Zic *et al.* (PPTA), (2023), arXiv:2306.16230 [astro-ph.HE].
- [29] D. J. Reardon *et al.*, *Astrophys. J. Lett.* **951**, L7 (2023), arXiv:2306.16229 [astro-ph.HE].
- [30] H. Xu *et al.*, *Res. Astron. Astrophys.* **23**, 075024 (2023), arXiv:2306.16216 [astro-ph.HE].
- [31] G. Franciolini, A. Iovino, Junior., V. Vaskonen, and H. Veermäe, (2023), arXiv:2306.17149 [astro-ph.CO].
- [32] L. Liu, Z.-C. Chen, and Q.-G. Huang, (2023), arXiv:2307.01102 [astro-ph.CO].
- [33] S. Wang, Z.-C. Zhao, J.-P. Li, and Q.-H. Zhu, (2023), arXiv:2307.00572 [astro-ph.CO].
- [34] Y.-F. Cai, X.-C. He, X. Ma, S.-F. Yan, and G.-W. Yuan, (2023), arXiv:2306.17822 [gr-qc].
- [35] K. Inomata, K. Kohri, and T. Terada, (2023), arXiv:2306.17834 [astro-ph.CO].
- [36] D. G. Figueroa, M. Pieroni, A. Ricciardone, and P. Simakachorn, (2023), arXiv:2307.02399 [astro-ph.CO].
- [37] Z. Yi, Q. Gao, Y. Gong, Y. Wang, and F. Zhang, (2023), arXiv:2307.02467 [gr-qc].
- [38] Q.-H. Zhu, Z.-C. Zhao, and S. Wang, (2023), arXiv:2307.03095 [astro-ph.CO].
- [39] S. Young, I. Musco, and C. T. Byrnes, *JCAP* **11**, 012 (2019), arXiv:1904.00984 [astro-ph.CO].
- [40] M. Biagetti, V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto, *Phys. Lett. B* **820**, 136602 (2021), arXiv:2105.07810 [astro-ph.CO].
- [41] M. Shibata and M. Sasaki, *Phys. Rev. D* **60**, 084002 (1999), arXiv:gr-qc/9905064.
- [42] T. Harada, C.-M. Yoo, T. Nakama, and Y. Koga, *Phys. Rev. D* **91**, 084057 (2015), arXiv:1503.03934 [gr-qc].
- [43] I. Musco, *Phys. Rev. D* **100**, 123524 (2019), arXiv:1809.02127 [gr-qc].
- [44] A. Escrivà, C. Germani, and R. K. Sheth, *Phys. Rev. D* **101**, 044022 (2020), arXiv:1907.13311 [gr-qc].
- [45] S. Young, *Int. J. Mod. Phys. D* **29**, 2030002 (2019), arXiv:1905.01230 [astro-ph.CO].
- [46] C. Germani and R. K. Sheth, *Phys. Rev. D* **101**, 063520 (2020), arXiv:1912.07072 [astro-ph.CO].
- [47] V. De Luca and A. Riotto, *Phys. Lett. B* **828**, 137035 (2022), arXiv:2201.09008 [astro-ph.CO].
- [48] G. Ferrante, G. Franciolini, A. Iovino, Junior., and A. Urbano, *Phys. Rev. D* **107**, 043520 (2023), arXiv:2211.01728 [astro-ph.CO].
- [49] A. D. Gow, H. Assadullahi, J. H. P. Jackson, K. Koyama, V. Vennin, and D. Wands, *EPL* **142**, 49001 (2023), arXiv:2211.08348 [astro-ph.CO].
- [50] C.-M. Yoo, T. Harada, J. Garriga, and K. Kohri, *PTEP* **2018**, 123E01 (2018), arXiv:1805.03946 [astro-ph.CO].
- [51] C.-M. Yoo, T. Harada, S. Hirano, and K. Kohri, *PTEP* **2021**, 013E02 (2021), arXiv:2008.02425 [astro-ph.CO].
- [52] C. Germani and I. Musco, *Phys. Rev. Lett.* **122**, 141302 (2019), arXiv:1805.04087 [astro-ph.CO].
- [53] V. De Luca, G. Franciolini, and A. Riotto, *Phys. Lett. B* **807**, 135550 (2020), arXiv:2001.04371 [astro-ph.CO].
- [54] I. Musco, V. De Luca, G. Franciolini, and A. Riotto, *Phys. Rev. D* **103**, 063538 (2021), arXiv:2011.03014 [astro-ph.CO].
- [55] A. R. Liddle and D. H. Lyth, *Cosmological inflation and large scale structure* (2000).
- [56] V. De Luca, G. Franciolini, A. Kehagias, M. Peloso, A. Riotto, and C. Ünal, *JCAP* **07**, 048 (2019), arXiv:1904.00970 [astro-ph.CO].
- [57] K. A. Malik and D. Wands, *Phys. Rept.* **475**, 1 (2009), arXiv:0809.4944 [astro-ph].
- [58] N. Bartolo, S. Matarrese, and A. Riotto, *JCAP* **01**, 003 (2004), arXiv:astro-ph/0309692.
- [59] K. Inomata, *JCAP* **03**, 013 (2021), arXiv:2008.12300 [gr-qc].
- [60] N. Bartolo, S. Matarrese, and A. Riotto, *JCAP* **01**, 019 (2007), arXiv:astro-ph/0610110.
- [61] N. Bartolo, S. Matarrese, and A. Riotto, *JCAP* **06**, 024 (2006), arXiv:astro-ph/0604416.
- [62] J. M. Bardeen, *Phys. Rev. D* **22**, 1882 (1980).

APPENDIX

The components of a perturbed spatially flat Robertson-Walker metric can be written, up to second-order and using conformal time, as (primes indicate here differentiation with respect to the conformal time τ) [60]

$$\begin{aligned} g_{00} &= -a^2(\tau) (1 + 2\phi_1 + \phi_2), \\ g_{0i} &= a^2(\tau) \left(\hat{\omega}_{1i} + \frac{1}{2} \hat{\omega}_{2i} \right), \\ g_{ij} &= a^2(\tau) \left[(1 - 2\psi_1 - \psi_2) \delta_{ij} + \left(\hat{\chi}_{1ij} + \frac{1}{2} \hat{\chi}_{2ij} \right) \right]. \end{aligned} \quad (\text{A.1})$$

The standard splitting of the perturbations into scalar, transverse (*i.e.* divergence-free) vector, and transverse trace-free tensor parts, with respect to the three-dimensional space with metric δ_{ij} , can be performed in the following way:

$$\begin{aligned} \hat{\omega}_i &= \partial_i \omega + \omega_i, \\ \hat{\chi}_{ij} &= D_{ij} \chi + \partial_i \chi_j + \partial_j \chi_i + \chi_{ij}, \end{aligned} \quad (\text{A.2})$$

where ω_i and χ_i are transverse vectors ($\partial^i \omega_i = \partial^i \chi_i = 0$), χ_{ij} is a symmetric transverse and trace-free tensor ($\partial^i \chi_{ij} = 0$, $\chi_i^i = 0$) and $D_{ij} = \partial_i \partial_j - (1/3) \delta_{ij} \partial^k \partial_k$ is a trace-free operator. Here and in the following latin indices are raised and lowered using δ^{ij} and δ_{ij} , respectively. For our purposes the metric in Eq. (A.1) can be simplified. In fact, first-order vector perturbations are not generated during inflation and tensor modes are negligible. Thus, in the following we can neglect ω_{1i} , χ_{1i} and χ_{1ij} . However the same is not true for the second-order perturbations. In the second-order theory the second-order vector and tensor contributions can be generated by the first-order scalar perturbations even if they are initially zero. Thus we have to take them into account and we shall use the metric

$$\begin{aligned} g_{00} &= -a^2(\tau) (1 + 2\phi_1 + \phi_2), \\ g_{0i} &= a^2(\tau) \left(\partial_i \omega_1 + \frac{1}{2} \partial_i \omega_2 + \frac{1}{2} \omega_{2i} \right), \\ g_{ij} &= a^2(\tau) \left[(1 - 2\psi_1 - \psi_2) \delta_{ij} + D_{ij} \left(\chi_1 + \frac{1}{2} \chi_2 \right) + \frac{1}{2} (\partial_i \chi_{2j} + \partial_j \chi_{2i} + \chi_{2ij}) \right]. \end{aligned} \quad (\text{A.3})$$

The first-order perturbations of the Einstein tensor components which we need are

$$\delta_1 G^0_0 = \frac{1}{a^2} \left[6 \left(\frac{a'}{a} \right)^2 \phi_1 + 6 \frac{a'}{a} \psi'_1 + 2 \frac{a'}{a} \partial_i \partial^i \omega_1 - 2 \partial_i \partial^i \psi_1 - \frac{1}{2} \partial_k \partial^i D^k_i \chi_1 \right], \quad (\text{A.4})$$

$$\delta_1 G^0_i = \frac{1}{a^2} \left(-2 \frac{a'}{a} \partial_i \phi_1 - 2 \partial_i \psi'_1 - \frac{1}{2} \partial_k D^k_i \chi'_1 \right). \quad (\text{A.5})$$

At second-order they are

$$\begin{aligned} \delta_2 G^0_0 &= \frac{1}{a^2} \left[3 \left(\frac{a'}{a} \right)^2 \phi_2 + 3 \frac{a'}{a} \psi'_2 - \partial_i \partial^i \psi_2 + \frac{a'}{a} \partial_i \partial^i \omega_2 - \frac{1}{4} \partial_k \partial_i D^{ki} \chi_2 - 12 \left(\frac{a'}{a} \right)^2 (\phi_1)^2 \right. \\ &\quad - 12 \frac{a'}{a} \phi_1 \psi'_1 - 3 \partial_i \psi_1 \partial^i \psi_1 - 8 \psi_1 \partial_i \partial^i \psi_1 - 3 (\psi'_1)^2 + 4 \frac{a'}{a} \phi_1 \partial_i \partial^i \omega_1 - 2 \frac{a'}{a} \partial_k \omega_1 \partial^k \phi_1 \\ &\quad - \frac{1}{2} \frac{a''}{a} \partial_k \omega_1 \partial^k \omega_1 + 12 \frac{a'}{a} \psi_1 \psi'_1 + \frac{1}{2} \partial_i \partial_k \omega_1 \partial^i \partial^k \omega_1 - 2 \frac{a'}{a} \partial_k \psi_1 \partial^k \omega_1 \\ &\quad + 4 \frac{a'}{a} \psi_1 \partial_i \partial^i \omega_1 - \frac{1}{2} \partial_k \partial^k \omega_1 \partial_k \partial^k \omega_1 - 2 \partial_k \omega_1 \partial^k \psi'_1 - 2 \psi'_1 \partial_i \partial^i \omega_1 - \phi_1 \partial_i \partial^k D^i_k \chi_1 \\ &\quad - 2 \psi_1 \partial_k \partial^i D^k_i \chi_1 + \partial_k \partial_i \psi_1 D^{ki} \chi_1 - 2 \frac{a'}{a} \partial_i \partial_k \omega_1 D^{ik} \chi_1 - 2 \frac{a'}{a} \partial_k \omega_1 \partial_i D^{ik} \chi_1 \\ &\quad - \partial_k \omega_1 \partial^i D^k_i \chi'_1 - \frac{1}{2} \partial_i \partial^i D_{mk} \chi_1 D^{km} \chi_1 + \partial_m \partial^k D_{ik} \chi_1 D^{im} \chi_1 + \frac{1}{2} \partial_k D^{km} \chi_1 \partial^i D_{mi} \chi_1 \\ &\quad \left. - \frac{1}{8} \partial^i D^{km} \chi_1 \partial_i D_{km} \chi_1 + \frac{1}{8} D^{ik} \chi'_1 D_{ki} \chi'_1 + \frac{a'}{a} D^{ki} \chi_1 D_{ik} \chi'_1 \right], \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned}
\delta_2 G^0_i = & \frac{1}{a^2} \left(-\frac{a'}{a} \partial_i \phi_2 - \partial_i \psi'_2 - \frac{1}{4} \partial_k D^k_i \chi'_2 + \frac{1}{4} \partial_k \partial^k \omega_{2i} + 8 \frac{a'}{a} \phi_1 \partial_i \phi_1 + 4 \phi_1 \partial_i \psi'_1 + 2 \psi'_1 \partial_i \phi_1 \right. \\
& - 4 \psi'_1 \partial_i \psi_1 - 4 \psi_1 \partial_i \psi'_1 + \partial_i \phi_1 \partial_i \partial^i \omega_1 - \partial_i \partial_k \omega_1 \partial^k \phi_1 + 8 \frac{a''}{a} \phi_1 \partial_i \omega_1 - 4 \left(\frac{a'}{a} \right)^2 \phi_1 \partial_i \omega_1 \\
& - 2 \frac{a'}{a} \partial^k \omega_1 \partial_i \partial_k \omega_1 + \partial_i \partial^i \psi_1 \partial_i \omega_1 + \partial^k \omega_1 \partial_i \partial_k \psi_1 - \partial_k \phi_1 D^k_i \chi'_1 + \frac{1}{2} \partial^k \phi_1 D_{ik} \chi'_1 \\
& - \psi_1 \partial_k D^k_i \chi'_1 + \frac{1}{2} \partial_k \psi_1 D^k_i \chi'_1 - \psi'_1 \partial_k D^k_i \chi_1 - \partial_k \psi'_1 D^k_i \chi_1 + \partial_i \omega_1 \partial_k \partial^m D^k_m \chi_1 \\
& + \left(\frac{a'}{a} \right)^2 \partial^k \omega_1 D_{ik} \chi_1 + \partial^k \omega_1 \partial_m \partial_i D^m_k \chi_1 - \frac{1}{2} \partial^m \omega_1 \partial_k \partial^k D_{im} \chi_1 + \frac{1}{2} \partial_k D^{km} \chi_1 D_{im} \chi'_1 \\
& \left. + \frac{1}{2} \partial_k D_{im} \chi'_1 D^{km} \chi_1 - \frac{1}{4} \partial_i D_{mk} \chi_1 D^{km} \chi'_1 - \frac{1}{2} \partial_i D_{mk} \chi'_1 D^{km} \chi_1 - 2 \frac{a''}{a} \partial^k \omega_1 D_{ik} \chi_1 \right). \tag{A.7}
\end{aligned}$$

Correspondingly, the first-order components of the energy-momentum tensor are

$$\begin{aligned}
\delta_1 T^0_0 &= -\delta \rho_1 \\
\delta_1 T^0_i &= (\rho_0 + P_0)(v_{1i} + \hat{\omega}_{1i}), \tag{A.8}
\end{aligned}$$

in terms of the background energy density ρ_0 and pressure P_0 . At second-order they read

$$\delta_2 T^0_0 = -\frac{1}{2} \delta \rho_2 - (\rho_0 + P_0) v_{1k} (v_1^k + \hat{\omega}_1^k), \tag{A.9}$$

$$\delta_2 T^0_i = \frac{1}{2} (\rho_0 + P_0) [(v_{2i} + \hat{\omega}_{2i} - 2\phi_1(v_{1i} + 2\hat{\omega}_{1i}) + 4\delta_{1i} g_{ik} v_1^k] + \frac{1}{2} (\delta \rho_1 + \delta P_1)(v_{1i} + \hat{\omega}_{1i}). \tag{A.10}$$

Einstein's equations, as written in a generic gauge, are automatically gauge invariant and one can check that they can be written in terms of gauge-invariant quantities. One is therefore free to evaluate them in the most convenient gauge on the left- and on the right-hand side. For instance, we evaluate the first-order Einstein tensor in the generalized longitudinal gauge (also called Poisson gauge) for which $\hat{\omega} = \chi = 0$, while we evaluate the energy-momentum tensor in the comoving orthogonal gauge for which $v_i = \hat{\omega}_i = 0$. In such a case, one easily finds by subtracting Eqs. (A.4) and (A.5),

$$\nabla^2 \Phi_1 = \frac{3}{2} \mathcal{H}^2 \delta_1^{\text{com}}, \tag{A.11}$$

where $\Psi_1 = \Phi_1$ are the Bardeen's gauge-invariant potentials [62], $\mathcal{H} = a'/a = aH$, and δ_1^{com} is the gauge-invariant density contrast in the comoving orthogonal gauge. Using the fact that in momentum space

$$\Phi_1(k, \tau) = -\frac{2}{3} T(k\tau) \zeta_1(k), \tag{A.12}$$

we get

$$\delta_1^{\text{com}}(k, \tau) = \frac{4}{9} \frac{k^2}{\mathcal{H}^2} T(k\tau) \zeta_1(k). \tag{A.13}$$

This equation is valid at all scales, not only on superhorizon scales. Similarly, at second-order one obtains the equations

$$\begin{aligned}
-\frac{3}{2} \mathcal{H}^2 \delta_2^{\text{com}} &= 3\mathcal{H}^2 \Phi_2 + 3\mathcal{H} \Psi'_2 - \partial_i \partial^i \Psi_2 - 12\mathcal{H}^2 (\Phi_1)^2 - 3 \partial_i \Phi_1 \partial^i \Phi_1 - 8 \Phi_1 \partial_i \partial^i \Phi_1 - 3 (\Phi'_1)^2, \\
0 &= -\mathcal{H} \partial_i \Phi_2 - \partial_i \Psi'_2 + 8\mathcal{H} \Phi_1 \partial_i \Phi_1 - 2 \Phi'_1 \partial_i \Phi_1, \tag{A.14}
\end{aligned}$$

which give

$$\frac{3}{2} \mathcal{H}^2 \delta_2^{\text{com}} = \nabla^2 \Psi_2 + 3 \partial_i \Phi_1 \partial^i \Phi_1 + 8 \Phi_1 \nabla^2 \Phi_1 + 3 (\Phi'_1)^2 + \frac{6\mathcal{H}}{\nabla^2} (\partial^i \Phi'_1 \partial_i \Phi_1 + \Phi_1' \nabla^2 \Phi_1), \tag{A.15}$$

which reproduces Eq. (24) of the main text. The equation of motion for the second-order gravitational potential can be found in Refs. [59, 60] and during the radiation phase it reads

$$\Psi_2'' + 4\mathcal{H} \Psi_2' - \frac{1}{3} \nabla^2 \Psi_2 = \frac{2}{3} \partial^i \Phi_1 \partial_i \Phi_1 + \frac{8}{3} \Phi_1 \nabla^2 \Phi_1 + 2 (\Phi'_1)^2 + \mathcal{H} N^j_i (A^i_j)' + \frac{1}{3} N^j_i (A^i_j)_{,k}^k, \tag{A.16}$$

$$A_2^i{}_j = 6 \partial^i \Phi_1 \partial_j \Phi_1 + \frac{2}{\mathcal{H}} (\partial^i \Phi_1 \partial_j \Phi_1)' + \frac{2}{\mathcal{H}^2} \partial^i \Phi_1' \partial_j \Phi_1', \quad (\text{A.17})$$

and

$$N^j{}_i = \frac{3}{2} \frac{1}{\nabla^2} \left(\frac{\partial^j \partial_i}{\nabla^2} - \frac{1}{3} \delta^j{}_i \right). \quad (\text{A.18})$$

Eq. (A.16) has to be solved with the superhorizon initial condition (with $\Phi_1(\mathbf{x}) \simeq -2\zeta_1(\mathbf{x})/3$) [59]

$$\Psi_2(\mathbf{x}) = -2\Phi_1^2(\mathbf{x}) + 2N^j{}_i (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) = -\frac{8}{9} \zeta_1^2(\mathbf{x}) + \frac{8}{9} N^j{}_i (\partial^i \zeta_1(\mathbf{x}) \partial_j \zeta_1(\mathbf{x})), \quad (\text{A.19})$$

and in momentum space the solution is the sum of two pieces

$$\Psi_2(\tau, \mathbf{k}) = \Psi_2(\mathbf{k}) T(k\tau) + \int \frac{d^3 p}{(2\pi)^3} uv I_\Psi(u, v, k\tau) \left(\frac{2}{3} \right)^2 \zeta_1(\mathbf{p}) \zeta_1(\mathbf{k} - \mathbf{p}), \quad (\text{A.20})$$

where $u = |\mathbf{k} - \mathbf{p}|/k$, $v = p/k$, and the function I_Ψ can be found in Refs. [59–61]. A detailed inspection of the time behaviour of $I_\Psi(u, v, k\tau)$ shows that $\Psi_2(\tau, \mathbf{k}) \propto a^{-2}$ due to the last term of Eq. (A.16) [59].

Similarly, one can introduce the second-order curvature perturbation ζ_2 , which in the Newtonian gauge takes the form [57–59]

$$\zeta_2 = -\Psi_2 + \frac{1}{4} \delta_2 + \frac{1}{16\mathcal{H}} (\delta_1^2)' - \frac{1}{9} \delta_1^2 - \frac{1}{17\mathcal{H}} \delta_1 (\Phi_1' + 2\mathcal{H}\Phi_1) - \frac{1}{32\mathcal{H}^2} \left[\delta_1^k \delta_{1,k} - \nabla^{-2} (\delta_{1,i} \delta_{1,j})^{,ij} \right]. \quad (\text{A.21})$$

On sub-horizon scales, where all gauges coincide and thus $\delta_1 = \delta_1^{\text{com}}$, this expression provides the behaviour for the curvature perturbation at second-order. The last term represents the leading contribution at late time and one can show that

$$\frac{2}{3} \nabla^2 \zeta_2(t, \mathbf{x}) \simeq -\frac{1}{48a^2 H^2} \sum_{i \neq j} \partial_i \partial_j \delta_1^{\text{com}}(t, \mathbf{x}) \partial^i \partial^j \delta_1^{\text{com}}(t, \mathbf{x}), \quad (\text{A.22})$$

as shown in Eq. (23) of the main text.

Finally, we provide the steps to estimate the second-order density contrast in the comoving gauge δ_2^{com} , as shown in Eq. (31) of the manuscript, which is needed to compute the corresponding impact on the threshold of PBH formation. Starting from Eq. (A.16) and treating the laplacian term $\nabla^2 \Psi_2$ as a perturbation, we can rewrite the equation by introducing a source term $\mathcal{S}(t, \mathbf{x})$ as $\ddot{\Psi}_2 + 5H\dot{\Psi}_2 = \mathcal{S}$ (working with cosmic time), such that its solution formally reads

$$\Psi_2(t, \mathbf{x}) = \int^t \frac{dt'}{a^5(t')} \int^{t'} dt'' a^5(t'') \frac{\mathcal{S}(\mathbf{x})}{a^2(t'')}, \quad (\text{A.23})$$

where the source has been evaluated with proper superhorizon initial conditions, since we are interested in evaluating the density profile at the horizon crossing time. The equation then simplifies to

$$\begin{aligned} \Psi_2(t, \mathbf{x}) &= \Psi_2(\mathbf{x}) + \int dt \left(\frac{1}{15} \frac{\nabla^2 \Psi_2}{a^2 H} + \frac{1}{a^2 H} \left[\frac{2}{15} \partial^i \Phi_1 \partial_i \Phi_1 + \frac{8}{15} \Phi_1 \nabla^2 \Phi_1 + \frac{2}{5} N^j{}_i \nabla^2 (\partial^i \Phi_1 \partial_j \Phi_1) \right] \right) \\ &= \Psi_2(\mathbf{x}) + \frac{1}{30} \frac{\nabla^2 \Psi_2(\mathbf{x})}{a^2 H^2} + \frac{1}{a^2 H^2} \left[\frac{1}{15} \partial^i \Phi_1 \partial_i \Phi_1 + \frac{4}{15} \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \frac{1}{5} N^j{}_i \nabla^2 (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) \right]. \end{aligned} \quad (\text{A.24})$$

By implementing the initial conditions shown in Eq. (A.19), we get after some simplifications

$$\begin{aligned} \Psi_2(t, \mathbf{x}) &= -2\Phi_1^2(\mathbf{x}) + 3 \frac{1}{\nabla^2} \left(\frac{\partial^j \partial_i}{\nabla^2} - \frac{1}{3} \delta^j{}_i \right) (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) - \frac{1}{5} \frac{1}{a^2 H^2} \partial^i \Phi_1 \partial_i \Phi_1 + \frac{1}{10} \frac{1}{a^2 H^2} \left(\frac{\partial^j \partial_i}{\nabla^2} \right) (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) \\ &\quad + \frac{2}{15} \frac{1}{a^2 H^2} \Phi_1 \nabla^2 \Phi_1 + \frac{3}{10} \frac{1}{\nabla^2} \frac{1}{a^2 H^2} (2\partial^j \nabla^2 \Phi_1 \partial_j \Phi_1 + \nabla^2 \Phi_1 \nabla^2 \Phi_1 + \partial^i \partial^j \Phi_1 \partial_i \partial_j \Phi_1), \end{aligned} \quad (\text{A.25})$$

where we have used that $\nabla^2 \Phi_1^2 = \partial^i \partial_i \Phi_1^2 = \partial^i (2\Phi_1 \partial_i \Phi_1) = 2\partial^i \Phi_1 \partial_i \Phi_1 + 2\Phi_1 \nabla^2 \Phi_1$. By taking a Laplacian, we get

$$\nabla^2 \Psi_2(t, \mathbf{x}) = -2\nabla^2 \Phi_1^2(\mathbf{x}) + 3 \frac{\partial^j \partial_i}{\nabla^2} (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) - \partial^i \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x})$$

$$\begin{aligned}
& -\frac{2}{5} \frac{1}{a^2 H^2} (\partial^i \nabla^2 \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) + \partial^i \partial_j \Phi_1(\mathbf{x}) \partial_i \partial^j \Phi_1(\mathbf{x})) \\
& + \frac{1}{10} \frac{1}{a^2 H^2} (2\partial^j \nabla^2 \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x}) + \nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \partial^i \partial^j \Phi_1(\mathbf{x}) \partial_i \partial_j \Phi_1(\mathbf{x})) \\
& + \frac{2}{15} \frac{1}{a^2 H^2} (\nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + 2\partial_j \Phi_1(\mathbf{x}) \partial^j \nabla^2 \Phi_1(\mathbf{x}) + \Phi_1(\mathbf{x}) \nabla^4 \Phi_1(\mathbf{x})) \\
& + \frac{3}{10} \frac{1}{a^2 H^2} (2\partial^j \nabla^2 \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x}) + \nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \partial^i \partial^j \Phi_1(\mathbf{x}) \partial_i \partial_j \Phi_1(\mathbf{x})), \tag{A.26}
\end{aligned}$$

such that

$$\begin{aligned}
\nabla^2 \Psi_2(t, \mathbf{x}) &= -2\nabla^2 \Phi_1^2(\mathbf{x}) + 3 \frac{\partial^j \partial_i}{\nabla^2} (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) - \partial^i \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) \\
&+ \frac{1}{a^2 H^2} \left(\frac{2}{3} \partial^i \nabla^2 \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) + \frac{8}{15} \nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \frac{2}{15} \Phi_1(\mathbf{x}) \nabla^4 \Phi_1(\mathbf{x}) \right). \tag{A.27}
\end{aligned}$$

The second-order density contrast then becomes

$$\begin{aligned}
\delta_2^{\text{com}} &= \frac{2}{3} \frac{1}{a^2 H^2} \nabla^2 \Psi^{(2)} + 2 \frac{1}{a^2 H^2} \partial_i \Phi^{(1)} \partial^i \Phi^{(1)} + \frac{16}{3} \frac{1}{a^2 H^2} \Phi^{(1)} \nabla^2 \Phi^{(1)} \\
&= \frac{1}{a^2 H^2} \left(-\frac{4}{3} \nabla^2 \Phi_1^2(\mathbf{x}) + 2 \frac{\partial^j \partial_i}{\nabla^2} (\partial^i \Phi_1(\mathbf{x}) \partial_j \Phi_1(\mathbf{x})) + \frac{4}{3} \partial^i \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) + \frac{16}{3} \Phi^{(1)} \nabla^2 \Phi^{(1)} \right) \\
&+ \frac{1}{a^4 H^4} \left(\frac{2}{3} \partial^i \nabla^2 \Phi_1(\mathbf{x}) \partial_i \Phi_1(\mathbf{x}) + \frac{8}{15} \nabla^2 \Phi_1(\mathbf{x}) \nabla^2 \Phi_1(\mathbf{x}) + \frac{2}{15} \Phi_1(\mathbf{x}) \nabla^4 \Phi_1(\mathbf{x}) \right), \tag{A.28}
\end{aligned}$$

that provides Eq. (31) of the main text.