Precision predictions for the primordial power spectra of scalar potential models of inflation

D. J. Brooker,[1,*] N. C. Tsamis,[2,†] and R. P. Woodard[1,‡]

[1]*Department of Physics, University of Florida, Gainesville, Florida 32611, USA*
[2]*Department of Physics, Institute of Theoretical Physics & Computational Physics, University of Crete, GR-710 03 Heraklion, Greece*

(Received 28 July 2015; published 4 February 2016)

We exploit a new numerical technique for evaluating the tree order contributions to the primordial scalar and tensor power spectra for scalar potential models of inflation. Among other things we use the formalism to develop a good analytic approximation which goes beyond generalized slow roll expansions in that (1) it is not contaminated by the physically irrelevant phase, (2) its 0th order term is exact for the constant first slow roll parameter, and (3) the correction is multiplicative rather than additive. These features allow our formalism to capture at first order effects which are higher order in other expansions. Although this accuracy is not necessary to compare current data with any specific model, our method has a number of applications owing to the simpler representation it provides for the connection between the power spectra and the expansion history of a general model.

DOI: 10.1103/PhysRevD.93.043503

I. INTRODUCTION

A central prediction of primordial inflation is the generation of a nearly scale invariant spectrum of tensor [1] and scalar [2] perturbations. These are increasingly recognized as quantum gravitational phenomena [3–5]. The scalar power spectra was first resolved in 1992 [6] and is by now observed with stunning accuracy by a variety of ground and space-based detectors [7–10]. The inflation community was transfixed with the March 2014 announcement that BICEP2 had resolved the tensor power spectrum [11]. Although subsequent work has shown this signal to be attributable to polarized dust emission [12,13], inclusion of the BICEP2 data gives the strongest upper bound on the tensor-to-scalar ratio [14], and so we have crossed the critical threshold at which the tensor power spectrum is more accurately constrained by polarization data than by temperature data. No one knows if the tensor signal is large enough to be resolved with current technology but many increasingly sensitive polarization experiments are planned, under way, or actually analyzing data, including POLARBEAR2 [15], PIPER [16], SPIDER [17], BICEP3 [18], and EBEX [19].

The triumphal progress of observational cosmology has not seen a comparable development of inflation theory. There are many, many theories for what caused primordial inflation, and all of them make different predictions for the tree order power spectra [20–22]. In most cases there is no precise analytic prediction [23] and numerical techniques must be employed instead [24–28]. At loop order the situation is even worse because there are excellent reasons for doubting that the naive correlators represent what is being observed [29–32], but there is no agreement on what should replace them [33–41]. Although defining loop corrections is by no means urgent, it may eventually become relevant with the full development of the data on the matter power spectrum which is potentially recoverable from highly redshifted 21 cm radiation [42,43].

We cannot do anything right now about the multiplicity of models, or about the ambiguity in how to define loop corrections for any one of these models. Our goal instead is to devise a good analytic prediction for the tree order power spectra from any scalar potential model. We will elaborate a numerical scheme developed previously [44,45], both to make the scheme even more efficient and to motivate what should be an excellent analytic approximation for the power spectra. One fascinating feature of this formalism is that the tensor power spectrum can easily be converted into its scalar cousin, so one need only work with the simpler tensor result. We use the numerical formalism to examine a wide variety of models with the aim of answering two questions:

(1) For models in which the first slow roll parameter $\epsilon(t)$ evolves, what value of the constant ϵ approximation gives the best fit to the actual power spectrum?
(2) How numerically accurate is our analytic formula for the nonlocal correction factor to the constant ϵ approximation?

It is useful to compare and contrast our formalism with the generalized slow roll approximation introduced by Stewart [46], and developed by Dvorkin and Hu [47], to deal with models for which $\epsilon(t)$ is small, but some of its derivatives are order one when expressed in Hubble units. In that technique one expands the mode function about its

*djbrooker@ufl.edu
†tsamis@physics.uoc.gr
‡woodard@phys.ufl.edu

de Sitter limit, using the de Sitter Green's function to develop a series of nonlocal corrections which depend upon the past history of $\epsilon(t)$. In contrast, our formalism is based on the norm squared of the mode function, which avoids having to keep track of the complicated and physically irrelevant phase. That allows our first order corrections to recover effects which are second order in the generalized slow roll approximation. Another difference is that our 0th order term is exact for arbitrary constant $\epsilon(t)$. Finally, our corrections are multiplicative rather than additive.

Although our formalism is more accurate, at the same order, than the generalized slow roll expansion, it is debatable whether current observations require greater accuracy for comparison with any specific model. Our motivation is rather to better understand how a *general* expansion history affects the power spectra. This has applications for the power spectra on three times scales: the interpretation of anomalies in current data; the next generation of observations which will reduce the error on n_s by a factor of 5 and might resolve the tensor power spectrum; and in the very long term, when the full development of 21 cm cosmology might provide enough data to resolve one loop corrections. These applications are as follows:

(i) To facilitate the deconvolution of anomalies in the power spectrum so as to identify the sorts of models which might have produced them;
(ii) To generalize the famous single-scalar consistency relation [48–50] so one can say something even with sparse data, before the tensor spectral index has been well measured; and
(iii) To understand whether loop corrections can receive significant contributions from early times when $\epsilon(t)$ was small and $H(t)$ was large.

Regarding the third point, one should note that the ζ-ζ propagator contains a factor of $1/\epsilon(t)$ which is usually assumed to be canceled by powers of $\epsilon(t)$ from the vertices [51]. However, it seems possible that the propagator—which depends nonlocally on $\epsilon(t)$—might receive significant contributions from small, early values of $\epsilon(t)$. If so, one might expect loop corrections from vertices at late times to be enhanced by large factors of $\epsilon_{\text{late}}/\epsilon_{\text{early}}$. A closely related issue is deciding what time best describes the putative loop counting parameter of $GH^2(t)$ [51].

A different sort of application concerns nonlocal modified gravity models of cosmology which are conjectured to represent quantum gravitational effects that became nonperturbatively strong during primordial inflation [52–54]. These quantum gravitational effects derive from secular growth in the graviton propagator which is known for de Sitter [55–57], but not for geometries in which $\epsilon(t)$ evolves [58]. Our formalism will facilitate better extrapolations of these growth factors to general geometries, which should motivate more realistic models.

This paper consists of six sections, of which the first is this Introduction. Section II reviews scalar potential models and the simple procedure for passing from the expansion history to the potential and vice versa. In Sec. III we define the two tree order power spectra, explain the relation between them, and give constant ϵ results. Our improved formalism is derived in Sec. IV, along with the analytic approximation. Section V presents numerical studies. Our conclusions comprise Sec. VI.

II. SCALAR POTENTIAL MODELS

The Lagrangian for a general scalar potential model is

$$\mathcal{L} = \frac{1}{16\pi G} R\sqrt{-g} - \frac{1}{2}\partial_\mu\varphi\partial_\nu\varphi g^{\mu\nu}\sqrt{-g} - V(\varphi)\sqrt{-g}. \tag{1}$$

We assume a homogeneous, isotropic, and spatially flat background characterized by $\varphi_0(t)$ and scale factor $a(t)$,

$$ds^2 = -dt^2 + a^2(t) d\vec{x}\cdot d\vec{x} \Rightarrow H(t) \equiv \frac{\dot{a}}{a}, \qquad \epsilon(t) \equiv -\frac{\dot{H}}{H^2}. \tag{2}$$

The nontrivial Einstein equations for this background are

$$3H^2 = 8\pi G\left[\frac{1}{2}\dot{\varphi}_0^2 + V(\varphi_0)\right], \tag{3}$$

$$-2\dot{H} - 3H^2 = 8\pi G\left[\frac{1}{2}\dot{\varphi}_0^2 - V(\varphi_0)\right]. \tag{4}$$

As long as the tensor power spectrum remains unresolved, there is no question that scalar potential models can be devised to fit the data because one can regard the observed scalar power spectrum as a first order differential equation for $H(t)$ [59]. Once $H(t)$ is known, there is a simple way of using Eqs. (3) and (4) to construct a potential $V(\varphi)$ which supports any function $H(t)$ that obeys $\dot{H}(t) < 0$ [60–64]. One first adds (3) and (4) to obtain an equation for the scalar background,

$$\varphi_0(t) = \varphi_i \pm \int_{t_i}^{t} dt' \sqrt{\frac{-\dot{H}(t')}{16\pi G}}. \tag{5}$$

By graphing this relation and then rotating the graph by 90° one can easily invert (5) to solve for time $t(\varphi)$. The final step is to subtract (4) from (3) to find the potential,

$$V(\varphi) = \frac{1}{8\pi G}[\dot{H}(t) + 3H^2(t)]_{t=t(\varphi)}. \tag{6}$$

Rather than specifying the expansion history $a(t)$ and using relations (5) and (6) to reconstruct the potential, it is more usual to specify the potential and then solve for the expansion history $a(t)$. This is greatly facilitated by making the slow roll approximation,

$$H \approx \sqrt{\frac{1}{3} 8\pi G V(\varphi)}, \qquad \epsilon \approx \frac{1}{16\pi G}\left[\frac{V'(\varphi)}{V(\varphi)}\right]^2. \tag{7}$$

It is desirable to express the scale factor $a = a_i e^N$ in terms of the number of e-foldings N from the beginning of inflation. Then one can use the slow roll approximation (7) to solve for the scalar's evolution from initial value φ_i by inverting the relation,

$$N = -8\pi G \int_{\varphi_i}^{\varphi} d\psi \frac{V(\psi)}{V'(\psi)}. \tag{8}$$

An important special case is power law potentials,

$$V(\varphi) = A\varphi^\alpha \Rightarrow \epsilon = \frac{\epsilon_i}{1 - \frac{4\epsilon_i}{\alpha}N}, \qquad H = H_i\left[1 - \frac{4\epsilon_i}{\alpha}N\right]^{\frac{\alpha}{4}}. \tag{9}$$

III. THE PRIMORDIAL POWER SPECTRA

The purpose of this section is to introduce notation to describe the scalar and tensor power spectra and review the local approximate formulas for them. We begin by defining the two spectra and explaining how the tensor result can be used to derive the scalar result. Then we consider the special cases of expansion histories with constant $\epsilon(t)$, and where $\epsilon(t)$ makes an instantaneous transition from one constant value of $\epsilon(t)$ to another.

A. Generalities

It is useful to define time dependent extensions of the scalar and tensor power spectra, $\Delta^2_{\mathcal{R}}(k)$ and $\Delta^2_h(k)$. At tree order these time dependent power spectra take the form of constants times the norm squared of the scalar and tensor mode functions $v(t,k)$ and $u(t,k)$,

$$\Delta^2_{\mathcal{R}}(t,k) = \frac{k^3}{2\pi^2} \times 4\pi G \times |v(t,k)|^2, \tag{10}$$

$$\Delta^2_h(t,k) = \frac{k^3}{2\pi^2} \times 32\pi G \times 2 \times |u(t,k)|^2. \tag{11}$$

The actual primordial power spectra are defined by evaluating these time dependent ones long after the time t_k of the first horizon crossing at which $k = H(t_k)a(t_k)$. After t_k the mode functions approach constants, and it is these constant values which define the predicted power spectra,

$$\Delta^2_{\mathcal{R}}(k) \equiv \Delta^2_{\mathcal{R}}(t,k)|_{t \gg t_k}, \qquad \Delta^2_{\mathcal{R}}(k) \equiv \Delta^2_{\mathcal{R}}(t,k)|_{t \gg t_k}. \tag{12}$$

The equations of motion and normalization conditions for the scalar and tensor mode functions are

$$\ddot{v} + \left(3H + \frac{\dot{\epsilon}}{\epsilon}\right)\dot{v} + \frac{k^2}{a^2}v = 0, \qquad v\dot{v}^* - \dot{v}v^* = \frac{i}{\epsilon a^3}, \tag{13}$$

$$\ddot{u} + 3H\dot{u} + \frac{k^2}{a^2}u = 0, \qquad u\dot{u}^* - \dot{u}u^* = \frac{i}{a^3}. \tag{14}$$

The full system (10)–(14) is frustrating because the phenomenological predictions (12) emerge from late times, whereas it is only at early times $k \gg H(t)a(t)$ at which one has a good asymptotic form for the mode functions,

$$v(t,k) \underset{k \gg Ha}{\longrightarrow} \frac{1}{\sqrt{2k\epsilon(t)a^2(t)}} \exp\left[-ik\int_{t_i}^{t} \frac{dt'}{a(t')}\right], \tag{15}$$

$$u(t,k) \underset{k \gg Ha}{\longrightarrow} \frac{1}{\sqrt{2ka^2(t)}} \exp\left[-ik\int_{t_i}^{t} \frac{dt'}{a(t')}\right]. \tag{16}$$

So these forms (15) and (16) serve to define the initial conditions, and one must then use Eqs. (13) and (14) to evolve $v(t,k)$ and $u(t,k)$ forward until well past the first horizon crossing, at which point the mode functions are nearly constant and one can use them in expressions (10)–(12) to compute the primordial power spectra. It is this cumbersome and highly model dependent procedure which we seek to simplify and systematize.

First, we take note of an important relation between the scalar and tensor systems. This is that the scalar relations (13) follow from the tensor ones (14) by simple changes of the scale factor and time [65],

$$a(t) \to \sqrt{\epsilon(t)} \times a(t), \qquad \frac{\partial}{\partial t} \to \frac{1}{\sqrt{\epsilon(t)}} \times \frac{\partial}{\partial t}. \tag{17}$$

We will therefore concentrate on the tensor system, and we do so in terms of the norm-squared tensor mode function,

$$M(t,k) \equiv |u(t,k)|^2. \tag{18}$$

B. The case of constant $\epsilon(t)$

An important special case is when $\epsilon(t)$ is constant, for which the Hubble parameter and scale factor are

$$\epsilon(t) = \epsilon_i \Rightarrow H(t) = \frac{H_i}{1 + \epsilon_i H_i \Delta t},$$
$$a(t) = [1 + \epsilon_i H_i \Delta t]^{\frac{1}{\epsilon_i}}, \tag{19}$$

where $\Delta t \equiv t - t_i$. Note that the combination $H(t)[a(t)]^\epsilon$ is constant. The appropriate tensor mode function for constant $\epsilon(t)$ is

$$u_0(t,k) = \frac{1}{a(t)\sqrt{2k}} \times \sqrt{\frac{\pi z}{2}} H_\nu^{(1)}(z),$$
$$z(t,k) \equiv \frac{k}{(1-\epsilon)Ha},$$
$$\nu \equiv \frac{3}{2} + \frac{\epsilon}{1-\epsilon}. \tag{20}$$

From the small argument expansion of the Hankel function we can infer the constant late time limit of (20),

$$u_0(t,k) \underset{k \ll Ha}{\longrightarrow} \sqrt{\frac{\pi z}{4a^2 k}} \times -\frac{i\Gamma(\nu)}{\pi}\left(\frac{2}{z}\right)^\nu, \tag{21}$$

$$= -\frac{i(1+\epsilon)\Gamma(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})}{\sqrt{2\pi k^3}}[2(1-\epsilon)]^{\frac{\epsilon}{1-\epsilon}}\left[\frac{H(t)a^\epsilon(t)}{k^\epsilon}\right]^{\frac{1}{1-\epsilon}}. \tag{22}$$

It is usual to evaluate the constant factor of $H(t)a^\epsilon(t)$ at horizon crossing,

$$u_0(t,k) \underset{k \ll Ha}{\longrightarrow} \frac{H(t_k)}{\sqrt{2k^3}} \times -\frac{i(1+\epsilon)\Gamma(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})}{\sqrt{\pi}}[2(1-\epsilon)]^{\frac{\epsilon}{1-\epsilon}}. \tag{23}$$

Substituting (23) in expressions (10)–(12) gives the famous constant ϵ predictions for the power spectra,

$$\Delta^2_{\mathcal{R}}(k)|_{\dot\epsilon=0} = \left(\frac{\hbar}{c^5}\right) \times \frac{GH^2(t_k)}{\pi\epsilon} \times \frac{(1+\epsilon)^2\Gamma^2(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})}{\pi}[2(1-\epsilon)]^{\frac{2\epsilon}{1-\epsilon}}, \tag{24}$$

$$\Delta^2_h(k)|_{\dot\epsilon=0} = \left(\frac{\hbar}{c^5}\right) \times \frac{16}{\pi} GH^2(t_k) \times \frac{(1+\epsilon)^2\Gamma^2(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})}{\pi}[2(1-\epsilon)]^{\frac{2\epsilon}{1-\epsilon}}. \tag{25}$$

The final factor in expressions (24) and (25) contains an ϵ-dependent correction which is not usually quoted because it is so near unity for small ϵ,

$$C(\epsilon) \equiv \frac{(1+\epsilon)^2\Gamma^2(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})}{\pi}[2(1-\epsilon)]^{\frac{2\epsilon}{1-\epsilon}}. \tag{26}$$

Figure 1 shows the dependence of $C(\epsilon)$ versus ϵ for the full inflationary range of $0 \le \epsilon < 1$. Note that $C(\epsilon)$ is a monotonically decreasing function of ϵ. In particular, it goes to zero for $\epsilon \to 1^-$. If we assume the single-scalar relation of $r = 16\epsilon$, then the current upper bound of $r < 0.09$ implies $\epsilon < 0.0056$. At this upper bound the constant ϵ correction factor is about 0.997. It would be even closer to unity for smaller vales of ϵ.

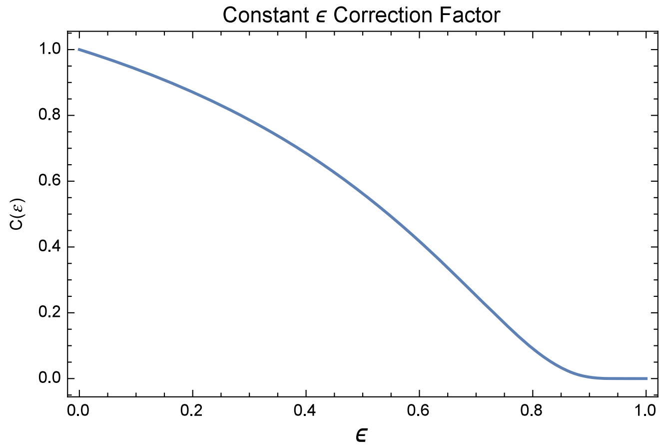

FIG. 1. Graph of $\frac{1}{\pi}(1+\epsilon)^2\Gamma^2(\frac{1}{2}+\frac{\epsilon}{1-\epsilon})[2(1-\epsilon)]^{\frac{2\epsilon}{1-\epsilon}}$ as a function of ϵ.

C. The case of a jump from $\epsilon(t) = \epsilon_1$ to $\epsilon(t) = \epsilon_2$

Suppose the Universe begins with constant $\epsilon(t) = \epsilon_1$, with initial values of the Hubble parameter and scale factor H_1 and a_1, respectively. At some time t_2 the first slow roll parameter makes an instantaneous transition to $\epsilon(t) = \epsilon_2 > \epsilon_1$. In both regions we express the scale factor in terms of the number of e-foldings N as $a(t) = a_1 e^N$. If the transition time $t = t_2$ corresponds to $N = N_2$, then we have

$$N < N_2 \Rightarrow \epsilon = \epsilon_1, \qquad H = H_1 e^{-\epsilon_1 N}, \tag{27}$$

$$N > N_2 \Rightarrow \epsilon = \epsilon_2, \qquad H = H_1 e^{\Delta\epsilon N_2 - \epsilon_2 N}, \tag{28}$$

where $\Delta\epsilon \equiv \epsilon_2 - \epsilon_1$.

It is useful to define mode functions assuming the two constant values of $\epsilon(t) = \epsilon_i$ had held for all time,

$$u_i(t,k) \equiv \frac{1}{\sqrt{2ka^2(t)}}\sqrt{\frac{\pi z}{2}} H^{(1)}_{\nu_i}(z),$$
$$z \equiv \frac{k}{(1-\epsilon_i)Ha},$$
$$\nu_i \equiv \frac{1}{2}\left(\frac{3-\epsilon_i}{1-\epsilon_i}\right). \tag{29}$$

The actual mode function after the transition is a linear combination of the positive and negative frequency solutions,

$$N < N_2 \Rightarrow u(t,k) = u_1(t,k), \tag{30}$$

$$N > N_2 \Rightarrow u(t,k) = \alpha u_2(t,k) + \beta u_2^*(t,k). \tag{31}$$

The combination coefficients are

$$\alpha = \frac{i\pi}{4}\Big[\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)[\sqrt{z_2}H^{(1)}_{\nu_2}(z_2)]^*_{,z_2} - [\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)]_{,z_1}\sqrt{z_2}H^{(1)*}_{\nu_2}(z_2)\Big], \tag{32}$$

$$\beta = \frac{i\pi}{4}\Big[-\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)[\sqrt{z_2}H^{(1)}_{\nu_2}(z_2)]_{,z_2} + [\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)]_{,z_1}\sqrt{z_2}H^{(1)}_{\nu_2}(z_2)\Big], \tag{33}$$

where z_1 and z_2 are

$$z_i \equiv \frac{1}{1-\epsilon_i}\frac{k}{H(t_2)a(t_2)} = \frac{1}{1-\epsilon_i}\frac{H(t_k)a(t_k)}{H(t_2)a(t_2)}. \tag{34}$$

We seek to understand the effect of varying the transition point N_2 relative to first horizon crossing N_k, with the important dimensional parameters k and $H(t_k)$ held fixed. Of course, this is accomplished by adjusting the initial values a_1 and H_1,

$$N_k < N_2 \Rightarrow H_1 = e^{\epsilon_1 N_k}H(t_k), \quad a_1 = \frac{k}{e^{N_k}H(t_k)}, \tag{35}$$

$$N_k > N_2 \Rightarrow H_1 = e^{\epsilon_2 N_k - \Delta\epsilon N_2}H(t_k), \qquad a_1 = \frac{k}{e^{N_k}H(t_k)}. \tag{36}$$

It is useful to express the late time limit of $M(t,k)$ in terms of the results M_i which would pertain if $\epsilon(t) = \epsilon_i$ for all time,

$$M_i \equiv \frac{H^2(t_k)}{2k^3} \times C(\epsilon_i), \tag{37}$$

where expression (26) gives the constant ϵ correction factor $C(\epsilon)$. The late time limit of the actual mode function $u(t,k)$ always derives from (31), but the late time limit of $u_2(t,k)$ depends upon whether the transition comes before or after the first horizon crossing,

$$N_k < N_2 \Rightarrow \lim_{t\to\infty} u_2(t,k) = -i\sqrt{M_2} \times e^{-\frac{\Delta\epsilon}{1-\epsilon_2}(N_k - N_2)}, \tag{38}$$

$$N_k > N_2 \Rightarrow \lim_{t\to\infty} u_2(t,k) = -i\sqrt{M_2}. \tag{39}$$

Hence the late time limit of $M(t,k) = |u(t,k)|^2$ is

$$N_k < N_2 \Rightarrow \lim_{t\to\infty} M(t,k) = |\alpha-\beta|^2 \times e^{-\frac{2\Delta\epsilon}{1-\epsilon_2}(N_k - N_2)} \times M_2, \tag{40}$$

$$N_k > N_2 \Rightarrow \lim_{t\to\infty} M(t,k) = |\alpha-\beta|^2 \times M_2. \tag{41}$$

Because only the difference of (32) and (33) enters the late time limit, the imaginary part of $H^{(1)}_{\nu_2}(z_2) = J_{\nu_2}(z_2) + iN_{\nu_2}(z_2)$ drops out,

$$\alpha - \beta = \frac{i\pi}{2}[\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)[\sqrt{z_2}J_{\nu_2}(z_2)]_{,z_2} - [\sqrt{z_1}H^{(1)}_{\nu_1}(z_1)]_{,z_1}\sqrt{z_2}J_{\nu_2}(z_2)]. \tag{42}$$

In evaluating the z_i one must distinguish between the cases for which the first horizon crossing occurs before and after the transition,

$$N_k < N_2 \Rightarrow z_i = \frac{1}{1-\epsilon_i}e^{(1-\epsilon_1)(N_k - N_2)}, \tag{43}$$

$$N_k > N_2 \Rightarrow z_i = \frac{1}{1-\epsilon_i}e^{(1-\epsilon_2)(N_k - N_2)}. \tag{44}$$

Figure 2 shows the late time limit of $M(t,k)$ for an instantaneous transition from $\epsilon_1 = \frac{1}{200}$ to $\epsilon_2 = \frac{1}{10}$ at $N = N_2$. For $\Delta N \equiv N_2 - N_k \lll -1$ the transition occurs long before the first horizon crossing so $M(t,k)$ approaches M_2, the result for a universe which has had $\epsilon(t) = \epsilon_2$ for all time. This follows from our analytic expressions because $z_i \gg 1$ in this regime, so we have

$$N_k \ll N_2 \Rightarrow \sqrt{\frac{\pi z_1}{2}}H^{(1)}_{\nu_1}(z_1) \to \exp\Big[iz_1 - i\Big(\nu_1 + \frac{1}{2}\Big)\frac{\pi}{2}\Big], \tag{45}$$

$$\Rightarrow \sqrt{\frac{\pi z_2}{2}}J_{\nu_2}(z_2) \to \cos\Big[z_2 - \Big(\nu_2 + \frac{1}{2}\Big)\frac{\pi}{2}\Big], \tag{46}$$

$$\Rightarrow \alpha - \beta \to \exp\Big[i(z_1 - z_2) + (\nu_2 - \nu_1)\frac{\pi}{2}\Big]. \tag{47}$$

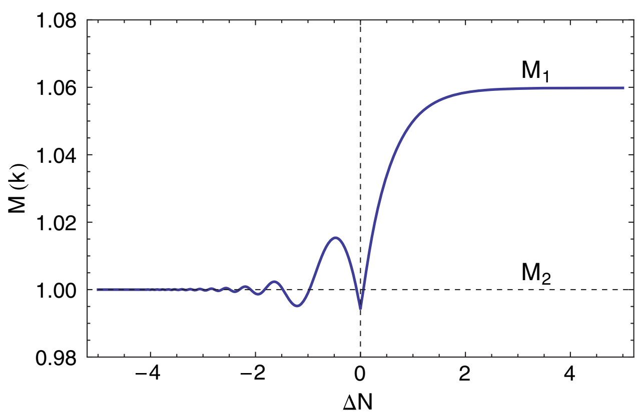

FIG. 2. Graph of $M(k) = \lim_{t\to\infty} M(t,k)$ in units of M_2 for an instantaneous transition from $\epsilon_1 = \frac{1}{200}$ to $\epsilon_2 = \frac{1}{10}$ as a function of the number of e-foldings $\Delta N \equiv N_2 - N_k$ from the first horizon crossing. The curve has a cusp at $\Delta N = 0$ because $H(t_k)$ and k are held fixed, whereas the way they depend upon the initial values of $H(t)$ and $a(t)$ changes from $\Delta N < 0$ to $\Delta N > 0$.

For $\Delta N \gg +1$ the transition occurs long after the first horizon crossing, which implies that the new value of $\epsilon(t) = \epsilon_2$ is irrelevant and $M(t,k)$ freezes in at the value M_1 that would pertain for a universe with $\epsilon(t) = \epsilon_1$ for all time. This is the regime of $z_i \ll 1$, for which our analytic expressions give

$$N_k \ll N_2 \Rightarrow \sqrt{\frac{\pi z_1}{2}} H^{(1)}_{\nu_1}(z_1) \to -\frac{\Gamma(\nu_1)}{\sqrt{\pi}}\left(\frac{2}{z_1}\right)^{\nu_1 - \frac{1}{2}}, \tag{48}$$

$$\Rightarrow \sqrt{\frac{\pi z_2}{2}} J_{\nu_2}(z_2) \to \frac{\sqrt{\pi}}{\Gamma(1+\nu_2)}\left(\frac{z_2}{2}\right)^{\nu_2 + \frac{1}{2}}, \tag{49}$$

$$\Rightarrow \alpha - \beta \to \sqrt{\frac{M_1}{M_2}} \exp\left[\frac{\Delta\epsilon(N_k - N_2)}{1-\epsilon_2}\right]. \tag{50}$$

Although the details depend upon the values of ϵ_1 and ϵ_2, Fig. 2 really is generic and has been known since the 1992 study of Starobinsky [66]. In particular, as N_2 approaches N_k from below there are always oscillations of decreasing frequency and increasing amplitude, the value at $N_2 = N_k$ always is somewhat below $M_2 < M_1$, and the value for $N_2 > N_k$ always rises monotonically to approach M_1. Similar results pertain for transitions of the inflaton potential [67,68].

IV. OUR EVOLUTION EQUATION

This is the main analytic portion of the paper. It begins by reviewing the derivation of an evolution equation for $M(t,k) \equiv |u(t,k)|^2$ [44,45]. We then factor out the main effect by writing $M(t,k) = M_0(t,k) \times \Delta M(t,k)$, where $M_0(t,k) \equiv |u_0(t,k)|^2$ is the constant ε result evaluated at the instantaneous $\epsilon(t)$. Next $M_0(t,k)$ is simplified and the asymptotic behaviors are discussed. By linearizing the equation for $\Delta M(t,k)$ we derive what should be an excellent approximation for $\Delta M(t,k)$ for a general inflationary expansion history.

A. An evolution equation for $M(t,k)$

The tensor power spectrum (11) depends upon the norm squared of the tensor mode function $u(t,k)$. It is numerically wasteful to follow the irrelevant phase using the tensor evolution equations (14), especially during the early time regime of $k \gg H(t)a(t)$ when oscillations are rapid. The better strategy is to use (14) to derive an equation for $M(t,k) \equiv |u(t,k)|^2$ directly. This is accomplished by computing the first two time derivatives,

$$\dot{M}(t,k) = u(t,k) \times \dot{u}^*(t,k) + \dot{u}(t,k) \times u^*(t,k), \tag{51}$$

$$\ddot{M}(t,k) = u(t,k) \times \ddot{u}^*(t,k) + 2\dot{u}(t,k) \times \dot{u}^*(t,k) + \ddot{u}(t,k) \times u^*(t,k). \tag{52}$$

Now use (14) to eliminate $\ddot{u}$ and $\ddot{u}^*$ in (52),

$$\ddot{M} = -3H\dot{M} - \frac{2k^2}{a^2}M + 2\dot{u}\dot{u}^*. \tag{53}$$

Squaring (51) and subtracting the square of the Wronskian (14) gives $\dot{u}\dot{u}^*$,

$$\dot{M}^2 = +u^2\dot{u}^{*2} + 2M\dot{u}\dot{u}^* + \dot{u}^2 u^{*2}, \tag{54}$$

$$\frac{1}{a^6} = -u^2\dot{u}^{*2} + 2M\dot{u}\dot{u}^* - \dot{u}^2 u^{*2}. \tag{55}$$

Hence the desired evolution equation for $M(t,k)$ is [44,45]

$$\ddot{M} + 3H\dot{M} + \frac{2k^2}{a^2}M = \frac{1}{2M}\left[\dot{M}^2 + \frac{1}{a^6}\right]. \tag{56}$$

As already noted, the transformation (17) converts (56) into an equation for the norm squared of the scalar mode function $N(t,k) \equiv |v(t,k)|^2$, so both power spectra follow from $M(t,k)$.

One indication of how much more efficient it is to evolve (56) than (14) comes from comparing the asymptotic expansions of $u(t,k)$ and $M(t,k)$ in the early time regime of $k \gg H(t)a(t)$. The expansion for $u(t,k)$ is in powers of $1/k$ and is not even local at first order,

$$u(t,k) = \left\{1 + \frac{i\alpha(t)}{k} + \frac{\beta(t)}{k^2} + O\left(\frac{1}{k^3}\right)\right\} \times \frac{\exp[-ik\int_{t_i}^t \frac{dt'}{a(t')}]}{\sqrt{2ka^2(t)}}, \tag{57}$$

$$\alpha(t) = \frac{1}{2}\int_{t_i}^t dt'[2-\epsilon(t')]H^2(t')a(t'), \tag{58}$$

$$\beta(t) = -\frac{1}{2}\alpha^2(t) + \frac{1}{4}[2-\epsilon(t)]H^2(t)a^2(t). \tag{59}$$

In contrast, $M(t,k)$ gives a series in $1/k^2$ which is local to all orders,

$$M(t,k) = \left\{1 + \frac{\bar{\alpha}(t)}{k^2} + \frac{\bar{\beta}(t)}{k^4} + O\left(\frac{1}{k^6}\right)\right\} \times \frac{1}{2ka^2(t)}, \tag{60}$$

$$\bar{\alpha}(t) = \left(1 - \frac{1}{2}\epsilon\right)H^2a^2, \tag{61}$$

$$\bar{\beta}(t) = \left[\frac{9}{4}\epsilon\left(1-\frac{2}{3}\epsilon\right)\left(1-\frac{1}{2}\epsilon\right)+\frac{9\dot{\epsilon}}{8H}-\frac{3\epsilon\dot{\epsilon}}{4H}+\frac{\ddot{\epsilon}}{8H^2}\right]H^4a^4. \tag{62}$$

Taking the norm squared of (57) helps to explain why our formalism is so much more accurate than the generalized slow roll approximation [46,47],

$$|u(t,k)|^2 = \left\{1+\frac{i\alpha}{k}+\frac{\beta}{k^2}+O\left(\frac{1}{k^3}\right)\right\} \times \left\{1-\frac{i\alpha}{k}+\frac{\beta}{k^2}+O\left(\frac{1}{k^3}\right)\right\}\times\frac{1}{2ka^2}, \tag{63}$$

$$= \left\{1+\frac{[\alpha^2+2\beta]}{k^2}+O\left(\frac{1}{k^4}\right)\right\}\times\frac{1}{2ka^2}. \tag{64}$$

Comparing (64) with (60) reveals that one must expand $u(t,k)$ to second order to recover the first order correction to $M(t,k)$. Dvorkin and Hu have noted (in the context of a late time expansion for the scalar mode functions, rather than this early time expansion for the tensor mode functions) that simply using the first order correction of the mode function to infer the power spectrum does not give a very accurate result [47]. From (64) we can see that it also gives the misleading impression that the correction to $M(t,k)$ is nonlocal, whereas one can see from expression (59) that part of the second order correction exactly cancels this, leaving a purely local correction to $M(t,k)$.

B. Factoring out the constant ϵ part

Reflection on the early time expansion (60)–(62) leads to the following form for the terms which include no derivatives of $\epsilon(t)$,

$$M_0(t,k) = \frac{1}{2ka^2(t)}\left\{1+\sum_{n=1}^{\infty}f_n(\epsilon(t))\left[\frac{H(t)a(t)}{k}\right]^{2n}\right\}, \tag{65}$$

$$f_n(\epsilon) \equiv \frac{(2n-1)!!}{(2n)!!}[(n+1)-n\epsilon][n-(n-1)\epsilon]\cdots[3-2\epsilon][2-\epsilon] \times[(n-1)\epsilon-(n-2)][(n-2)\epsilon-(n-3)]\cdots[\epsilon-0][0-(-1)]. \tag{66}$$

This is just $M_0(t,k) = |u_0(t,k)|^2$, where $u_0(t,k)$ is the constant ϵ solution (20) evaluated at the instantaneous value of $\epsilon(t)$. The evolution of $\epsilon(t)$ is so slow in most cases that it makes sense to factor $M_0(t,k)$ out of the result and derive an equation for the more sedate evolution of the residual amplitude.

We begin by writing

$$M(t,k) \equiv M_0(t,k)\times\Delta M(t,k), \quad M_0(t,k)\equiv|u_0(t,k)|^2. \tag{67}$$

Differentiating (67) results in the relations

$$\dot{M} = \dot{M}_0\times\Delta M + M_0\times\Delta\dot{M}, \tag{68}$$

$$\ddot{M} = \ddot{M}_0\times\Delta M + 2\dot{M}_0\times\Delta\dot{M} + M_0\Delta\ddot{M}, \tag{69}$$

$$\frac{\dot{M}^2}{2M} = \frac{\dot{M}_0^2}{2M_0}\times\Delta M + \dot{M}_0\times\Delta\dot{M} + M_0\times\frac{\Delta\dot{M}^2}{2\Delta M}, \tag{70}$$

$$\frac{1}{2a^6M} = \frac{1}{2a^6M_0}\times\Delta M + \frac{1}{2a^6M_0}\times\left[-\Delta M+\frac{1}{\Delta M}\right]. \tag{71}$$

Substituting relations (68)–(71) into (56) and dividing by $M(t,k)$ gives

$$\frac{\Delta\ddot{M}}{\Delta M}+\left[3H+\frac{\dot{M}_0}{M_0}\right]\frac{\Delta\dot{M}}{\Delta M}-\frac{1}{2}\left(\frac{\Delta\dot{M}}{\Delta M}\right)^2 + \frac{1}{2a^6M_0^2}\left[1-\frac{1}{\Delta M^2}\right] = -\frac{\ddot{M}_0}{M_0}-3H\frac{\dot{M}_0}{M_0}-\frac{2k^2}{a^2}+\frac{1}{2}\left(\frac{\dot{M}_0}{M_0}\right)^2+\frac{1}{2a^6M_0^2} \equiv S(t,k). \tag{72}$$

This is an evolution equation for $\Delta M(t,k)$, which is driven by a source $S(t,k)$. From (60)–(62) we see that the early time expansion of $\Delta M(t,k)$ is

$$\Delta M(t,k) = 1+\left[\frac{9\dot{\epsilon}}{8H}-\frac{3\epsilon\dot{\epsilon}}{4H}+\frac{\ddot{\epsilon}}{8H^2}\right]\left(\frac{aH}{k}\right)^4+O\left(\frac{a^6H^6}{k^6}\right). \tag{73}$$

C. Simplifications

Because $M_0(t,k)$ is an exact solution for constant $\epsilon(t)$, it must be that the source $S(t,k)$ is proportional to $\dot{\epsilon}$ and $\ddot{\epsilon}$. This is not obvious from expression (72) because of the complicated way $M_0(t,k)$ depends upon time explicitly through $a(t)$ and implicitly through $z(t,k)$ and $\nu(t)$,

$$M_0(t,k) = \frac{\frac{\pi z}{2} \times |H_\nu^{(1)}(z)|^2}{2ka^2(t)},$$
$$z(t,k) \equiv \frac{k}{(1-\epsilon)Ha},$$
$$\nu(t) \equiv \frac{1}{2} + \frac{1}{1-\epsilon}. \tag{74}$$

In Appendix A we make the following simplifications:

(1) Define the z and ν dependent part of M_0 as $\sigma(z,\nu) \equiv \ln[2ka^2(t) \times M_0(t,k)]$;

(2) Use the chain rule to express time derivatives of $M_0(t,k)$ as z and ν derivatives of $\sigma(z,\nu)$ multiplied by time derivatives of $z(t,k)$ and $\nu(t)$;

(3) Use Bessel's equation to eliminate the second z derivative;

(4) Change variables in $\sigma(z,\nu)$ from z to $\zeta \equiv \ln(z)$ and from $\nu \equiv \frac{1}{2} + \Delta\nu$ to $\xi \equiv \ln[\Delta\nu]$;

(5) Change the evolution variable from comoving time t to the number of e-foldings $N \equiv \ln[a(t)/a_i]$; and

(6) Express $\Delta M(t,k)$ in terms of a new dependent variable $h(t,k)$ as $\Delta M(t,k) \equiv \exp[-\frac{1}{2}h(t,k)]$.

When all of these things are done, Eq. (72) takes the form

$$\begin{aligned}
&\partial_N^2 h - \left[\frac{1}{2}\partial_N h\right]^2 + [1-\epsilon+\partial_N\sigma]\partial_N h + [2(1-\epsilon)e^{\zeta-\sigma}]^2[e^h-1] \\
&= 2\left[\frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\frac{\partial\sigma}{\partial\zeta} + 2\left[\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\frac{\partial\sigma}{\partial\xi} \\
&\quad + 4\left[-\partial_N\epsilon + \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\left[\frac{\partial^2\sigma}{\partial\zeta\partial\xi} + \frac{1}{2}\frac{\partial\sigma}{\partial\zeta}\frac{\partial\sigma}{\partial\xi}\right] + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\left[\frac{\partial^2\sigma}{\partial\xi^2} - \frac{\partial\sigma}{\partial\xi} + \frac{1}{2}\left(\frac{\partial\sigma}{\partial\xi}\right)^2\right] \\
&\quad + 4\left[-2\partial_N\epsilon + \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\left[\frac{(2-\epsilon)}{(1-\epsilon)^2} + e^{2\zeta}(e^{-2\sigma}-1)\right].
\end{aligned} \tag{75}$$

If desired, the derivative of σ with respect to N on the first line of (75) can be expressed like the terms on the right hand side of the equation,

$$\partial_N\sigma = \left[-(1-\epsilon) + \frac{\partial_N\epsilon}{1-\epsilon}\right]\frac{\partial\sigma}{\partial\zeta} + \frac{\partial_N\epsilon}{1-\epsilon}\frac{\partial\sigma}{\partial\xi}. \tag{76}$$

If N_k represents the e-folding at which the first horizon crossing occurs, then one can express the scale factor in terms of $\Delta N \equiv N - N_k$,

$$a = a_i e^N = a_i e^{N_k} \times e^{\Delta N} = \frac{k e^{\Delta N}}{H(t_k)}. \tag{77}$$

Hence we have

$$\begin{aligned}
M(t,k) &= \frac{e^{\sigma-\frac{1}{2}h}}{2ka^2} \\
&= \frac{H^2(t_k)C(\epsilon_k)}{2k^3} \times \exp\left[\sigma - \ln[C(\epsilon_k)] - 2\Delta N - \frac{1}{2}h\right],
\end{aligned} \tag{78}$$

where (26) gives $C(\epsilon)$. The correction to the constant ϵ prediction we are seeking is the late time limit of the exponential factor in expression (78).

D. Asymptotic analysis

In using Eq. (75) it is important to understand its limiting forms for early times ($k \gg Ha$) and for late times ($k \ll Ha$). At early times $z(t,k)$ is large and $h(t,k)$ is small. In Appendix B we expand each of the factors of equation (75) to show that its early time limiting form is

$$\begin{aligned}
&\partial_N^2 h + (1-\epsilon)\partial_N h + 4(1-\epsilon)^2 z^2 h + O(z^0 \times h) \\
&= -\left[2(\nu+3)\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon}\right]\frac{(\nu-\frac{1}{2})}{z^2} + O\left(\frac{1}{z^4}\right).
\end{aligned} \tag{79}$$

Equation (79) represents a damped, driven oscillator with

$$\text{Friction Force} \Rightarrow -(1-\epsilon) \times \partial_N h, \tag{80}$$

$$\text{Restoring Force} \Rightarrow -4(1-\epsilon)^2 z^2 \times h, \tag{81}$$

$$\text{Driving Force} \Rightarrow -\left[2(\nu+3)\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon}\right]\frac{(\nu-\frac{1}{2})}{z^2}. \tag{82}$$

The restoring force (81) pushes $h(t,k)$ down to zero if it ever gets displaced. The driving force (82) does push $h(t,k)$ away from zero, but its coefficient falls like $1/z^2$, whereas the restoring force grows like z^2. The "time" (that is, N) derivatives are irrelevant at leading order in z, so the result in this regime is just the local "tracking relation" we noted in expression (73),

$$h(t,k) = -\left[2(\nu+3)\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon}\right]\frac{(\nu-\frac{1}{2})}{4(1-\epsilon)^2 z^4} + O\left(\frac{1}{z^6}\right), \tag{83}$$

$$= -\frac{1}{4}[(9-7\epsilon)\partial_N\epsilon + \partial_N^2\epsilon]\left(\frac{Ha}{k}\right)^4 + O\left(\frac{H^6a^6}{k^6}\right). \tag{84}$$

This explains why the early time expansion is local to all orders. It also explains the striking property of Fig. 2 that an instantaneous jump in $\epsilon(t)$—which makes the source *diverge*—has negligible effect until just a few e-foldings before the first horizon crossing. One consequence is that we may as well begin numerical evolution at $N = N_k - 7$ using expansion (84) to determine the initial values of $h(t,k)$ and $\partial_N h(t,k)$.

At late times $z(t,k)$ is small, but $h(t,k)$ can grow to reach significant values. In Appendix C we expand the various factors of (75) to show that the late time limiting form is

$$\partial_N^2 h - \left[\frac{1}{2}\partial_N h\right]^2 + [2\partial_N\epsilon\Delta\nu^2 F + 3 - \epsilon]\partial_N h$$
$$= 4\partial_N\epsilon\Delta\nu[(2\Delta\nu+1)F + 1] + 4\left(\frac{\partial_N^2\epsilon}{1-\epsilon}\right)\Delta\nu F$$
$$+ 4\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\Delta\nu\left[\Delta\nu F^2 + 2F - 1 + \Delta\nu\psi'\left(\frac{1}{2}+\Delta\nu\right)\right]$$
$$+ O(z^2). \tag{85}$$

Here $F(t,k)$ stands for the quantity

$$F \equiv -1 - \ln\left(\frac{z}{2}\right) + \psi\left(\frac{1}{2}+\Delta\nu\right)$$
$$= \Delta N - 1 + \ln\left[\frac{2(1-\epsilon)H}{H(t_k)}\right] + \psi\left(\frac{1}{2}+\frac{1}{1-\epsilon}\right). \tag{86}$$

The late time equation (85) implies

$$\partial_N h = 4\partial_N\epsilon\Delta\nu^2 F + O(z^2), \tag{87}$$

$$\partial_N^2 h = 4\left[\frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\Delta\nu F + 4\partial_N\epsilon\Delta\nu$$
$$+ 4\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\Delta\nu\left[-1 + \Delta\nu\psi'\left(\frac{1}{2}+\Delta\nu\right)\right]$$
$$+ O(z^2). \tag{88}$$

Hence the asymptotic form of $h(t,k)$ at late times is

$$h(t,k) = 4\Delta\nu\epsilon\Delta N + 4\Delta\nu\ln\left[\frac{H}{H(t_k)}\right] + 2\ln\left[\frac{C(\epsilon)}{C(\epsilon_k)}\right]$$
$$- 2\ln[\mathcal{C}(k)] + O(z^2). \tag{89}$$

Comparison with (78) reveals the unknown constant $\mathcal{C}(k)$ as the correction factor we seek to the constant ϵ prediction for the tensor power spectrum.

E. An analytic approximation for $\Delta M(t,k) = e^{-\frac{1}{2}h(t,k)}$

The behaviors we noted in the previous section are generic, and they imply that we only need to bridge a small range of e-foldings around first horizon crossing N_k to carry the early form (84) into the late form (89). In this region $h(t,k)$ is small and we can linearize Eq. (75),

$$\partial_N^2 h + [1-\epsilon+\partial_N\sigma]\partial_N h + [2(1-\epsilon)e^{\zeta-\sigma}]^2 h \approx \mathcal{S}(N,N_k), \tag{90}$$

where $\mathcal{S}(N,N_k)$ is the full source term on the right hand side of (75). Just like the early time form (79), Eq. (90) is a damped, driven harmonic oscillator. Figure 3 shows the friction term for the $V = \frac{1}{2}m^2\varphi^2$ model. Figure 4 gives log and linear plots of the restoring force for the same model.

It is easy to develop a Green's function solution to (90). Note that the homogeneous equation takes the form

$$\chi'' - \frac{\omega'}{\omega}\chi' + \omega^2\chi = 0,$$
$$\chi' \equiv \partial_N\chi(N,N_k),$$
$$\omega' \equiv \partial_N\omega(N,N_k), \tag{91}$$

where the frequency is

$$\omega(N,N_k) \equiv 2(1-\epsilon)e^{\zeta-\sigma}. \tag{92}$$

The two linearly independent solutions of (91) can be expressed in terms of the integral of $\omega(N,N_k)$,

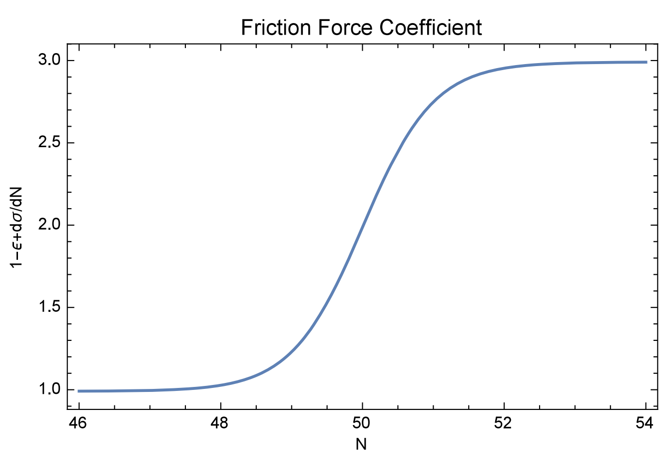

FIG. 3. Graph of $(1-\epsilon+\partial_N\sigma)$ as a function of N, assuming $N_k = 50$ and $\epsilon(N) = \frac{1}{200-2N}$, which corresponds to $V(\varphi) \propto \varphi^2$.

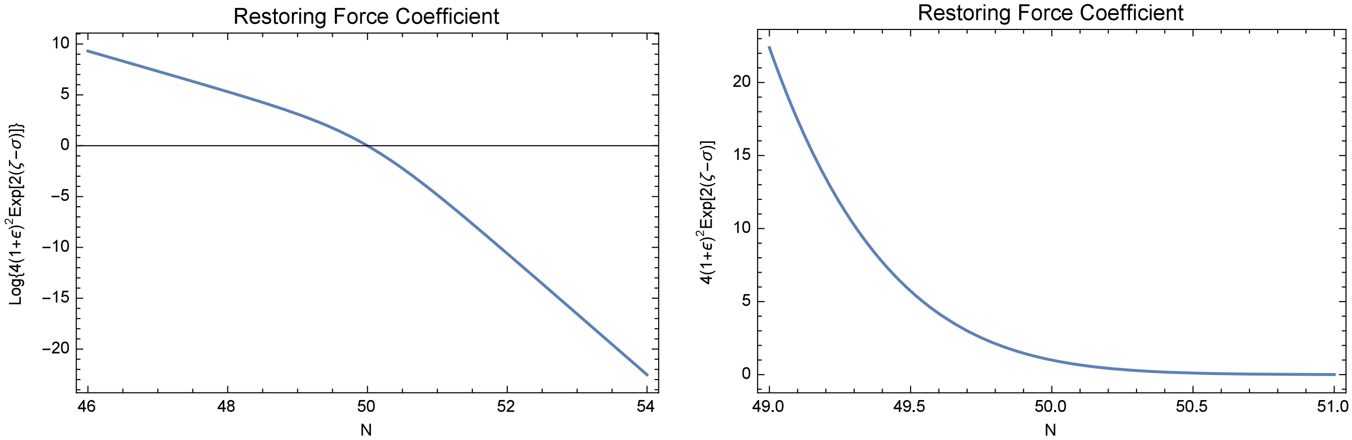

FIG. 4. Log (left) and linear (right) plots of $[2(1-\epsilon)e^{\zeta-\sigma}]^2$ as a function of N, assuming $N_k = 50$ and $\epsilon(N) = \frac{1}{200-2N}$, which corresponds to $V(\varphi) \propto \varphi^2$.

$$\chi_\pm(N, N_k) = \exp\left[\pm i \int_0^N dn\omega(n, N_k)\right] \Rightarrow \chi'_+\chi_- - \chi_+\chi'_- = 2i\omega. \tag{93}$$

Hence the retarded Green's function we seek is

$$G(N; N') = \frac{\theta(N - N')}{\omega(N', N_k)} \sin\left[\int_{N'}^N dn\omega(n, N_k)\right]. \tag{94}$$

And the Green's function solution to (90) is

$$h(t, k) = \int_0^N dn \sin\left[\int_n^N dn'\omega(n', N_k)\right] \frac{\mathcal{S}(n, N_k)}{\omega(n, N_k)}. \tag{95}$$

The asymptotic expansion (84) is so accurate at early times that one may as well begin the evolution at some point near to the first horizon crossing, say $N_1 = N_k - 7$. Then the Green's function solution takes the form

$$h(t, k) = \cos\left[\int_{N_1}^N dn\omega(n, N_k)\right] \times h(t_1, k) + \sin\left[\int_{N_1}^N dn\omega(n, N_k)\right] \times \frac{\partial_N h(t_1, N_k)}{\omega(N_1, N_k)} + \int_{N_1}^N dn \sin\left[\int_n^N dn'\omega(n', N_k)\right] \frac{\mathcal{S}(n, N_k)}{\omega(n, N_k)}. \tag{96}$$

The initial values $h(t_1, k)$ and $\partial_N h(t_1, k)$ can be either computed from (84) or simply approximated as zero.

Whether one uses expression (95) or (96), the goal is to evolve it to some point safely after the first horizon crossing, say $N_2 = N_k + 7$. Then the nonlocal correction factor $\mathcal{C}(k)$ can be estimated by ignoring the order z^2 terms in expression (89),

$$\mathcal{C}(k) \approx \exp\left[2\Delta\nu_2\epsilon_2\Delta N_2 + 2\Delta\nu_2 \ln\left[\frac{H(t_2)}{H(t_k)}\right] + \ln\left[\frac{C(\epsilon_2)}{C(\epsilon_k)}\right] - \frac{1}{2}h(t_2, k)\right]. \tag{97}$$

Expression (97) is radically different from other numerical schemes for computing the tensor power spectrum in that it gives an approximate but closed form expression for arbitrary first slow roll parameter $\epsilon(N)$. One consequence is that we can use the transformation (17) to immediately read off the analogous correction to the constant ϵ prediction (24) for the scalar power spectrum. Expression (97) is also the best way of deconvolving features in the power spectrum [69,70] to reconstruct the geometrical conditions which produced them.

V. NUMERICAL ANALYSES

The purpose of this section is to support various conclusions using numerical solutions of our full Eq. (75) for $h(t, k)$. Recall that the full amplitude is given by $M(t, k) = M_0(t, k) \times \exp[-\frac{1}{2}h(t, k)]$, where $M_0(t, k)$ is the known constant ϵ solution (74). Recall also that the ultimate observable is the correction factor $\mathcal{C}(k)$—inferred from $h(t, k)$ using expression (89)—to the constant ϵ approximation (25) for the tensor power spectrum.

A. $\mathcal{C}(k) - 1$ is small for smooth models

It has long been obvious the constant ϵ approximations (24) and (25) are wonderfully accurate for models in which ϵ is small and varies smoothly near the first horizon crossing [24]. Figure 5 confirms this for two simple monomial potentials,

$$V(\varphi) \propto \varphi^2 \Rightarrow \epsilon(N) = \frac{1}{200 - 2N}, \tag{98}$$

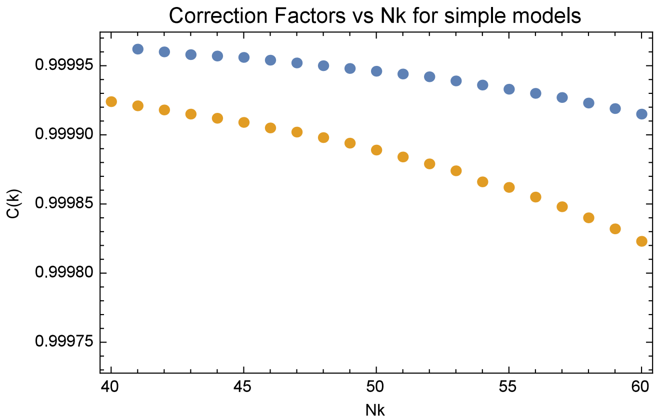

FIG. 5. Correction factors $\mathcal{C}(k)$ to the constant ϵ approximation for $\epsilon(N) = [200 - 2N]^{-1}$ (blue dots), corresponding $V(\varphi) \propto \varphi^2$, and for $\epsilon(N) = [100 - N]^{-1}$ (yellow dots), corresponding to $V(\varphi) \propto \varphi^4$.

$$V(\varphi) \propto \varphi^4 \Rightarrow \epsilon(N) = \frac{1}{100 - N}. \tag{99}$$

Figure 5 also answers the first of the questions posed at the end of the Introduction: it seems that the constant ϵ approximation is most accurate for ϵ near to ϵ_k. One can see this by comparing the value of the correction factor $C(\epsilon)$, defined in (26), with the nonlocal correction factor $\mathcal{C}(k)$ shown in Fig. 5, over the 20 e-foldings of the first horizon crossing ($40 < N_k < 60$) depicted,

$$V(\varphi) \propto \varphi^2 \Rightarrow \left\{ \begin{array}{c} 0.99546 < C(\epsilon_k) < 0.99317 \\ 0.99996 < \mathcal{C}(k) < 0.99991 \end{array} \right\}, \tag{100}$$

$$V(\varphi) \propto \varphi^4 \Rightarrow \left\{ \begin{array}{c} 0.99084 < C(\epsilon_k) < 0.98619 \\ 0.99992 < \mathcal{C}(k) < 0.99982 \end{array} \right\}. \tag{101}$$

There is about 50 times more variation in $C(\epsilon_k)$ than in $\mathcal{C}(k)$, limiting the potential improvement to a positive offset of about $\Delta N \approx \frac{20}{50} = 0.4$. Because other models show $\mathcal{C}(k) > 1$ there is actually no preference for shifting the point at which ϵ is evaluated.

B. $\mathcal{C}(k) - 1$ significant for changes near horizon crossing

It has also long been understood that the constant ϵ formulas require significant corrections when $\epsilon(N)$ suffers large variation within several e-foldings of the first horizon crossing [66,67]. We already saw this in the exact results depicted in Fig. 2 for an instantaneous jump in ϵ. Figure 6 makes the same point for two smooth transitions. The left hand graph shows the effect on $\mathcal{C}(k)$ of a smooth transition from $\epsilon = 0$ to $\epsilon = \frac{1}{2}$ via a logistic function centered at a critical value N_c,

$$\epsilon(N) = \frac{0.5}{1 + e^{N_c - N}}. \tag{102}$$

The right hand graph shows $\mathcal{C}(k)$ for a $V(\varphi) \propto \varphi^2$ model which experiences a Gaussian bump, centered at N_c, which actually induces a brief deceleration,

$$\epsilon(N) = \frac{1}{200 - 2N} + \exp[-10(N - N_c)^2]. \tag{103}$$

This is one of the models for which $\mathcal{C}(k)$ is larger than one.

C. Equation (96) is quite accurate near horizon crossing

The previous two points were known before in general terms. Our contributions in this paper are the following:

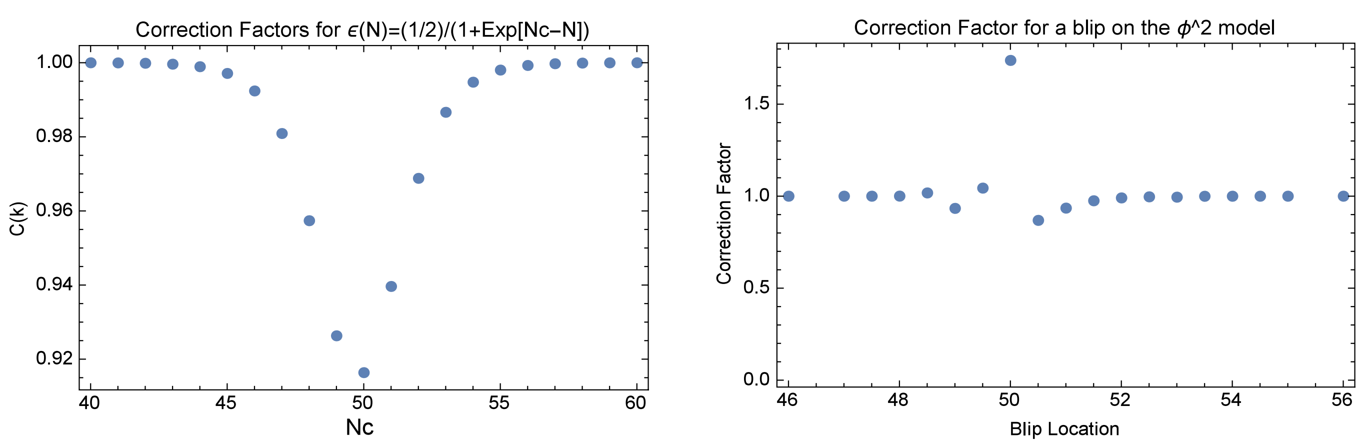

FIG. 6. Correction factors $\mathcal{C}(k)$ to the constant ϵ approximation for two models with smooth transitions centered at an arbitrary point N_c. The first model has $\epsilon(N) = \frac{1}{2[1+e^{N_c - N}]}$, corresponding to the left hand graph. The right hand graph corresponds to a $V(\varphi) \propto \varphi^2$ which experiences a Gaussian "blip" defined by (103). In each case horizon crossing is fixed at $N_k = 50$, and the graph shows how $\mathcal{C}(k)$ changes as N_c varies.

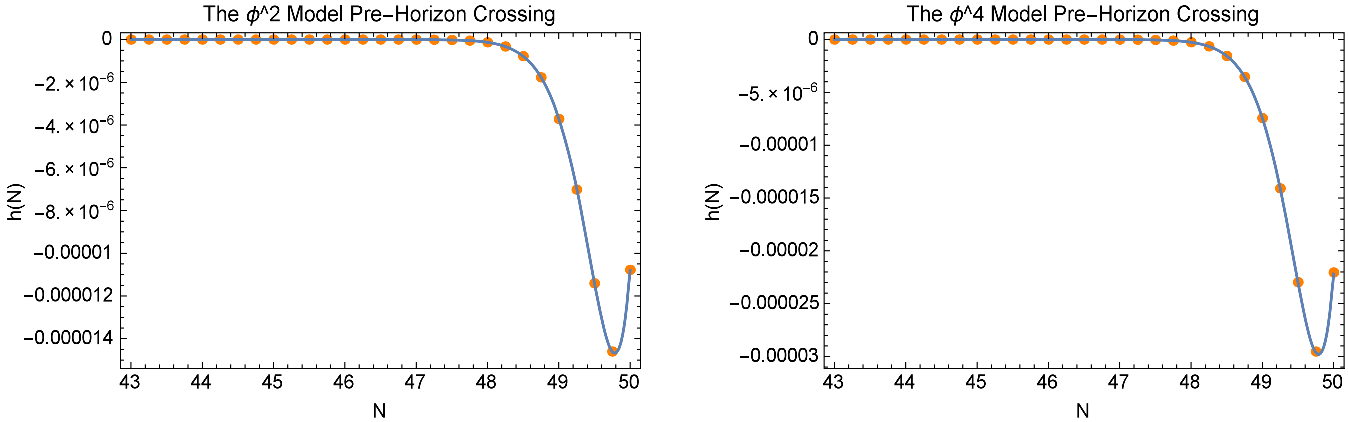

FIG. 7. The prehorizon crossing regime of $h(t,k)$ for two simple models. The left hand graph shows $\epsilon(N) = [200 - 2N]^{-1}$, corresponding to $V(\varphi) \propto \varphi^2$, and the right hand graph shows $\epsilon(N) = [100 - N]^{-1}$, corresponding to $V(\varphi) \propto \varphi^4$. In each case the continuous blue line represents numerical evolution of the full nonlinear equation (75), and the yellow dots give the analytic approximation (96).

(1) An analytic quantification—through the asymptotic expansions (84) and (89)—of when to expect significant corrections to the constant ϵ approximation; and

(2) An analytic approximation (96) of the function $h(t,k)$ which gives us the nonlocal correction factor through expression (97).

Figure 7 shows just how accurate our approximation is in the period before the first horizon crossing. It even catches the turning points at $N \sim 49.8$.

Our analytic approximation (96) continues to be very accurate after the first horizon crossing for models in which there is no significant evolution of $\epsilon(t)$ at late times. Figure 8 illustrates this by showing $h(t,k)$ versus N for a class of models in which ϵ makes a transition (centered about the horizon crossing of $N_k = 50$) from an early value of $\epsilon = \frac{1}{200}$ to a late value of $\epsilon = \frac{1}{10}$ through a logistic function with steepness parameter $K = 1,\ 2,\ 10,$

$$\epsilon(N) = \frac{1}{200} + \frac{\frac{19}{200}}{1 + \exp[-K(N - N_k)]}. \tag{104}$$

In each case the horizon crossing value is $\epsilon_k = \frac{41}{400}$.

Note from Fig. 8 that the final value of $h(t,k)$ is largest when the transition is most gradual. Because the full amplitude is $M(t,k) = M_0(t,k) \times \exp[-\frac{1}{2} h(t,k)]$, one might expect that $M(t,k)$ therefore freezes into a smaller amplitude for a more gradual transition. In fact, the reverse is true because only the value of ϵ near the horizon crossing is relevant, so making ϵ stay small longer causes the freeze-in amplitude to be larger. This is evident from the nonlocal correction factors $\mathcal{C}(k)$ for the three cases,

$$K = 1 \Rightarrow \mathcal{C}(k) = 0.993556 \qquad (0.994121), \tag{105}$$

$$K = 2 \Rightarrow \mathcal{C}(k) = 0.989201 \qquad (0.989418), \tag{106}$$

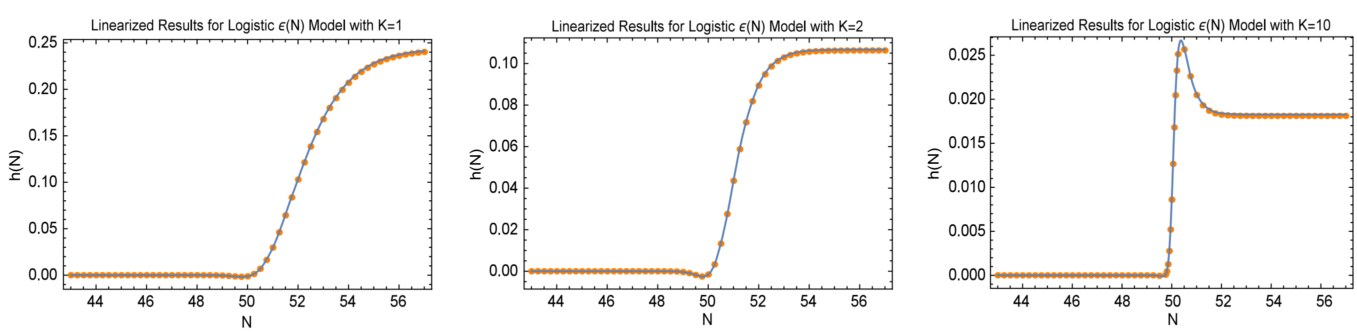

FIG. 8. The pre- and posthorizon crossing regimes for $\epsilon(N) = 0.005 + 0.095 \times (1 + \exp[-K \times (N - N_k)])^{-1}$, for $K = 1$ (left), $K = 2$ (center), and $K = 10$ (right). Each of these models interpolates between $\epsilon = 0.005$ at early times to $\epsilon = 0.100$ at late times, with $\epsilon_k = 0.0525$. The continuous blue line represents numerical evolution of the full nonlinear equation (75), and the yellow dots give the analytic approximation (96).

$$K = 10 \Rightarrow \mathcal{C}(k) = 0.975152 \qquad (0.975230). \tag{107}$$

(The parenthesized values are for the linearized approximation, which shows that it is indeed quite good.) We fixed the values of $H(t_k)$ and $a(t_k)$ to be the same for each model, so these correction factors give the relative freeze-in amplitudes for $M(t,k)$.

The much larger and opposite-sense effect which is evident in the asymptotic values of $h(t,k)$ of Fig. 8 is needed to compensate for the factor $M_0(t,k)$. Recall from Sec. III B that if $\epsilon(t)$ becomes constant at ϵ_1 for times $t > t_1$, then we can write

$$\epsilon(t) = \epsilon_1 \Rightarrow H(t)a^{\epsilon_1}(t) = H_1 a_1^{\epsilon_1}, \tag{108}$$

where $H_1 \equiv H(t_1)$ and $a_1 \equiv a(t_1)$. Each of the three models has effectively reached this condition by 10 e-foldings after the first horizon crossing, but the values of H_1 and a_1 are smaller the steeper the transition. That affects the factor of $M_0(t,k)$, which approaches a constant given by (22),

$$M_0(t,k) \to \frac{H^2(t_k)}{2k^3} \times C(\epsilon_1) \times \left[\frac{H_1 a_1^{\epsilon_1}}{H(t_k)a^{\epsilon_1}(t_k)}\right]^{\frac{2}{1-\epsilon_1}}. \tag{109}$$

The final factor of (109) is significantly larger for more gradual transitions, which is mostly canceled by the larger asymptotic values of $h(t,k)$, to leave the small effect evident in the nonlocal correction factors (105)–(107).

The same late time effect of $h(t,k)$ partially compensating for changes in $M_0(t,k)$ is evident from the results for $V(\varphi) \propto \varphi^2$ and $V(\varphi) \propto \varphi^4$ models displayed in Fig. 9. In this case $\epsilon(t)$ continues to evolve after the first horizon crossing. Considered as a function of N we have $\partial_N \ln[H(N)] = -\epsilon(N)$, so the asymptotic form (89) can be reexpressed as

$$h(t,k) = \frac{4}{1-\epsilon(N)} \int_{N_k}^{N} dn \Delta n \epsilon'(n) + 2\ln\left[\frac{C(\epsilon(N))}{C(\epsilon_k)}\right] - 2\ln[\mathcal{C}(k)] + O(e^{-2\Delta N}), \tag{110}$$

where $\Delta n \equiv n - N_k$ and $\Delta N \equiv N - N_k$. Because ϵ typically grows slowly with N (as it does for both of the models in Fig. 9), the integral grows and dominates the slowly falling logarithm of (110), so that $h(t,k)$ grows like ΔN^2. This growth is evident for both models in Fig. 9.

D. Problems long after horizon crossing

Of course, too much growth endangers the linearized approximation we made in passing from the full equation (75) to (90). Recall that this entails changing two terms,

$$-\left[\frac{1}{2}\partial_N h\right]^2 \to 0, \tag{111}$$

$$\exp\left[h(t,k)\right] - 1 \to h(t,k). \tag{112}$$

There is never any problem with (112) because $h(t,k)$ is small before the first horizon crossing and the coefficient of this term is minuscule after the first horizon crossing. The problematic approximation is (111), although only in the region after the first horizon crossing for models in which ϵ evolves at very late times. One can see from expression (87) that two terms contribute to provide the factor of $F^2 \sim \Delta N^2$ [recall the definition (86) of F] which is evident in the late time evolution equation (85),

$$-\left[\frac{1}{2}\partial_N h\right]^2 \to -[2\partial_N \epsilon \Delta\nu^2 F]^2, \tag{113}$$

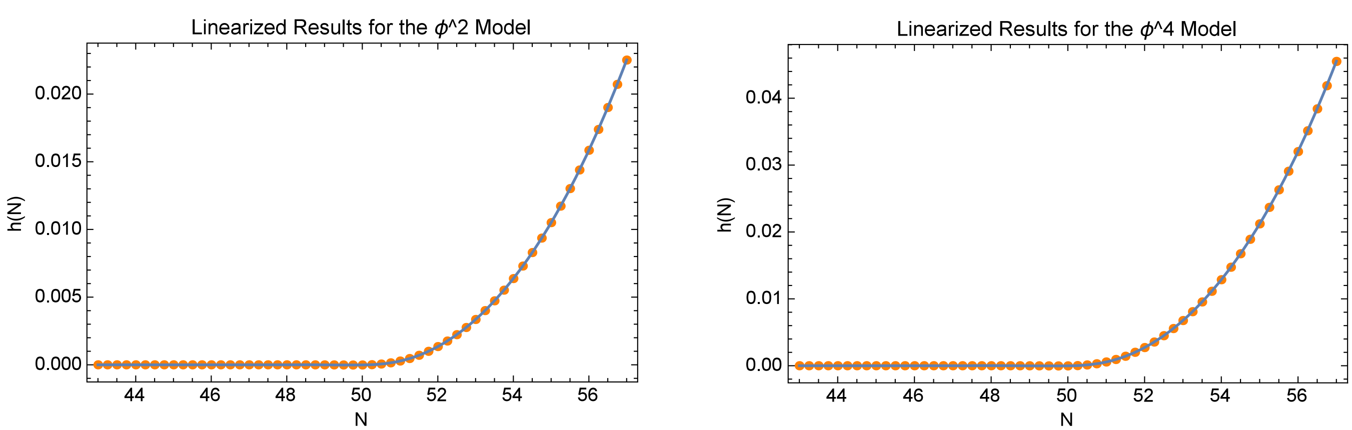

FIG. 9. The pre- and posthorizon crossing regimes of $h(t,k)$ for two simple models. The left hand graph concerns $\epsilon(N) = [200 - 2N]^{-1}$, corresponding to $V(\varphi) \propto \varphi^2$, and the right hand graph concerns $\epsilon(N) = [100 - N]^{-1}$, corresponding to $V(\varphi) \propto \varphi^4$. In each case the continuous blue line represents numerical evolution of the full nonlinear equation (75), and the yellow dots give the analytic approximation (96).

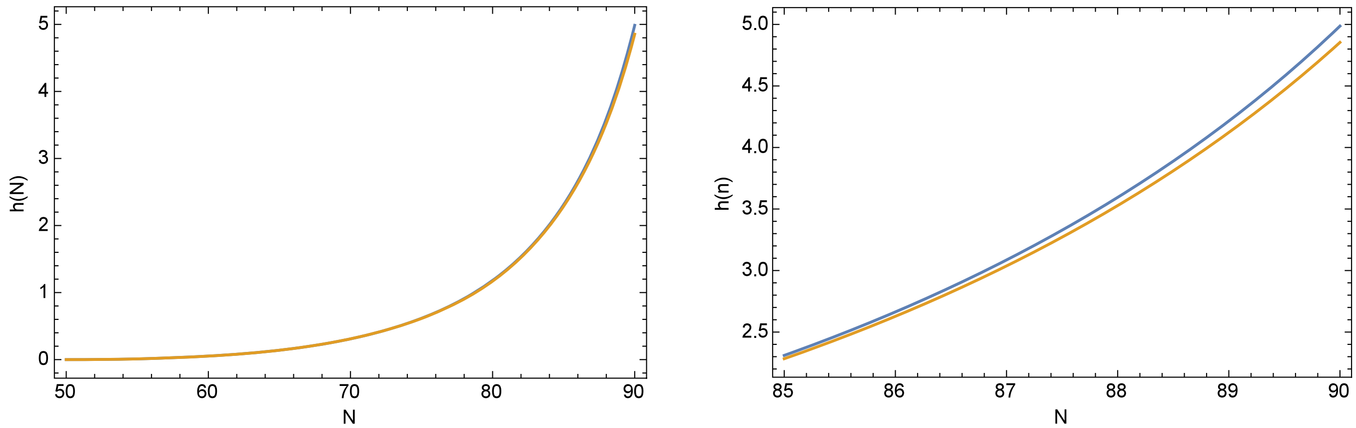

FIG. 10. Each graph shows the model with $\epsilon(N) = [200 - 2N]^{-1}$, corresponding to $V(\varphi) \propto \varphi^2$. Horizon crossing is at $N_k = 50$ and inflation ends at $N = 99.5$. In each case the blue line represents numerical evolution of the full nonlinear equation (75), and the yellow line gives our analytic approximation (96).

$$2\partial_N \epsilon \Delta\nu^2 F \times \partial_N h \to +2[2\partial_N \epsilon \Delta\nu^2 F]^2. \qquad (114)$$

These terms are enhanced by the factor of $F^2 \sim \Delta N^2$ but suppressed by $(\partial_N \epsilon)^2$. In the full nonlinear equation (113) cancels half of (114), but this cancellation does not happen in the linearized equation because (113) is not present. Hence we expect the very late time growth of the linearized approximation (96) to be less than what it is for the actual solution. This is barely evident in Fig. 10 for the $V(\varphi) \propto \varphi^2$ model at very late times.

The problem we have just described might seem serious, but it is not. The full amplitude $M(t,k) = M_0(t,k) \times \exp[-\frac{1}{2} h(t,k)]$ really does become constant shortly after the first horizon crossing. The growth of $h(t,k)$ is only an artifact of being factored out by $M_0(t,k)$, which also grows for models in which ϵ increases at late times. Because the problem has such a simple origin, there are two easy fixes:

(1) Either evaluate $\mathcal{C}(k)$ using expression (97) at some point N_2 before nonlinear effects become important or
(2) Subtract the right hand side of (113) from the source $\mathcal{S}(N, N_k)$ in the Green's function solution.

VI. EPILOGUE

The full scalar and tensor power spectra can be expressed in terms of two amplitudes $N(t,k)$ and $M(t,k)$,

$$\Delta^2_{\mathcal{R}}(k) = \frac{2Gk^3}{\pi} \times N(t \gg t_k, k) \times \{1 + O(GH^2)\}, \qquad (115)$$

$$\Delta^2_h(k) = \frac{32Gk^3}{\pi} \times M(t \gg t_k, k) \times \{1 + O(GH^2)\}. \qquad (116)$$

If the one loop corrections of order $GH^2 \lesssim 10^{-11}$ are ever to be resolved, we must have precise predictions for the two amplitudes. Part of this problem entails finding a unique model for primordial inflation, which is beyond the scope of our present effort. We have instead focused on predicting how the amplitudes depend upon the inflationary expansion history $a(t)$. Our analysis is based on earlier work in which nonlinear equations for the two amplitudes were derived [44,45].

Because the transformation (17) carries $M(t,k)$ into $N(t,k)$, we worked with the simpler tensor amplitude. We express its late time limiting form as

$$M(t \gg t_k, k) = \frac{H^2(t_k)}{2k^3} \times C(\epsilon(t_k)) \times \mathcal{C}(k), \qquad (117)$$

where $C(\epsilon)$ was defined in expression (26) and graphed in Fig. 1. Our numerical studies show that this factor really does need to be present, and it is best to evaluate it at the time t_k of the first horizon crossing. The remaining factor $\mathcal{C}(k)$ represents nonlocal effects from the expansion history before and after the first horizon crossing. It has long been clear that this factor is close to unity for models in which $\epsilon(t)$ is smooth near the first horizon crossing, but $\mathcal{C}(k)$ can give significant corrections when there are large changes within a few e-foldings of the first horizon crossing [66,67].

Our key results (89) and (97) give, for the first time ever, a good analytic approximation for the nonlocal correction factor $\mathcal{C}(k)$. Our technique was to first get close to the exact solution by factoring out the known, constant ϵ solution $M_0(t,k)$,

$$M(t,k) = M_0(t,k) \times \exp\left[-\frac{1}{2} h(t,k)\right]. \qquad (118)$$

Of course, this means that the evolution equation (75) for $h(t,k)$ is driven by a source term which vanishes whenever $\epsilon(t)$ is constant. From the equation's asymptotic early time form (79) we can see that $h(t,k)$ behaves as a damped, driven harmonic oscillator. For more than a few e-foldings before the first horizon crossing the restoring force (81) is so large that $h(t,k)$ is both small and completely determined by local conditions according to a wonderfully convergent expansion (84). That is evident from Fig. 2 even for an instantaneous jump in $\epsilon(t)$.

As long as $\partial_N h(t,k)$ remains small, the full Eq. (75) can be linearized to a form (90) for which we were able to derive a Green's function solution (96). It cannot be overstressed that this solution pertains for an *arbitrary* inflationary expansion history. The assumption of linearity on which it is based should be valid long before the first horizon crossing. It can break down long after the first horizon crossing but in a way which is very simple to repair.

Our formalism differs from the generalized slow roll approximation [46,47] in three ways:

(1) Instead of correcting the mode function $u(t,k)$ and then inferring how this affects $M(t,k) \equiv |u(t,k)|^2$, we correct $M(t,k)$ directly;
(2) Our 0th order term is exact for arbitrary constant $\epsilon(t)$; and
(3) Our corrections are multiplicative rather than additive.

As the early time expansions (64) and (60) show, our formalism captures effects at first order which require going to second order in the generalized slow roll expansion. Given a specific model, the additional accuracy of our formalism is not required for the analysis of current data. Its advantage derives rather from the more explicit connection it makes between data and a general, initially unknown model. This has potential applications for the power spectra on three time scales:

(1) For current data it facilitates the process of inferring the sorts of models which might explain anomalies;
(2) For next generation data, which might begin resolving the tensor power spectrum, it permits exploitation of the general relation (17) between the tensor and scalar power spectra to develop a version of the single scalar consistency relation [48–50] that could be employed before the tensor spectral index has been well measured; and
(3) For far future data, when the full development of 21 cm cosmology might permit loop corrections to be resolved, it elucidates both when the loop counting parameter of $GH^2(t)$ should be evaluated, and whether there can be enhancements of the form $\epsilon_{\rm late}/\epsilon_{\rm early}$.

Our work also has three more general applications. First, there is a close relation between $M(t,k)$ and the vacuum expectation value of a massless, minimally coupled (MMC) scalar,

$$\langle\Omega|\varphi^2(t,\vec{x})|\Omega\rangle = \int \frac{dkk^2}{2\pi^2} M(t,k). \qquad (119)$$

This relation should allow us to estimate the secular growth for an arbitrary inflationary expansion history, which is an important step in building nonlocal models to represent the quantum gravitational backreaction on inflation [52–54,71]. Second, note that our transformation (17) could be used to convert the propagator of a MMC scalar into the propagator for the scalar perturbation field $\zeta(t,\vec{x})$ for an arbitrary inflationary expansion history. Of course, we do not have a MMC scalar propagator for arbitrary $a(t)$, but perhaps the transformation could be used to derive relations between loops involving gravitons and loops involving ζ. Finally, our technique for passing from the oscillatory mode functions to their norm squared [44,45] can be applied for any perturbations whose mode functions obey second order equations. It would be interesting to see what it gives for Higgs inflation and for $f(R)$ models of inflation.

ACKNOWLEDGMENTS

We are grateful for conversations and correspondence on this subject with P. K. S. Dunsby, S. Odintsov, L. Patino, M. Romania, S. Shandera, and M. Sloth. This work was partially supported by the European Union (European Social Fund, ESF) and Hellenic national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) under the "Θαλής" action MIS-375734, under the "Αριστεια" action, under the "Funding of proposals that have received a positive evaluation in the 3rd and 4th Call of ERC Grant Schemes"; by NSF Grants No. PHY-1205591 and No. PHY-1506513; and by the Institute for Fundamental Theory at the University of Florida.

APPENDIX A: SIMPLIFYING EQ. (72)

The first time derivative of $M_0(t,k)$ is

$$\dot{M}_0 = -2HM_0 + \dot{z}M_0' + \dot{\nu}\partial_\nu M_0, \qquad (A1)$$

where a prime stands for the derivative with respect to z and ∂_ν denotes differentiation with respect to ν. It is best to postpone using the explicit expressions for $\dot{z}$ and $\dot{\nu}$,

$$\dot{z} = -\frac{k}{a} + \frac{\dot{\epsilon}}{1-\epsilon} \times z, \qquad \dot{\nu} = \frac{\dot{\epsilon}}{(1-\epsilon)^2}. \qquad (A2)$$

The time second derivative of $M_0(t,k)$ is

$$\ddot{M}_0 = (4+2\epsilon)H^2 M_0 + (-4H\dot{z} + \ddot{z})M_0' + (-4H\dot{\nu} + \ddot{\nu})\partial_\nu M_0 + \dot{z}^2 M_0'' + 2\dot{z}\,\dot{\nu}\,\partial_\nu M_0' + \dot{\nu}^2\partial_\nu^2 M_0. \qquad (A3)$$

Bessel's equation implies that M_0'' can be eliminated using

$$M_0'' + 2M_0 - 2(2-\epsilon)\frac{H^2a^2}{k^2}M_0 = \frac{1}{2M_0}\left[M_0'^2 + \frac{1}{k^2a^4}\right]. \tag{A4}$$

The other derivatives we require are

$$3H\dot{M}_0 = -6H^2M_0 + 3H\dot{z}M_0' + 3H\dot{\nu}\partial_\nu M_0, \tag{A5}$$

$$-\frac{\dot{M}_0^2}{2M_0} = -2H^2M_0 + 2H\dot{z}M_0' + 2H\dot{\nu}\partial_\nu M_0 - \frac{(\dot{z}M_0' + \dot{\nu}\partial_\nu M_0)^2}{2M_0}. \tag{A6}$$

Substituting everything into the definition (72) of $S(t,k)$ gives

$$S(t,k) = -(H\dot{z} + \ddot{z})\frac{M_0'}{M_0} - (H\dot{\nu} + \ddot{\nu})\frac{\partial_\nu M_0}{M_0} - 2\dot{z}\,\dot{\nu}\left[\frac{\partial_\nu M_0'}{M_0} - \frac{M_0'}{2M_0}\frac{\partial_\nu M_0}{M_0}\right] - \dot{\nu}^2\left[\frac{\partial_\nu^2 M_0}{M_0} - \frac{1}{2}\left(\frac{\partial_\nu M_0}{M_0}\right)^2\right] - \left(\dot{z}^2 - \frac{k^2}{a^2}\right) \times \left\{\frac{(4-2\epsilon)}{(1-\epsilon)^2z^2} - 2 + \frac{1}{2k^2a^4M_0^2}\right\}. \tag{A7}$$

Each of the five terms on the right hand side of (A7) is proportional to at least one derivative of $\epsilon(t)$. Before exhibiting this it is desirable to isolate the ϵ-dependent part of the index ν, and to change from comoving time t to the number of e-foldings since the beginning of inflation $N \equiv \ln[a(t)/a_i]$,

$$\Delta\nu \equiv \frac{1}{1-\epsilon}, \qquad \frac{d}{dt} = H\frac{d}{dN}, \qquad \frac{d^2}{dt^2} = H^2\left[\frac{d^2}{dN^2} - \epsilon\frac{d}{dN}\right]. \tag{A8}$$

With this notation the five prefactors from the right hand side of (A7) are

$$H\dot{z} + \ddot{z} = \left[\frac{H\epsilon\dot{\epsilon}}{1-\epsilon} + \frac{\ddot{\epsilon}}{1-\epsilon} + 2\left(\frac{\dot{\epsilon}}{1-\epsilon}\right)^2\right]z = H^2\left[\frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]z, \tag{A9}$$

$$H\dot{\nu} + \ddot{\nu} = \left[\frac{H\dot{\epsilon}}{1-\epsilon} + \frac{\ddot{\epsilon}}{1-\epsilon} + 2\left(\frac{\dot{\epsilon}}{1-\epsilon}\right)^2\right]\Delta\nu = H^2\left[\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\Delta\nu, \tag{A10}$$

$$2\dot{z}\,\dot{\nu} = \left[-2H\dot{\epsilon} + 2\left(\frac{\dot{\epsilon}}{1-\epsilon}\right)^2\right]z\Delta\nu = H^2\left[-2\partial_N\epsilon + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]z\Delta\nu, \tag{A11}$$

$$\dot{\nu}^2 = \left(\frac{\dot{\epsilon}}{1-\epsilon}\right)^2\Delta\nu^2 = H^2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\Delta\nu^2, \tag{A12}$$

$$\dot{z}^2 - \frac{k^2}{a^2} = \left[-2H\dot{\epsilon} + \left(\frac{\dot{\epsilon}}{1-\epsilon}\right)^2\right]z^2 = H^2\left[-2\partial_N\epsilon + \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]z^2. \tag{A13}$$

In comparing expressions (A9)–(A13) with (A7) it will be seen that every derivative with respect to z is paired with one factor of z, and every derivative with respect to ν is paired with one factor of $\Delta\nu$. This suggests differentiating with respect to the logarithms,

$$\zeta \equiv \ln(z) \Rightarrow z\frac{\partial}{\partial z} = \frac{\partial}{\partial\zeta}, \qquad \xi \equiv \ln(\Delta\nu) \Rightarrow \Delta\nu\frac{\partial}{\partial\nu} = \frac{\partial}{\partial\xi}. \tag{A14}$$

From (A7) it is also apparent that the factor of $2ka^2$ in the denominator of $M_0(t,k)$ is always canceled, by either ratios or explicit factors. It is best to define a new variable for the logarithm of the factor in the numerator (74),

$$\sigma(z,\nu) \equiv \ln\left[\frac{\pi z}{2}|H_\nu^{(1)}(z)|^2\right] = \ln\left[\frac{\pi z}{2}[J_\nu^2(z) + N_\nu^2(z)]\right]. \tag{A15}$$

Our final form for the source is therefore

$$\frac{S(t,k)}{H^2} = -\left[\frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\frac{\partial\sigma}{\partial\zeta} - \left[\partial_N\epsilon + \frac{\partial_N^2\epsilon}{1-\epsilon} + 2\left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\frac{\partial\sigma}{\partial\xi}$$
$$-2\left[-\partial_N\epsilon + \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\left[\frac{\partial^2\sigma}{\partial\zeta\partial\xi} + \frac{1}{2}\frac{\partial\sigma}{\partial\zeta}\frac{\partial\sigma}{\partial\xi}\right] - \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\left[\frac{\partial^2\sigma}{\partial\xi^2} - \frac{\partial\sigma}{\partial\xi} + \frac{1}{2}\left(\frac{\partial\sigma}{\partial\xi}\right)^2\right]$$
$$-2\left[-2\partial_N\epsilon + \left(\frac{\partial_N\epsilon}{1-\epsilon}\right)^2\right]\left[\frac{(2-\epsilon)}{(1-\epsilon)^2} + e^{2\zeta}(e^{-2\sigma}-1)\right]. \tag{A16}$$

We obviously want to make the same changes on the left hand side of (A7). It is also desirable to change the dependent variable from ΔM to $h(t,k) \equiv -2\ln[\Delta M(t,k)]$, all of which implies

$$\frac{\Delta\ddot{M}}{\Delta M} + \left[3H + \frac{\dot{M}_0}{M_0}\right]\frac{\Delta\dot{M}}{\Delta M} - \frac{1}{2}\left(\frac{\Delta\dot{M}}{\Delta M}\right)^2 + \frac{1}{2a^6M_0^2}\left[1 - \frac{1}{\Delta M^2}\right]$$
$$= -\frac{H^2}{2}\left\{\partial_N^2 h - \left[\frac{1}{2}\partial_N h\right]^2 + [1-\epsilon+\partial_N\sigma]\partial_N h + [2(1-\epsilon)e^{\zeta-\sigma}]^2[e^h - 1]\right\}. \tag{A17}$$

Equating (A16) to (A17) and dividing out the common factor of $-\frac{1}{2}H^2$ gives the final form (75) of our evolution equation.

APPENDIX B: EQUATION (75) AT EARLY TIMES

In the early time regime the parameter $z(t,k) \equiv \frac{k}{(1-\epsilon)Ha}$ is large which implies

$$\frac{\pi z}{2}|H_\nu^{(1)}(z)|^2 = 1 + \frac{(\nu^2-\frac{1}{4})}{2z^2} + \frac{3(\nu^2-\frac{1}{4})(\nu^2-\frac{9}{4})}{8z^4} + O\left(\frac{1}{z^6}\right). \tag{B1}$$

Hence the early time expansion for $\sigma(z,\nu)$ is

$$\sigma(z,\nu) = \frac{(\nu^2-\frac{1}{4})}{2z^2} + \frac{(\nu^2-\frac{1}{4})(\nu^2-\frac{13}{4})}{4z^4} + \dots. \tag{B2}$$

Expression (B2) implies the following expansions for the various σ-dependent factors in (75):

$$[1-\epsilon+\partial_N\sigma] = 1-\epsilon+O\left(\frac{1}{z^2}\right), \tag{B3}$$

$$[2(1-\epsilon)e^{\zeta-\sigma}]^2 = 4(1-\epsilon)^2z^2 + O(1), \tag{B4}$$

$$\frac{\partial\sigma}{\partial\zeta} = -\frac{(\nu^2-\frac{1}{4})}{z^2} + O\left(\frac{1}{z^4}\right),$$
$$\frac{\partial\sigma}{\partial\xi} = \frac{(\nu^2-\frac{1}{2}\nu)}{z^2} + O\left(\frac{1}{z^4}\right), \tag{B5}$$

$$\left[\frac{\partial^2\sigma}{\partial\zeta\partial\xi} + \frac{1}{2}\frac{\partial\sigma}{\partial\zeta}\frac{\partial\sigma}{\partial\xi}\right] = -\frac{(2\nu^2-\nu)}{z^2} + O\left(\frac{1}{z^4}\right), \tag{B6}$$

$$\left[\frac{\partial^2\sigma}{\partial\xi^2} - \frac{\partial\sigma}{\partial\xi} + \frac{1}{2}\left(\frac{\partial\sigma}{\partial\xi}\right)^2\right] = \frac{(\nu-\frac{1}{2})^2}{z^2} + O\left(\frac{1}{z^4}\right), \tag{B7}$$

$$\left[\frac{(2-\epsilon)}{(1-\epsilon)^2} + e^{2\zeta}(e^{-2\sigma}-1)\right] = \frac{3(\nu^2-\frac{1}{4})}{2z^2} + O\left(\frac{1}{z^4}\right). \tag{B8}$$

Substituting these expansions in (75) and additionally neglecting nonlinear terms in $h(t,k)$ gives Eq. (79).

APPENDIX C: EQUATION (75) AT LATE TIMES

In the late time regime of $z(t,k) \ll 1$, but still with $0 \le \epsilon(t) < 1$, it is the small argument expansion of the Neumann function in (A15) which controls the behavior of $\sigma(z,\nu)$,

$$\sigma = \ln\left[\frac{\Gamma^2(\nu)}{\pi}\left(\frac{2}{z}\right)^{2\nu-1}\right] + O(z^2)$$
$$= 2\Delta\nu\left[\Delta N + \ln\left[\frac{H}{H(t_k)}\right]\right] + \ln[C(\epsilon)] + O(z^2). \tag{C1}$$

Its derivative involves the digamma function $\psi(z) \equiv \frac{d}{dz}\ln[\Gamma(z)]$,

$$\partial_N\sigma = 2 + 2\Delta\nu\frac{\partial_N\epsilon}{1-\epsilon}\left[-1-\ln\left(\frac{z}{2}\right) + \psi\left(\frac{1}{2}+\Delta\nu\right)\right] + O(z^2). \tag{C2}$$

The term in square brackets is defined in expression (86) and grows roughly linearly in N. The seven σ-dependent factors of expression (75) have the expansions

$$[1-\epsilon+\partial_N\sigma] = 3-\epsilon+\frac{\partial_N\epsilon}{1-\epsilon}\times 2\Delta\nu F + O(z^2), \tag{C3}$$

$$[2(1-\epsilon)e^{\zeta-\sigma}]^2 = 4(1-\epsilon)^2z^2\times\left(\frac{z}{2}\right)^{4\Delta\nu}\frac{\pi^2}{\Gamma^4(\nu)} + O(z^{4+4\Delta\nu}), \tag{C4}$$

$$\frac{\partial\sigma}{\partial\zeta} = -2\Delta\nu + O(z^2),$$
$$\frac{\partial\sigma}{\partial\xi} = 2\Delta\nu(F+1) + O(z^2), \tag{C5}$$

$$\left[\frac{\partial^2\sigma}{\partial\zeta\partial\xi} + \frac{1}{2}\frac{\partial\sigma}{\partial\zeta}\frac{\partial\sigma}{\partial\xi}\right] = -2\Delta\nu^2(F+1) - 2\Delta\nu + O(z^2), \tag{C6}$$

$$\left[\frac{\partial^2\sigma}{\partial\xi^2} - \frac{\partial\sigma}{\partial\xi} + \frac{1}{2}\left(\frac{\partial\sigma}{\partial\xi}\right)^2\right] = 2\Delta\nu^2\left[(F+1)^2 + \psi'\left(\frac{1}{2} + \Delta\nu\right)\right] + O(z^2), \tag{C7}$$

$$\left[\frac{(2-\epsilon)}{(1-\epsilon)^2} + e^{2\zeta}(e^{-2\sigma} - 1)\right] = \Delta\nu + \Delta\nu^2 + O(z^2). \tag{C8}$$

From (C4) the restoring force drops out of (75), but all the other terms contribute to give the late time limiting form (85).

[1] A. A. Starobinsky, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 719 (1979) [JETP Lett. **30**, 682 (1979)].
[2] V. F. Mukhanov and G. V. Chibisov, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 549 (1981) [JETP Lett. **33**, 532 (1981)].
[3] R. P. Woodard, Rep. Prog. Phys. **72**, 126002 (2009).
[4] A. Ashoorioon, P. S. Bhupal Dev, and A. Mazumdar, Mod. Phys. Lett. A **29**, 1450163 (2014).
[5] L. M. Krauss and F. Wilczek, Phys. Rev. D **89**, 047501 (2014).
[6] G. F. Smoot *et al.*, Astrophys. J. **396**, L1 (1992).
[7] G. Hinshaw *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **208**, 19 (2013).
[8] Z. Hou *et al.*, Astrophys. J. **782**, 74 (2014).
[9] J. L. Sievers *et al.* (Atacama Cosmology Telescope Collaboration), J. Cosmol. Astropart. Phys. 10 (2013) 060.
[10] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **571**, A16 (2014).
[11] P. A. R. Ade *et al.* (BICEP2 Collaboration), Phys. Rev. Lett. **112**, 241101 (2014).
[12] R. Adam *et al.* (Planck Collaboration), arXiv:1409.5738.
[13] P. A. R. Ade *et al.* (BICEP2 and Planck Collaborations), Phys. Rev. Lett. **114**, 101301 (2015).
[14] P. A. R. Ade *et al.* (Planck Collaboration), arXiv: 1502.01589.
[15] K. Hattori *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **732**, 299 (2013).
[16] J. Lazear *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9153**, 91531L (2014).
[17] A. S. Rahlin *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9153**, 915313 (2014).
[18] Z. Ahmed *et al.* (BICEP3 Collaboration), Proc. SPIE Int. Soc. Opt. Eng. **9153**, 91531N (2014).
[19] K. MacDermid *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9153**, 915311 (2014).
[20] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).
[21] A. R. Liddle and D. H. Lyth, Phys. Rep. **231**, 1 (1993).
[22] J. E. Lidsey, A. R. Liddle, E. W. Kolb, E. J. Copeland, T. Barreiro, and M. Abney, Rev. Mod. Phys. **69**, 373 (1997).
[23] E. D. Stewart and D. H. Lyth, Phys. Lett. B **302**, 171 (1993).
[24] L. M. Wang, V. F. Mukhanov, and P. J. Steinhardt, Phys. Lett. B **414**, 18 (1997).
[25] J. Martin and D. J. Schwarz, Phys. Rev. D **62**, 103520 (2000).
[26] R. Easther, H. Finkel, and N. Roth, J. Cosmol. Astropart. Phys. 10 (2010) 025.
[27] M. J. Mortonson, H. V. Peiris, and R. Easther, Phys. Rev. D **83**, 043505 (2011).
[28] R. Easther and H. V. Peiris, Phys. Rev. D **85**, 103533 (2012).
[29] Y. Urakawa and T. Tanaka, Prog. Theor. Phys. **122**, 779 (2009).
[30] Y. Urakawa and T. Tanaka, Prog. Theor. Phys. **122**, 1207 (2009).
[31] S. B. Giddings and M. S. Sloth, J. Cosmol. Astropart. Phys. 01 (2011) 023.
[32] S. B. Giddings and M. S. Sloth, J. Cosmol. Astropart. Phys. 07 (2010) 015.
[33] C. T. Byrnes, M. Gerstenlauer, A. Hebecker, S. Nurmi, and G. Tasinato, J. Cosmol. Astropart. Phys. 08 (2010) 006.
[34] Y. Urakawa and T. Tanaka, Phys. Rev. D **82**, 121301 (2010).
[35] M. Gerstenlauer, A. Hebecker, and G. Tasinato, J. Cosmol. Astropart. Phys. 06 (2011) 021.
[36] S. B. Giddings and M. S. Sloth, Phys. Rev. D **84**, 063528 (2011).
[37] S. P. Miao and R. P. Woodard, J. Cosmol. Astropart. Phys. 07 (2012) 008.
[38] T. Tanaka and Y. Urakawa, Prog. Theor. Exp. Phys. **2013**, 083E01 (2013).
[39] T. Tanaka and Y. Urakawa, Prog. Theor. Exp. Phys. **2013**, 063E02 (2013).
[40] S. P. Miao and S. Park, Phys. Rev. D **89**, 064053 (2014).
[41] T. Tanaka and Y. Urakawa, Prog. Theor. Exp. Phys. **2014**, 073E01 (2014).
[42] A. Loeb and M. Zaldarriaga, Phys. Rev. Lett. **92**, 211301 (2004).
[43] S. Furlanetto, S. P. Oh, and F. Briggs, Phys. Rep. **433**, 181 (2006).
[44] M. G. Romania, N. C. Tsamis, and R. P. Woodard, Classical Quantum Gravity **30**, 025004 (2013).
[45] M. G. Romania, N. C. Tsamis, and R. P. Woodard, J. Cosmol. Astropart. Phys. 08 (2012) 029.
[46] E. D. Stewart, Phys. Rev. D **65**, 103508 (2002).
[47] C. Dvorkin and W. Hu, Phys. Rev. D **81**, 023518 (2010).

[48] D. Polarski and A. A. Starobinsky, Phys. Lett. B **356**, 196 (1995).
[49] J. Garcia-Bellido and D. Wands, Phys. Rev. D **52**, 6739 (1995).
[50] M. Sasaki and E. D. Stewart, Prog. Theor. Phys. **95**, 71 (1996).
[51] R. P. Woodard, Int. J. Mod. Phys. D **23**, 1430020 (2014).
[52] N. C. Tsamis and R. P. Woodard, Phys. Rev. D **80**, 083512 (2009).
[53] N. C. Tsamis and R. P. Woodard, Phys. Rev. D **81**, 103509 (2010).
[54] M. G. Romania, N. C. Tsamis, and R. P. Woodard, Lect. Notes Phys. **863**, 375 (2013).
[55] A. Vilenkin and L. H. Ford, Phys. Rev. D **26**, 1231 (1982).
[56] A. D. Linde, Phys. Lett. **116B**, 335 (1982).
[57] A. A. Starobinsky, Phys. Lett. **117B**, 175 (1982).
[58] A. Dolgov and D. N. Pelliccia, Nucl. Phys. **B734**, 208 (2006).
[59] D. S. Salopek, J. R. Bond, and J. M. Bardeen, Phys. Rev. D **40**, 1753 (1989).
[60] N. C. Tsamis and R. P. Woodard, Ann. Phys. (N.Y.) **267**, 145 (1998).
[61] T. D. Saini, S. Raychaudhury, V. Sahni, and A. A. Starobinsky, Phys. Rev. Lett. **85**, 1162 (2000).
[62] S. Capozziello, S. Nojiri, and S. D. Odintsov, Phys. Lett. B **634**, 93 (2006).
[63] R. P. Woodard, Lect. Notes Phys. **720**, 403 (2007).
[64] Z. K. Guo, N. Ohta, and Y. Z. Zhang, Mod. Phys. Lett. A **22**, 883 (2007).
[65] N. C. Tsamis and R. P. Woodard, Classical Quantum Gravity **21**, 93 (2004).
[66] A. A. Starobinsky, Pis'ma Zh. Eksp. Teor. Fiz. **55**, 477 (1992) [JETP Lett. **55**, 489 (1992)].
[67] V. F. Mukhanov and M. I. Zelnikov, Phys. Lett. B **263**, 169 (1991).
[68] J. A. Adams, B. Cresswell, and R. Easther, Phys. Rev. D **64**, 123514 (2001).
[69] D. J. H. Chung, E. W. Kolb, A. Riotto, and I. I. Tkachev, Phys. Rev. D **62**, 043508 (2000).
[70] O. Elgaroy, S. Hannestad, and T. Haugboelle, J. Cosmol. Astropart. Phys. 09 (2003) 008.
[71] N. C. Tsamis and R. P. Woodard, J. Cosmol. Astropart. Phys. 09 (2014) 008.