OPEN ACCESS
IOP Publishing

Classical and Quantum Gravity

Class. Quantum Grav. 39 (2022) 245002 (25pp)

https://doi.org/10.1088/1361-6382/aca066

Universal signature of quantum entanglement across cosmological distances

Suddhasattwa Brahma*, **Arjun Berera and Jaime Calderón-Figueroa**

Higgs Centre for Theoretical Physics, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom

E-mail: suddhasattwa.brahma@gmail.com

Received 15 July 2022; revised 17 October 2022
Accepted for publication 4 November 2022
Published 16 November 2022

Abstract

Although the paradigm of inflation has been extensively studied to demonstrate how macroscopic inhomogeneities in our Universe originate from quantum fluctuations, most of the established literature ignores the crucial role that *entanglement* between the modes of the fluctuating field plays in its observable predictions. In this paper, we import techniques from quantum information theory to reveal hitherto undiscovered predictions for inflation which, in turn, signals how quantum entanglement across cosmological scales can affect large scale structure. Our key insight is that observable long-wavelength modes must be part of an *open quantum system*, so that the quantum fluctuations can decohere in the presence of an environment of short-wavelength modes. By assuming the simplest model of single-field inflation, and considering the leading order interaction term from the gravitational action, we derive a *universal lower bound* on the observable effect of such inescapable entanglement. Although this signal is too weak for direct detection in the foreseeable future, we discuss the importance of its theoretical implications.

Keywords: inflation, entanglement, open quantum system, primordial vacuum fluctuations, scalar perturbations

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.

1. Introduction

Non-local behaviour, manifested through quantum entanglement, is a characteristic trait of quantum theory. An implication of this is that physics on different scales are able to influence each other. Conversely, the Wilsonian prescription tells us that it is possible to derive effective field theories (EFTs) which can describe physics below a particular energy scale by integrating out high-energy degrees of freedom (dofs). Therefore, it is important to understand what are the conditions under which a quantum system allows for a derivative expansion, or some analogous controlled approximation, for having the usual EFT description.

An early epoch of accelerated expansion [1–6]—inflation—presents a rare example of interplay between physics on microscopic and macroscopic scales since it explains the rich structure of late-time inhomogeneities, observed in the distributions of galaxies and the temperature variations in the cosmic microwave background (CMB), as being sourced by minuscule quantum mechanical vacuum fluctuations [7]. Primordial quantum fluctuations, whose physical wavelengths are stretched by the background expansion, become larger than the comoving Hubble radius at some point in the remote past and then re-enter as classical perturbations which we observe today. However, since the modes on super-Hubble scales alone form our system, we need to figure out which is the best EFT that describes the dynamics of these long-wavelength dofs.

Our main findings are that there are non-negligible corrections to the power spectrum arising from this primordial entanglement due to UV-IR mode-mixing which are, in principle, detectable from future observations. Although the corrections to the power spectrum and the spectral index is very small for slow roll inflation, nevertheless the effect of such corrections on the running of the spectral index, or the running of the running, is detectable in principle. Secondly, in order for the perturbative description to be valid, inflation cannot be semi-infinite in the past since the small corrections from primordial entanglement build up with time. A by-product of our work is that it serves as an indirect test for cubic non-Gaussianities in vanilla slow-roll models of inflation, which cannot otherwise be detected directly from any observation [8, 9]. Since these leading order nonlinearities necessarily act as an interaction term responsible for mode-couplings between the system and environment dofs, verification of our prediction shall indirectly validate their existence. (More complicated models of inflation, which allow for detectable amounts of non-Gaussianity, will only enhance this effect.) Finally, since entanglement is a purely quantum phenomenon, our results present a smoking gun for the quantum origin of inflation. In principle, present observations of macroscopic inhomogeneities can be modelled by a classical probability distribution and an unequivocal test of their quantum origin (such as cosmological Bell inequalities [10–13] or specific features in the non-Gaussianity functions [14]) are typically too small for realistic models of inflation [15, 16]. A null detection of our predictions, from future observations, will definitively rule out a quantum origin for inflation since we give a universal lower bound of a verifiable effect which must be present for quantum fluctuations resulting from a single scalar field dynamics. Our methods, however, are completely general and lay the foundation for calculating the same effect for any theory of the early-Universe, e.g. *Ekpyrosis* [17, 18], which sources inhomogeneities from quantum fluctuations.

The paper is organized as follows. After explaining in a bit more detail why standard EFT does not apply to inflation in the next section, we describe our setup and how an open quantum system approach applies to it. We go on to calculate the dynamics of the reduced density matrix of the system dofs by solving the kernel, and then go on to calculate the power spectrum using this reduced density matrix. Finally, we conclude by emphasizing the main ramifications of our findings. We have relegated the details of solving the master equation, and related issues of non-perturbative resummation and renormalization, to the appendices.

2. Inflation as an open quantum system

As with any system in physics, it is customary to treat inflationary dynamics as that of a system of the infrared modes which are of observational interest to us. Of course, this is simply a restatement of the fact that we do not have observational access to the entire Hilbert space of fluctuations. In this sense, there is nothing unusual or exotic in treating primordial perturbations as an open system in inflation, and this is what is customarily done in any case. The interesting new aspect is a rather technical one—how does one consider an effective theory of these low-energy (super-Hubble) perturbations? Typically, one would turn to Wilsonian EFT to evaluate the correlation functions of interest and then calculate quantum effects in the form of loop corrections coming from higher energy modes. However, for a cosmological setup, such a clean distinction between low and high energy dofs is not possible as we show below. Instead, one needs to use the setup of an open quantum system, and calculate the evolution of the reduced density matrices for the modes of interest. This is the novelty of our work since it shows how this method is much more general than evaluating standard loop corrections, if deviations from Markovian behaviour, and other such subtleties which appear in gravitational systems, are to be captured effectively.

The crucial thing to note in this case is that the standard Wilsonian effective action does not exist since the sub-Hubble modes, which are integrated out, are not excluded by any conservation law [19, 20]. This is contrary to traditional EFTs in which energy conservation ensures that high-energy dofs cannot be part of the system if they were not initially present, and the entire dynamics of the low-energy system can be described by an effective Lagrangian consisting only of the light dofs. In the cosmological system, the long wavelength modes interact and exchange energy with the sub-Hubble ones and therefore, the time evolution of the system under consideration cannot be described by unitary evolution of some low-energy Hamiltonian. In fact, these long wavelength dofs need to be treated as an open quantum system rather than an isolated one [21], which is coupled to an environment of unobservable short wavelength fluctuations, thereby allowing for a pure state to evolve into a mixed one [22]. The non-Hamiltonian evolution underlying this process is the main physical insight which allows physics on the shortest scales to affect structure formation on the largest scales of our Universe, going against intuition derived from Wilsonian EFT since we need to study an *out-of-equilibrium* system in this case. The key question, which we address in this article, is what are the observable consequences of this primordial entanglement between the long and short wavelength modes?

The simplest model of inflation consists of a (minimally-coupled) single scalar field, with a postulated form for its potential to allow for an accelerated expansion and then a graceful exit from it. Gravitational nonlinearities, arising from the Einstein–Hilbert action [23], lead to coupling between the momentum modes (in Fourier space) of cosmological perturbations. We consider the coupling between the band of super-Hubble modes and unobservable short-wavelength ones to be given by gravity and, is thus, *inescapable*. From more general considerations, one expects cosmological perturbations to be gravitationally coupled to other dofs present in the Universe, and this shall only result in a larger amount of entanglement, and a subsequent magnification of our result. Consequently, we present a *universal lower bound* on the observable consequences of primordial quantum entanglement—universal since this is due to interactions arising from pure gravity (and must, therefore, be present in any theory which has general relativity as its low energy limit) and it is valid for *any* model of inflation, since additional fields will only lead to a bigger effect (see e.g. [24–26]).

The novelty of our work is twofold: Firstly, we consider corrections to the primordial spectra of observable modes coming from quantum entanglement due to considering an *open EFT* and secondly, we consider interactions between long and short dofs (where the Hubble horizon of the quasi-de Sitter (dS) background sets the reference scale) of the inflation fluctuations themselves (as opposed to considering additional postulated fields) due to gravitational non-linearities. Even on a conceptual level, the quantum-to-classical transition of cosmological perturbations necessarily relies on treating them as part of an open system, thereby allowing for decoherence. This is why we need to use techniques for open EFTs in which one studies the (non-unitary) evolution of the reduced density matrix obtained by tracing over the unobservable dofs. These methods are prevalent in other branches of physics, and we adapt them here for cosmology, in the hope of measuring the degree of entanglement, between our super-Hubble system modes with its environment (see [27, 28] for entanglement entropy calculations in this setup), through observable predictions. Therefore, we strengthen the bonds between cosmology and quantum information theory by uncovering new observable effects in the former, which have remained out of bounds of closed quantum system treatments with Wilsonian EFTs, by importing mathematical tools from the latter.

2.1. *Failure of Wilsonian effective action*

Beside the general statements made above, let us give a more quantitative argument as to why the Wilsonian effective action is not applicable for inflation, and can only be used as a first approximation. We will follow [30], and reproduce their analysis, to convey our main message.

Let us break down the modes of cosmological perturbations in momentum space schematically as follows:

$$\varphi \sim \varphi_{k>aH} \otimes \varphi_{k<aH}, \tag{1}$$

where the first term is to be understood as the short-wavelength modes which form the environment for our system modes (denoted abstractly by the second term). Throughout we will assume that our Hilbert space has a UV cutoff (something akin to the Planck scale), below which our effective description makes sense.

The standard starting point for calculating the Wilsonian effective action is the Euclidean path integral, defined as:

$$e^{-S_{\mathrm{W}}(\varphi_{k<aH})} := \int D\varphi_{k>aH}(\tau)\, e^{-S_{\mathrm{E}}(\varphi_{k<aH},\varphi_{k>aH})}, \tag{2}$$

where τ is Euclidean time. Interestingly, this Wilsonian effective action is related to the low-energy density matrix is the following way [30, 31] (where $\mathcal{N}$ is just some normalization factor):

$$\left\langle \varphi^{(f)}_{k<aH} \middle| \rho_{k<aH} \middle| \varphi^{(i)}_{k<aH} \right\rangle = \mathcal{N} \int_{\varphi_{k<aH}(\tau=0^-):=\varphi^{(i)}_{k<aH}}^{\varphi_{k<aH}(\tau=0^+):=\varphi^{(f)}_{k<aH}} D\varphi_{k<aH}(\tau)\, e^{-S_{\mathrm{W}}(\varphi_{k<aH})}. \tag{3}$$

The crucial thing to realize is how one needs to put periodic boundary conditions in the Euclidean action, for the high-energy modes, in order to arrive at the above expression (3) for the low-energy density matrix. This can be seen most straightforwardly from the fact that the low-energy density matrix is given by:

$$\left\langle \varphi^{(f)}_{k<aH} \middle| \rho_{k<aH} \middle| \varphi^{(i)}_{k<aH} \right\rangle = \mathcal{N} \int_{\varphi_{k<aH}(\tau=0^-):=\varphi^{(i)}_{k<aH}}^{\varphi_{k<aH}(\tau=0^+):=\varphi^{(f)}_{k<aH}} D\varphi_{k<aH}(\tau) \times D\varphi_{k>aH}(\tau) e^{-S_{\mathrm{E}}}, \tag{4}$$

where the fact that the high-energy modes are periodic over the limits of the path-integral is implied above. We can rewrite the above expression as:

$$\left\langle \varphi^{(f)}_{k<aH} \middle| \rho_{k<aH} \middle| \varphi^{(i)}_{k<aH} \right\rangle = \int D\varphi_{k>aH}(\tau) \left\langle \varphi^{(f)}_{k<aH}\, \tilde{\varphi}_{k>aH} \middle| \rho \middle| \varphi^{(i)}_{k<aH}\, \tilde{\varphi}_{k>aH} \right\rangle, \tag{5}$$

where ρ denotes the full density matrix. All we have done is to show that the low-energy density matrix can be written in terms of the Wilsonian effective action, provided we choose periodic boundary conditions for the high-energy modes that we are integrating out. From (5), it is clear that, when continued onto real time, these periodic boundary conditions translates to fixing *both* the initial and the final states of the high-energy modes to be in the ground state.

This is the key insight which we want to harp upon a bit longer—it is equivalent to work with the low-energy density matrix or use the Wilsonian effective action only if we assume the integrated out modes to remain in their vacuum state. However, during inflation, not only is the Hamiltonian time-dependent but also the boundary separating system modes from the environment ones are time-dependent itself. In other words, a mode which is environment at one time, need not be so at a later point. This is why the Wilsonian effective action is only a (very rough) first approximation to deal with quantum fields on an inflationary background. There is no reason to assume that the initial and final state of the environment degrees of freedom are both the ground state and, in fact, what we show in this paper is that entanglement builds up between the UV and IR modes in such a way that this standard Wilsonian picture is unable to capture. In other words, even when we assume that the initial state of the system is factorized between the ground state of the short wavelength modes and the ground state of the long wavelength ones, due to the dynamics, this factorization does *not* hold at a later time. There is a path-integral method of dealing with this—instead of the Wilsonian effective action, one needs to invoke the Feynman-Vernon influence functional [30]. However, in this paper we shall work in the canonical picture where we shall write down the dynamics of the reduced density matrix in terms of a Hamiltonian part and a non-Hamiltonian evolution (denoted by Lindblad operators). What is important to emphasize is that although we will assume Markovian dynamics for our system density matrix, this is still an approximation which we will seek to relax in future work.

The bottomline is then as follows: calculating standard radiative corrections assuming a Wilsonian effective action is insufficient for a time-dependent background in which our observable degrees of freedom only forms a subsystem of the entire Hilbert space. What is instead required is an open EFT approach to the problem—one in which we calculate the dynamics of the long wavelength density matrix, which can deviate (albeit by a small amount due to the slow-roll nature of the background) from the standard Wilsonian low-energy effective action. This deviation captures the measure of entanglement, and the associated dissipative effects, which builds up between the long and short wavelength modes. This latter quantity is necessarily missed in standard Wilsonian EFT and something which we will quantify here in this work.

3. Setting up the system

The quasi-dS space can be described by the homogeneous metric: $ds^2 = -dt^2 + a^2(t)d\mathbf{x}^2 = -a^2(\tau)\left(d\tau^2 - d\mathbf{x}^2\right)$ where the scale factor a, Hubble parameter $H = \dot{a}/a$ and the slow-roll parameters $\epsilon := -\dot{H}/H^2$, $\eta := \dot{\epsilon}/(H\epsilon)$ describe evolution of the background. Overdots and primes denote derivatives with respect to cosmic time 't' and conformal time 'τ', respectively, which are related by $ad\tau = dt$. Throughout, we shall assume that H is almost constant and the slow-roll parameters remain small during the entire evolution. We denote the comoving curvature perturbation as ζ, and the canonical Mukhanov-Sasaki variable as $\chi = z(\tau)\zeta$, where $z^2 = 2\epsilon a^2 M_{\mathrm{Pl}}^2$, where M_{Pl} is the reduced Planck mass. The underlying idea is that this comoving curvature perturbation, which is a combination of the quantum fluctuations of matter and the linearized gravitational field, can be described as a set of harmonic oscillators to leading order (in Fourier space), accentuated by interaction terms which result due to the non-linearity of the gravitational action.

The quadratic action for the canonical variable reads:

$$\mathcal{L}^{(2)} = \int d^3x \left[(\chi')^2 - (\partial\chi)^2 + \frac{z''}{z}\chi^2\right],$$

the solutions for which describe the dynamics of the mode functions to be used below. This is nothing but the Lagrangian for a harmonic oscillator with a time-dependent mass term. Going to Fourier space and introducing the usual ladder operators, the quadratic Hamiltonian:

$$\hat{H}^{(2)} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\left(k\left[\hat{c}_{\mathbf{k}}\hat{c}_{\mathbf{k}}^\dagger + \hat{c}_{-\mathbf{k}}\hat{c}_{-\mathbf{k}}^\dagger\right] - i\frac{z'}{z}\left[\hat{c}_{\mathbf{k}}\hat{c}_{-\mathbf{k}} - \hat{c}_{\mathbf{k}}^\dagger\hat{c}_{-\mathbf{k}}^\dagger\right]\right), \tag{6}$$

shows that there are two distinct terms which govern the free evolution of these modes. The first term is the usual Hamiltonian for a massless scalar field in flat space (a collection of harmonic oscillators) whereas the second term is the squeezing interaction, characteristic of the curved space background. Since the background is time-dependent, it sources zero-momentum correlated pairs of the canonical field. This term dominates when $k \ll z'/z \approx aH$, i.e. it determines the evolution once the physical wavelengths of these modes become super-Hubble whereas the sub-Hubble modes $k \gg aH$ are in their quantum vacuum, as shown by the first term above. We will assume that the modes are in the Bunch-Davies vacuum, so the mode functions take the form:

$$\chi_k(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{7}$$

In the presence of these terms in the quadratic Hamiltonian alone, the evolution of the quantum vacuum to the squeezed state is evidently a unitary one [32].

The story turns interesting once we introduce the cubic non-Gaussianities which imply that there must be higher order interaction terms which lead to mode-coupling between the quantum fluctuations. The cubic action for scalar perturbations contains several terms [23, 33], with the leading-order one given by $\mathcal{L}^{(3)} = \int d^3x\, a\epsilon^2\zeta(\partial\zeta)^2$. Going to the interaction picture, operators are written as $\hat{\mathcal{O}}_I = \hat{U}_0^\dagger\hat{\mathcal{O}}\hat{U}_0$, where $\hat{U}_0$ is the unitary operator corresponding to the quadractic Hamiltonian (6). In this formalism, the interaction Hamiltonian for the Mukhanov-Sasaki variable is given by:

$$\hat{H}_{\mathrm{I}}(\tau) = \lambda(\tau)\int d^3x\, \hat{\chi}(\tau,\mathbf{x})(\partial\hat{\chi}(\tau,\mathbf{x}))^2,$$

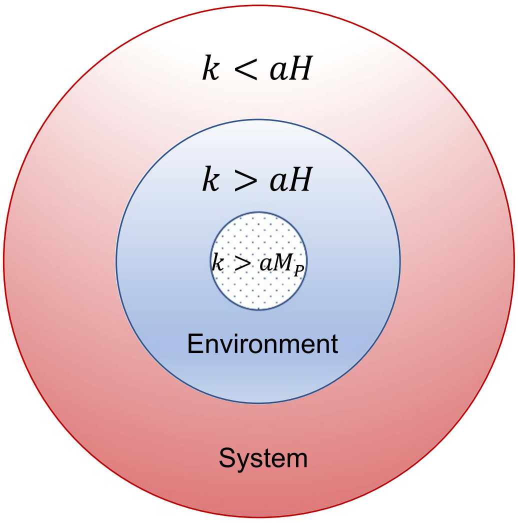


Figure 1. Illustration of the bath and environment modes for this problem.

where $\lambda \equiv -a^2\epsilon^2/z^3$, i.e. it is time-dependent. In the interaction picture, time evolution of the density matrix is governed by the von-Neumann equation:

$$\begin{aligned}\frac{d\rho_I}{d\tau} &= -i\left[\hat{H}_I(\tau), \rho_I(\tau)\right] \\ &= -i[\hat{H}_I(\tau), \rho_I(\tau_0)] - \int_{\tau_0}^{\tau} d\tau' \left\{\hat{H}_I(\tau)\hat{H}_I(\tau')\rho_I(\tau') - \hat{H}_I(\tau)\rho_I(\tau')\hat{H}_I(\tau')\right. \\ &\quad \left. -\hat{H}_I(\tau')\rho_I(\tau')\hat{H}_I(\tau) + \rho_I(\tau')\hat{H}_I(\tau')\hat{H}_I(\tau)\right\}, \end{aligned} \tag{8}$$

where, in the second line we have used the formal solution of the density matrix and replaced it back into the equation of motion to bring it to a more amenable form.

In order to describe the cosmological perturbations as part of an open quantum system, we need to write the density matrix in terms of system ($\mathcal{S}$) and environment ($\mathcal{E}$) dofs. To do so, we define a Hilbert space of quantum fluctuations separated into two regions—the super- and sub-Hubble mode spaces $\mathcal{H}_{\mathcal{S}}(t)$ and $\mathcal{H}_{\mathcal{E}}(t)$, respectively. Each of these are tensor products of the Fock space corresponding to the wavenumber k, i.e. $\mathcal{H}_{\mathcal{S}} = \prod \mathcal{H}_k,\ k < aH$ and similarly for $\mathcal{H}_{\mathcal{E}}$, with $k > aH$. These spaces are depicted in figure (1). We assume that the modes in $\mathcal{H}_{\mathrm{UV}}$ (i.e. the trans-Planckian modes) can be accounted for by using usual renormalization techniques, and remain agnostic about the quantum gravity theory at play. Note that the important feature is that the boundary between $\mathcal{H}_{\mathcal{S}}$ and $\mathcal{H}_{\mathcal{E}}$ depends on the dynamically expanding background, leading to a time-dependent Hilbert space for the system modes, $\mathcal{H}_{\mathcal{S}}$, and this is how non-unitarity creeps into the system.

This splitting also naturally extends for states. For instance, before considering any interactions, the full-state in the Schrödinger picture is given by $|\Psi(\tau)\rangle = |SQ(\tau)\rangle_{k<aH} \otimes |0\rangle_{k>aH}$, where $|SQ(\tau)\rangle$ denotes the squeezed state of super-Hubble modes induced by the second term on the r.h.s of equation (6), and $|0\rangle$ represents the Bunch-Davies vacuum. Once we go to the interaction picture, the same delimitation between system and environment holds, but the squeezed states take a different functional form since the time-evolution of the wavefunction is not determined by the free-theory unitary operator, but by the interaction Hamiltonian instead.

In any case, the form of the squeezed state in the interaction picture is not relevant for our purposes, since we shall take a perturbative approach where we operate on the initial states, which are identical in both the Schrödinger and interaction pictures. Thus, noticing that observationally relevant modes are those that have crossed the Hubble horizon, it is clear that every mode of interest has started out in the Bunch-Davies state, or, in other words, these modes started out as being part of the environment.

Likewise, operators are also split in the same fashion according to the comoving momenta of their Fourier modes. Hence, the interaction Hamiltonian can be expressed as (see the appendix for more details):

$$\hat{H}_I(\tau) = \lambda(\tau)\int d^3x\, \hat{\chi}^{\mathcal{S}}(\tau,\mathbf{x})\left(\partial\hat{\chi}^{\mathcal{E}}(\tau,\mathbf{x})\right)^2$$
$$= -\lambda(\tau)\int_{\Delta_k}(\mathbf{k}_2\cdot\mathbf{k}_3)\hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau)\hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau)\hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau), \tag{9}$$

where $\int_{\Delta_k} := \int d^3k_1\int d^3k_2\int d^3k_3\ (2\pi)^{-6}\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)$ ensures translation invariance. This particular combination of system and environment modes is the predominant one, since the derivatives favour larger momenta modes (i.e. environment modes) and by conservation of momentum, the remaining mode should belong to the system[1].

4. Master equation and kernel

Going back to the density matrix, we now assume that at τ_0 the system and environment are not entangled. This, together with a *weak coupling* between $\mathcal{S}$ and $\mathcal{E}$, implies $\rho_I(\tau) = \rho_{\mathcal{S}}(\tau)\otimes\rho_{\mathcal{E}}(\tau_0)$. Then, one can trace out the environmental dofs to find the *reduced density matrix*, $\rho_r(\tau) = \mathrm{Tr}_{\mathcal{E}}\left[\rho_I(\tau)\right]$, and its evolution through equation (8), such that:

$$\rho_r'(\tau) = \int\frac{d^3p}{(2\pi)^3}\,\lambda(\tau)\int_{\tau_0}^{\tau}d\tau'\,\lambda(\tau')\left\{\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\rho_r(\tau')K_p(\tau,\tau')\right.$$
$$-\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)\rho_r(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')K_p^*(\tau,\tau') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\rho_r(\tau')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)K_p(\tau,\tau')$$
$$\left.+\rho_r(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)K_p^*(\tau,\tau')\right\}, \tag{10}$$

where we have considered the Fourier expansion of the different fields, together with $\rho_{\mathcal{E}}(\tau_0) = |0\rangle\langle 0|$, to arrive at the *master equation* above (the *born approximation*). The function $K_p(\tau,\tau')$ is typically known as the *kernel*, which arises from tracing out and integrating over the environment dofs. Details of the calculation are shown in appendix B, where we show the kernel to be:

$$K_{p_1}(\tau,\tau') = -2\int\frac{d^3p_2}{(2\pi)^3}\left(\mathbf{p}_2\cdot\mathbf{p}_3\right)^2\chi^{\varepsilon}_{p_2}(\tau)\chi^{\varepsilon}_{p_2}(\tau')^*\chi^{\varepsilon}_{p_3}(\tau)\chi^{\varepsilon}_{p_3}(\tau')^*;$$
$$\mathbf{p}_3 = -(\mathbf{p}_1+\mathbf{p}_2). \tag{11}$$

Clearly the integrals depend sensitively on the Bunch-Davies mode functions assumed to characterize the initial state. In performing the integration and considering the leading-order behavior in τ, we get:

[1] However, note that this does not imply we are in the 'squeezed limit' configuration (in the common lore of cosmologists). In fact, it turns out that the maximum contribution to the momentum integrals come when all the modes are near-horizon.

$$K_p(\tau,\tau') \approx -\frac{e^{2i(\tau-\tau')/\tau}\left[1-e^{-ip(\tau-\tau')}\right]\left[\tau-(1-i)\tau'\right]^2}{8\pi^2 p\tau^4(\tau')^2(\tau-\tau')^2}. \tag{12}$$

The effect from modes in the far-UV (upper limit of the integral) is dismissed by assuming τ' has a small imaginary part, which can be thought of as a damping factor to make the term convergent in this limit[2]. This is the same prescription as for any interacting field theory, including the *in-in* formalism.

Finally, it is also worth commenting on the *Born-Markov approximation*, which assumes the temporal locality of the correlations of environment dofs. This allows one to write the master equation in terms of Lindblad operators and depict the non-unitary evolution of the density matrix, and are well-suited to study the suppression of coherence [34, 35]. Since these can be interpreted as randomly varying terms, they are equivalent to the stochastic inflation formalism [36] and is, thus, a complementary way to study open quantum systems. In general, non-unitarity dynamics are introduced by tracing out environment dofs, such that dissipation-like terms can be extracted from the master equation and its solution [37, 38]. This behaviour cannot be captured by, and is not equivalent to, calculating loop corrections in an ordinary QFT. Although we shall follow a perturbative method to solve for the master equation, one could have calculated Lindblad operators in a similar way for our cosmological setup [19, 21, 39].

5. Corrections to the power spectrum

In order to compute the corrections to the power spectrum due to entanglement we need to solve the master equation (10) through some suitable approximation(s). First, let us show how one gets an almost scale-invariant power spectrum produced by quantum fluctuations from the zeroth-order approximation, as follows:

$$\begin{aligned}\Delta_\zeta^2(q) &= \frac{q^3}{2\pi^2 z^2}\left\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\right\rangle \\ &= \frac{q^3}{2\pi^2 z^2}\mathrm{Tr}\left[\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\rho_r(\tau_0)\right] \approx \frac{1}{2\epsilon M_{\mathrm{Pl}}^2}\left(\frac{H}{2\pi}\right)^2,\end{aligned} \tag{13}$$

where $\rho_r(\tau_0) = |0\rangle\langle 0|$. The first-order corrections to the power spectrum come from the perturbative solution of the reduced density matrix equation (10), while also substituting the explicit form of the kernel equation (12) in the final expression. In doing so we get:

$$\begin{aligned}\rho_r(\tau) \approx \rho_r(\tau_0) &+ \sum_{\mathbf{p}}\int_{\tau_0}^{\tau} d\tau'\lambda(\tau')\int_{\tau_0}^{\tau'} d\tau''\lambda(\tau'') \\ &\times \Big\{\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)K_p(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'') \\ &\times K_p^*(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K_p(\tau',\tau'') \\ &+ \rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K_p^*(\tau',\tau'')\Big\},\end{aligned} \tag{14}$$

[2] In other words, this is the standard $i\epsilon$ prescription since for large p_2, one has to deal with terms of the form $e^{-ip_2(\tau-\tau')}$.

where we have taken $\rho_r(\tau'') \to \rho_r(\tau_0)$, as appropriate for this order of approximation. Then, we can use this expression to compute the two-point correlation function:

$$\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\rangle = \frac{1}{2q}\left(1+\frac{1}{(q\tau)^2}\right) + 2\int_{-1/q}^{\tau} d\tau'\lambda(\tau')\int_{-1/q}^{\tau'} d\tau''\lambda(\tau'') \times \left\{K_q(\tau',\tau'')\left[(\chi_q(\tau))^2\chi_q^*(\tau')\chi_q^*(\tau'') - |\chi_q(\tau)|^2\chi_q(\tau')\chi_q^*(\tau'')\right] + \text{c.c.}\right\}, \tag{15}$$

which ultimately leads to the power spectrum, including entanglement effects. Intermediate steps to arrive to this expression are shown in appendix C. Notice that the first term on the r.h.s. corresponds to the tree-level value, whereas entanglement effects are buried in the integrals. The lower integration limits are a consequence of the step functions filtering the environment mode functions. These limits imply that the environment will influence a particular mode only after it crosses the horizon, as one would expect.

The integration over τ'' in equation (15), or in equation (14) for that matter, can be performed analytically, however, the subsequent integration over τ' has to be done numerically since an analytical approximation could not be made. Furthermore, several terms in the final answer diverge, as a result of taking the equal time limit in the kernel equation (12). These divergences are manifestations of taking the coincidence limit for the power spectrum, when radiative corrections are included (see [40] for a detailed discussion of this), and one needs to focus on the finite terms alone. See the appendix for details of why the divergences are spurious E. The finite corrections are thus of the form:

$$\Delta^2_\zeta(q\tau) = \frac{1}{2\epsilon M_{\text{Pl}}^2}\left(\frac{H}{2\pi}\right)^2 f(N_c)\,, \tag{16}$$

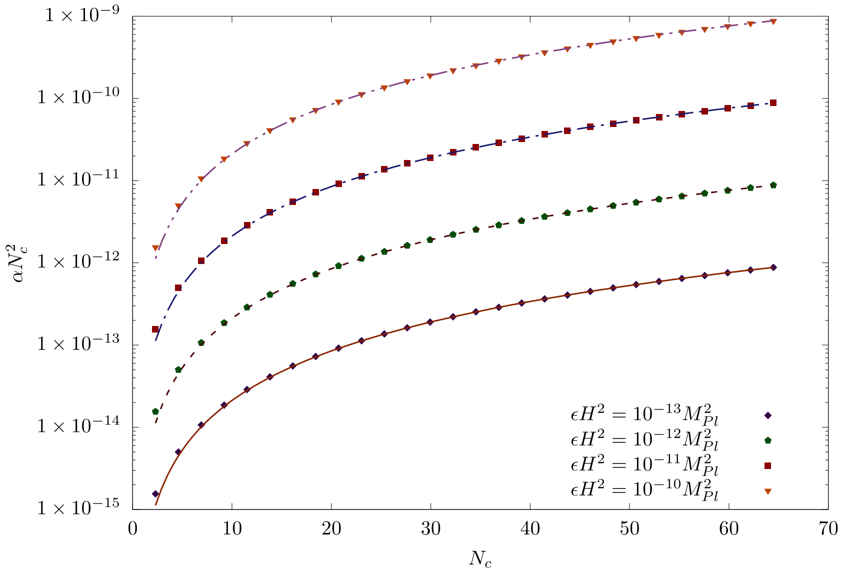
where $N_c := \ln(-1/q\tau)$ specifies the number of $e-$folds the Universe has expanded after a Fourier mode crossed the horizon ($q\tau_* = -1$) and became part of the system. The 'entanglement function' $f(N_c)$, corresponding to the perturbative solution (14), was found to be:

$$f(N_c) = 1 - \alpha N_c^2\,, \qquad \alpha \approx 0.00211886\,\frac{\epsilon H^2}{2M_{\text{Pl}}^2}, \tag{17}$$

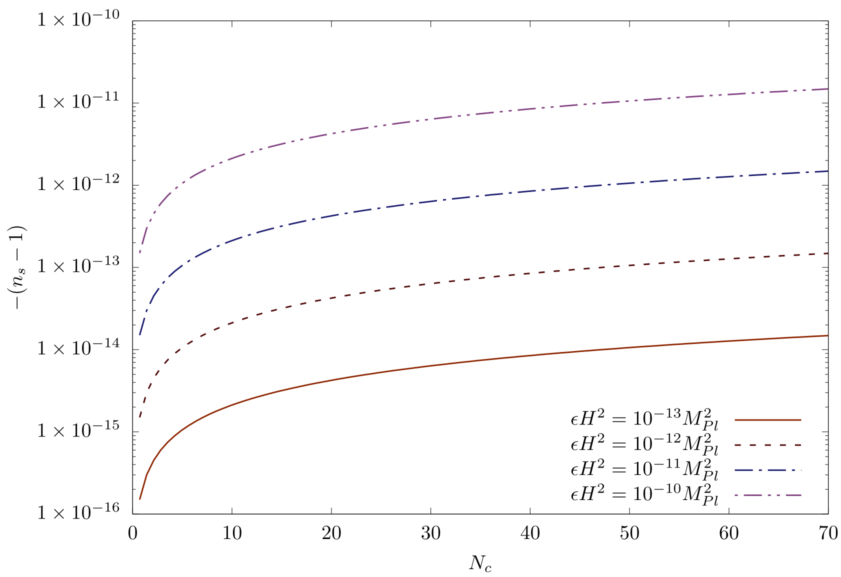
through the numerical integrations carried out in evaluating the two-point function using the reduced density matrix. Figure 2(a) shows the fit for different combinations of the parameters ϵH^2,[3] where is clear that the quadratic polynomial in N_c captures the behaviour of the correction to the power spectrum at least for $N_c \lesssim 100$.

Notice equation (16) is valid for any mode, from the very first sub-Hubble mode that crossed the horizon (for which N_c also denotes the total amount of inflation), all the way to the observable modes relevant for the CMB. The first thing to note is that the correction term is suppressed by a factor of $\epsilon H^2/M_{\text{Pl}}^2$, which could have been estimated from power counting as loop corrections to the two-point function of curvature perturbations, once everything is normalized and factors of the vertex operators (involving factors of the interaction parameter λ of the cubic term under consideration) and propagators are taken into account. The dependence on N_c^2 can be intuitively understood as making this correction larger with time—as entanglement entropy builds with more and more modes turning superhorizon [28], with the precise numerical prefactor in α resulting from numerical estimations in our open EFT and, thus, shows that our

[3] The factor of α here comes from fitting the curves shown in the plot. As will be shown later, we obtain a similar value for α when we perform the non-perturbative resummation in appendix D in equation (53).



(a)



(b)

Figure 2. (a) Deviation from 'tree-level' spectrum due to entanglement effects. For each case, a quadratic polynomial in N_c fits the data. The free parameter α is found to be determined by the product ϵH^2. The region below the dashed line (second parameter) is consistent with $r(=16\epsilon) < 0.036$, as found in [29]. (b) Deviation from scale invariance due to entanglement between sub-Hubble and super-Hubble modes. On a non-logarithmic scale, the functions are linear, with α corresponding to the running of the spectral index in this range.

calculation involves dissipative effects over and above straightforward radiative corrections [41–43] for the power spectrum. This brings up an interesting feature of our result—to be within the *regime of perturbative validity*, we find an upper bound on the duration of inflation. Since this implies a limit of $N_c \lesssim \alpha^{-1/2}$, using the specific values displayed in figure (2), one finds a bound of order $10^6 - 10^8$ *e*-folds, which although sufficiently long to solve the standard cosmological puzzles, puts an upper limit coming from perturbativity [44].

On the other hand, such secular divergences in a QFT in quasi-dS backgrounds usually arise from infrared effects [20, 45, 46] (see [47, 48] for some contrarian views). This naturally lead us to consider resummation of the master equation, yielding a nonperturbative result for the entanglement function given by $f(N_c) \simeq e^{-\alpha N_c^2}$. An exact analytic calculation using dynamical renormalization group [49] is difficult to perform in the absence of exact dS. Nevertheless, one can estimate the resummed function and we sketch the details of the calculation in appendix D. Other approaches to nonperturbative resummations have also yielded similar results for analogous systems [24, 49]. Interestingly, note that this result still carries signatures of entanglement in the form of the resummed function, as opposed to standard one-loop corrections in which one does not expect any additional time-dependence to appear [41, 42, 50]. The nonperturbative result simply ensures that the time-dependence is milder, and not secularly-divergent, as perturbative theory would seem to imply.

IR effects are mostly ignored for inflationary calculations—we provide a fresh perspective on this by showing how such effects cannot be ignored when the evolution of the reduced density matrix for the system modes is calculated by solving the master equation. In light of the above discussion, one is naturally led to the question—what is it that we are resumming and if these IR effects are anything that can be observed? The answer to both these questions lies in the realization that there are non-unitary contributions from the dissipation to the master equation, which cannot be thought of contributing to any particular diagram. In other words, it is impossible to perform a similar result from the loop corrections alone, as it does not take into account any procedure of coarse-graining the system dofs from the global wavefunction. The dissipation between the system and environment modes cannot be captured in the loop expansion alone. This, of course, implies that these type of IR effects cannot be distinguished in any local observation, although it can lead to detectable signatures in some global observables of the interacting system. This conclusion is in agreement with existing literature [51], and we are pointing out in this paper how entanglement between different modes has an IR part which leaves its mark in the CMB.

6. Discussion

Although inflation is consistent with observations, the data so far has not shown the type of smoking gun evidence that uniquely distinguishes its dynamical characteristics viz a viz other possible alternatives. What this paper has shown is that, in the next phase for theory in studying this process, the high precision demands now reach a non-Wilsonian regime of quantum field theory. It is crucial to realize that the standard loop calculations from Wilsonian EFT would never be able to capture the same physics as discussed here. There are of course similar looking features to the calculations done here and those for loops in standard Wilsonian calculations since, for slowly rolling backgrounds, the zeroth order approximation is to assume standard EFT. However, it is in the subtle difference where the precision from the calculations done here are realized.

To this end, in this work we have considered the effects on the primordial spectrum coming from the minimum amount of entanglement expected from single-field inflation. To do so we followed an open QFT approach, which captures the non-trivial interaction between system and environment dofs, the possibility of the later to be part of the former, as well as the fact that we have only access to the system. For these reasons, standard Wilsonian EFTs do not capture the entire dynamics of inflationary perturbations, in particular the resulting entanglement and its effects on observable quantities.

These corrections, illustrated in figure (2), are independent of whether we consider the resummed function or simply the first-order correction. Importantly, this is exactly opposite to the predictions for the spectral tilt $(n_s - 1)$, and its running, for standard inflationary models *without considering entanglement* [52]: In that case, both of these quantities typically decrease (in magnitude) with increasing amounts of N_c for the modes of interest [53]. We conclude that other classical effects that mimic this are unlikely.

One might be tempted to think that the numerical estimates for our correction, as shown in figure (2), can be made negligibly small by fine-tuning the value of ε to be arbitrarily tiny. However, this is not quite the case since apart from the cubic interaction term considered in this work, there exists another exactly similar term but now with the coefficient $\varepsilon\eta$. Given the current value of spectral tilt as measured by the Planck team, it implies that either of ε or η must be an $\mathcal{O}(0.01)$ number. It is, therefore, clear that our result would hold even for such fine-tuning. This point has been emphasized in the context of decoherence (and effective stochastic dynamics) [39], where this effect is caused by the same leading-order terms we consider in this work for calculating the entanglement between the modes. Therefore, our result reinforces the idea that decoherence is not only important for conceptual reasons but also has significance for cosmological observations [54].

Another point of contention can come from comparing our result with that expected from loop corrections, in particular the (lack of) time dependence of the spectrum of single–clock inflation. It has been argued in [42] that the time dependence emerging from the interaction here considered is cancelled by a quartic term of the form $-\zeta^2(\partial\zeta)^2$. Such a term is invoked under the argument of imposing diffeomorphism invariance under a time–dependent rescaling and translation of the spatial coordinates. However, this point of view has been recently challenged in [55, 56], where it is argued (in the context of non-Gaussianities) that one can only gauge away these effects in the exact limit $k=0$. Clearly, this is part of a larger discussion which is beyond the scope of this paper, and it would be interesting to compute the effects of the quartic interaction using open quantum system techniques. Alternatively, one could look at other observables with manifestly gauge invariant interactions, such as the tensor spectrum [57], which has the additional advantage of allowing to compare the Wilsonian and Open EFT approaches without these concerns. Moreover, this also puts into light the potential importance of open EFTs to study topics such as the IR divergences of dS spacetimes or the decoupling of UV modes.

What our work here has highlighted can be extrapolated beyond inflation and it can be argued that any early Universe mechanism that exclusively has a quantum origin of the Universe where all the disparate scales are produced (e.g. Ekpyrosis, modified gravity, etc), will have entanglement between the different modes, similar to what we found here from our explicit inflation calculations. In this work we examined the consequences of this entanglement between scales to the scalar power spectrum. This was done in the simplest single scalar field model, where the quantitative effect was not significant. However, as shown above, the calculation clearly points to the type of inflation models where the effect will not be negligible and the precision corrections could be observable with future data (e.g. for DBI inflation or curvation and other models with large f_{NL}). We simply provide a model-independent lower bound for this effect in inflation due to the nonlinearity of gravity alone. To reiterate, the generality of our effect is highlighted by the fact that this should be present for any early-Universe scenario which posits that macroscopic inhomogeneities originate from quantum fluctuations. For any such scenario, the Hubble horizon will always act as the delimiter between the system and unobservable environment modes and calculations of primordial statistics should be done within an open quantum system. If no such corrections are observed in the future, it would unequivocally state that cosmological perturbations cannot have a quantum origin (but rather have a different mechanism, e.g. warm inflation [58, 59]). Since dynamics of the background will affect the exact form of the mode functions and also enter the pre-factor of the interaction term, we cannot estimate the magnitude of the corrections in other such scenarios without an explicit computation. Nevertheless we can confidently state their existence. Moreover, it is straightforward to extend our computation to primordial gravitational waves, the leading order interaction term for which is independent of slow-roll parameters and therefore, the expected correction to the tensor power spectrum due to entanglement may be larger than the scalar one (in relative magnitude to the tree-level value). We have already shown the emergence of non-Markovian behaviour for the tensor modes [57] and the non-perturbative resummation for these will be pursued in the future and this work forms the foundation of exploring such hitherto undiscovered quantum effects in cosmology.

The effect found in this paper shows for a inflation with a quantum origin, the Universe would hold a memory of the primordial physics that connects the different scales. Beyond just the precision corrections it produces in the power spectrum, we will explore in future work other possible conceptual consequences of this entanglement that connects various disparate modes in the Universe.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

SB is supported by the Higgs Fellowship. AB is partially supported by STFC. JCF is supported by the Secretary of Higher Education, Science, Technology and Innovation of Ecuador (SENESCYT). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Appendix A. System and environment Hilbert spaces

To begin with, let us set the stage by expressing the operators in terms of the system ($\mathcal{S}$) and environment ($\mathcal{E}$) modes. For this, we filter the sub and super-horizon modes through a window function, where the simplest choice is the Heaviside step function, H_θ. Then, the Fourier expansion of any given operator, like the comoving curvature perturbation, is:

$$\hat{\zeta}(\tau,\mathbf{x}) = \frac{1}{z}\int \frac{d^3k}{(2\pi)^3}\left[\chi_k(\tau)\hat{c}_{\mathbf{k}} + \chi_k^*(\tau)\hat{c}^\dagger_{-\mathbf{k}}\right]\left[H_\theta(aH-k) + H_\theta(k-aH)\right]e^{i\mathbf{k}\cdot\mathbf{x}}$$

$$= \frac{1}{z}\int \frac{d^3k}{(2\pi)^3}\left[\left(\chi_k^{\mathcal{S}}(\tau) + \chi_k^{\mathcal{E}}(\tau)\right)\hat{c}_{\mathbf{k}} + \left(\chi_k^{\mathcal{S}}(\tau)^* + \chi_k^{\mathcal{E}}(\tau)^*\right)\hat{c}^\dagger_{-\mathbf{k}}\right]e^{i\mathbf{k}\cdot\mathbf{x}} \quad (18)$$

$$= \hat{\zeta}^{\mathcal{S}}(\tau,\mathbf{x}) + \hat{\zeta}^{\mathcal{E}}(\tau,\mathbf{x}), \quad (19)$$

where

$$\chi_k^{\mathcal{S}}(\tau) = \chi_k(\tau) H_\theta(aH - k) = \chi_k(\tau) H_\theta(1 + k\tau), \tag{20}$$

$$\chi_k^{\mathcal{E}}(\tau) = \chi_k(\tau) H_\theta(k - aH) = \chi_k(\tau) H_\theta(-1 - k\tau). \tag{21}$$

For later convenience, let us also rewrite the mode functions (7), obtained by fixing Bunch-Davies initial states:

$$\chi_k(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right). \tag{22}$$

A.1. Interaction Hamiltonian

Once we have split the fields into operators acting on system and environment states, the interaction Hamiltonian (in the interaction picture) can thus be written as:

$$\hat{H}_I(\tau) = \lambda(\tau) \int d^3x \left[\hat{\chi}^{\mathcal{E}}(\tau, \mathbf{x}) + \hat{\chi}^{\mathcal{S}}(\tau, \mathbf{x})\right] \left[\partial \left(\hat{\chi}^{\mathcal{E}}(\tau, \mathbf{x}) + \hat{\chi}^{\mathcal{S}}(\tau, \mathbf{x})\right)\right]^2, \tag{23}$$

where there are 8 different combinations of system and environment terms (64 considering the Fourier expansion of each field). As explained in the main text, the dominant terms in the integral are associated to the derivatives of environment operators, which can be seen from its Fourier transform. Moreover, since we are ultimately interested in the effect of the environment on the system (and because of momentum conservation), the remaining field left in the cubic term should act on $\mathcal{H}_{\mathcal{S}}$. In an alternative approach to that presented in this work, one can work out the Dyson expansion to compute perturbatively the state vectors and then construct the reduced density matrix. In doing so, it is easy to prove that at the same order in perturbation theory, some of the other combinations cancel out. In particular, this happens for $\hat{H}_I(\tau) \sim \hat{\chi}^{\mathcal{E}}(\tau, \mathbf{x})^3$, which is not important as far as corrections to the power spectrum go—since no system modes are involved—but it does prove that the kernel does not influence environment modes, or in other words, environment modes are not affected by the (rest of the) environment.

With these considerations, the interaction Hamiltonian we will work with is given by (as shown in equation (9)):

$$\begin{aligned}\hat{H}_I(\tau) &= \lambda(\tau) \int d^3x\, \hat{\chi}^{\mathcal{S}}(\tau, \mathbf{x}) \left(\partial \hat{\chi}^{\mathcal{E}}(\tau, \mathbf{x})\right)^2 \\ &= -\lambda(\tau) \int_{\Delta_k} (\mathbf{k}_2 \cdot \mathbf{k}_3) \hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau),\end{aligned} \tag{24}$$

where

$$\int_{\Delta_k} := \int \frac{d^3k_1}{(2\pi)^3} \int \frac{d^3k_2}{(2\pi)^3} \int \frac{d^3k_3}{(2\pi)^3}\, (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3).$$

Appendix B. Evolution equation and kernel calculation

Having defined system and environment fields, we shall now focus on the density matrix. In order to make contact with observations one must trace out the high-energy degrees of freedom (dof), which renders a (reduced) density matrix ρ_r acting on $\mathcal{H}_{\mathcal{S}}$. This operation is applied to both sides of equation (8), yielding the master equation that dictates the evolution of ρ_r. However, before doing so, we need to invoke some typical approximations, which are briefly mentioned in the main text.

First, it is assumed that at the beginning of inflation system and environment are not correlated, so the initial state reads $|\Psi(\tau_0)\rangle = |\mathcal{S}_0\rangle \otimes |\mathcal{E}_0\rangle$, and the density matrix $\rho_I(\tau_0) = \rho_{\mathcal{S}}(\tau_0) \otimes \rho_{\mathcal{E}}(\tau_0)$. A second assumption is that the environment does not change due to its interaction with the system, or in other words, it behaves as a proper environment as is typically required for decoherence. This implies $\rho_{\mathcal{E}}(\tau) \approx \rho_{\mathcal{E}}(\tau_0)$, which is possible due to the weak coupling between system and environment. Bringing these facts together, the density matrix at later times reads $\rho_I(\tau) = \rho_{\mathcal{S}}(\tau) \otimes \rho_{\mathcal{E}}(\tau_0)$, which facilitates taking the trace over $\mathcal{H}_{\mathcal{E}}$ on equation (8). In doing so, we get:

$$\begin{aligned}
\rho_r'(\tau) = &-\lambda(\tau) \int_{\tau_0}^{\tau} d\tau'\ \lambda(\tau') \int_{\Delta_p} \int_{\Delta_k} (\mathbf{p}_2 \cdot \mathbf{p}_3)(\mathbf{k}_2 \cdot \mathbf{k}_3) \left\{ \hat{\chi}^{\mathcal{S}}_{\mathbf{p}_1}(\tau) \hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau') \right. \\
&\times \mathrm{Tr}_{\mathcal{E}} \left[\hat{\chi}^{\mathcal{E}}_{\mathbf{p}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_3}(\tau)\ \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau') \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau') \rho_{\mathcal{E}}(\tau_0) \right] \rho_r(\tau') \\
&- \hat{\chi}^{\mathcal{S}}_{\mathbf{p}_1}(\tau) \mathrm{Tr}_{\mathcal{E}} \left[\hat{\chi}^{\mathcal{E}}_{\mathbf{p}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_3}(\tau) \rho_{\mathcal{E}}(\tau_0) \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau') \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau') \right] \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau') \\
&- \hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau') \mathrm{Tr}_{\mathcal{E}} \left[\hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau') \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau') \rho_{\mathcal{E}}(\tau_0) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_3}(\tau) \right] \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{\mathbf{p}_1}(\tau) \\
&\left. + \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{\mathbf{k}_1}(\tau') \mathrm{Tr}_{\mathcal{E}} \left[\rho_{\mathcal{E}}(\tau_0) \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_2}(\tau') \hat{\chi}^{\mathcal{E}}_{\mathbf{k}_3}(\tau') \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_3}(\tau) \right] \hat{\chi}^{\mathcal{S}}_{\mathbf{p}_1}(\tau) \right\},
\end{aligned} \tag{25}$$

where the first order effects arising from $\mathrm{Tr}_{\mathcal{E}}([H_I(\tau), \rho_I(\tau_0)])$ are considered null. To see this as a valid assumption, break the commutator and notice the resulting elements have the form $\sim \hat{\chi}^{\mathcal{S}}_{\mathbf{p}_1}(\tau) \rho_r(\tau_0) \mathrm{Tr}_{\mathcal{E}} \left(\hat{\chi}^{\mathcal{E}}_{\mathbf{p}_2}(\tau) \hat{\chi}^{\mathcal{E}}_{\mathbf{p}_3}(\tau) |0\rangle \langle 0| \right)$, where a non-zero trace requires $\mathbf{p}_2 = -\mathbf{p}_3$, and by conservation of momentum, $\mathbf{p}_1 = 0$. The contribution from any other combination of momenta is null, and by taking an IR cutoff (which could be given by a small mass induced by the renormalization procedure), the first order terms can be dismissed overall.

Next, considering the Fourier expansion of the different fields, together with $\rho_{\mathcal{E}}(\tau_0) = |0\rangle \langle 0|$, we arrive at the concise form of the master equation, as presented in equation (10) of the main text:

$$\begin{aligned}
\rho_r'(\tau) = &\int \frac{d^3p}{(2\pi)^3}\ \lambda(\tau) \int_{\tau_0}^{\tau} d\tau'\ \lambda(\tau') \left\{ \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau) \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau') \rho_r(\tau') K_p(\tau, \tau') \right. \\
&- \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau) \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau') K^*_p(\tau, \tau') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau') \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau) K_p(\tau, \tau') \\
&\left. + \rho_r(\tau') \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau') \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau) K^*_p(\tau, \tau') \right\},
\end{aligned} \tag{26}$$

where we have introduced the *kernel* as:

$$\begin{aligned}
&K_{p_1}(\tau, \tau') = -2 \int \frac{d^3p_2}{(2\pi)^3} (\mathbf{p}_2 \cdot \mathbf{p}_3)^2 \chi^{\varepsilon}_{p_2}(\tau) \chi^{\varepsilon}_{p_2}(\tau')^* \chi^{\varepsilon}_{p_3}(\tau) \chi^{\varepsilon}_{p_3}(\tau')^*; \\
&\mathbf{p}_3 = -(\mathbf{p}_1 + \mathbf{p}_2).
\end{aligned} \tag{27}$$

In order to compute this integral, we first need to work out the integration limits which are determined by the step functions associated to the fields. For this, let us write the kernel as:

$$\begin{aligned}
K_{p_1} \sim &\int_0^{\infty} dp_2 \int_{-1}^{1} d(\cos\theta) F(p_1, p_2, \theta) H_\theta(p_2 - aH) H_\theta(p_2 - (aH)') \\
&\times H_\theta(p_3 - aH) H_\theta(p_3 - (aH)'),
\end{aligned}$$

where F encompasses the product of Bunch-Davies functions equation (7), together with the momenta-dependent pre-factors. Further, without loss of generality, we have aligned $\mathbf{p}_1$ to the $\mathbf{z}$ direction, so θ denotes the angle between $\mathbf{p}_2$ and $\mathbf{p}_1$ and $p_3^2 = p_1^2 + p_2^2 + 2p_1p_2\cos\theta$.

Next, notice that $\tau' < \tau$ implies $(aH)' < aH$, so the regions in momentum space that contribute to the kernel are solely delimited by the product $H_\theta(p_2 - aH)H_\theta(p_3 - aH)$. Then, the angular region where $\mathbf{p}_3$ is sub-horizon is delimited by:

$$p_1^2 + p_2^2 + 2p_1p_2\cos\theta > (aH)^2 \implies \cos\theta > \frac{1}{2p_1p_2}\left[(aH)^2 - (p_1^2 + p_2^2)\right] \equiv \omega.$$

This condition sets the lower limit of the integral over the (cosine of the) polar angle, where we also have to account for the possibility $-1 > \omega$ as follows:

$$\begin{aligned} K_{p_1} &\sim \int_{aH}^{\infty} dp_2 \int_{\max\{-1,\omega\}}^{1} d(\cos\theta) F(p_1,p_2,\theta) \\ &\sim \int_{aH}^{\infty} dp_2 \left\{ \int_{-1}^{1} d(\cos\theta) F(p_1,p_2,\theta) H_\theta(-1-\omega) \right. \\ &\quad \left. + \int_{\omega}^{1} d(\cos\theta) F(p_1,p_2,\theta) H_\theta(1+\omega) \right\}, \end{aligned}$$

where,

$$\begin{aligned} H_\theta(-1-\omega) &\implies -1 \geqslant \omega \implies p_2 \geqslant aH + p_1\,, \\ H_\theta(1+\omega) &\implies -1 \leqslant \omega \implies p_2 \leqslant aH + p_1. \end{aligned}$$

In consequence, the kernel integral is conveniently split as:

$$\begin{aligned} K_{p_1} \sim & \int_{aH+p_1}^{\infty} dp_2 \int_{-1}^{1} d(\cos\theta) F(p_1,p_2,\theta) \\ & + \int_{aH}^{aH+p_1} dp_2 \int_{\omega}^{1} d(\cos\theta) F(p_1,p_2,\theta). \end{aligned} \tag{28}$$

In this way, the polar angle θ and the magnitude of p_2 are constrained such that both p_2 and p_3 are sub-horizon. For small values of p_1—which already has to satisfy $p_1 < aH$—the second integral becomes negligibly small as the lower and upper integration (momentum) limits come closer.

Performing both integrals, the leading-order behaviour of the kernel is given by (as presented in equation (12)):

$$K_{p_1}(\tau,\tau') \approx -\frac{e^{2i(\tau-\tau')/\tau}\left[1 - e^{-ip_1(\tau-\tau')}\right]\left[\tau - (1-i)\tau'\right]^2}{8\pi^2 p_1 \tau^4 (\tau')^2 (\tau-\tau')^2}. \tag{29}$$

Appendix C. Perturbative solution to the master equation

Having derived the explicit expression for the kernel, we can now extract the corrections to the power spectrum through the perturbative solution for the reduced density matrix. For convenience, we rewrite the master equation (10):

$$\begin{aligned} \rho_r'(\tau) = \sum_{\mathbf{p}} \lambda(\tau) \int_{\tau_0}^{\tau} d\tau'\, \lambda(\tau') \big\{ & \hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\rho_r(\tau')K_p(\tau,\tau') \\ & - \hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\rho_r(\tau')\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')K_p^*(\tau,\tau') - \hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\rho_r(\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)K_p(\tau,\tau') \\ & + \rho_r(\tau')\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)K_p^*(\tau,\tau') \big\}, \end{aligned} \tag{30}$$

where for future convenience we have substituted the integral over p for a sum, and a $1/V$ factor is implicit in front of the summation symbol. In order to find a first-order approximation to the solution, we take $\rho_r(\tau'') \to \rho_r(\tau_0) = |\mathcal{S}_0\rangle\langle\mathcal{S}_0| = |0\rangle\langle 0|$, which yields:

$$\rho_r(\tau) \approx \rho_r(\tau_0) + \sum_{\mathbf{p}} \int_{\tau_0}^{\tau} d\tau' \lambda(\tau') \int_{\tau_0}^{\tau'} d\tau'' \lambda(\tau'') \times \Big\{ \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)K_p(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')K^*_p(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K_p(\tau',\tau'') + \rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K^*_p(\tau',\tau'') \Big\}. \tag{31}$$

Finally, we provide the details for the calculation of the corrections to the power spectrum. In general, it is given by:

$$\Delta^2_\zeta(q\tau) = \frac{q^3}{2\pi^2 z^2} \left\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\right\rangle, \tag{32}$$

where the two-point correlation function is computed through the reduced density matrix technology as follows:

$$\left\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\right\rangle = \mathrm{Tr}\left[\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\rho_r(\tau)\right]. \tag{33}$$

In equation (13), we showed as a sanity check that the zeroth-order approximation of the reduced density matrix reproduces the well-known scale invariant spectrum. Considering the entanglement effects buried in equation (14), the correlator is given by:

$$\left\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\right\rangle = \mathrm{Tr}\left[\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)|0\rangle\langle 0|\right] + \mathrm{Tr}\Bigg[\int_{\tau_0}^{\tau} d\tau' \lambda(\tau') \int_{\tau_0}^{\tau'} d\tau'' \lambda(\tau'')\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau) \times \sum_{\mathbf{p}} \Big\{ \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)K_p(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')K^*_p(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K_p(\tau',\tau'') + \rho_r(\tau_0)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau'')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau')K^*_p(\tau',\tau'') \Big\}\Bigg], \tag{34}$$

where the corrections are clearly dictated by the second trace. Then, focusing on this bit, we take out of the sum the terms with momentum $p = q$, rendering:

$$\mathrm{Tr}\Bigg[\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\Big\{\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau')\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau'')|0\rangle\langle 0|K_q(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau')|0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau'')K^*_q(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau'')|0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau')K_q(\tau',\tau'') + |0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau'')\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau')K^*_q(\tau',\tau'') + \hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau')\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau'')|0\rangle\langle 0|K_q(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau')|0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau'')K^*_q(\tau',\tau'') - \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau'')|0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau')K_q(\tau',\tau'') + |0\rangle\langle 0|\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau'')\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau')K^*_q(\tau',\tau'') + \sum_{\mathbf{p}\neq\pm\mathbf{q}} \dots \Big\}\Bigg]. \tag{35}$$

It is easy to show that the last sum vanishes, whereas the other terms, which involve momenta with magnitude q, lead to equation (15),

$$\begin{aligned}\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\rangle &= \frac{1}{2q}\left(1+\frac{1}{(q\tau)^2}\right)+2\int_{-1/q}^{\tau} d\tau'\lambda(\tau')\int_{-1/q}^{\tau'} d\tau''\lambda(\tau'') \\ &\times\Big\{K_q(\tau',\tau'')\left[(\chi_q(\tau))^2\chi_q^*(\tau')\chi_q^*(\tau'')\right. \\ &\left.-|\chi_q(\tau)|^2\chi_q(\tau')\chi_q^*(\tau'')\right]+\text{c.c.}\Big\}. \end{aligned} \tag{36}$$

Introducing the change of variable $w=q\tau$, the scalar power spectrum including entanglement effects can be written as:

$$\begin{aligned}\Delta_\zeta^2(w) &= \frac{1}{2\epsilon M_{\text{Pl}}^2}\left(\frac{H}{2\pi}\right)^2\left[1+w^2+\frac{\epsilon H^2}{2M_{\text{Pl}}^2}w^2\int_{-1}^{w}dw'w'\int_{-1}^{w'}dw''w''\right. \\ &\times\Big\{K(w',w'')\left[(\chi(w))^2\chi^*(w')\chi^*(w'')\right. \\ &\left.\left.-|\chi(w)|^2\chi(w')\chi^*(w'')\right]+\text{c.c.}\Big\}\right], \end{aligned} \tag{37}$$

where any q-dependence on the functions is either absorbed by the new variable or by the momenta in the power spectrum formula. As a point of detail, notice that the (isolated) factor of w^2 corresponds to the decaying mode, which was seen to be subdominant with respect to every other term (after a few e-folds after horizon crossing).

As mentioned on the main text, the integration over w' (or z') has to be performed numerically after dealing with equal-time divergences (to be discussed soon after). In doing so, one finds that:

$$\Delta_\zeta^2(w) = \frac{1}{2\epsilon M_{\text{Pl}}^2}\left(\frac{H}{2\pi}\right)^2(1-\alpha N_c^2), \tag{38}$$

where $N_c=-\ln(-w)$ denotes the number of e-folds after horizon crossing, and the constant is given by (decaying mode not considered):

$$\alpha\approx 0.00211886\frac{\epsilon H^2}{2M_{\text{Pl}}^2}. \tag{39}$$

Appendix D. Nonperturbative resummation

The validity of equation (38) (equation (17) in the main text) relies on the validity of perturbation theory. Here, we will present an attempt to resum this result following [24]. For this, we will have to make more assumptions, so that, even though we are not dealing with exactly the same case, the following procedure presents a strong evidence that the full power spectrum can be indeed resummed.

First, the mode functions are split into a growing and a decaying mode,

$$\hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)=\chi_q^+(\tau)\hat{P}_{\mathbf{q}}-\chi_q^-(\tau)\hat{X}_{\mathbf{q}}, \tag{40}$$

where $\chi_q^{+(-)}$ and $\hat{P}_{\mathbf{q}}(\hat{X}_{\mathbf{q}})$ denote the growing (decaying) mode function and operator respectively, and are given by:

$$\chi_q^+(\tau)=\sqrt{\frac{-\pi\tau}{2}}Y_{3/2}(|q\tau|),\qquad \chi_q^-(\tau)=\sqrt{\frac{-\pi\tau}{2}}J_{3/2}(|q\tau|), \tag{41}$$

$$\hat{P}_{\mathbf{q}} = \frac{i}{\sqrt{2}}\left(\hat{c}^{\dagger}_{-\mathbf{q}} - \hat{c}_{\mathbf{q}}\right), \qquad \hat{X}_{\mathbf{q}} = \frac{1}{\sqrt{2}}\left(\hat{c}_{\mathbf{q}} + \hat{c}^{\dagger}_{-\mathbf{q}}\right). \tag{42}$$

The operators satisfy the usual commutation relation $\left[\hat{X}_{\mathbf{q}}, \hat{P}_{\mathbf{k}}\right] = i\delta_{\mathbf{q},-\mathbf{k}}$, with vanishing commutators for equal operators.

Then, at leading order in $q\tau$, the power spectrum of the growing mode operator captures the behaviour of the full power spectrum, since:

$$\left\langle \hat{\chi}^{\mathcal{S}}_{\mathbf{q}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{q}}(\tau)\right\rangle \approx \left[\chi^{+}_{q}(\tau)\right]^{2}\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle \approx \frac{1}{2q^{2}\tau^{3}}, \tag{43}$$

in agreement with the leading order term of equation (36). Then, for our purposes it proves easier to deal with this magnitude as opposed to the full power spectrum, which presents an 'extra' time dependence.

Next, we shall re-write the master equation in terms of the growing and decaying mode operators. However, before doing so, it is useful to take the Markov approximation, for which under the assumption of weak coupling one can take $\rho_r(\tau') \to \rho_r(\tau)$, rendering:

$$\begin{aligned}\rho_r'(\tau) = \sum_{\mathbf{p}} \lambda(\tau)\int_{\tau_0}^{\tau} d\tau'\ \lambda(\tau')\,\big\{&\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\rho_r(\tau)K_p(\tau,\tau')\\ &-\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)\rho_r(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')K_p^{*}(\tau,\tau') - \hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\rho_r(\tau)\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)K_p(\tau,\tau')\\ &+\rho_r(\tau)\hat{\chi}^{\mathcal{S}}_{-\mathbf{p}}(\tau')\hat{\chi}^{\mathcal{S}}_{\mathbf{p}}(\tau)K_p^{*}(\tau,\tau')\big\}.\end{aligned} \tag{44}$$

In this way one can compute the integrals over τ' which only include the Bunch-Davies functions and the kernel (or its conjugate). In performing the integrations one finds equal-time divergences (see next section), but for now we shall focus on the late time leading-order behaviour of the integrals which we will exploit in order to cast the master equation in a more convenient way. The integrals are:

$$\begin{aligned}\mathcal{Y}^{+}_{q}(\tau) &= \int_{\tau_0}^{\tau} d\tau'\lambda(\tau')\chi^{+}_{q}(\tau')K_q(\tau,\tau') \sim -\int_{\tau_0}^{\tau} d\tau'\lambda(\tau')\chi^{+}_{q}(\tau')K_q^{*}(\tau,\tau')\\ &\sim \frac{\epsilon^{1/2}H}{2^{3/2}M_{\rm Pl}}\frac{i}{8\pi^{2}}\frac{\ln(|q\tau|)}{q^{3/2}\tau^{4}},\end{aligned} \tag{45}$$

$$\begin{aligned}\mathcal{Y}^{-}_{q}(\tau) &= \int_{\tau_0}^{\tau} d\tau'\lambda(\tau')\chi^{-}_{q}(\tau')K_q(\tau,\tau') \sim -\int_{\tau_0}^{\tau} d\tau'\lambda(\tau')\chi^{-}_{q}(\tau')K_q^{*}(\tau,\tau')\\ &\sim \frac{\epsilon^{1/2}H}{2^{3/2}M_{\rm Pl}}\frac{i}{24\pi^{2}}\frac{q^{5/2}\ln(|q\tau|)}{\tau},\end{aligned} \tag{46}$$

where we have omitted a subdominant real part, which would become imaginary in the power spectrum, leading to an oscillatory term that is of no interest in the present discussion. With these considerations, the master equations can be written as:

$$\begin{aligned}\rho_r'(\tau) = \lambda(\tau)\sum_{\mathbf{p}}\Big\{&\chi^{+}_{p}(\tau)\mathcal{Y}^{+}_{p}(\tau)\left[\hat{P}_{\mathbf{p}}\hat{P}_{-\mathbf{p}}\rho_r(\tau) - \rho_r(\tau)\hat{P}_{\mathbf{p}}\hat{P}_{-\mathbf{p}}\right]\\ &+\chi^{-}_{p}(\tau)\mathcal{Y}^{-}_{p}(\tau)\left[\hat{X}_{\mathbf{p}}\hat{X}_{-\mathbf{p}}\rho_r(\tau) - \rho_r(\tau)\hat{X}_{\mathbf{p}}\hat{X}_{-\mathbf{p}}\right] - \chi^{+}_{p}(\tau)\mathcal{Y}^{-}_{p}(\tau)\\ &\times\left[\hat{P}_{\mathbf{p}}\hat{X}_{-\mathbf{p}}\rho_r(\tau) - \hat{X}_{-\mathbf{p}}\rho_r(\tau)\hat{P}_{\mathbf{p}} - \rho_r(\tau)\hat{X}_{-\mathbf{p}}\hat{P}_{\mathbf{p}} + \hat{P}_{\mathbf{p}}\rho_r(\tau)\hat{X}_{-\mathbf{p}}\right] - \mathcal{Y}^{+}_{p}(\tau)\chi^{-}_{p}(\tau)\\ &\times\left[\hat{X}_{\mathbf{p}}\hat{P}_{-\mathbf{p}}\rho_r(\tau) - \hat{P}_{-\mathbf{p}}\rho_r(\tau)\hat{X}_{\mathbf{p}} - \rho_r(\tau)\hat{P}_{-\mathbf{p}}\hat{X}_{\mathbf{p}} + \hat{X}_{\mathbf{p}}\rho_r(\tau)\hat{P}_{-\mathbf{p}}\right]\Big\}.\end{aligned} \tag{47}$$

Then, one can compute the corrections to this power spectrum, taking advantage of the following equality:

$$\frac{d}{d\tau}\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle = \mathrm{Tr}\left[\hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\rho_r'(\tau)\right]. \tag{48}$$

The trace can be computed by using the commutation relations of $\hat{X}$ and $\hat{P}$, together with the properties of the trace. In doing so, we find that:

$$\mathrm{Tr}\left[\hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\rho_r'(\tau)\right] \approx 4i\lambda(\tau)\mathcal{Y}_q^{+}(\tau)\chi_q^{-}(\tau)\mathrm{Tr}\left[\hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\rho_r(\tau)\right], \tag{49}$$

where we are ignoring a subdominant contribution from the product of two decaying modes (the others are null). Hence, we have that:

$$\frac{d}{d\tau}\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle = -\frac{\epsilon H^2}{48\pi^2 M_{\mathrm{Pl}}^2}\frac{\ln(|q\tau|)}{\tau}\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle. \tag{50}$$

Performing the integration, one arrives to:

$$\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle = f(\tau)\left\langle \hat{P}_{\mathbf{q}}\hat{P}_{-\mathbf{q}}\right\rangle(\tau_*), \tag{51}$$

where

$$\begin{aligned} f(\tau) &= \exp\left[-\frac{\epsilon H^2}{48\pi^2 M_{\mathrm{Pl}}^2}\int_{-1/q}^{\tau} d\tau'\frac{\ln(-q\tau)}{\tau}\right] \\ &= \exp\left[-\frac{\epsilon H^2}{96\pi^2 M_{\mathrm{Pl}}^2}\ln^2(|q\tau|)\right]. \end{aligned} \tag{52}$$

Notice that the pre-factor of the logarithm above can be written as:

$$\alpha = 0.00211086\frac{\epsilon H^2}{2M_{\mathrm{Pl}}^2}, \tag{53}$$

which is remarkably close to the parameter α obtained through the perturbative expansion, as shown in equation (39). This shows that long after horizon crossing there are other terms that take over the secular term, rendering a well-behaved power spectrum, whose final form is:

$$\Delta_\zeta^2 = \frac{1}{2\epsilon M_{\mathrm{Pl}}^2}\left(\frac{H}{2\pi}\right)^2 e^{-\alpha N_c^2}, \tag{54}$$

where α has been estimated numerically (equation (39)) and analytically (equation (53)).

Let us highlight the approximations we had to make in order to arrive at this result. Firstly, as opposed to [24], we are dealing with the long and short wavelength modes of the curvature perturbations themselves. This implies working with mode functions which are much more complicated, and therefore, one has to resort to numerical estimations for the integrals involved while working in the perturbative approach. More importantly, recall that our secular divergence is for an open EFT in quasi-dS background. This requires calculating kernels, as opposed to loop corrections for scalar fields on dS, and thus, an application of the dynamical renormalization group is not entirely straightforward [49]. However, interestingly, we find that the resummed 'entanglement function' takes a form that is remarkable similar to what one would expect from the dynamical renormalization group analyses [49]. In conclusion, even though we arrive at the late-time value of the resummed function within the Markovian approximation, its final form reassures us about its validity.

Appendix E. Renormalization of equal-time divergences

When taking the equal time limit in the kernel, while evaluating the necessary integrals involved in calculating the corrections to the power spectrum, we found divergences appearing in the master equation which need to be dealt with. The most typical way to deal with such divergences for usual QFTs is to determine the counterterms which have to be introduced in the action. However, note that our case is somewhat different and, in essence, one does not have to add such counterterms in order to preserve the symmetries of the background [40]. We are only adding such local counterterms to show that these divergences are spurious, and the finite terms are the only ones which are physically meaningful and contribute to the spectrum [40].

Let us begin with the master equation once again:

$$\rho_r'(\tau) = \sum_{\mathbf{p}} \lambda(\tau) \int_{\tau_0}^{\tau} d\tau' \, \lambda(\tau') \left\{ \hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\rho_r(\tau')K_p(\tau,\tau') \right.$$
$$-\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\rho_r(\tau')\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')K_p^*(\tau,\tau') - \hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\rho_r(\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)K_p(\tau,\tau')$$
$$\left. +\rho_r(\tau')\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)K_p^*(\tau,\tau') \right\}. \tag{55}$$

The perturbative expansion instructs $\rho_r(\tau') \to \rho_r(\tau_0)$, so that applying the operators at τ' renders:

$$\rho_r'(\tau) = \sum_{\mathbf{p}} \lambda(\tau) \int_{\tau_0}^{\tau} d\tau' \, \lambda(\tau') \left\{ K_p(\tau,\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\chi_p^{s}(\tau')^* \left|1_{\mathbf{p}}\right\rangle \left\langle 0\right| \right.$$
$$-K_p^*(\tau,\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\left|0\right\rangle \left\langle 1_{-\mathbf{p}}\right| \chi_p^{s}(\tau') - K_p(\tau,\tau')\chi_p^{s}(\tau')^* \left|1_{\mathbf{p}}\right\rangle \left\langle 0\right| \hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)$$
$$\left. +K_p^*(\tau,\tau')\left|0\right\rangle \left\langle 1_{-\mathbf{p}}\right| \chi_p^{s}(\tau')\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau) \right\}. \tag{56}$$

Then, we may perform the integrals over τ', where we will only pay attention to the upper limit, which is the one leading to the *equal-time* divergences. The integrals we need to compute are:

$$\int_{\tau_0}^{\tau} d\tau'\tau' K_p(\tau,\tau')\chi_p^{s}(\tau')^* , \qquad \int_{\tau_0}^{\tau} d\tau'\tau' K_p^*(\tau,\tau')\chi_p^{s}(\tau'), \tag{57}$$

which can be done analytically. In order to isolate the divergence, we take the upper limit as $\tau' \to \tau - i\delta$, with $0 < \delta \ll 1$, to then Taylor expand the resulting expression around $\delta = 0$. This yields:

$$\int_{\tau_0}^{\tau - i\delta} d\tau'\tau' K_p(\tau,\tau')\chi_p^{s}(\tau')^* \sim -\frac{i}{8\pi^2}\frac{\chi_p^{s}(\tau)^*}{\tau^3}\ln\left(\frac{\delta}{\mu}\right) + \mathcal{O}(\delta^0), \tag{58}$$

$$\int_{\tau_0}^{\tau - i\delta} d\tau'\tau' K_p^*(\tau,\tau')\chi_p^{s}(\tau') \sim \frac{i}{8\pi^2}\frac{\chi_p^{s}(\tau)}{\tau^3}\ln\left(\frac{\delta}{\mu}\right) + \mathcal{O}(\delta^0), \tag{59}$$

where μ is a renormalization scale. Next, plugging these expressions back to (56) and recovering the full operators, we arrive to:

$$\rho_r'(\tau) = -\frac{i}{8\pi^2}\frac{\epsilon H^2}{8M_{\mathrm{Pl}}^2\tau^2}\ln\left(\frac{\delta}{\mu}\right)\sum_{\mathbf{p}} \left\{ \hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau)\rho_r(\tau_0) \right.$$
$$\left. +\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\rho_r(\tau_0)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau) - \hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau)\rho_r(\tau_0)\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau) - \rho_r(\tau_0)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau)\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau) \right\}$$
$$+\mathcal{O}(\delta^0). \tag{60}$$

Notice this can be further simplified by taking the second and fourth terms on the r.h.s. with opposite momenta (taking advantage of the sum), so the second and third terms cancel out such that:

$$\rho_r'(\tau) = -\frac{i}{8\pi^2}\frac{\epsilon H^2}{8M_{\text{Pl}}^2\tau^2}\ln\left(\frac{\delta}{\mu}\right)\sum_{\mathbf{p}}\left[\hat{\chi}_{\mathbf{p}}^{\mathcal{S}}(\tau)\hat{\chi}_{-\mathbf{p}}^{\mathcal{S}}(\tau),\rho_r(\tau_0)\right] + \mathcal{O}(\delta^0). \tag{61}$$

Compare this to the first order approximation of the Liouville–von Neumann equation,

$$\rho_I'(\tau) \approx -i\left[H_I(\tau),\rho_I(\tau_0)\right], \tag{62}$$

which suggests the need of a mass counterterm,

$$\delta\mathcal{L} = -\delta H = \frac{\delta m^2}{2}\sum_{\mathbf{p}}\hat{\chi}_{\mathbf{p}}(\tau)\hat{\chi}_{-\mathbf{p}}(\tau), \tag{63}$$

With:

$$\delta m^2 = \frac{1}{8\pi^2}\frac{\epsilon H^2}{4M_{\text{Pl}}^2\tau^2}\ln\left(\frac{\delta}{\mu}\right). \tag{64}$$

By introducing this counterterm we may just ignore any divergences found in the computation of the correction of the power spectrum found in equation (37). The other divergences encountered can be taken care of by similar counterterms and shows that these divergences are spurious and only the finite term arising from this calculation contributes to the power spectrum, as has been shown in the main text. Finally, notice that even though there are no finite contributions at first order, there are infinite terms like those emerging from the IR or the computed counterterm.

ORCID iD

Suddhasattwa Brahma https://orcid.org/0000-0003-4241-6701

References

[1] Starobinsky A A 1980 A new type of isotropic cosmological models without singularity *Phys. Lett.* B **91** 99–102

[2] Fang L Z 1980 Entropy generation in the early universe by dissipative processes near the Higgs phase transitions *Phys. Lett.* B **95** 154–6

[3] Guth A H 1981 Inflationary universe: a possible solution to the horizon and flatness problems *Phys. Rev.* D **23** 347

[4] Sato K 1981 First order phase transition of a vacuum and expansion of the universe *Mon. Not. Roy. Astron. Soc.* **195** 467–79

[5] Linde A D 1982 A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems *Phys. Lett.* B **108** 389–93

[6] Albrecht A and Steinhardt P J 1982 Cosmology for grand unified theories with radiatively induced symmetry breaking *Phys. Rev. Lett.* **48** 1220

[7] Mukhanov V F and Chibisov G 1982 The vacuum energy and large scale structure of the universe *Sov. Phys. JETP* **56** 258–65

[8] Chen X 2010 Primordial non-gaussianities from inflation models *Adv. Astron.* **2010** 638979

[9] Pajer E, Schmidt F and Zaldarriaga M 2013 The observed squeezed limit of cosmological three-point functions *Phys. Rev.* D **88** 083502

[10] Campo D and Parentani R 2006 Inflationary spectra and violations of Bell inequalities *Phys. Rev.* D **74** 025001

[11] Campo D and Parentani R 2005 Inflationary spectra and partially decohered distributions *Phys. Rev.* D **72** 045015

[12] Maldacena J 2016 A model with cosmological Bell inequalities *Fortsch. Phys.* **64** 10–23
[13] Martin J and Vennin V 2016 Bell inequalities for continuous-variable systems in generic squeezed states *Phys. Rev.* A **93** 062117
[14] Green D and Porto R A 2020 Signals of a quantum universe *Phys. Rev. Lett.* **124** 251302
[15] Martin J and Vennin V 2017 Obstructions to Bell CMB experiments *Phys. Rev.* D **96** 063501
[16] Morse M J 2020 Statistical Bounds on CMB Bell Violation (available at: https://arxiv.org/abs/2003.13562)
[17] Khoury J, Ovrut B A, Seiberg N, Steinhardt P J and Turok N 2002 From big crunch to big bang *Phys. Rev.* D **65** 086007
[18] Khoury J, Ovrut B A, Steinhardt P J and Turok N 2001 The Ekpyrotic universe: colliding branes and the origin of the hot big bang *Phys. Rev.* D **64** 123522
[19] Burgess C, Holman R, Tasinato G and Williams M 2015 EFT Beyond the horizon: stochastic inflation and how primordial quantum fluctuations go classical *J. High Energy Phys.* JHEP03(2015)090
[20] Burgess C, Holman R and Tasinato G 2016 Open EFTs, IR effects & late-time resummations: systematic corrections in stochastic inflation *J. High Energy Phys.* JHEP01(2016)153
[21] Shandera S, Agarwal N and Kamal A 2018 Open quantum cosmological system *Phys. Rev.* D **98** 083535
[22] Burgess C P, Holman R and Hoover D 2008 Decoherence of inflationary primordial fluctuations *Phys. Rev.* D **77** 063534
[23] Maldacena J M 2003 Non-Gaussian features of primordial fluctuations in single field inflationary models *J. High Energy Phys.* JHEP05(2003)013
[24] Boyanovsky D 2015 Effective field theory during inflation: reduced density matrix and its quantum master equation *Phys. Rev.* D **92** 023527
[25] Boyanovsky D 2018 Imprint of entanglement entropy in the power spectrum of inflationary fluctuations *Phys. Rev.* D **98** 023515
[26] Boyanovsky D 2018 Information loss in effective field theory: entanglement and thermal entropies *Phys. Rev.* D **97** 065008
[27] Sharman J W and Moore G D 2007 Decoherence due to the Horizon after Inflation *J. Cosmol. Astropart. Phys.* JCAP11(2007)020
[28] Brahma S, Alaryani O and Brandenberger R 2020 Entanglement entropy of cosmological perturbations *Phys. Rev.* D **102** 043529
[29] Ade P A R *et al* 2021 Improved constraints on primordial gravitational waves using planck, WMAP and BICEP/Keck observations through the 2018 observing season *Phys. Rev. Lett.* **127** 151301
[30] Agon C, Balasubramanian V, Kasko S and Lawrence A 2018 Coarse grained quantum dynamics *Phys. Rev.* D **98** 025019
[31] Balasubramanian V, McDermott M B and Van Raamsdonk M 2012 Momentum-space entanglement and renormalization in quantum field theory *Phys. Rev.* D **86** 045014
[32] Albrecht A, Ferreira P, Joyce M and Prokopec T 1994 Inflation and squeezed quantum states *Phys. Rev.* D **50** 4807–20
[33] Adshead P, Hu W, Dvorkin C and Peiris H V 2011 Fast computation of bispectrum features with generalized slow roll *Phys. Rev.* D **84** 043519
[34] Puri R R 2001 *Mathematical Methods of Quantum Optics* vol 79 (Springer Science & Business Media) (available at: https://link.springer.com/book/10.1007/978-3-540-44953-9)
[35] Joos E, Zeh H D, Kiefer C, Giulini D J, Kupsch J and Stamatescu I-O 2003 *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer Science & Business Media) (available at: https://link.springer.com/book/10.1007/978-3-662-05328-7)
[36] Starobinsky A A 1986 Stochastic de Sitter (Inflationary) stage in the early universe *Lect. Notes Phys.* **246** 107–26
[37] Prokopec T and Rigopoulos G I 2007 Decoherence from isocurvature perturbations in inflation *J. Cosmol. Astropart. Phys.* JCAP11(2007)029
[38] Koksma J F, Prokopec T and Schmidt M G 2011 Decoherence in quantum mechanics *Annals Phys.* **326** 1548–76
[39] Nelson E 2016 Quantum decoherence during inflation from gravitational nonlinearities *J. Cosmol. Astropart. Phys.* JCAP03(2016)022
[40] Fröb M B, Roura A and Verdaguer E 2012 One-loop gravitational wave spectrum in de Sitter spacetime *J. Cosmol. Astropart. Phys.* JCAP08(2012)009
[41] Senatore L and Zaldarriaga M 2010 On loops in inflation *J. High Energy Phys.* JHEP12(2010)008

[42] Pimentel G L, Senatore L and Zaldarriaga M 2012 On loops in inflation III: time independence of zeta in single clock inflation *J. High Energy Phys.* JHEP07(2012)166

[43] Seery D 2007 One-loop corrections to a scalar field during inflation *J. Cosmol. Astropart. Phys.* JCAP11(2007)025

[44] Assassi V, Baumann D and Green D 2012 On soft limits of inflationary correlation functions *J. Cosmol. Astropart. Phys.* JCAP11(2012)047

[45] Starobinsky A A and Yokoyama J 1994 Equilibrium state of a selfinteracting scalar field in the De Sitter background *Phys. Rev.* D **50** 6357–68

[46] Gautier F and Serreau J 2013 Infrared dynamics in de Sitter space from Schwinger-Dyson equations *Phys. Lett.* B **727** 541–7

[47] Urakawa Y and Tanaka T 2010 IR divergence does not affect the gauge-invariant curvature perturbation *Phys. Rev.* D **82** 121301

[48] Woodard R P 2016 Some inconvenient truths *J. High Energy Phys.* JHEP05(2016)152

[49] Burgess C P, Leblond L, Holman R and Shandera S 2010 Super-Hubble de Sitter fluctuations and the dynamical RG *J. Cosmol. Astropart. Phys.* JCAP03(2010)033

[50] Senatore L and Zaldarriaga M 2013 The constancy of ζ in single-clock Inflation at all loops *J. High Energy Phys.* JHEP09(2013)148

[51] Giddings S B and Sloth M S 2011 Cosmological observables, IR growth of fluctuations and scale-dependent anisotropies *Phys. Rev.* D **84** 063528

[52] Kosowsky A and Turner M S 1995 CBR anisotropy and the running of the scalar spectral index *Phys. Rev.* D **52** 1739–43

[53] Easther R and Peiris H 2006 Implications of a running spectral index for slow roll inflation *J. Cosmol. Astropart. Phys.* JCAP09(2006)010

[54] Martin J and Vennin V 2018 Observational constraints on quantum decoherence during inflation *J. Cosmol. Astropart. Phys.* JCAP05(2018)063

[55] Matarrese S, Pilo L and Rollo R 2021 Resilience of long modes in cosmological observables *J. Cosmol. Astropart. Phys.* JCAP01(2021)062

[56] Bartolo N, Carollo G B, Matarrese S, Pilo L and Rollo R 2022 The physical content of long tensor modes in cosmology (arXiv:2208.00075)

[57] Brahma S, Berera A and Calderón-Figueroa J 2022 Quantum corrections to the primordial tensor spectrum: open EFTs & Markovian decoupling of UV modes *J. High Energy Phys.* JHEP08(2022)225

[58] Berera A and Fang L-Z 1995 Thermally induced density perturbations in the inflation ERA *Phys. Rev. Lett.* **74** 1912–15

[59] Berera A 1995 Warm inflation *Phys. Rev. Lett.* **75** 3218–21