Published for SISSA by Springer

Received: *August 23, 2023*
Accepted: *November 11, 2023*
Published: *December 12, 2023*

The cosmological tree theorem

Santiago Agüí Salcedo[a] **and Scott Melville**[b,a]

[a] *DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.*
[b] *Astronomy Unit, Queen Mary University of London, Mile End Road, London, E1 4NS, U.K.*

E-mail: santiago.agui@protonmail.com, s.melville@qmul.ac.uk

Abstract: A number of diagrammatic "cutting rules" have recently been developed for the wavefunction of the Universe which determines cosmological correlation functions. These leverage perturbative unitarity to relate particular "discontinuities" in Feynman-Witten diagrams (with cosmological boundary conditions) to simpler diagrams, in much the same way that the Cutkosky rules relate different scattering amplitudes. In this work, we make use of a further causality condition to derive new cutting rules for Feynman-Witten diagrams on any time-dependent spacetime background. These lead to the cosmological analogue of Feynman's tree theorem for amplitudes, which can be used to systematically expand any loop diagram in terms of (momentum integrals of) tree-level diagrams. As an application of these new rules, we show that certain singularities in the wavefunction cannot appear in equal-time correlators due to a cancellation between "real" and "virtual" contributions that closely parallels the KLN theorem. Finally, when combined with the Bunch-Davies condition that certain unphysical singularities are absent, these cutting rules completely determine any tree-level exchange diagram in terms of simpler contact diagrams. Altogether, these results remove the need to ever perform nested time integrals when computing cosmological correlators.

Keywords: Cosmological models, Effective Field Theories, de Sitter space, Scattering Amplitudes

ArXiv ePrint: 2308.00680

https://doi.org/10.1007/JHEP12(2023)076

JHEP12(2023)076

Contents

1 Introduction **1**
- 1.1 Notation and conventions 5
- 1.2 Summary of main results 9

2 Wavefunction coefficients **13**
- 2.1 Causality and a tree-level bootstrap 13
- 2.2 Loops and the cosmological tree theorem 20
  - 2.2.1 Some examples 20
  - 2.2.2 Proof of the general theorem at one loop 27

3 Correlation functions **31**
- 3.1 IR singularities and the cosmological KLN theorem 34
- 3.2 Some examples 40

4 Discussion **44**

A Comparison with previous cutting rules **46**

B Non-Gaussianity at next-to-next-to-leading order **52**
- B.1 Bispectrum at one loop 53
- B.2 Field vev at two loops 57

C Wavefunction symmetry factors **57**

D Further identities **60**
- D.1 Symmetric version of the cutting rules and tree theorem 60
- D.2 Higher loops 62
- D.3 Finite time and restricted integration regions 65

E Cosmological loop-tree duality **69**

JHEP12(2023)076

1 Introduction

Loops are hard. It is an unhappy fact of life: to make precise predictions in most quantum field theories, one must compute Feynman diagrams that contain loops. Even on Minkowski spacetime, this is difficult due to momentum integrals that often diverge. On a cosmological (time-dependent) spacetime background, things are even worse: loop diagrams contain potentially divergent integrals both over momenta and over time. Since loop corrections can

play an important role in a wide variety of cosmological settings — including precise calculations of the primordial power spectrum [1–18] and non-Gaussianity [19–31]; inflationary features that may seed primordial black holes [32–41] or produce a stochastic gravitational wave background [42–46]; de Sitter holography [47–50] and the IR stability of de Sitter spacetime [51–56] — it is important to develop tools for evaluating (and better understanding) these loop diagrams. Our goal in this work is to take a step in this direction, and in particular address the issue of the time integrals. The main result is a *cosmological tree theorem* which can be used to expand any loop diagram in terms of (momentum integrals of) tree-level diagrams, which are comparatively much easier to evaluate. This theorem applies on any time-dependent spacetime background, and therefore can play a useful role in future studies of the cosmological implications of loop corrections.

The cosmological wavefunction. The quantum-mechanical object that we focus on computing is the wavefunction of the Universe [57]. This wavefunction characterises the state of the quantum fields at a given time, and has recently played an increasingly central role in inflationary cosmology. Any equal-time correlation function can be calculated from it using the Born rule, and these cosmological correlators encode valuable information about the new high-energy fields that can be excited during inflation [58]. But beyond efficiently encoding all correlators, use of the wavefunction has been driven by its many parallels with scattering amplitudes. The S-matrix programme of the 1960's developed a suite of constraints and calculational tools for amplitudes using the fundamental axioms of unitarity, causality and locality [59], and the modern incarnation of these ideas includes powerful bootstrap techniques (see e.g. [60–63] for recent reviews). Proceeding along similar lines has led to a variety of "bootstrap" approaches for efficiently computing the cosmological wavefunction.

Cosmological bootstrap(s). One influential approach uses the de Sitter isometries [64–67], building on earlier work [68–71] by introducing the requirement that certain unphysical singularities must vanish for a Bunch-Davies initial state (see [72] for a recent review). Another symmetry-based approach is the cosmological scattering equation [73, 74], or the analytic continuation from (Euclidean) AdS where harmonic analysis is possible [75–82], including recent progress towards a non-perturbative Kallen-Lehmann representation of the two-point function [83–85]. Other bootstrap approaches take a general ansatz with the desired analytic structure and apply various known limits and consistency relations [86–89], without the need for full de Sitter symmetry, or build the desired coefficient by gluing together simpler seed functions [90–94]. Altogether, these important advances represent an "NLO revolution" for cosmological correlators: they have enabled the computation of primordial non-Gaussianity and other observables in a range of different models beyond the leading-order contact diagrams [95–102], much like the NNLO revolution currently taking place for scattering amplitudes (where on-shell unitarity methods and factorisation have enabled an explosion of new two-loop computations, see e.g. [103]).

JHEP12(2023)076

Unitarity. One aspect of many of these approaches is the use of perturbative unitarity. Much like for amplitudes, unitary time evolution is encoded in perturbation theory[1] as a set of "cutting rules" which can relate any Feynman-Witten diagram for the wavefunction to simpler diagrams with fewer internal lines [104–107] (see also [108] for recent cutting rules in AdS).[2] In the context of amplitudes, pairing unitarity and the optical theorem with causality and analyticity leads to powerful UV/IR relations that connect the coefficients in any low-energy Effective Field Theory with properties of the underlying high-energy completion (see e.g. [116–123] for an incomplete list). Despite some progress in this direction [77, 124–127], the application of analogous UV/IR relations in cosmology has mostly been held back by our understanding of causality/analyticity on time-dependent spacetime backgrounds.

Causality. Relativistic causality is the central pillar on which the Minkowski S-matrix programme is built. This is the postulate that a local observer can only access information from inside their past lightcone and affect events inside their future lightcone — or stated mathematically: all local operators must commute at space-like separations. At the level of the S-matrix, this property implies[3] a particular analytic structure for $2 \to 2$ scattering in the complex s-plane at fixed t, where s and t are the usual Mandelstam variables. On a curved spacetime background, the implications of relativistic causality are less clear due to the different local lightcones (mathematically, there is no analogue of s and t), although see [128–135] for recent progress in implementing causality as a constraint in gravitational field theories.

An alternative way forward is to consider the weaker condition of *non-relativistic* causality. This is the requirement that any source which is local-in-time can only affect its future, and is the property that underpins the classic Kramers-Kronig relations for the dispersion of light in a medium. In a cosmological context, this non-relativistic causality was recently used in [127] to find analogous dispersion relations for wavefunction coefficients. Here, we are going to explore another consequence of non-relativistic causality: namely that it forbids loops (closed time-like curves).

Feynman tree theorem. In the context of amplitudes, this aspect of non-relativistic causality was exploited by Feynman long ago [136, 137] to derive what is often dubbed "Feynman's tree theorem". This theorem provides a systematic way to cut open any closed loop in a Feynman diagram and replace it by a sum over simpler diagrams with fewer loops. It makes manifest the fact that, while drawing loops is often helpful to visualise certain quantum corrections, since no physical signal could ever travel on such a trajectory they must be expendable. Feynman's theorem has since been developed in various

[1] Beyond the diagrammatic expansion, these unitarity relations can be understood as an infinite number of conserved charges which follow from the total probability (the norm of the wavefunction) remaining constant in time [104].

[2] A more geometric formulation of the wavefunction is provided by the cosmological polytope [109–114], and [115] recently derived a number of new cutting rules in that language.

[3] The connection between causality and analyticity is fairly rigorous in simple theories, e.g. a real scalar field with a mass gap, although in a completely general setting one often postulates analyticity in place of causality since it can be precisely implemented at the level of the S-matrix.

JHEP12(2023)076

ways [138–140], and is now widely used in the literature, often together with on-shell unitarity and recursion relations (see e.g. [141, 142] and [143–148] for some early work in that direction). One prominent incarnation of Feynman's theorem is the "loop-tree duality" of [149, 150] (see also [151–155]), which can be used to efficiently implement IR subtractions numerically [156–160]. A version of Feynman's tree theorem has even been applied to the two-point function on a quasi-de Sitter background in order to determine the signs of higher-loop diagrams without explicitly computing them [161]. In short, Feynman's remarkably simple relation has had a remarkable impact on how we compute and understand loop corrections.

Main results. In this work, we consider the constraints imposed by non-relativistic causality for the cosmological wavefunction. Concretely, our main results are to:

(i) derive a cosmological analogue of the Feynman tree theorem, which can be used to express a Feynman(-Witten) diagram with any number of loops in terms of purely tree-level diagrams,

(ii) use this tree theorem to determine, from purely tree-level data, simple expressions for the one-loop wavefunction coefficients and the corresponding one-, two- and three-point cosmological correlators,

(iii) demonstrate a cosmological analogue of the KLN theorem, namely that particular singularities in the wavefunction cancel out in cosmological correlators,

(iv) describe a novel bootstrap procedure which combines unitarity cutting rules, the absence of unphysical singularities and our causality condition in order to determine any tree-level exchange diagram from simpler cut diagrams with no internal lines.

As a result, *any* Feynman-Witten diagram (with an arbitrary number of edges and loops, and on an arbitrary time-dependent spacetime background) can now be expressed in terms of the tree-level single-vertex diagrams of the theory.

As an example of the utility of our tree theorem, we consider the problem of determining from the wavefunction the equal-time correlation functions relevant for cosmology. We find that many (otherwise mysterious) cancellations between loop- and tree-level contributions can be explained by our Cosmological Tree Theorem, which provides a systematic way of simplifying the map from wavefunction to correlators. For example, for the one-loop power spectrum (or one-loop bispectrum), the standard approach via the Born rule generates 7 (or 13) separate Feynman-Witten wavefunction diagrams. Applying the Cosmological Tree Theorem reduces this to just 2 (or 5) tree-level diagrams for massless fields. The cancellations which take place between loop- and tree-level coefficients are analogous to the cancellation of IR divergences that takes place when computing an observable cross-section from a scattering amplitude: a phenomenon often called the "KLN theorem" [162, 163] (though see the recent discussion in [164, 165] for several subtleties in this cancellation). The Cosmological Tree Theorem (2.60) for wavefunction coefficients and the resulting Cosmological KLN theorem (together with the explicit results (3.23) and (B.7) for the one-loop power spectrum and bispectrum) are the main results of this work.

JHEP12(2023)076

Synopsis. We end this introduction with a description of our notation and a technical summary of the above results (i-iv). Then in section 2.1 we begin in earnest by deriving simple consequences of causality for tree-level diagrams and describing a bootstrap procedure for determining an arbitrary exchange diagram from its cut contact diagrams. In section 2.2 we move on to loop diagrams and illustrate our new Cosmological Tree Theorem with several simple examples, before providing a general proof for an arbitrary diagram. In section 3, we show how the Cosmological Tree Theorem can be used to identify the analytic structure of both the wavefunction and the corresponding in-in correlators, and give a simplified mapping from the wavefunction coefficients to the one-loop power spectrum (the longer expression for other correlators can be found in appendix B). Finally, we conclude in section 4 with a discussion of the future directions that this opens up.

1.1 Notation and conventions

While the majority of our notation is fairly standard, we collect it here for ease of reference. A reader already familiar with recent literature on the wavefunction of the Universe may wish to simply note the definition (1.15) of our disc/Disc operations and their graphical representation (1.16) and then proceed to our main results in section 1.2.

Time and momenta. We will consider quantum field theories on an isotropic time-dependent spacetime background, which can be written as $ds^2 = a^2(t)\left(-dt^2 + d\mathbf{x}^2\right)$ in conformal coordinates. We label each field $\phi_{\mathbf{k}}(t)$ by its temporal location t and spatial momentum $\mathbf{k}$ using the spatial Fourier transformation,

$$f(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, f_{\mathbf{k}}\, \exp(i\mathbf{k}\cdot\mathbf{x})\,, \tag{1.1}$$

which commutes with time derivatives and integrals. We adopt the following shorthands for time and momenta integrals, and frequently absorb factors of $(2\pi)^3$ into the Dirac delta function,

$$\int_t \equiv \int_{-\infty}^0 dt\,, \quad \int_{\mathbf{p}} \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3}\,, \quad \tilde{\delta}^3\left(\mathbf{k}\right) \equiv (2\pi)^3\delta^3\left(\mathbf{k}\right)\,. \tag{1.2}$$

Bold type always refers to spatial vectors, and we write their magnitude as $k \equiv |\mathbf{k}| \equiv +\sqrt{\delta^{ij}\mathbf{k}_i\mathbf{k}_j} \equiv \sqrt{\mathbf{k}\cdot\mathbf{k}}$.

Wavefunction coefficients. The central object of our study is the wavefunction of the Universe, Ψ, which describes how an initial Bunch-Davies vacuum[4] state $|\Omega\rangle$ evolves in time from the far past to conformal time $t = 0$. This time evolution is implemented by the unitary operator $\hat{U}$, and we project the resulting state onto the field eigenstate $|\phi\rangle$ at $t = 0$, which gives the wavefunction,

$$\Psi[\phi] \equiv \langle\phi|\hat{U}|\Omega\rangle \equiv \exp\left[+\sum_n^\infty \int_{\mathbf{k}_1,\ldots,\mathbf{k}_n} \frac{1}{n!}\psi_{\mathbf{k}_1\ldots\mathbf{k}_n}\phi_{\mathbf{k}_1}\cdots\phi_{\mathbf{k}_n}\right]\,. \tag{1.3}$$

[4]When discussing a general conformally flat spacetime, we refer to the vacuum which coincides with the instantaneous ground state in the far past as the Bunch-Davies vacuum.

JHEP12(2023)076

The $\psi_{\mathbf{k}_1\ldots\mathbf{k}_n}$ appearing in the exponential parameterisation of the wavefunction are the *wavefunction coefficients*. We occasionally write ψ_n as shorthand for $\psi_{\mathbf{k}_1\ldots\mathbf{k}_n}$.

Propagators. To develop a perturbative expansion for these coefficients, we define the usual *bulk-to-bulk* and *bulk-to-boundary* propagators,

$$\begin{aligned}\langle\phi=0|i\hat{\Pi}_{\mathbf{k}'}\hat{\phi}_{\mathbf{k}}(t)|\Omega\rangle &\equiv K_k(t)\,\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)\\ \langle\phi=0|T\hat{\phi}_{\mathbf{k}}(t_1)\hat{\phi}_{\mathbf{k}'}(t_2)|\Omega\rangle &\equiv G_k(t_1,t_2)\,\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)\end{aligned} \tag{1.4}$$

where we have written the field and its conjugate momentum in the Heisenberg picture, and when no time argument is given these are to be evaluated at $t=0$. T denotes time-ordering. In the free theory (1.8), these are related by,[5]

$$G_k(t_1,t_2)=iP_k\left[\Theta\left(t_1-t_2\right)K_k^*(t_1)K_k(t_2)+\Theta\left(t_2-t_1\right)K_k(t_1)K_k^*(t_2)-K_k(t_1)K_k(t_2)\right] \tag{1.5}$$

where P_k is the free-theory power spectrum, related to the wavefunction by $2\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{k}'}\equiv -\delta^d\left(\mathbf{k}+\mathbf{k}'\right)/P_k$. To prevent a proliferation of δ functions, it will often be convenient to write the power spectrum as,

$$P_{\mathbf{k}\mathbf{k}'}\equiv P_k\,\tilde{\delta}^d\left(\mathbf{k}+\mathbf{k}'\right)\ . \tag{1.6}$$

For concreteness, when we give examples below for a massless scalar field on Minkowki or de Sitter, the corresponding power spectrum and propagators are given by,

$$\begin{aligned}&\text{Minkowski:} && a(t)=1\,, && P_k=\frac{1}{2k}\,, && K_k(t)=e^{ikt}\\ &\text{de Sitter:} && a(t)=\frac{-1}{Ht}\,, && P_k=\frac{H^2}{2k^3}\,, && K_k(t)=e^{ikt}\left(1-ikt\right)\end{aligned} \tag{1.7}$$

together with (1.5) for $G_k(t_1,t_2)$.

Perturbation theory. We consider theories with a (generally time-dependent) Hamiltonian that can be separated into a solvable quadratic part plus an interacting part which can be treated perturbatively. For the quadratic part, a scalar field would have the Lagrangian,

$$\mathcal{L}_{\text{free}}(t)=\int d^3\mathbf{x}\,\sqrt{-g}\left(\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi-\frac{1}{2}m^2\phi^2\right) \tag{1.8}$$

where $g_{\mu\nu}(t)$ is the time-dependent background metric.[6] Our only requirement on the interactions is that the Hamiltonian $\mathcal{H}_{\text{int}}(t)$ be Hermitian (i.e. time evolution remains unitary in the interacting theory), although we also focus mostly on parity-invariant interactions for the sake of more concise expressions.

Since each wavefunction coefficient in (1.3) corresponds to the connected part of,

$$\frac{\langle\phi=0|i\hat{\Pi}_{\mathbf{k}_1}\ldots i\hat{\Pi}_{\mathbf{k}_n}\hat{U}|\Omega\rangle}{\langle\phi=0|\hat{U}|\Omega\rangle}=\frac{1}{\Psi[0]}\frac{\delta^n\Psi[\phi]}{\delta\phi_{\mathbf{k}_1}\ldots\delta\phi_{\mathbf{k}_n}}\bigg|_{\phi=0}\,, \tag{1.9}$$

[5] Note that $G_k(t_1,t_2)$ differs from the usual Feynman propagator by the final term in (1.5), which is a result of the different bra boundary condition (which is the zero-field eigenstate rather than the vacuum).

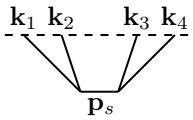
[6] We focus on the evolution of quantised fluctuations on a fixed time-dependent background (i.e. a decoupling limit $M_P\to\infty$ in which backreaction can be neglected).

JHEP12(2023)076

we can represent this matrix element as a diagrammatic series over Feynman(-Witten) diagrams in which,

- Each "external" line (which connects an interaction vertex at time t to the boundary) carrying momentum $\mathbf{k}$ represents a factor of the free-theory $K_k(t)$,
- Each "internal" line (which connects two interaction vertices at times t_1 and t_2) carrying momentum $\mathbf{p}$ represents a factor of the free-theory $G_p(t_1, t_2)$,
- Each vertex which connects n lines of momenta[7] $\{\mathbf{k}_1, \ldots, \mathbf{k}_n\}$ at time t corresponds to a factor of $\delta^n \mathcal{L}_{\rm int}(t)/\delta\phi_{\mathbf{k}_1} \ldots \phi_{\mathbf{k}_n}$,
- All vertex times and internal momenta are then integrated over, i.e. perform $\int_{-\infty}^{0} dt$ for every vertex and $\int_{\mathbf{p}}$ for every internal line,
- Finally, our normalisations are such that we include a factor of $+i$ for every vertex and $-i$ for every propagator: this leads to an overall i^{1-L}, where L is the number of loops in the diagram [106].

For instance, a theory in which $\mathcal{L}_{\rm int}$ contains the interaction $\frac{1}{3!}\lambda(t)\phi^3$ would also contain the diagram,



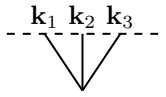
$$\equiv i \int_{t_1 t_2} \lambda(t_1)\lambda(t_2)\; K_{k_1}(t_1) K_{k_2}(t_1) G_{p_s}(t_1, t_2) K_{k_3}(t_2) K_{k_4}(t_2)\; \tilde{\delta}\left(\sum_{b=1}^{4} \mathbf{k}_b\right) \tag{1.10}$$

where $\mathbf{p}_s = \mathbf{k}_1 + \mathbf{k}_2$ is fixed by momentum conservation. Expanding (1.9) order by order in $\mathcal{H}_{\rm int}$ then gives each wavefunction coefficient as an expansion in Feynman diagrams with increasing numbers of loops,

$$\psi_n = \psi_n^{\text{tree}} + \psi_n^{\text{1-loop}} + \psi_n^{\text{2-loop}} + \ldots \tag{1.11}$$

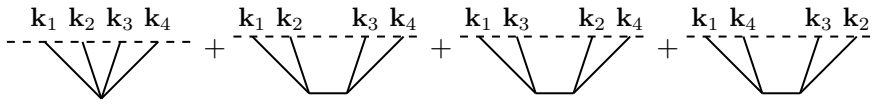
Explicitly, in the main text we make use of,

$$\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} =$$

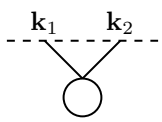


, (1.12)

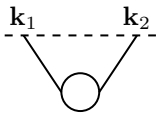
$$\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} =$$



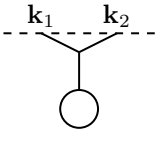
$$\psi^{\text{1-loop}}_{\mathbf{k}_1\mathbf{k}_2} = \frac{1}{2}$$



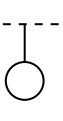
$$+ \frac{1}{2}$$



$$+ \frac{1}{2}$$



, $\psi^{\text{1-loop}}_{\mathbf{k}} = \frac{1}{2}$



[7] We use a condensed notation in which $\mathbf{k}$ represents the momentum *and also* any other relevant quantum numbers which distinguish the fields, e.g. their species (if more than one is present), their helicity (if spin is non-zero), etc. The $\delta/\delta\phi_{\mathbf{k}}$ then corresponds in the obvious way to taking a functional derivative with respect to whatever field is identified by $\mathbf{k}$.

JHEP12(2023)076

where the factors of 1/2 for the loop diagrams are the usual symmetry factors which appear when expanding a matrix element of the form (1.9).[8] In appendix B we give the analogous expansions for ψ_5^{tree}, $\psi_3^{1-\text{loop}}$ and $\psi_1^{2-\text{loop}}$ as further examples.

Discontinuities. The wavefunction coefficients will, in general, be non-analytic functions of the spatial momenta due to the dependence of $K_k(t)$ on the "energy" ω_k of each external line. On Minkowski, $\omega_k = +\sqrt{k^2+m^2}$ is given by the usual on-shell condition for free propagation. On a general time-dependent spacetime, we identify ω_k with the phase of $K_k(t)$ in the far past, $K_k(t) \sim e^{+i\omega_k t}$: for quasi-de Sitter spacetimes, this gives $\omega_k = k$ for any finite[9] m^2/H^2. As discussed in [104–107], we can exploit this particular non-analyticity to project any K_k or G_p within a diagram onto its real or imaginary part. The essential idea is to analytically continue a particular energy (or subset of energies) to negative values and then exploit the Hermitian analyticity of the propagators,

$$K_k(t)\big|_{\omega_k\to-\omega_k} = K_k^*(t)\,, \qquad\qquad G_p(t_1,t_2)\big|_{\omega_p\to-\omega_p} = G_p^*(t_1,t_2)\,. \tag{1.13}$$

This is usually achieved by continuing $k \to -k$ in order to cross a branch cut (e.g. sending $+\sqrt{k^2+m^2} \to -\sqrt{k^2+m^2}$).

In this work, we view each wavefunction coefficient or Feynman diagram as a function of both the momenta $\{\mathbf{k}\}$ and their energies $\{\omega_k\}$, and make use of the following two "discontinuity" operations,

$$\begin{aligned}
\underset{k_1,\ldots,k_j}{\text{disc}}\,[F(\omega_{k_1},\ldots,\omega_{k_n};\{\mathbf{k}\})] &\equiv F(\omega_{k_1},\ldots,\omega_{k_n};\{\mathbf{k}\}) - F\Big(-\omega_{k_1},\ldots,-\omega_{k_j},\omega_{k_{j+1}},\ldots,\omega_{k_n};\{\mathbf{k}\}\Big)\,,\\
\underset{k_1,\ldots,k_j}{\text{Disc}}\,[F(\omega_{k_1},\ldots,\omega_{k_n};\{\mathbf{k}\})] &\equiv F(\omega_{k_1},\ldots,\omega_{k_n};\{\mathbf{k}\}) - F^*\Big(\omega_{k_1},\ldots,\omega_{k_j},-\omega_{k_{j+1}},\ldots,-\omega_{k_n};\{-\mathbf{k}\}\Big)\,.
\end{aligned} \tag{1.14}$$

In words: disc corresponds to analytically continuing all indicated energies and is used to extract the imaginary part of external lines. On the other hand, Disc corresponds to analytically continuing all energies *except* those indicated, and is used to extract the imaginary part of internal lines. For example,

$$\underset{k_1}{\text{disc}}\,[G_{p_s}K_{k_1}K_{k_2}] = 2iG_{p_s}K_{k_2}\,\text{Im}\,K_{k_1}\,, \quad \underset{k_2}{\text{Disc}}\,[G_{p_s}K_{k_1}K_{k_2}] = 2iK_{k_1}\,\text{Im}\,(G_{p_s}K_{k_2})\,. \tag{1.15}$$

Note that for contact Feynman diagrams with no internal lines, $\underset{k}{\text{disc}}\,[i\psi] = \underset{k}{\text{Disc}}\,[i\psi]$.

In contrast to previous works,[10] we make use of these two different operations in order to avoid ever analytically continuing an internal energy. As far as we are concerned, each ψ_n depends only on the $\{\mathbf{k}\}$ and $\{\omega_k\}$ of the external momenta, and we may only analytically

[8] These symmetry factors appear in precisely the same way for scattering amplitudes, and we discuss them further in appendix C.

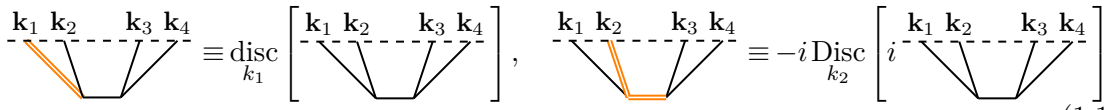
[9] Note that in the flat space limit, $H \to 0$ with m, k fixed, $K_k(t)$ reduces to the Minkowski propagator and ω_k becomes $\sqrt{k^2+m^2}$.

[10] In particular, while our Disc coincides with the Disc of [106] and the $\widetilde{\text{Disc}}$ of [67], the use of disc appears novel.

JHEP12(2023)076

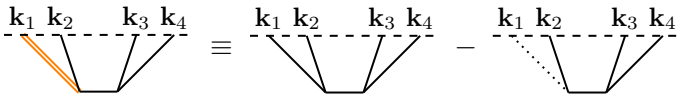
continue those (i.e. internal energies like $p_s = |\mathbf{k}_1 + \mathbf{k}_2|$ are treated as a function of $\mathbf{k}_1$ and $\mathbf{k}_2$ due to momentum conservation).[11]

To represent the discontinuities diagrammatically, we introduce graphs with highlighted lines to denote taking a single discontinuity with respect to those momenta. If only external lines are highlighted, it is the disc operation, while if any internal line is highlighted it is the Disc operation. For instance,[12]



(1.16)

If no lines are highlighted, no discontinuity is to be taken. It will also be convenient to use a dotted external line to denote the analytically continued propagator $K_k^*(t)$, since then the disc corresponds to,



(1.17)

Finally, a comment about the terminology "discontinuity". The name is inspired by the analogy with scattering amplitudes, where the discontinuity across e.g. the s-channel branch cut can be written in two equivalent ways,

$$\lim_{\epsilon\to 0}\left[\mathcal{A}_{12\to 34}(s+i\epsilon,t)-\mathcal{A}_{12\to 34}(s-i\epsilon,t)\right]=\mathcal{A}_{12\to 34}(s,t)-\mathcal{A}^*_{34\to 12}(s,t)\,. \tag{1.18}$$

Morally, the left-hand-side corresponds to our disc operation (since $s \pm i\epsilon$ gives an energy $\omega_s = \pm\sqrt{s}$) and the right-hand-side corresponds to our Disc operation (since by CPT the time-reversed process can be viewed as flipping the signs of all momenta). To make the analogy more concrete, we showed in [127] that the Minkowski wavefunction in the complex $\omega_1^2 = k_1^2 + m_1^2$ plane has a branch cut only along the positive real axis, and $\underset{\omega_1}{\text{disc}}$ is precisely the discontinuity across this cut.

1.2 Summary of main results

To complement our more thorough and pedagogical presentation in sections 2 and 3, here we provide a self-contained summary of the main narrative.

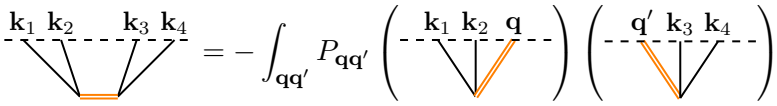
Review of cutting rules from unitarity. Existing cutting rules leverage the *unitarity* of time evolution in the interacting theory (together with the Hermitian analyticity of the free-theory propagators (1.13)) in order to reduce the number of internal lines in a Feynman

[11] This was recently made precise in [127] through a particular "off-shell" extension of the wavefunction on Minkowski. An analogous construction for general time-dependent spacetime backgrounds will be reported elsewhere [166].

[12] The factors of i for highlighted internal lines are such that the Disc selects $\text{Im}\, G_p$ at tree-level and $\text{Re}\, G_p$ for a one-loop diagram (given the factor of i^{1-L} in our Feynman rules).

JHEP12(2023)076

diagram. For instance, the simplest such rule is [106, 107],

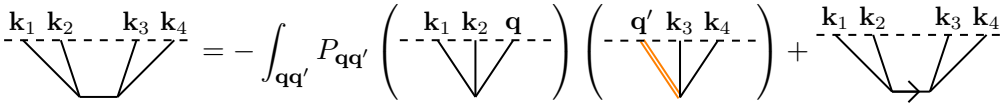
Unitarity $\Rightarrow$  (1.19)

which says that the Disc operation (depicted by a highlighted line) effectively "cuts" an internal line, brings the two resulting half-edges to the boundary and multiplies by a factor of the boundary power spectrum $P_{\mathbf{q}\mathbf{q}'}$.[13] This identity can then be used to fix the discontinuity of ψ_4 in terms of its cuts into $\psi_3 \times \psi_3$,

$$-i\,\mathrm{Disc}\left[i\,\psi^{\mathrm{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}\right] = \sum_{\mathrm{perm.}}^{3} \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\, \underset{q}{\mathrm{disc}}\left[\psi^{\mathrm{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}}\right] \underset{q'}{\mathrm{disc}}\left[\psi^{\mathrm{tree}}_{\mathbf{k}_3\mathbf{k}_4\mathbf{q}'}\right] . \tag{1.20}$$

in direct analogy with the usual Cutkosky cutting rules for scattering amplitudes.

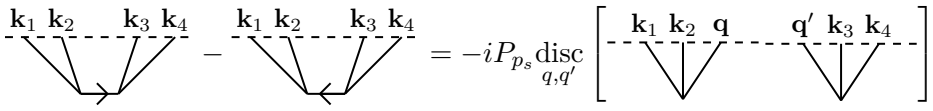
New cutting rules from causality. In this work, we leverage *causality* of the free theory to further constrain the perturbative wavefunction coefficients. For instance, the entire exchange diagram above can be written as,

 (1.21)

where a directed arrow from t_1 to t_2 represents the retarded propagator,[14]

$$G^R_p(t_1,t_2) \equiv 2P_p \mathrm{Im}\left[K_p(t_1)K^*_p(t_2)\right]\Theta(t_1-t_2) = G_p(t_1,t_2) - 2P_pK_p(t_1)\mathrm{Im}\left[K_p(t_2)\right] . \tag{1.22}$$

Since G^R_p is real its discontinuity vanishes, so taking Disc of (1.21) immediately reproduces (1.19). However, while the final term in (1.21) cannot be constrained by unitarity alone, it can be constrained by causality. In particular, we show that,

Causality $\Rightarrow$  (1.23)

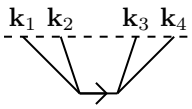
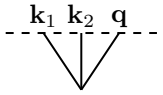
as a consequence of the fact that $G^R_p(t_1,t_2) - G^R_p(t_2,t_1)$ is a smooth function of t_1 and t_2 and so can be written in terms of K_p's without any step functions. This represents a qualitatively new way to cut diagrams, and demonstrates how causality can complement unitarity in fixing wavefunction coefficients. Furthermore, we find that demanding the

[13] For later convenience, we have chosen to express these cuts using an integral over the resulting external momenta of the cut line. In tree-level examples such as this one, these momenta integrals are trivial thanks to the δ functions inside $P_{\mathbf{q}\mathbf{q}'}$ and each wavefunction coefficient.

[14] G^R_p corresponds to the retarded part of the two-point function (A.1), as described in appendix A.

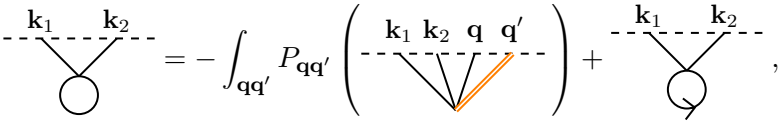
JHEP12(2023)076

absence of certain unphysical ("folded") singularities can be used to completely fix the remaining freedom in the retarded exchange diagram,

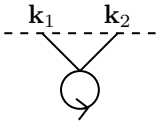
Analyticity (no folded singularities) $\Rightarrow$  is fixed by  , (1.24)

which is ultimately a re-writing of the bootstrap approach developed in [88] — the main difference[15] is that our retarded exchange is guaranteed to be an even function of the exchanged momenta, which can simplify the bootstrap procedure. Altogether, the combination of unitarity, causality and a particular analytic structure are enough to completely determine this (and indeed *any* tree-level) exchange diagram in terms of its cut diagrams with fewer internal lines.

Causality at loop level. The consequences of causality are even more striking for loop diagrams. While unitarity alone can determine the Disc of an arbitrary loop diagram in terms of diagrams with fewer internal lines [106], the combination of unitarity and causality immediately fixes the entire diagram (both Disc and non-Disc parts). The simplest example of this phenomenon is the following one-loop diagram,

 (1.25)

which has been re-written in terms of G_p^R using (1.22). Taking the Disc again removes the final term and reproduces the unitarity cutting rule of [106]. The new observation that we exploit here is that,

Causality $\Rightarrow$  (1.26)

Causality forbids loops (closed time-like curves), and hence once any loop diagram is expanded in terms of retarded G_p^R propagators *all of the loops must vanish*. Consequently, expressions like (1.25) can express any loop diagram in terms of (momentum integrals of) tree-level diagrams.

Cosmological tree theorem. Our main result is to do this systematically for any closed loop within a diagram D, and hence prove the *cosmological tree theorem* (2.60). Schematically, this takes the form,

$$-D \;=\; \sum_{\substack{\text{cuts}\\ C}} \int_{\substack{\text{cut line}\\ \text{momenta}}} \prod_{\substack{\text{subdiagrams}\\ n}} \text{disc}\left[D_C^{(n)}\right] \tag{1.27}$$

[15] There is another, more subtle, difference: the bootstrap procedure we describe here based on (1.21) and G_p^R does not require $\mathcal{H}_{\rm int}$ to be unitarity, it relies only on unitarity/causality of the free theory.

JHEP12(2023)076

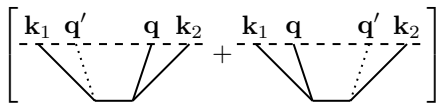
where the sum is over all possible ways of cutting one or more internal lines in the loop, and as a result of the cuts C the diagram may split into disconnected components which we label $D_C^{(n)}$. As a result of the cuts, every term on the right-hand-side has at least one fewer loops than the original diagram. Applied recursively to every closed loop in the diagram, this can be used to replace any arbitrary loop-diagram with (momentum integrals of) tree-level diagrams. As a simple example, the $\psi_2^{1-\text{loop}}$ given in (1.12) can be written as,

$$-2\psi^{\text{1-loop}}_{\mathbf{k}_1\mathbf{k}_2} = \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\, \underset{q'}{\text{disc}}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}\mathbf{q}'}\right] + \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'}\, \underset{q_2'}{\text{disc}}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}_2'}\right] \underset{q_1'}{\text{disc}}\left[\psi^{\text{tree}}_{\mathbf{k}_2\mathbf{q}_2\mathbf{q}_1'}\right] \,. \tag{1.28}$$

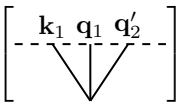
While the momenta integrals may still pose a challenge, this has achieved the following important simplification: *there are no longer any time integrals*. In practice, this means that once the tree-level wavefunction coefficients have been determined to a sufficiently high number of external legs, then the loop momenta integrands are completely fixed by the above tree theorem. This is an important step towards a general Landau analysis of the singularities in ψ_n, since from ψ_n^{tree} alone we can now determine all possible poles in the loop integrand.

Cosmological KLN theorem. Finally, as an application of our new relations, we consider how causality constrains cosmological correlators. While the standard Born rule mapping from wavefunction to correlator introduces many terms at loop-level, our tree theorem can be used to considerably simplify this map. In particular, we show that the power spectrum of massless fields at one-loop order can be written as (a momentum integral of) just two tree-level diagrams,

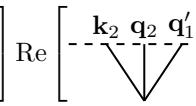
$$\frac{\langle\Omega|\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}|\Omega\rangle}{P_{k_1}P_{k_2}} = \frac{P_{\mathbf{k}_1\mathbf{k}_2}}{P_{k_1}P_{k_2}} + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\,\text{Re}\left[\right. \tag{1.29}$$



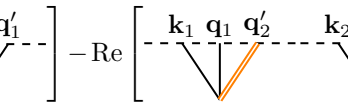
$$\left.\right] + \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \left(2\text{Re}\left[\right.\right.$$



$$\left.\left.\right] \text{Re}\left[\right.\right.$$



$$\left.\left.\right] - \text{Re}\left[\right.\right.$$



$$\left.\left.\right]\right)$$

where the dotted line represents the analytic continuation to negative energy $\omega_{q'}$ and P_k is the free-theory power spectrum. The analogous expression for the one-loop bispectrum is given in (B.7), (B.10), (B.11). Interestingly, we find that the pattern of analytic continuations and cuts is always such that no vertex (or connected set of vertices) can have a total energy which depends on both the total external energy and the loop momenta.[16] In the wavefunction coefficients, such vertices generically do appear and lead to branch cuts in the total external energy once the loop integration is performed. These branch cuts do not appear in equal-time correlators, as noted in [31, 127] in various examples, and our tree theorem makes their cancellation manifest. We demonstrate that this cancellation is not confined to the power spectrum or bispectrum, but in fact takes place very generally. Given the close analogy with the cancellation of IR divergences in amplitudes, we refer to this result as the *cosmological KLN theorem*.

[16] For instance in (1.29), there is no vertex (or set of vertices) which depends on both $\omega_{k_1} + \omega_{k_2}$ and $\mathbf{q} = -\mathbf{q}'$.

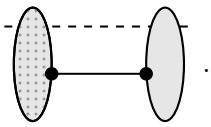
JHEP12(2023)076

2 Wavefunction coefficients

In this section, we describe how the causal properties of the retarded propagator G_p^R can be used to constrain the wavefunction coefficients defined in (1.3). We begin with a discussion of tree-level diagrams in subsection 2.1, and then move on to loops in subsection 2.2.

2.1 Causality and a tree-level bootstrap

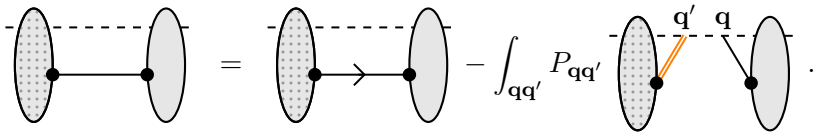
Here we consider the constraints that causality can place on a general tree-level exchange diagram, which we represent diagrammatically as,[17]



(2.1)

Each gray blob represents a particular subdiagram with arbitrarily many external and internal lines (i.e. a general function of K_k and G_p and their conjugates), and they are connected by the single internal line shown. We have added a dotted pattern to the left blob to indicate that it need not be the same subdiagram as the right blob.

Introducing directed edges. From the definition (1.22) of the retarded propagator and the definition (1.16) of the discontinuity, we can trade any internal line (factor of G_p) for a directed line (factor of G_p^R) as follows,

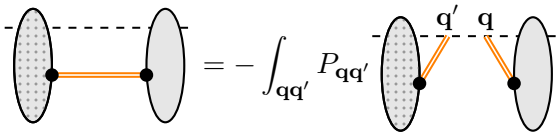


(2.2)

Note that since G_R is real,

$$\text{Disc}\left[iG_p^R(t_1,t_2)\right] = 0\,, \tag{2.3}$$

i.e. a directed internal line will vanish when highlighted. Taking the Disc of (2.2) therefore produces,



(2.4)

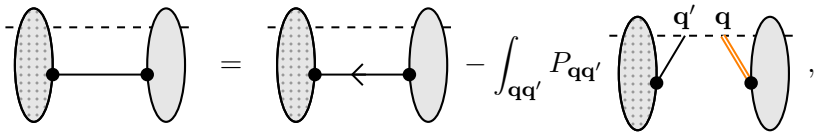
which are the tree-level unitarity cutting rules of [106, 107]. Note that while (2.4) requires Hermiticity of the interaction Hamiltonian (i.e. real coupling constants so the discontinuity on the left-hand-side implements $G_p \to 2i\,\text{Im}\,G_p$), the causal representation (2.2) does not (since disc does not require any complex conjugation, there is no reality restriction on the couplings).[18]

[17] As described in section 1.1, the dashed horizontal line represents the time $t = 0$ at which the wavefunction is evaluated: that the gray blob extend above this line simply reflects that they may contain arbitrarily many external (as well as internal) lines terminating at $t = 0$, and in particular does not represent any interaction or time evolution beyond $t = 0$.

[18] As described in appendix A, the definition (1.22) which underpins (2.2) follows from the unitarity and causality of the free theory (1.8) only.

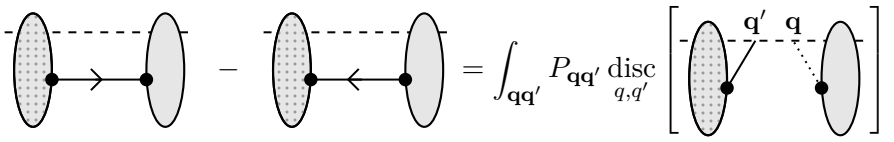
JHEP12(2023)076

Cutting rules from causality. But now notice that since $G_p(t_1, t_2)$ is symmetric in t_1 and t_2, we could equally have written,



(2.5)

using $G_p^R(t_2, t_1)$, which is essentially the advanced propagator from t_1 to t_2. Additional cutting rules therefore follow from comparing these two different expressions: for instance their difference is,

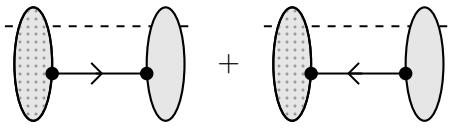


(2.6)

where the dotted external line indicates the analytically continued propagator $K_q^*(t)$. So by combining retarded and advanced propagation (or equivalently $\psi(\{\mathbf{k}_L\}, \{\mathbf{k}_R\}) - \psi(\{\mathbf{k}_R\}, \{\mathbf{k}_L\})$, where $\{\mathbf{k}_L\}$ and $\{\mathbf{k}_R\}$ are the external momenta of the left- and right-hand blobs) we can effectively cut the internal line and produce two disconnected subdiagrams. This cutting rule is tied to causality, since it can be written as the propagator identity,

$$\begin{aligned} iG_p^R(t_1, t_2) - iG_p^R(t_2, t_1) &= 2iP_p \mathrm{Im}\left[K_p(t_1)K_p^*(t_2)\right](\Theta(t_1 - t_2) + \Theta(t_2 - t_1)) \\ &= P_p \operatorname*{disc}_p \left[K_p(t_1)K_p^*(t_2)\right] \end{aligned} \tag{2.7}$$

where the right-hand-side can only be written in terms of K_p alone since G_p^R is proportional to the Heaviside step function $\Theta(t_1 - t_2)$. In an acausal theory, with no such retarded propagator (no way of separating sources in the future from responses in the past), it would not be possible to construct a cutting rule of the form (2.6).[19] As mentioned in the introduction, these causal cutting rules complement previous unitarity cutting rules in that they constrain the part of the diagram with vanishing Disc. What we have shown is that unitarity and causality together can fix the diagram up to a residual freedom which corresponds to adding,



(2.8)

We give an overview and explicit comparison of the different unitarity and causality cutting rules in appendix A, where we also make precise the connection with unitarity / causality of the two-point function.

[19] Note that analogous cutting rules exist for S-matrix elements, in which case the internal line of the Feynman diagram would be the usual Feynman propagator and the highlighted line would correspond to simply K_p^* rather than $\mathrm{Im}K_p$.

JHEP12(2023)076

Causal bootstrap. The cutting rule (2.6) is written in terms of graphs with directed edges, but ideally we would phrase such constraints in terms of the original diagram (2.2) with undirected edges since this is what determines the wavefunction coefficients / cosmological correlators. To do this requires the symmetric part (2.8) of the retarded propagator. Remarkably, we have found that in practice it is not necessary to compute this object explicitly, since it can inferred from the unitarity/causality cuts by demanding the absence of certain unphysical singularities. This provides yet another way to "bootstrap" these wavefunction coefficients (i.e. determine them without actually computing nested time integrals like (1.10)). Concretely, we input the following three ingredients:

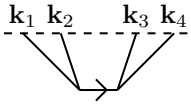
(i) *Causality of the free theory*, so that each bulk-to-bulk $G_p(t_1,t_2)$ may be expanded in terms of a retarded $G_p^R(t_1,t_2) = i\Delta_p^S(t_1,t_2)\Theta(t_1-t_2)$ for some Δ_p^S,

(ii) *Unitarity of the free theory*, which fixes $\Delta_p^S(t_1,t_2) = 2P_p\text{Im}\left[K_p(t_1)K_p^*(t_2)\right]$ in terms of the bulk-to-boundary propagator,

(iii) *Bunch-Davies initial state*, so that the only singularities are at kinematics for which the total energy entering one or more vertex vanishes,

and the output is a procedure for determining any tree-level exchange diagram from its cut diagrams with fewer internal lines. It was previously shown in [88] that the absence of certain unphysical ("folded") singularities can be combined with unitarity of the interacting theory to bootstrap an exchange diagram from its unitarity cuts. The approach described here is essentially the same, but we apply this bootstrap to determine the retarded part of the exchange (2.8) rather than the full diagram: this means that we only require unitarity of the free theory (not necessarily of the interacting theory) and in practice can exploit the fact that retarded propagator is a real, even function of p to simplify the computations and remove the need for any further conditions (locality, etc.) to completely fix the exchange diagram up to contact terms.

An example. Before describing the general procedure, let us give a simple example. Suppose we have determined the contact ψ_3^{tree} diagram in (1.12), and wish to determine from it the exchange diagrams in ψ_4^{tree}. Replacing the internal line with a retarded propagator using (1.22) produces (1.21), in which the first term is determined by ψ_3^{tree} but the second term,



$$\equiv P_{p_s}\underset{p_s}{\text{disc}}\left[\int_{t_1t_2} K_{k_1}(t_1)K_{k_2}(t_1)K_{k_3}(t_2)K_{k_4}(t_2)\,K_{p_s}(t_1)K_{p_s}^*(t_2)\Theta\left(t_1-t_2\right)\right] \tag{2.9}$$

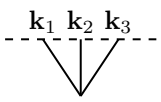
seems to require a nested time integral. Rather than compute this explicitly, we can notice that since it takes the form $P_{p_s}\underset{p_s}{\text{disc}}\left[f(p_s)\right]$, it must be an *even* function of p_s.[20] We can

[20] This follows from the fact that $\underset{p_s}{\text{disc}}$ is clearly odd by definition (1.15), and the power spectrum P_{p_s} is either odd or can be made odd by a suitable renormalisation (since any even p_s^{2n} term can be absorbed into the local counterterm $(\partial^n\phi)^2$).

JHEP12(2023)076

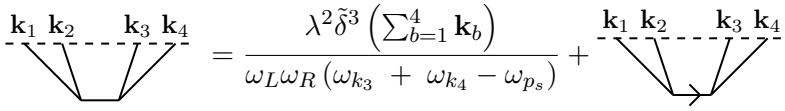
therefore play the following trick. On general grounds, the Bunch-Davies wavefunction coefficient may only have singularities when the total energy flowing into one or more vertices vanishes. In this example, that corresponds to the total energy $\omega_T = \omega_{k_1} + \omega_{k_2} + \omega_{k_3} + \omega_{k_4}$ and the partial energies $\omega_L = \omega_{k_1} + \omega_{k_2} + \omega_{p_s}$ and $\omega_R = \omega_{k_3} + \omega_{k_4} + \omega_{p_s}$. Since the ψ_3^{tree} discontinuity in the first term of (1.21) introduce unphysical singularities at $\omega_{p_s} = \omega_{k_1} + \omega_{k_2}$ and $\omega_{k_3} + \omega_{k_4}$, we can demand that these are cancelled by the even function (2.9). This effectively fixes (2.9) up to terms which are analytic in p_s^2 (and such terms can be absorbed into contact diagrams).

For example, for a $\frac{\lambda}{3!}\phi^3$ interaction on a Minkowski spacetime background that produces,



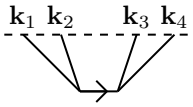
$$= \frac{\lambda \tilde{\delta}^3 \left(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3\right)}{\omega_{k_1} + \omega_{k_2} + \omega_{k_3}}, \tag{2.10}$$

the equation (1.21) becomes,



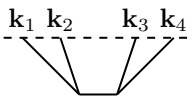
$$= \frac{\lambda^2 \tilde{\delta}^3 \left(\sum_{b=1}^4 \mathbf{k}_b\right)}{\omega_L \omega_R \left(\omega_{k_3} + \omega_{k_4} - \omega_{p_s}\right)} + \tag{2.11}$$

The only even function of p_s which cancels the unphysical singularity (and does not introduce any further unphysical singularities) is,



$$= \frac{\lambda^2 \tilde{\delta}^3 \left(\sum_{b=1}^4 \mathbf{k}_b\right)}{\omega_T} \frac{1}{\omega_{p_s}^2 - (\omega_{k_3} + \omega_{k_4})^2} \tag{2.12}$$

up to an analytic function of p_s^2 which can be absorbed into contact interactions. Substituting this back into (2.11), we arrive at the simple answer,



$$= \frac{\lambda^2 \tilde{\delta}^3 \left(\sum_{b=1}^4 \mathbf{k}_b\right)}{\omega_T \omega_L \omega_R} \,. \tag{2.13}$$

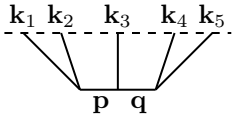
Notice that we have therefore determined this exchange diagram *without doing the nested time integrals* in (1.10). Of course, in this simple Minkowski example the time integrals are fairly straightforward (and indeed produce (2.13)). But this same approach can be applied more generally to *any* tree-level diagram on *any* time-dependent spacetime background.

General procedure. For any tree-level diagram D, we can split it into a retarded and a cut part as shown in (2.2): we will write this as $D = D_R + D_C$. Now the bootstrap proceeds as follows. Label the partial energies of the original diagram D which depend on the momentum of the internal line's energy ω_p as $\{E_1(\omega_p), \ldots, E_n(\omega_p)\}$. D will generically contain singularities when any $E_j(\omega_p) = 0$. Due to the analytic continuation, the cut D_C terms will contain additional singularities at $E_j(-\omega_p) = 0$. These unphysical singularities are not present in D, and so they must exactly cancel with singularities in D_R. However, since D_R is proportional to $P_p \underset{p}{\text{disc}}$ it must be an even function of p, and so if it contains

JHEP12(2023)076

a singularity at $E_j(-\omega_p)$ it must also contain one at $E_j(+\omega_p)$ with the same residue. Proceeding in this way, one can determine all of the singularities at $E_j(\pm\omega_p)$ and their residues, which fixes D_R up to a remainder that is analytic in ω_p^2. This remainder, if non-zero, can be absorbed into the local interaction which resembles D but with the internal line collapsed into a single vertex. This represents the freedom to add any contact diagram to D — this will not affect the cut diagrams D_C and is ultimately related to our freedom to perform field redefinitions.

Another Minkowski example. As a second simple example of this procedure, consider the following tree-level diagram,

$D =$ 

(2.14)

The partial energies flowing into each vertex or collection of vertices are,

$$\begin{aligned} &E_1(\omega_p) = \omega_{k_1} + \omega_{k_2} + \omega_p\,, && E_2(\omega_p) = \omega_{k_3} + \omega_p + \omega_q\,, && E_3 = \omega_{k_4} + \omega_{k_5} + \omega_q\,, \\ &E_4(\omega_p) = \omega_{k_3} + \omega_{k_4} + \omega_{k_5} + \omega_p\,, && E_5 = \omega_{k_1} + \omega_{k_2} + \omega_{k_3} + \omega_q\,, && \omega_T = \omega_{k_1} + \omega_{k_2} + \omega_{k_3} + \omega_{k_4} + \omega_{k_5} \end{aligned} \tag{2.15}$$

where we have highlighted the p dependence of E_1, E_2 and E_3. Replacing this internal line with retarded propagators as in (2.2) produces the following cut contribution,[21]

$$D_C = -\frac{\lambda^3}{2\omega_p E_3}\left[\frac{1}{E_1(\omega_p)E_2(\omega_p)E_4(\omega_p)} - \frac{1}{E_1(\omega_p)E_2(-\omega_p)E_4(-\omega_p)}\right], \tag{2.16}$$

for the cubic interaction on Minkowski (2.10) (and the 4-point exchange we bootstrapped from it in (2.13)). For now we suppress the overall momentum-conserving δ-function. D_C has unphysical singularities at both $E_2(-\omega_p)$ and $E_4(-\omega_p) = 0$, which cannot appear in D. Consequently, we can construct D_R using the ansatz $\sum_j Z_j/(E_j(\omega_p)E_j(-\omega_p))$ and fix each Z_j so that these folded singularities cancel. The result is,

$$\begin{aligned} D_R = -\frac{\lambda^3}{E_3}\Bigg[&\frac{1}{E_2(\omega_p)E_2(-\omega_p)E_1(\omega_{k_3}+\omega_q)E_4(-\omega_{k_3}-\omega_q)} \\ &+ \frac{1}{E_4(\omega_p)E_4(-\omega_p)E_1(\omega_{k_3}+\omega_{k_4}+\omega_{k_5})E_2(-\omega_{k_3}-\omega_{k_4}-\omega_{k_5})}\Bigg] \end{aligned} \tag{2.17}$$

so that the total $D_C + D_R$ is simply,

$$D = \frac{\lambda^3}{\omega_T E_1(\omega_p)E_2(\omega_p)E_3}\left(\frac{1}{E_4(\omega_p)} + \frac{1}{E_5}\right). \tag{2.18}$$

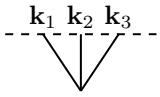
Again, notice how there was no need to ever perform a nested time integral: the retarded part D_R was fixed entirely by the singularities of D_C.

[21] Note that we could have cut $\mathbf{q}$ and taken the disc of a single line in the resulting ψ_3^{tree}: since this would exactly mimic the previous example, here we have chosen to cut $\mathbf{p}$ and take the disc of the resulting 4-point exchange diagram.

JHEP12(2023)076

In this simple Minkowski setting, it is straightforward to verify (2.18) by explicitly performing the time integrals. The virtue of fixing D using causality is that we remove the need to perform any time integrals, once the basic $\psi_3^{\rm tree}$ vertex is known. This comes at the cost of introducing unphysical singularities in the intermediate steps which are ultimately absent in the final result (2.18). The simple structure of (2.18), and indeed of all Minkowski wavefunction coefficients, can be made manifest using the "old-fashioned perturbation theory" based on the Lippmann-Schwinger equation (see e.g. [91, 112, 127] for recent reviews). This approach also removes the need for time integrals, and introduces only the physical singularities in the partial energies. However, the utility of OFPT is largely limited to Minkowski spacetime. So the second virtue of fixing D using causality as described above, is that this method can be applied to any time-dependent background.

De Sitter example. To demonstrate this, consider the same exchange diagram (2.11) but this time for the interaction $\mathcal{L}_{\rm int} = \frac{\lambda}{3!}\dot\pi^3 a(t)$ for a massless scalar field π on de Sitter. The contact diagram is readily determined by doing a single time integral of the K_k propagator in (1.7),



$$= -\frac{2\lambda k_1^2 k_2^2 k_3^2}{H(k_1+k_2+k_3)^3}\,, \tag{2.19}$$

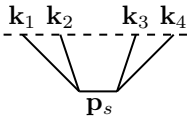
where for now we will omit the $\tilde\delta^3$ functions. The cut contributions to the exchange diagram are then,

$$D_C = -\frac{2\lambda^2 k_1^2 k_2^2 k_3^2 k_4^2 p}{(k_{12}+p_s)^3}\left[\frac{1}{(k_{34}+p_s)^3} - \frac{1}{(k_{34}-p_s)^3}\right] \tag{2.20}$$

where $k_{ij} \equiv k_i + k_j$. Proceeding as above by making a general ansatz $D_R = \sum_{i,j} Z_j/(E_j(p)E_j(-p))^i$, we can immediately determine the retarded contributions,

$$D_R = -\frac{4\lambda^2 k_1^2 k_2^2 k_3^2 k_4^2}{k_T^5}\left[\frac{6k_{34}^2}{(p_s^2-k_{34}^2)} - \frac{6p_s^2 k_T k_{34}}{(p_s^2-k_{34}^2)^2} + \frac{p_s^2 k_T^2\,(3k_{34}^2+p_s^2)}{(p_s^2-k_{34}^2)^3}\right]\,, \tag{2.21}$$

again up to an analytic function of p_s^2. Summing D_C and D_R gives,



$$= -\frac{4\lambda^2 k_1^2 k_2^2 k_3^2 k_4^2}{k_T^5}\left(\frac{4p_s^4 k_T^2}{s^3} + \frac{6p_s^3(k_T^3 - s k_T)}{s^3} + \frac{3p_s^2}{s^3}(2s^2 - s k_T^2 + k_T^4) + 6\right)\,, \tag{2.22}$$

for two $\dot\pi^3$ interactions on de Sitter, where we have used the variables,

$$k_T \equiv k_1+k_2+k_3+k_4 \quad \text{and} \quad s \equiv -(k_1+k_2+p_s)(k_3+k_4+p_s)\,. \tag{2.23}$$

In our opinion, this bootstrap route is far simpler than performing the nested time integration in (1.10). We verified (2.22) by carrying out the time integration explicitly, and

JHEP12(2023)076

also by checking that it correctly factorises in limits[22] as $k_T \to 0$, $E_L \to 0$ or $E_R \to 0$. It also matches the result given in [88, (6.47)].

While the three examples given in this section are fairly simple diagrams for which the nested time integration can be performed explicitly, the point that we wish to stress is that this procedure can be applied to *any* tree-level diagram on *any* time-dependent spacetime, including those for which the time integrals become arduous or intractable. This removes the need to ever perform a nested time integral in cosmology at tree-level. By repeatedly applying (2.2) and fixing the D_R part by matching against the unphysical singularities in D_C, all internal lines can be removed and any D can be expressed as a product of simple contact diagrams (which require only a single time integral) and their discontinuities.

Manifest locality. There is a further difference between the bootstrap procedure described above and that introduced in [88].[23] In [88], the ansatz used for the exchange diagram is only partially fixed by demanding the absence of unphysical singularities: there is a remainder which must be fixed by demanding that $\partial_{k_b}\psi_n|_{k_b=0}$ vanishes for each of the external lines (assumed massless) with energy $\omega_{k_b} = k_b$. This follows from the property of massless mode functions,

$$\frac{\partial}{\partial k}\left[K_k(t)\right]|_{k=0} = 0\,. \tag{2.25}$$

In the bootstrap described above, both D_C and D_R automatically satisfy $\partial_{k_b} D_C|_{k_b=0} = \partial_{k_b} D_R|_{k_b=0} = 0$ since they can be explicitly represented as integrals of $K_k(t)$ propagators (e.g. as in (2.9)). In the language of [88], the additional contribution $B_{\rm MLT}$ is always zero in our construction.

This is best illustrated by the examples of a $\pi\dot{\pi}^2$ interaction for massless fields in de Sitter, for which the $\psi_3^{\rm tree}$ input for the boostrap would be,

$$\psi^{\rm tree}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} \propto -\frac{e_3 e_2 + k_T e_2^2 - 2k_T^2 e_3}{k_T^2}, \tag{2.26}$$

or a $\dot{\pi}(\partial_i\pi)^2$ interaction in de Sitter, for which the $\psi_3^{\rm tree}$ input for the boostrap would be,

$$\psi^{\rm tree}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} \propto -\frac{k_T^6 - 3k_T^4 e_2 + 11k_T^3 e_3 - 4k_T^2 e_2^2 - 4k_T e_2 e_3 + 12 e_3^2}{2k_T^3}\,, \tag{2.27}$$

where we have introduced the symmetric polynomials,

$$e_3 = k_1 k_2 k_3\,, \qquad e_2 = k_1 k_2 + k_2 k_3 + k_3 k_1\,, \qquad k_T = k_1 + k_2 + k_3\,. \tag{2.28}$$

[22] Note that in the amplitude limit, $k_T \to 0$, the variable $s \to (k_1+k_2)^2 - p_s^2$ and becomes the usual Mandelstam invariant. In fact, in that limit,

$$(2.22) \sim -\frac{24\lambda^2 k_1^2 k_2^2 k_3^2 k_4^2}{k_T^5}\,\frac{(k_1+k_2)^2}{s} \propto \frac{\mathcal{A}}{k_T^5 \prod_{a=1}^4 k_a^2 P_{k_a}} \tag{2.24}$$

and indeed coincides with the expected amplitude $\mathcal{A}$ for an $\dot{\pi}^3 \times \dot{\pi}^3$ interaction (see e.g. [87, 105] for a description of the proportionality constant in (2.24)).

[23] We thank David Stefanyszyn for pointing this out to us.

JHEP12(2023)076

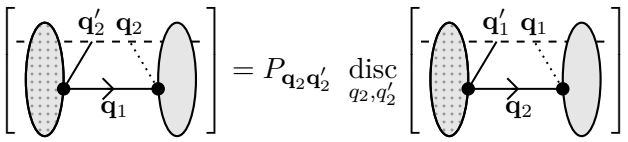
In both cases, D_R can be fixed (up to an analytic remainder which can be absorbed into contact interactions) by demanding that it cancels the unphysical poles appearing in D_C and that it is an even function of p_s, and the resulting $D = D_C + D_R$ satisfies $\partial_{k_b} D|_{k_b=0}$ since both D_R and D_C satisfy this property separately. It would therefore be interesting to recast this causal bootstrap as a BCFW-type recursion relation, since the boundary term from integrating an arc at infinity seems simpler than in previous approaches [67, 88].

Further identities. Finally, we remark that (2.6) is by no means the only constraint which follows from the properties of G_p^R. Unlike the bulk-to-bulk G_p, the retarded G_p^R is simply proportional to $\Theta(t_1 - t_2)$. As a result, we have propagator identities of the form,

$$\mathrm{Im}\left[K_{p_1}(t_1)K^*_{p_2}(t_2)\right] G^R_{p_2}(t_1,t_2) = G^R_{p_1}(t_1,t_2)\,\mathrm{Im}\left[K_{p_2}(t_1)K^*_{p_1}(t_2)\right] \tag{2.29}$$

which corresponds to the diagrammatic identity,

$$P_{\mathbf{q}_1\mathbf{q}'} \underset{q_1,q_1'}{\mathrm{disc}} \left[\right. \quad = P_{\mathbf{q}_2\mathbf{q}_2'} \underset{q_2,q_2'}{\mathrm{disc}} \left[\right. \quad . \tag{2.30}$$



In words, these identities say that in any diagram with *colinear* momenta (i.e. a pair of external legs with $\mathbf{q} + \mathbf{q}' = 0$), the disc with respect to these momenta can be used to *exchange* the momenta of external and internal lines. This is a further constraint which all Feynman-Witten diagrams must satisfy due to the causal properties of the free theory. Unlike the above causal cutting rules, these identities only apply to particular colinear configurations of momenta. At present, we have not found any particularly useful application for them (except that they guarantee the consistency of different loop-level cutting rules discussed in the next section). There may be some connection to the factorisation in non-local soft limits recently studied in [94]. It would be interesting to explore these relations further in future.

2.2 Loops and the cosmological tree theorem

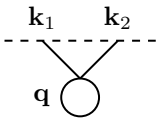
Now we turn our attention to diagrams containing loops. In particular, we introduce and then prove our Cosmological Tree Theorem, which can replace any loop diagram with a sum over (momentum integrals of) tree diagrams. To build some intuition for how this works, we begin by outlining several simple examples in subsection 2.2.1, postponing the general statement of the theorem to subsection 2.2.2. While we initially focus on scalar fields, we explain the straightforward generalisation to spinning fields at the end of this section.

2.2.1 Some examples

First we show, simply by inspection, that the one-loop correction to the wavefunction coefficients which determine the power spectrum and the bispectrum can be written in terms of simpler tree-level diagrams.

JHEP12(2023)076

One vertex. The simplest one-loop diagram contains only a single vertex. For instance, a $\frac{\lambda}{4!}\phi^4$ interaction contributes to the quadratic wavefunction coefficient through the diagram,

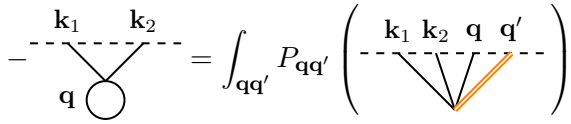


$$= \int_{\mathbf{q}} \int_t \lambda\, K_{k_1}(t) K_{k_2}(t) G_q(t,t)\; \tilde{\delta}\left(\mathbf{k}_1 + \mathbf{k}_2\right) \ . \tag{2.31}$$

The coupling λ can depend on both time and, in the case of derivative interactions, all three momenta entering the vertex. Notice that since the bulk-to-bulk propagator in this diagram is evaluated at coincident times, it no longer contains any $\Theta(t_1 - t_2)$ step functions. It can be written simply as,

$$G_q(t,t) = 2P_q K_q(t) \mathrm{Im}(K_q(t)). \tag{2.32}$$

This allows us to express the momentum integrand in (2.31) in terms of a tree-level diagram without any internal lines, using the disc operation (1.15) to extract the required imaginary part of an external propagator. In this example, the identity (2.32) implies,

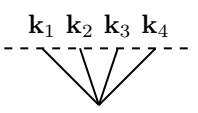


$$= \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \left(\right) \tag{2.33}$$

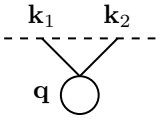
where the highlighted line corresponds to taking a disc, as defined in (1.16). The identity (2.33) is the simplest example of our tree theorem. It shows that the time integration in (2.31) can be replaced by cutting open the loop and taking a discontinuity. This turns out to be a very general property of loop diagrams and is connected to the causal structure of the propagator (anticipated in (1.26)).

Before moving on to our next example, let us make a few remarks about (2.31).

(i) For massless fields on a Minkowski spacetime background, these diagrams are particularly simple,



$$= \frac{\lambda\, \tilde{\delta}^4\left(\sum_{b=1}^4 \mathbf{k}_b\right)}{\omega_{k_1} + \omega_{k_2} + \omega_{k_3} + \omega_{k_4}} \ ,$$



$$= \frac{\lambda \tilde{\delta}^4\left(\mathbf{k}_1 + \mathbf{k}_2\right)}{\omega_{k_1} + \omega_{k_2}} \int_{\mathbf{q}} \frac{1}{\omega_{k_1} + \omega_{k_2} + 2\omega_q}, \tag{2.34}$$

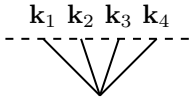
and it is straightforward to confirm that (2.33) holds, essentially as a consequence of the partial fraction identity,

$$\frac{1}{(\omega_{k_1} + \omega_{k_2})(\omega_{k_1} + \omega_{k_2} + 2\omega_q)} = -\frac{1}{2\omega_q}\left(\frac{1}{\omega_{k_1} + \omega_{k_2} + \omega_q + \omega_{q'}} - \frac{1}{\omega_{k_1} + \omega_{k_2} + \omega_q - \omega_{q'}}\right)\Big|_{q'=q} . \tag{2.35}$$

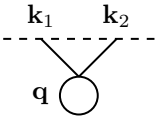
JHEP12(2023)076

The simplicity of the integrand in (2.34) can again be understood from old-fashioned perturbation theory and its resulting recursion relation. One way to think of (2.33) is that it gives up some of that simplicity (by splitting the single rational integrand (2.34) into the two separate terms in (2.35)) in order to buy a greater generality: (2.33) can be immediately applied to fields of any mass on any time-dependent spacetime background.[24]

(ii) As an example of this generality, consider the contact diagram given by the interaction $\frac{\lambda}{4!}\dot{\pi}^4$ on a fixed de Sitter spacetime. In that case, performing the explicit time integration leads to,



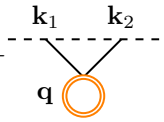
$$= \lambda\tilde{\delta}^3\left(\sum_{b=1}^{4}\mathbf{k}_b\right)\frac{24k_1^2k_2^2k_3^2k_4^2}{(k_1+k_2+k_3+k_4)^5},$$

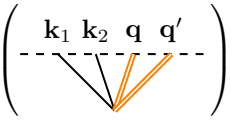


$$= -\lambda\tilde{\delta}^3\left(\mathbf{k}_1+\mathbf{k}_2\right)\int_{\mathbf{q}}\left(\frac{24k_1^2k_2^2q^4}{(k_1+k_2+2q)^5}-\frac{24k_1^2k_2^2q^4}{(k_1+k_2)^5}\right), \tag{2.36}$$

which again satisfy (2.33).

(iii) In many ways the cutting rule (2.33) is a completion of those developed in [106], since it expresses both the real *and* imaginary part of the loop diagram in terms of tree-level cuts. It is straightforward to recover the results of [106] by taking a further discontinuity of (2.33),

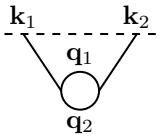
$$-$$ 

$$= \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}$$ 

(2.37)

where we have used that $\underset{q}{\text{Disc}}\left[\underset{q'}{\text{disc}}\left[K_q(t)K_{q'}(t)\right]\right] = \underset{q,q'}{\text{disc}}\left[K_q(t)K_{q'}(t)\right]$. Note that the unitarity cutting rule (2.37) requires unitarity time evolution in the fully interacting theory (so that λ commutes with the Disc operation defined in (1.15)), whereas the tree theorem (2.33) uses only the unitarity/causality of the free theory.[25]

Two vertices. The next-simplest one-loop diagrams contains two vertices. For instance, a $\frac{\lambda}{3!}\phi^3$ interaction contributes to the quadratic wavefunction coefficient through the diagram,



$$= \int_{\mathbf{q}_1\mathbf{q}_2}\int_{t_1,t_2}\lambda(t_1)\lambda(t_2)K_{k_1}(t_1)G_{q_1}(t_1,t_2)G_{q_2}(t_2,t_1)K_k(t_2)$$
$$\times\,\tilde{\delta}^3(\mathbf{k}_1+\mathbf{q}_1-\mathbf{q}_2)\tilde{\delta}^3(\mathbf{k}_2-\mathbf{q}_1+\mathbf{q}_2)\,. \tag{2.38}$$

[24] We will return to the comparison with OFPT at the end of this section.

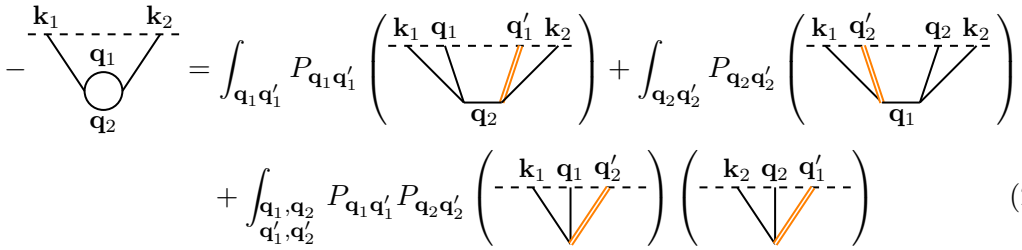
[25] Since $\partial_t \underset{p}{\text{disc}}\left[K_p(t)\right] = \underset{p}{\text{disc}}\left[\partial_t K_p(t)\right]$, λ automatically commutes with disc for interactions with time-derivatives.

JHEP12(2023)076

Ordinarily, a product of two bulk-to-bulk propagators would contain two Heaviside Θ functions. However, when evaluated at cyclic arguments as in (2.38), this product only contains a single Θ function. It can therefore be written in terms of a single bulk-to-bulk propagator,

$$\begin{aligned} G_{p_1}(t_1,t_2)G_{p_2}(t_2,t_1) = {} & 2P_{p_1}\mathrm{Im}(K_{p_1}(t_2))G_{p_2}(t_2,t_1)K_{p_1}(t_1) + 2P_{p_2}\mathrm{Im}(K_{p_2}(t_1))G_{p_1}(t_1,t_2)K_{p_2}(t_2) \\ & - 4P_{p_1}P_{p_2}K_{p_1}(t_1)\mathrm{Im}(K_{p_2}(t_1))K_{p_2}(t_2)\mathrm{Im}(K_{p_1}(t_2)). \end{aligned} \tag{2.39}$$

This suggests that it should be possible to express the momentum integrand of (2.38) using diagrams with at most one internal line (and therefore tree-level). In fact, it is straightforward to verify that the three terms of (2.39) can be written as,



(2.40)

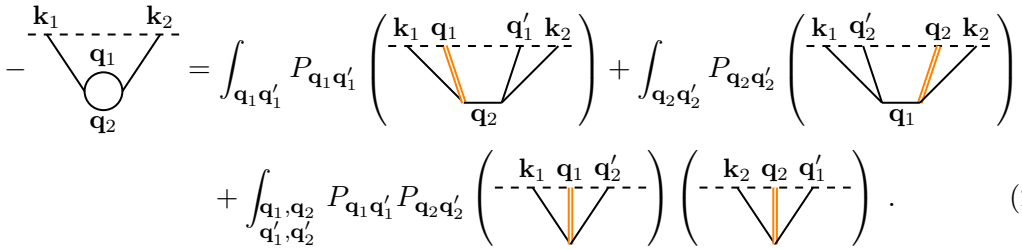
where again we have used coloured lines to denote the imaginary part of a propagator.

The identity (2.40) is another example of our tree theorem. It shows, as in the previous example, that the time integrals appearing in the conventional expression for this loop diagram (2.38) can be exchanged for suitable discontinuities of tree-level diagrams. The qualitative difference with (2.33) is that now the loop consists of two lines, and we must sum over cutting either/both of these lines. This need to sum over all possible cuts was also found in the cutting rules previously developed in [106, 107].

The other distinct feature of a loop with more than one edge is that it can be oriented in either of two ways (i.e. loop momenta can flow clockwise or anticlockwise around the loop). This manifests as a second identity for the bulk-to-bulk propagator,

$$\begin{aligned} G_{p_1}(t_1,t_2)G_{p_2}(t_2,t_1) = {} & 2P_{p_1}K_{p_1}(t_2)G_{p_2}(t_2,t_1)\mathrm{Im}(K_{p_1}(t_1)) + 2P_{p_2}K_{p_2}(t_1)G_{p_1}(t_1,t_2)\mathrm{Im}(K_{p_2}(t_2)) \\ & - 4P_{p_1}P_{p_2}\mathrm{Im}(K_{p_1}(t_1))K_{p_2}(t_1)\mathrm{Im}(K_{p_2}(t_2))K_{p_1}(t_2). \end{aligned} \tag{2.41}$$

which corresponds to cutting open the loop with the opposite orientation, producing an identity which differs from (2.40) only in the colouring of the lines,[26]



(2.42)

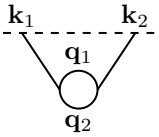
[26] If it seems that (2.42) is a trivial relabelling of (2.40), note that the two internal lines could correspond to different fields.

JHEP12(2023)076

That there are two distinct ways of cutting open a loop into trees will turn out to be quite general, and we show in subsection 2.2.2 below that this is tied to the existence of two causal propagators (i.e. retarded/advanced).

Some comments about (2.40) and (2.42):

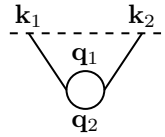
(i) It is instructive to compare with the simple expressions for a massless scalar on Minkowski. The relevant tree-level diagram $\psi_3^{\rm tree}$ was given in (2.10), and from it we constructed the ψ_4 exchange diagram (2.13) using the tree-level cutting rules. Now, we can use both of those expressions as input for (2.40) or (2.42), which determines the loop integrand to be,



$$= \frac{\lambda}{\omega_{k_1}+\omega_{k_2}} \int_{\mathbf{q}_1,\mathbf{q}_2} \left(\frac{1}{\omega_{k_1}+\omega_{k_2}+2\omega_{q_1}} + \frac{1}{\omega_{k_1}+\omega_{k_2}+2\omega_{q_2}} \right) \times \frac{\tilde{\delta}^3(\mathbf{k}_1+\mathbf{q}_1-\mathbf{q}_2)\tilde{\delta}\left(\mathbf{k}_2-\mathbf{q}_1+\mathbf{q}_2\right)}{(\omega_{k_1}+\omega_{q_1}+\omega_{q_2})(\omega_{k_2}+\omega_{q_1}+\omega_{q_2})} \,. \tag{2.43}$$

This is consistent with explicitly performing the nested time integrals in (2.38), but notice that using causality we required only $\psi_3^{\rm tree}$ as an input (i.e. a single time integral over only K_k's).

(ii) To illustrate this result on a non-trivial background (and for a derivative interaction), consider the interaction $\frac{\lambda}{3!}\dot{\pi}^3 a(t)$ on a fixed de Sitter background. This is typically the dominant source of primordial non-Gaussianity in the EFT of inflation [167]. At tree-level, the seed wavefunction coefficient $\psi_3^{\rm tree}$ is given by (2.19), and we used this to construct the four-point exchange diagram in (2.22). Using these as the input for (2.40) or (2.42) gives,

$-$ 

$$= \frac{\lambda^2 H^2}{4} \int_{\mathbf{q}} \left(\frac{3q_1}{k} - \frac{3k^4 q_1}{(k+q_1)^5} + (q_1 \leftrightarrow q_2) \right. \tag{2.44}$$

$$\left. + F(q_1,q_2,q_1+q_2) - F(q_1,q_2,-q_1+q_2) - F(q_1,q_2,q_1-q_2) + F(q_1,q_2,-q_1-q_2) \right)$$

where,

$$F(q_1,q_2,p) = k^4 q_1 q_2 \frac{5(k+p)(k+q_1+q_2)+(k+p)^2+10(k+q_1+q_2)^2}{(k+q_1+q_2)^3(2k+p+q_1+q_2)^5} \tag{2.45}$$

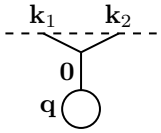
and we have omitted the overall $\tilde{\delta}^3\left(\mathbf{k}_1+\mathbf{k}_2\right)$ and enforced $|\mathbf{k}_1| = |\mathbf{k}_2| = k$ and $\mathbf{q}_2 = \mathbf{q}_1 + \mathbf{k}$. It is remarkable that this result can be obtained immediately from the simple $\psi_3^{\rm tree}$ in (2.19), without performing any further time integration.[27]

[27] Note that this differs from [106, (C.4)] by the additional contact terms in the second line, which are discussed further in appendix D.3.

JHEP12(2023)076

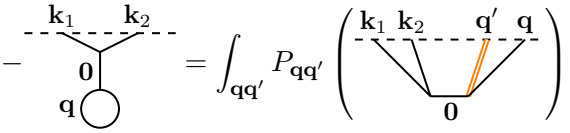
(iii) Note that while the bulk-to-bulk propagators are symmetric in their arguments, $G_p(t_1,t_2) = G_p(t_2,t_1)$, this symmetry is not manifest in the identities (2.39) or (2.41). However, if we consider the difference of these identities, for instance the difference of (2.40) and (2.42), we find that it vanishes thanks to the non-trivial tree-level identity (2.30). The consistency of different one-loop cutting rules is therefore guaranteed by the tree-level identities. This turns out to be a general pattern: the consistency of cutting higher-loop diagrams is guaranteed by lower-loop identities. We describe these consistency relations further in appendix D.2.

Finally, note that at this order in perturbation theory there is also a 1PI-reducible diagram that contributes to ψ_2,



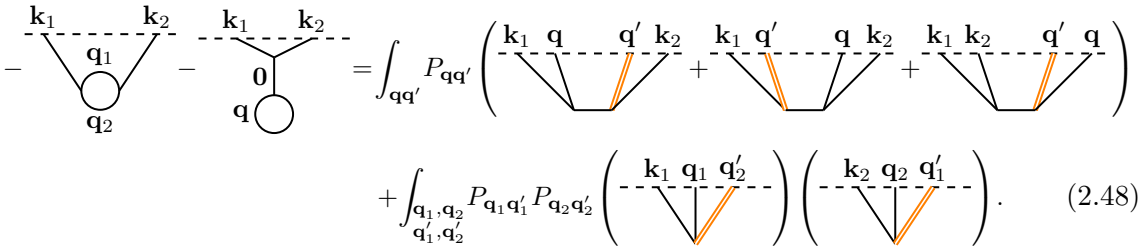
$$= \int_{\mathbf{q}} \int_{t_1,t_2} \lambda_1\lambda_2 K_{k_1}(t_1) K_{k_2}(t_1) G_0(t_1,t_2) G_q(t_2,t_2) \tilde{\delta}^3(\mathbf{k}_1+\mathbf{k}_2)\,. \tag{2.46}$$

This can be expressed in terms of a tree-level diagram as in the one-vertex example using (2.32),



(2.47)

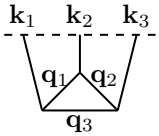
In particular, note that the total one-loop correction sourced by this cubic interaction can be written as,



(2.48)

where the first line now contains all three exchange channels of ψ_4^{tree}. Together with (2.33), this establishes the relation (1.28) given in the introduction, which expresses the full $\psi_2^{1-\text{loop}}$ coefficient in terms of ψ_4^{tree} and ψ_3^{tree}. So although we are deriving these identities by looking at particular diagrams, they naturally apply to the entire wavefunction coefficient (at a fixed order in perturbation theory).

Three-vertex loop. As a final example, we consider the one-loop diagram from three $\frac{\lambda}{3!}\phi^3$ vertices,



$$= \int_{\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3} \tilde{\delta}(\mathbf{k}_1+\mathbf{q}_3-\mathbf{q}_1)\tilde{\delta}(\mathbf{k}_2+\mathbf{q}_1-\mathbf{q}_2)\tilde{\delta}(\mathbf{k}_1+\mathbf{q}_1-\mathbf{q}_2) \tag{2.49}$$

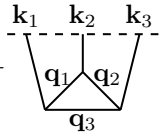
$$\times \int_{\substack{t_1 t_2 \\ t_3}} \lambda^3 K_{k_1}(t_1) G_{q_1}(t_1,t_2) K_{k_2}(t_2) G_{q_2}(t_2,t_3) K_{k_3}(t_3) G_{q_3}(t_3,t_1)\,.$$

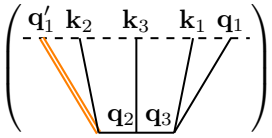
JHEP12(2023)076

While a product of three G_p propagators generally introduces three Θ functions in the integrand, for this particular cyclic identification of the arguments we find that only two Θ functions are necessary. As in the above examples, this means that it is possible to expand this product in terms of fewer bulk-to-bulk propagators. For instance,

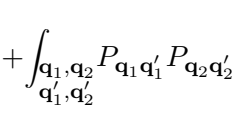
$$\begin{aligned} G_{q_1}(t_1,t_2)G_{q_2}(t_2,t_3)G_{q_3}(t_3,t_1) &= 2P_{q_1}K_{q_1}(t_1)\mathrm{Im}(K_{q_1}(t_2))G_{q_2}(t_2,t_3)G_{q_3}(t_3,t_1) + 2 \text{ perm.} \\ &\quad - 4P_{q_1}P_{q_2}K_{q_1}(t_1)K_{q_2}(t_2)\mathrm{Im}(K_{q_1}(t_2))\mathrm{Im}(K_{q_2}(t_3))G_{q_3}(t_3,t_1) + 2 \text{ perm.} \\ &\quad + 8P_{q_1}P_{q_2}P_{q_3}K_{q_1}(t_1)K_{q_2}(t_2)K_{q_3}(t_3)\mathrm{Im}(K_{q_1}(t_2))\mathrm{Im}(K_{q_2}(t_3))\mathrm{Im}(K_{q_3}(t_1)), \end{aligned} \tag{2.50}$$

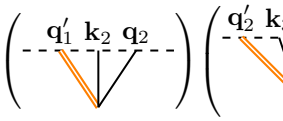
It is straightforward to verify that each of these seven terms corresponds to a particular discontinuity of a tree-level Feynman-Witten diagram, and consequently (2.49) can be written as,

$$-$$ 

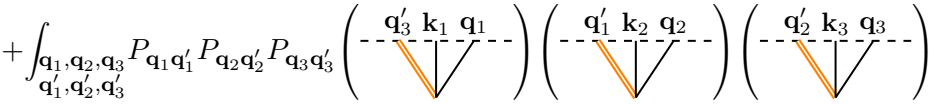
$$= \int_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_1\mathbf{q}_1'}$$ 

$$+2 \text{ perm.}$$

$$+ \int_{\substack{\mathbf{q}_1,\mathbf{q}_2 \\ \mathbf{q}_1',\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}$$ 



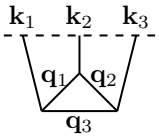
$$+2 \text{ perm.}$$

$$+ \int_{\substack{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3 \\ \mathbf{q}_1',\mathbf{q}_2',\mathbf{q}_3'}} P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}P_{\mathbf{q}_3\mathbf{q}_3'}$$ 

(2.51)

where the permutations indicated correspond to the other ways of cutting the internal lines of the loop. Note that if we fix the orientation of the loop as above, the disc is always taken of the clockwise-most half-edge after the cut. Similarly to the two-vertex loop, the triangle graph (2.49) loop can be oriented either clockwise or anticlockwise and there is therefore a second cutting rule which corresponds to summing over all ways of cutting the loop but instead taking the disc of the other (anticlockwise) half-edges after the cut. The difference between these two cutting rules again vanishes thanks to a collinear tree-level identity which generalises (2.30) to two internal edges.

Returning again to the example of $\frac{\lambda}{3!}\phi^3$ on Minkowski — for which ψ_3^{tree} was given in (2.10) and used to bootstrap both the four- and five-point exchange diagrams, (2.13) and (2.18) — we now see that this input can be used in (2.51) to completely determine the loop momentum integrand for the triangle graph. For that example, (2.51) evaluates to,



$$= \lambda^3 \int_{\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3} \tilde{\delta}(\mathbf{k}_1+\mathbf{q}_3-\mathbf{q}_1)\tilde{\delta}(\mathbf{k}_2+\mathbf{q}_1-\mathbf{q}_2)\tilde{\delta}(\mathbf{k}_1+\mathbf{q}_1-\mathbf{q}_2)$$

$$\times \left(\frac{1}{\omega_T e_1 e_2 e_3}\left(\frac{1}{E_3(k_T+2q_3)}+\frac{1}{E_1(k_T+2q_3)}\right)+2 \text{ perm.}\right)$$

(2.52)

where $\omega_T = \omega_{k_1}+\omega_{k_2}+\omega_{k_3}$ and we have introduced the partial energies,

$$\begin{aligned} e_1 &= \omega_{k_1}+\omega_{q_3}+\omega_{q_1}\,, & e_2 &= \omega_{k_2}+\omega_{q_1}+\omega_{q_2}\,, & e_3 &= \omega_{k_3}+\omega_{q_2}+\omega_{q_3}\,, \\ E_1 &= \omega_{k_2}+\omega_{k_3}+\omega_{q_1}+\omega_{q_3}\,, & E_2 &= \omega_{k_3}+\omega_{k_1}+\omega_{q_2}+\omega_{q_1}\,, & E_3 &= \omega_{k_1}+\omega_{k_2}+\omega_{q_3}+\omega_{q_2}. \end{aligned} \tag{2.53}$$

JHEP12(2023)076

This agrees with the result from old-fashioned perturbation theory [109, 127], or from doing the three nested time integrals in (2.49).

2.2.2 Proof of the general theorem at one loop

Having shown how a suitable expansion of bulk-to-bulk propagator products can be used to express the momentum integrand of some simple one-loop diagrams in terms of cut tree-level diagrams, the natural question is: how general is this phenomenon? In this subsection, we show that in fact *any* one-loop diagram can be decomposed in terms of trees, and we give an explicit expression for the cutting rule that achieves this.

Closed time-like curves. The general proof of our tree theorem rests on causality: namely the free propagation described by the classical G_p^R does not allow for closed time-like curves. Specifically, G_p^R is defined as the Green function for the classical equation motion of the free theory with causal boundary conditions $G_p^R(t_1, t_2) = 0$ if $t_2 > t_1$. For a quadratic Lagrangian like (1.8), such a propagator can be explicitly constructed from the mode functions and is given by (1.22). We reproduce this equation here for convenience,

$$G_p^R(t_1, t_2) = 2P_p \mathrm{Im}\left(K_p(t_1)K_p^*(t_2)\right)\Theta(t_1 - t_2)\,. \tag{2.54}$$

The key observation that translates this causal propagator into cutting rules for loop diagrams comes originally from Feynman [136], who pointed out that a product of N retarded propagators obeys,

$$\prod_{a=1}^{N} G_{p_a}^R(t_a, t_{a+1}) = 0 \quad \text{when} \quad t_{N+1} = t_1\,. \tag{2.55}$$

This is simply a consequence of being unable to order $t_1 > t_2 > \ldots > t_{N+1}$ if the vertices form a closed loop.

Propagator identities. Since we can write the bulk-to-bulk propagator (1.5) in terms of the retarded propagator (1.22), Feynman's identity (2.55) implies that the combinations,

$$\mathcal{L}_{p_1\ldots p_N}(t_1, t_2, \ldots, t_N) \equiv \prod_{a=1}^{N}\left(G_{p_a}(t_a, t_{a+1}) - 2P_{p_a}K_{p_a}(t_a)\mathrm{Im}(K_{p_a}(t_{a+1}))\right)\Big|_{t_{N+1}=t_1} \tag{2.56}$$

all vanish identically. As in [106, 107], and described in appendix A, for every propagator identity that relates different powers of G_p and K_p there is a cutting rule which relates Feynman diagrams with different numbers of internal lines.

Cosmological tree theorem. To prove the tree theorem at one loop, consider the general one-loop diagram composed of N interactions,

$$D = \tilde{\delta}\left(\sum_{a=1}^{N}\mathbf{k}_a\right)\int_{\mathbf{q}_1}\left[\prod_{a=1}^{N}\int_{t_a}\lambda_a K_{k_a}(t_a)G_{q_a}(t_a, t_{a+1})\right] \tag{2.57}$$

where $\mathbf{k}_a$ is the total momentum flowing into each vertex from the boundary, momentum conservation at each vertex requires that $\mathbf{q}_a - \mathbf{q}_{a+1} = \mathbf{k}_b$, and again we identify $t_{N+1} = t_1$

JHEP12(2023)076

and $\mathbf{q}_{N+1} = \mathbf{q}_1$ for notational convenience. λ_a is a vertex factor that can depend on each momenta flowing into the vertex and also the time t_a.

We can use the fact that (2.56) vanishes to write this diagram D as a sum over cuts. More precisely, we "cut" a line by making the replacement,

$$G_{q_a}(t_a, t_{a+1}) \to \int_{\mathbf{q}'_a} P_{\mathbf{q}_a \mathbf{q}'_a} \; K_{q_a}(t_a) K_{q'_a}(t_{a+1}) \tag{2.58}$$

in (2.57), which corresponds diagrammatically to replacing an internal line with two external half-edges. To formalise the sum over cuts, we will denote the set of internal lines that make up the loop by I, and the set of lines which have been cut by C (with sizes $|I|$ and $|C|$ respectively). Cutting the lines C in diagram D produces a new diagram D_C, which may no longer be connected. We use $D_C^{(n)}$ to refer to the connected subdiagrams of D_C. With that notation, the causal identity,

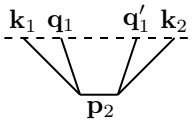
$$\left[\prod_{a=1}^{N} \int_{t_a} K_{k_a}(t_a)\right] \mathcal{L}_{p_1 \ldots p_N}(t_1, \ldots, t_N) = 0 \tag{2.59}$$

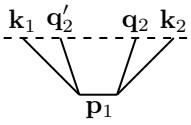
from (2.56) becomes the Cosmological Tree Theorem,

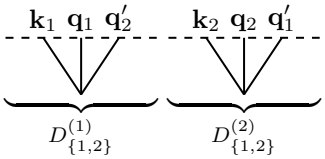
$$-D = \sum_{\substack{C \subseteq I \\ C \neq \{\}}}^{2^{|I|}-1} \left[\prod_{a \in C}^{|C|} \int_{\mathbf{q}_a \, \mathbf{q}'_a} P_{\mathbf{q}_a \mathbf{q}'_a}\right] \prod_n \operatorname*{disc}_{\{q'_a\}} \left[D_C^{(n)}\right] \; . \tag{2.60}$$

Eq. (2.60) is a more precise version of the schematic (1.27), and is the central result of this work.

An example. The previous examples in subsection 2.2.1 follow immediately from the general theorem (2.60) with $|I| = 1, 2$ and 3 internal lines.[28] For instance, for the two-vertex loop, we have two internal lines, labelled by $I = \{1, 2\}$, and therefore 3 cut diagrams,

$D_{\{1\}} =$ 

, $D_{\{2\}} =$ 

, $D_{\{1,2\}} =$ 

.

Note that $D_{\{1\}}$ and $D_{\{2\}}$ are connected, while $D_{\{1,2\}}$ contains the two connected subdiagrams indicated. Taking the discontinuities specified in (2.60) then reproduces the coloured diagrams on the right-hand-side of (2.40).

Further identities. What about the other identities discussed in subsection 2.2.1? Well, note that since (2.60) uses discontinuities with respect to the q'_a only, it is sensitive to how we have oriented the loop. Had we instead written each internal line as $G_{p_a}(t_{a+1}, t_a)$, then

[28] Note that $|I| = 1$ is somewhat subtle because it requires the evaluation of $\Theta(0)$. Given our definitions of $G_p^R(t_1, t_2)$ in (1.22) and $G_p(t, t)$ in (2.32), notice that $G_p^R(t, t) = 0$ and hence the directed loop shown in (1.26) indeed vanishes.

JHEP12(2023)076

applying the cuts as in (2.58) would have resulted in (2.60) with $\underset{\{q_a'\}}{\text{disc}}$ replaced by $\underset{\{q_a\}}{\text{disc}}$ (where the convention is that $\mathbf{q}_a'$ is the clockwise-most half-edge),

$$-D = \sum_{\substack{C\subseteq I\\ C\neq\{\}}}^{2^{|I|}-1} \left[\prod_{a\in C}^{|C|} \int_{\mathbf{q}_a\,\mathbf{q}_a'} P_{\mathbf{q}_a\mathbf{q}_a'}\right] \prod_n \underset{\{q_a\}}{\text{disc}}\left[D_C^{(n)}\right] \ . \tag{2.61}$$

At first sight, this may seem like a trivial re-labelling of (2.60), however it is a non-trivial consequence of the fact that $\mathcal{L}_{p_N\ldots p_1}(t_1, t_N, \ldots, t_{N-1})$ also vanishes, which follows from Feynman's identity (2.55) for the advanced propagator. So the Cosmological Tree Theorem for one-loop graphs is always, in effect, a *pair* of relations, which correspond to cutting the loop open with either orientation.

In fact, comparing the two tree theorems, we see that,

$$\sum_{\substack{C\subseteq I\\ C\neq\{\}}}^{2^{|I|}-1} \left[\prod_{a\in C}^{|C|} \int_{\mathbf{q}_a\,\mathbf{q}_a'} P_{\mathbf{q}_a\mathbf{q}_a'}\right] \left(\prod_n \underset{\{q_a\}}{\text{disc}}\left[D_C^{(n)}\right] - \prod_n \underset{\{q_a'\}}{\text{disc}}\left[D_C^{(n)}\right]\right) = 0 \tag{2.62}$$

These integrands correspond to the tree-level colinear identities discussed at the end of subsection 2.1.

The fact that the tree-level causality conditions guarantee consistency of the different one-loop cutting rules turns out to be a general trend. Although we have presented a general argument for one-loop graphs, it is straightforward to replace $\prod_a K_{k_a}(t_a)$ above with any desired function of bulk-to-boundary and bulk-to-bulk propagators. Eq. (2.60) is therefore a general result which can be applied to any closed loop within a diagram. In appendix D.2 we give some examples at two loops. Since each loop can be oriented either clockwise or anticlockwise, there can be several different ways to cut a two-loop diagram into tree diagrams. The difference between these identities corresponds precisely to the one-loop identities discussed in subsection 2.2.1 above. The general pattern is that the L-loop identities guarantee the consistency of the different ways of cutting open an $L+1$-loop diagram.

Including spin. So far we have focused on the application of the Cosmological Tree Theorem to scalar fields. However, the underlying conditions (unitarity/causality of free propagators) that we have used do not specify the mass or the spin of the field. Indeed, the hermitian analyticity of the bulk-to-bulk propagators of spinning fields was proven in [107]. Our Cosmological Tree Theorem can therefore be applied more widely to fields of any spin.

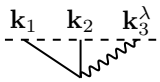
To extend the Cosmological Optical Theorem from scalar to any other field content, all that is required is to decorate the internal lines to account for all of the internal quantum numbers which characterise each field. For instance, the tree theorem representation of the two-point wavefunction coefficient becomes:

$$\begin{aligned}-2\psi^{\alpha_1\alpha_2}_{\mathbf{k}_1\mathbf{k}_2} = &\sum_{\lambda_1\lambda_1'} \int_{\mathbf{q}_1\mathbf{q}_1'} P^{\lambda\lambda'}_{\mathbf{q}_1\mathbf{q}_1'} \underset{q_1'}{\text{disc}}\left[\psi^{\alpha_1\lambda_1\lambda_1'\alpha_2}_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}_1'\mathbf{k}_2}\right] + \sum_{\lambda_2\lambda_2'} \int_{\mathbf{q}_2\mathbf{q}_2'} P^{\lambda_2\lambda_2'}_{\mathbf{q}_2\mathbf{q}_2'} \underset{q_2'}{\text{disc}}\left[\psi^{\alpha_1\lambda_2\lambda_2'\alpha_2}_{\mathbf{k}_1\mathbf{q}_2'\mathbf{q}_2\mathbf{k}_2}\right] \\ &+ \sum_{\substack{\lambda_1\lambda_1'\\ \lambda_2\lambda_2'}} \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P^{\lambda_1\lambda_1'}_{\mathbf{q}_1\mathbf{q}_1'} P^{\lambda_2\lambda_2'}_{\mathbf{q}_2\mathbf{q}_2'} \underset{q_2'}{\text{disc}}\left[\psi^{\alpha_1\lambda_2'\lambda_1}_{\mathbf{k}_1\mathbf{q}_2'\mathbf{q}_1}\right] \underset{q_1'}{\text{disc}}\left[\psi^{\alpha_2\lambda_2\lambda_1'}_{\mathbf{k}_2\mathbf{q}_2\mathbf{q}_1'}\right] \ .\end{aligned} \tag{2.63}$$

JHEP12(2023)076

where the superscript on each wavefunction coefficient indicates the intrinsic quantum numbers of each field (e.g. the spin and helicity).

As a concrete example, consider the interaction, $\mathcal{L}_{\rm int} = -\frac{1}{2}g(t)h^{ij}\partial_i\phi\partial_j\phi$, between the scalar ϕ and a metric fluctuation h_{ij} (in a transverse, traceless gauge). This produces a contact diagram:

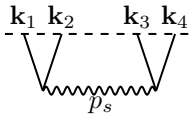


$$= \int_t ig(t)\, V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) K_{k_1}(t) K_{k_2}(t) K^\lambda_{k_3}(t) \tag{2.64}$$

The vertex function $V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ is proportional to the momentum of the scalar lines and the polarisation tensor[29] of the graviton with helicity λ,

$$V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \propto (\mathbf{k}_1)_i (\mathbf{k}_1)_j \epsilon^\lambda(\mathbf{k}_3)_{ij}. \tag{2.65}$$

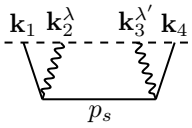
This interaction leads to two different exchange diagrams — one with the exchange of a graviton,



(2.66)

$$= \sum_{\lambda,\lambda'} \int_{t_1 t_2} ig(t_1)g(t_2) V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{p}_s) K_{k_1}(t_1) K_{k_2}(t_1) G^{\lambda\lambda'}_{p_s}(t_1,t_2) K_{k_3}(t_2) K_{k_4}(t_2) V^{\lambda'}(\mathbf{k}_3,\mathbf{k}_4,-\mathbf{p}_s)$$

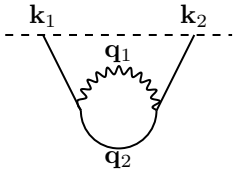
where the λ, λ' indices are summed over, and one with the exchange of a scalar,



(2.67)

$$= \int_{t_1 t_2} ig(t_1)g(t_2) V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{p}_s) K_{k_1}(t_1) K^\lambda_{k_2}(t_1) G_{p_s}(t_1,t_2) K^{\lambda'}_{k_3}(t_2) K_{k_4}(t_2) V^{\lambda'}(\mathbf{k}_3,\mathbf{k}_4,-\mathbf{p}_s)$$

where the indices are not summed (they are external kinematic data). The loop correction to ψ_2 from this interaction contains one graviton and one scalar as internal lines,



$$= \sum_{\lambda,\lambda'} \int_{t_1 t_2} g(t_1)g(t_2) \tag{2.68}$$

$$\times \int_{\mathbf{q}_1\mathbf{q}_2} K_{k_1}(t_1) G^{\lambda\lambda'}_{q_1}(t_1,t_2) G_{q_2}(t_1,t_2) K_{k_2}(t_2) V^\lambda(\mathbf{k}_1,\mathbf{q}_1,\mathbf{q}_2) V^{\lambda'}(\mathbf{k}_2,-\mathbf{q}_1,-\mathbf{q}_2)$$

Since both $G^{\lambda\lambda'}$ and $P^{\lambda\lambda'}$ are $\propto \delta^{\lambda\lambda'}$, we see that indeed (2.63) is satisfied by the above four diagrams.

[29] The polarisation tensors $\epsilon^\lambda(\mathbf{k})_{ij}$ are transverse, symmetric and traceless. They are also conjugate under parity $(\epsilon^\lambda(\mathbf{k})_{ij})^* = \epsilon(-\mathbf{k})^\lambda_{ij}$. This leads to the vertex function $V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ being hermitian analytic, $V^\lambda(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = (V^\lambda(-\mathbf{k}_1,-\mathbf{k}_2,-\mathbf{k}_3))^*$. While not required for the tree theorem, which uses only disc's, this is crucial for other unitarity cutting rules that involve Disc operations.

JHEP12(2023)076

Comparison with OFPT. Finally, note that on a Minkowski background one can use old-fashioned perturbation (OFPT) of the Lippmann-Schwinger equation to similarly express every wavefunction coefficient without the need for explicit time integrals. Let us close this section by making some explicit comparisons between our Cosmological Tree Theorem and the use of OFPT. There are three key differences between the two approaches:

(i) *Different inputs.* OFPT uses only $\psi_1 = 1/\omega_k$ as input, and builds all higher-point coefficients by exploiting the time translation Ward identity of Minkowski. On the other hand, the Cosmological Tree Theorem makes use of all cut diagrams, which can be bootstrapped by unitarity and causality from all contact ψ_n.

(ii) *Different terms.* This is related to the different inputs: OFPT expands a diagram in terms of ψ_1, while the Cosmological Tree Theorem expands in a more varied basis of ψ_n^{contact} so often requires fewer terms overall. This is summarised in table 1. For OFPT, the number of terms required is $E!$, where E being the number of edges.[30] On the other hand, explicitly performing the nested time integrals in what is traditionally called time-orderd perturbation theory (TOPT) requires 3^E terms (since each internal propagator $G_p(t_1, t_2)$ contains the 3 terms shown in (1.5)). The Cosmological Tree Theorem is an improvement over both approaches when it comes to the number of terms, since it decomposes the E edges loop into just $2^E - 1$ terms that we can relate to the discontinuity of different tree-level wavefunction coefficients.

(iii) *Different applicability.* OFPT becomes difficult beyond Minkowski (although see [91] for recent progress), but the Cosmological Tree Theorem can be applied on any time-dependent spacetime background. This makes it well-suited for computing cosmological correlators: for instance we show in section 3.2 below that it leads to a simple expression for the one-loop integrand for the inflationary power spectrum in the EFT of inflation.

3 Correlation functions

Having explored the consequences of causality for the wavefunction coefficients and derived a number of useful identities — including the tree-level cutting rule (2.6), the colinear identity (2.30) and the Cosmological Tree Theorem (2.60) for any loop diagram — we now turn to the question of how these can be used to constrain inflationary observables. In particular, we consider the *equal-time correlation functions* of a weakly coupled scalar field ϕ on a generic time-dependent background. In the context of the EFT of inflation, this background is quasi-de Sitter and ϕ represents the Goldstone of broken time translation: its correlators are closed related to the comoving curvature perturbations which later re-enter the horizon and ultimately seed the initial conditions for the Cosmic Microwave Background and Large Scale Structure [167].

[30] This can be seen by induction: since the OFPT representation of a E edges diagram has E single-cut subdiagrams, that have $E-1$ cut subdiagrams each, that have $E-2$ cut subdiagrams each, and so on... we can follow this procedure until one reaches the single vertex diagram that is only $1/\Omega_k$.

JHEP12(2023)076

	OFPT	TOPT	Cosmological Tree Theorem
$E=1$	1	1	1
$E=2$	2	9	3
$E=3$	6	27	7
$E=4$	24	81	15
$\vdots$	$\vdots$	$\vdots$	$\vdots$
E	$E!$	3^E	2^E-1

Table 1. Number of terms in the different representations of a one-loop diagram with E internal edges.

First, we briefly review the standard map from wavefunction to correlators via the Born rule. Then in subsection 3.1, we show on general grounds (for any interactions and spacetime background) that the Cosmological Tree Theorem implies a delicate cancellation between different wavefunction contributions to cosmological correlators. This is the cosmological analogue of the KLN theorem for amplitudes and cross sections, and in particular it leads to a complete cancellation of certain total energy singularities. Then in subsection 3.2, we provide a simplified expression for the one- and two-point function in any theory of massless fields at one-loop, exploiting this cancellation to remove all loops and redundant terms which vanish in dimensional regularisation. Finally, we specialise to the EFT of inflation and use this expression to evaluate the inflationary power spectrum. We give analogous simplified expressions for the one-loop corrections to primordial non-gaussianity (the bispectrum) in appendix B.

From wavefunction to correlators. From the wavefunctional $\Psi[\phi]$, one can extract any desired equal-time correlation function using the (functional version of the) Born rule,[31]

$$\left\langle \Omega \right| \mathcal{O}\left(\hat{\phi}, i\hat{\Pi}\right) \left| \Omega \right\rangle = \int \mathcal{D}\phi \; \Psi^*[\phi] \; \mathcal{O}\left(\phi, \frac{\delta}{\delta\phi}\right) \Psi[\phi] \tag{3.1}$$

which follows immediately from inserting a complete set of ϕ eigenstates to the left of the operator $\mathcal{O}$. The $\mathcal{D}\phi$ represents a functional integral over all $\phi_{\mathbf{k}}$ modes (this is not a path integral, but rather an integral over all field configurations at a fixed time).

In practice, we evaluate (3.1) by treating any non-Gaussianity in $\Psi[\phi]$ as a small perturbation. Concretely, we expand the wavefunction as in (1.3), namely,

$$\Psi[\phi] = \exp\left(\sum_{n=0}^{\infty} \int_{\mathbf{k}_1\ldots\mathbf{k}_n} \frac{\psi_{\mathbf{k}_1\ldots\mathbf{k}_n}}{n!} \, \phi_{\mathbf{k}_1} \ldots \phi_{\mathbf{k}_n}\right) \tag{3.2}$$

[31] We have implicitly normalised the state / wavefunctional so that $\langle 1 \rangle = 1$. Without this condition, the well-defined observables in this field theory are the ratios $\langle \mathcal{O} \rangle / \langle 1 \rangle$, which introduces an explicit normalisation factor of $1/\int \mathcal{D}\phi |\Psi[\phi]|^2$ on the right-hand-side. Diagrammatically, the role of this normalisation is to cancel all vacuum bubble contributions.

JHEP12(2023)076

where ψ_0 is simply an overall normalisation, and in perturbation theory we treat each $\psi_n \sim \mathcal{O}\left(g_*^{n-2}\right)$, except for the tadpole ψ_1 which $\sim \mathcal{O}\left(g_*\right)$, where g_* is a small power counting parameter.

For instance, suppose we define the power spectrum of the full interacting theory via,

$$\langle\Omega|\ \hat{\phi}_{\mathbf{k}}\hat{\phi}_{\mathbf{k}'}\ |\Omega\rangle = \mathcal{P}_k\ \tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right) \tag{3.3}$$

where we use a caligraphic $\mathcal{P}_k$ to distinguish this from the free-theory power spectrum P_k. A perturbative expansion of the Born rule (3.1) then fixes $\mathcal{P}_k$ in terms of the wavefunction coefficients of the previous section. As is often the case with quadratic correlators (e.g. the propagator), it is simpler to give the perturbative expansion for the inverse,

$$\begin{aligned}-\frac{\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)}{\mathcal{P}_k} = 2\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{k}'} + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\left(\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{k}'\mathbf{q}\mathbf{q}'} + 4\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{k}'\mathbf{q}}\mathrm{Re}\,\psi_{\mathbf{q}'}\right) \\ + 2\int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}\left(\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{q}_1\mathbf{q}_2'}\mathrm{Re}\,\psi_{\mathbf{k}'\mathbf{q}_1'\mathbf{q}_2} + \mathrm{Re}\,\psi_{\mathbf{k}\mathbf{k}'\mathbf{q}_1}\mathrm{Re}\,\psi_{\mathbf{q}_1\mathbf{q}_2\mathbf{q}_2'}\right) + \mathcal{O}\left(g_*^3\right)\end{aligned} \tag{3.4}$$

where we have treated all $\hat{\phi}$ fields as indistinguishable (and hence ψ_n is a symmetric function of its arguments).

Similar expansions can be given for all other n-point functions. The other correlator we consider in the main text will be the one-point function,

$$\frac{\langle\hat{\phi}_{\mathbf{k}}\rangle}{\mathcal{P}_k} = v\,\tilde{\delta}\left(\mathbf{k}\right) \tag{3.5}$$

where we have normalised by $\mathcal{P}_k$ in order to cancel various contributions which correspond diagrammatically to corrections that only affect the propagation of a single leg as it propagates to the boundary (i.e. diagrams of the form (B.2)). The Born rule (3.1) can be used to determine v perturbatively in the non-Gaussianity,

$$v\,\tilde{\delta}^3\left(\mathbf{k}\right) = 2\mathrm{Re}\,\psi_{\mathbf{k}} + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\,\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{q}\mathbf{q}'} + \mathcal{O}\left(g_*^3\right)\ . \tag{3.6}$$

Loop expansion. The rationale for the power counting $\psi_n \sim \mathcal{O}(g_*^{n-2})$ is that each wavefunction coefficient stems from a weakly coupled Lagrangian of the form, $g_*^{n-2}\mathcal{L}\left[g_*\phi\right]$, where $g_* \ll 4\pi$ suppresses Feynman diagrams which contain loops. As a result, in the loop expansion (1.11) of the wavefunction coefficients, we should treat,

$$\psi_n^{L\text{-loop}} \sim \mathcal{O}\left(g_*^{2L+n-2}\right)\ . \tag{3.7}$$

Equations (3.4) and (3.6) then have expansions in g_*, for instance,

$$v = v^{(1)} + \mathcal{O}\left(g_*^3\right)\ , \qquad\qquad \mathcal{P}_k = \mathcal{P}_k^{(0)} + \mathcal{P}_k^{(2)} + \mathcal{O}\left(g_*^4\right) \tag{3.8}$$

where $\mathcal{P}_k^{(0)} = P_k$ is the free-theory power spectrum and $\mathcal{P}_k^{(2)}$ is determined by $\psi_2^{1-\text{loop}}$, ψ_4^{tree} and $\psi_3^{\text{tree}} \times \psi_3^{\text{tree}}$. Note that we have defined the free theory such that $\psi_1^{\text{tree}} = 0$ and hence the leading order $v^{(1)}$ is determined by ψ_3^{tree} and $\psi_1^{1-\text{loop}}$.

JHEP12(2023)076

Notice that each n-point correlator starts at $\mathcal{O}\left(g_*^{n-2}\right)$, so at leading order (i.e. $\mathcal{O}(g_*)$) the only non-zero correlators are the one-loop correction to the vev and the tree-level bispectrum. At next-to-leading order (i.e. $\mathcal{O}(g_*^2)$), the non-zero correlators are the one-loop correction to the power spectrum and the tree-level trispectrum. These are the objects we focus on in the main text. At next-to-next-to-leading order (i.e. $\mathcal{O}(g_*^3)$), the non-zero correlators are the two-loop vev, the one-loop bispectrum and the tree-level five-point function. We describe these in appendix B.

Finally, although we focus predominantly on ϕ correlators, since these give the largest signal at the end of inflation, we have checked that similar conclusions can also be drawn for mixed correlators which also contain the conjugate momenta Π.

3.1 IR singularities and the cosmological KLN theorem

Before analysing any particular n-point function, in this subsection we give a general argument about how the Cosmological Tree Theorem will affect a generic correlator. In short, once our tree theorem replaces loop wavefunction diagrams with tree-level ones, it makes manifest the cancellations that can take place between e.g. $\psi_2^{\text{1-loop}}$, ψ_4^{tree} and $\psi_3^{\text{tree}} \times \psi_3^{\text{tree}}$ in the power spectrum. This generalises some of the important observations made recently in [31] about which wavefunction singularities can appear in correlators.

Analytic structure of wavefunction. Earlier in section 2.1, we made use of the condition that tree-level Feynman-Witten diagrams for the Bunch-Davies initial state can only have singularities when the total energy flowing into one or more vertices vanishes. The idea is that, for a Bunch-Davies initial state, the bulk-to-boundary propagators appearing in any contact diagram have the asymptotic behaviour,[32]

$$K_{k_1}(t) \dots K_{k_n}(t) \sim e^{i\omega_T t} \tag{3.9}$$

at large t. Integrating this from $t = -\infty$ is what produces singularities when $\omega_T \to 0$. Physically, this divergence corresponds to the interaction becoming long-lived: when there is no longer any $e^{i\omega_T t}$ to suppress the vertex in the far past, it gives an infinite contribution to the wavefunction. Since these contact diagrams can be used to construct any tree-level exchange diagram, we similarly conclude that those will have singularities when the energy at any vertex (or collection of vertices) vanishes.[33] This typically occur in regions of parameter space that are not observationally accessible: for instance when some of the ω_k are negative.

Recently, [127] studied the analytic structure of the wavefunction on Minkowski, and used old-fashioned perturbation theory to show that loop integrands similarly can only have singularities when the total energy flowing into one or more vertices vanishes (the "energy conservation condition"). As a result, a simple Landau analysis shows that ψ_n has branch

[32] On Minkowski, a massive field would have $\Omega_k = \sqrt{k^2 + m^2}$ in place of k in this expression. On a quasi-de Sitter background, however, since k blueshifts in the far past any finite mass parameter becomes negligible and the propagators scale like (3.9).

[33] The singularities when a collection of vertex energies vanish arise naturally in the bootstrap procedure of section 2.1 from the matching of D_R onto the unphysical singularities of D_C: for instance the terms $1/E_1(\omega_{k_3} + \omega_q)$ and $1/E_1(\omega_{k_3} + \omega_{k_4} + \omega_{k_5})$ in (2.17) are producing from the single-vertex function E_1 poles which depends on the total energy of two or three vertices.

JHEP12(2023)076

points at certain "thresholds": the minimum energy at which this energy conservation condition can be met.

The Cosmological Tree Theorem can be used to extend this argument to other spacetime backgrounds. Thanks to (2.60), a general loop integrand can be decomposed into tree-level diagrams, which have known singularities (when the total energy flowing into one or more vertices vanishes). The subtlety is that (2.60) involves discontinuities that change the sign of some energies and produce unphysical singularities, which ultimately cancel one another in the final sum. Applying the energy conservation condition of [127] directly to (2.60) will therefore produce a list of *possible* thresholds, but in practice we expect several of them to cancel out. We leave a systematic study of these cancellations for the future. For now, all we need is the idea that the singularities of a diagram arise when the total energy flowing into one or more vertex vanishes. Together with the Cosmological Tree Theorem, this is sufficient to demonstrate the cancellation of certain singular points which takes place for equal-time correlators.

Tree-level singularities. At their first non-trivial order, equal-time correlators are determined simply by tree-level wavefunction coefficients. There are three kinds of singularities which can occur in these objects when a total or partial energy vanishes:

(i) *Poles.* As can be seen in the various examples of section 2, for massless fields on both Minkowski and de Sitter there are generally poles in ψ_n whenever the energy flowing into a connected set of vertices vanishes. On Minkowski, these poles are the finite-time avatar of the energy-conserving δ-functions which appear in asymptotic observables such as the S-matrix (indeed, pushing the wavefunction to $t \to +\infty$ would replace each simple pole with such a δ function). On de Sitter, the order of the pole is fixed by dilation invariance.

(ii) *Late-time logs.* But poles are not the only kind of singularity that can appear at tree-level. Unlike for amplitudes, which are always finite at tree-level (thanks to the energy-conserving δ functions), when it comes to finite-time wavefunction coefficients particular Feynman-Witten diagrams can diverge and require renormalisation even at tree-level. This happens whenever the masses of the fields satisfy certain sum rules, which leads to them persisting at arbitrarily late times (by contrast, the simple poles above correspond to the early-time limit). For instance, for the cubic interaction $\frac{1}{3!}a^4(t)\phi^3$ in which each field has mass $m^2 = 2H^2$ (so that the total conformal dimension $3\Delta = 3$ matches the spacetime dimension), the $\psi_3^{\rm tree}$ coefficient

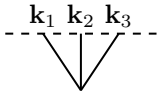
$$= \int_{-\infty}^{t_0} \frac{dt}{t} e^{ik_T t} \sim \log(k_T t_0) \tag{3.10}$$

and diverges as we approach the $t_0 = 0$ conformal boundary. As discussed in [71, 104], these can be renormalised into the Boundary Operator Expansion (in the same way that UV divergences in ordinary QFT are absorbed into the Operator Product Expansion), which effectively replaces $t_0 \to 1/\mu$ with some sliding RG scale. This is essentially the phenomenon of holographic renormalisation (see e.g. [168]), and

JHEP12(2023)076

in the context of de Sitter is perhaps most usefully viewed in terms of dynamical RG [22–24].

(iii) *Massive branch cuts.* The other way in which these simple poles become branch points is for fields with non-zero mass. Even on Minkowski, if we consider the complex k plane for massive fields we would encounter a branch point at $k = -m$ rather than a simple pole due to the square root in $\omega_k = \sqrt{k^2 + m^2}$. On de Sitter, we similarly find branch cuts in the complex k plane whenever massive fields are involved. For instance, consider the interaction $\frac{1}{2}a^4(t)\phi^2\sigma$ between two fields of mass $m^2 = 2H^2$ and one field of mass $m^2 = \frac{9}{4} + \nu^2$. The corresponding $\psi_3^{\rm tree}$ is,[34]

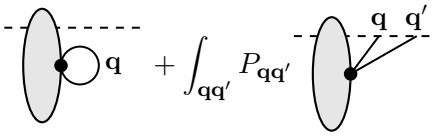


$$\propto \int_{-\infty}^{0} dt\, e^{ik_{12}t}\ (-t)^{-1/2}\ H^{(2)}_{i\nu}(-k_3 t) \propto P_{i\nu-\frac{1}{2}}\left(\frac{k_{12}}{k_3}\right) \tag{3.11}$$

This Legendre function has a *branch point* at $z = k_{12}/k_3 = -1$, which conventionally runs along the negative real axis to $z = -\infty$.

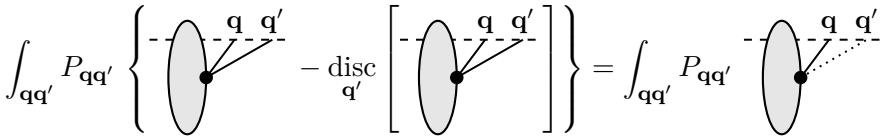
Since we are primarily interested in corrections to inflationary correlators, which are dominated by an approximately massless Goldstone mode, we will mostly focus on the pole singularities (i). However, it is important to mention singularities (ii) and (iii) since these would play an important role in future applications to the cosmological collider signals produced by heavy states during inflation.

One-vertex loop. The first kind of cancellation which takes place due to our tree theorem is for loops containing a single vertex. These always appear together with a partner contribution,



(3.12)

where the grey blob represents any particular set of other ψ diagrams and their complex conjugates. The two terms in (3.12) are the in-in analogue of the *virtual* and *real* emission contributions to the inclusive cross-section in particle physics. By applying the tree theorem, the one-loop virtual emission diagram can be replaced by a disc, so that (3.12) becomes,



(3.13)

where the dotted line represents an analytic continuation of the corresponding energy. Crucially, since a diagram can only contain singularities when the total energy flowing into

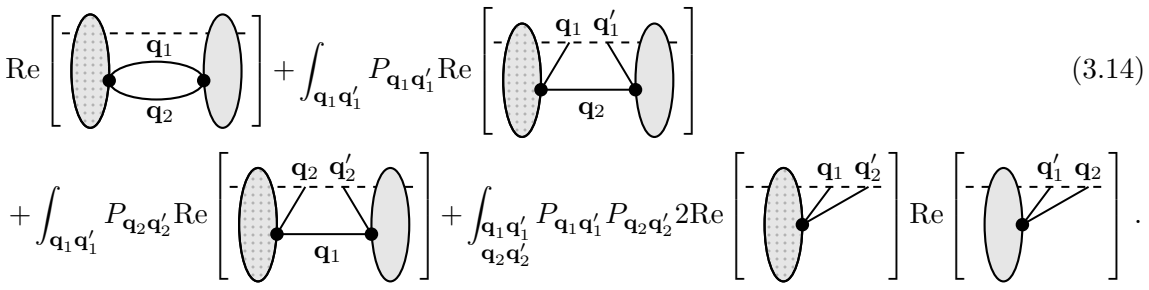
[34] While $\psi_n \sim t_0^\alpha$ for such fields near the boundary, one can extract the finite $\psi_n t_0^{-\alpha}$ by a suitable redefinition of the boundary operators in the theory [104], and this is what is shown in (3.11).

JHEP12(2023)076

a vertex vanishes, we see that this final diagram cannot contain any singularities which depend on both q and $\omega_{\rm blob}$, since the total energy at the shown vertex is $\omega_q - \omega_{q'} + \omega_{\rm blob} = \omega_{\rm blob}$ (where $\omega_{\rm blob}$ is the total energy flowing into the blob from the boundary). As a result, for massless fields this integral takes the simple form $\int_q P_q \, \text{Poly}(q)$ and such ("scale-free") integrals *vanish* in dimensional regularisation. This is perhaps the most striking consequence of the tree theorem: it guarantees that the combination of real and virtual emission shown in (3.13) exactly cancel when evaluated in dim reg. In the analogous calculation of the inclusive cross-section, this cancellation is the so-called KLN theorem. The fact that in-in correlators enjoy a "cosmological KLN theorem", thanks to causality and the tree theorem (2.60), explains why several recent calculations of equal-time correlation functions have found a simpler analytic structure than the corresponding wavefunction coefficients [31, 127].

Note that for massive fields, the cancellation shown in (3.13) still takes place (leaving just a single analytically continued diagram with total energy $\omega_{\rm blob}$), however the resulting integral is no longer scale-free and need not vanish. For instance on Minkowski $2P_q = 1/\sqrt{q^2+m^2}$ and hence gives a finite integral (which $\sim m^{d-1}$). On de Sitter, contact diagrams with two heavy external legs typically $\sim P^j_{i\nu-\frac{1}{2}}(z)\,(1-z^2)^{j/2}$ where $z = -1 + \left((q+q')^2 - \omega_{\rm blob}^2\right)/(2qq')$ and again $z=-1$ corresponds to the branch point: so although $q+q'=0$ in (3.13) (and hence the branch point becomes simply $\omega_{\rm blob}=0$), this function nonetheless produces a non-zero integral over q. However, note that in both of these cases the kind of singularity which emerges from performing the loop integral is *the same* as the singularity in $\omega_{\rm blob}$ already present in the tree-level integrand. This is guaranteed by a simple Landau analysis: the only way to increase the order of the singularity (e.g. to produce dilogs) from a singularity in the integrand is if this integrand has a singular point that mixes $\omega_{\rm blob}$ and q. Since our tree theorem shows this cannot happen, in effect it has shown that the analytic structure of the equal-time correlator remains that of the tree-level wavefunction (even at loop-level).

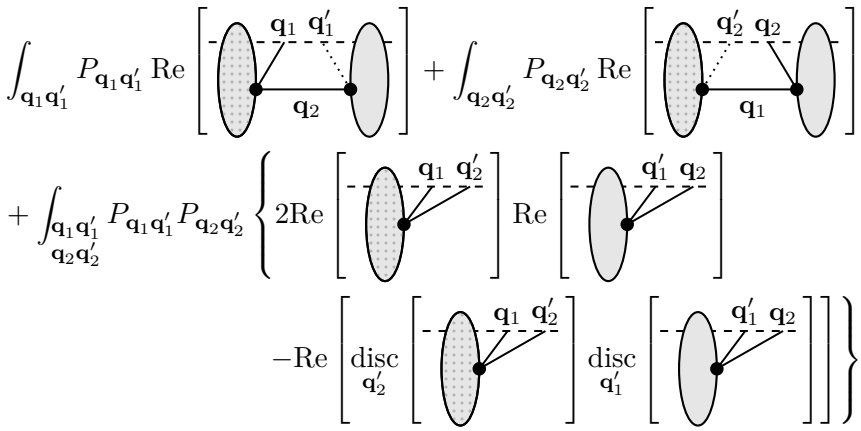
Two-vertex loop. This KLN cancellation is not limited to loops with a single vertex. An analogous cancellation takes place for loops containing two vertices, which always appear together with three other contributions,



(3.14)

As above, each gray blob represents a particular set of ψ_n diagrams (or their conjugates) and we have added a hatched pattern to indicate that there could be different sets attached to the left and right vertices. The first of these terms represents the "virtual" emission, while the latter three terms represent the "real" emission. Again we notice the close parallel

between extracting an in-in correlator from the wavefunction and extracting an inclusive cross-section from an amplitude. Applying our tree theorem replaces the virtual one-loop emission with disc's of real tree-level emission,



(3.15)

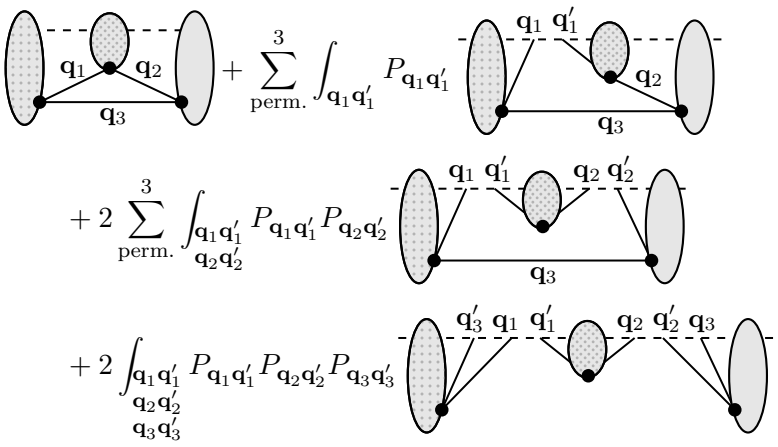
where again the dotted line represents an analytic continuation to negative energy of that particular field. The first line no longer vanishes for massless fields in dim reg, but it does have a simpler singularity structure than the original wavefunction coefficients. Explicitly, suppose that ω_L and ω_R are the total energies flowing into the left and right blobs from the boundary. This integrand can have singularities when the energy at either vertex vanishes, which would lead to singularities in ω_L or ω_R separately. But since the total energy flowing into both vertices is independent[35] of the loop momenta, we find that there is no thershold in $\omega_L + \omega_R$. As a result, the integral on the first line of (3.15) cannot produce any branch cut singularity in the total energy $\omega_L + \omega_R$ if the tree-level diagram contains only poles (and it cannot increase the order of the branch point if the tree-level diagram already contains a branch cut). In the second line there is a partial cancellation between the various disconnected diagrams, but the important observation is that these integrals also cannot produce any singularity in the total $\omega_L + \omega_R$, since each disconnected factor can only have singularities in ω_L or ω_R separately. So overall, we conclude that the virtual and real emissions shown in (3.14) combine in such a way that all branch cut singularities in $\omega_L + \omega_R$ exactly cancel for massless fields. This is a direct consequence of causality and the tree theorem. It also further strengthens the analogy with the KLN cancellation of IR divergences in amplitudes: the KLN theorem often implies that the soft divergences $\sim 1/\omega$ in the amplitude will cancel out between the real and virtual contributions to the inclusive cross-section, rendering the latter an IR-safe observable. Here we similarly find that while the wavefunction may contain higher-order singularities as the total energy $\omega_L + \omega_R \to 0$, these are guaranteed to cancel out when computing an in-in correlator.

Three-vertex loop. This pattern of cancellations continues for loops with arbitrarily many vertices. For instance, a loop with three vertices always comes with real emission

[35] Explicitly, the total energy flowing into both vertices is $\omega_{q_1} - \omega_{q_1'} + \omega_L + \omega_R = \omega_L + \omega_R$.

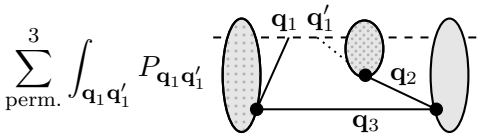
JHEP12(2023)076

diagrams corresponding to the different ways of cutting the loop,



(3.16)

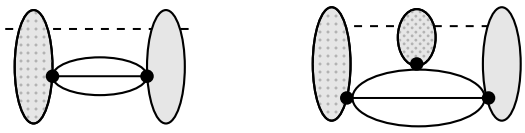
where the three gray blob represent any three particular sets of ψ_n diagrams and their conjugates. Applying the tree theorem, we again find that the connected terms on the first line combine to give,



(3.17)

plus the disc of disconnected diagrams like those on the second line. As before, if ω_L, ω_C and ω_R denote the total energies flowing into the left-, central- and right-blob, we find that the analytic continuation of the energy in (3.17) means that this integral cannot produce branch cut singularities in the total $\omega_L + \omega_C + \omega_R$ that were not already present in the tree-level wavefunction[36] i. It can, however, produce partial energy singularities in any single ω_i or pair $\omega_i + \omega_j$, as can the other disconnected diagrams.

Beyond one loop. Now we briefly discuss higher loop diagrams. In all of the preceding diagrams, the grey blobs may contain further loops: this does not affect our argument and in such cases the total energy singularity still cancels. The qualitative difference when going to higher loops is the presence of overlapping loops such as,



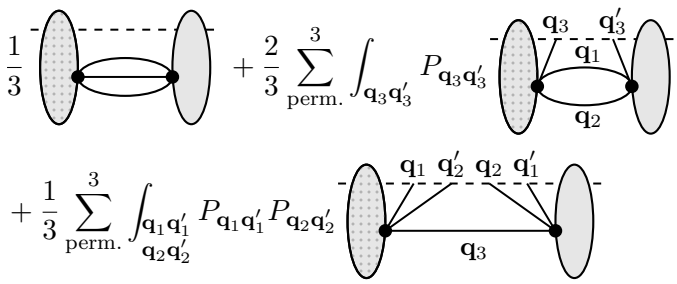
(3.18)

The tree theorem for such two-loop diagrams is described in appendix D.2. We find that the same pattern of KLN cancellations continues for these higher loop corrections. For

[36] The argument is the same as before: only the combination of all three vertices could produce the required threshold, but the total energy flowing into all three vertices is independent of the loop momentum.

JHEP12(2023)076

instance, the diagram on the left always appears with the following terms,



(3.19)

Once we apply the tree theorem the remaining connected terms all take the same tree-level form but with different analytic continuations of the external legs. Once combined, we find that the only terms which survive have two pairs of identified external energies with opposite signs, so that again the total energy flowing into both vertices is independent of the loop momenta. An explicit example of this can be found in appendix B for the two-loop field vev $v^{(3)}$.

Our main conclusion from this general discussion may be stated as:

> **Cosmological KLN theorem (Massive fields):** *Any total-energy branch point produced in the wavefunction by loop integration will not appear in the corresponding equal-time correlation functions.*

In particular, since for massless fields (with derivative interactions) the only tree-level singularities in the wavefunction are poles, we have a stronger corollary in that case,

> **Cosmological KLN theorem (Massless fields):** *Equal-time correlation functions may only have poles in the total energy.*

Our argument above uses unitarity and causality to prove this theorem in perturbation theory for all one-loop diagrams with up to three vertices, and it also holds true in all of the other examples we have checked.

Note although we have focussed on what happens in dimensional regularisation (where various scale-free integrals vanish), this KLN theorem also applies to other regularisation schemes. For instance, with a hard cut-off, integrals of the form $\int_q P_q \,\text{Poly}(q, \Omega_T)$ will not vanish, but rather produce a polynomial $\text{Poly}(\Lambda, \omega_T)$, which is also free of branch cuts at $\omega_T = 0$ for massless fields.

3.2 Some examples

Following our general discussion, we now focus on some specific examples to better illustrate the cancellations and their physical/computational significance.

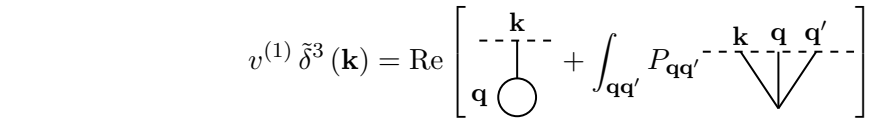
One-loop vev. The simplest example of the above KLN cancellation is for the one-point function (3.5), i.e. the vacuum expectation value (vev) of the field, v. Note that we have defined the free theory such that $v = 0$ in the absence of any interactions (so that physically, ϕ represents the fluctuations around the classical trajectory). At next-to-leading order in g_*, a cubic interaction can produce corrections to v via "tadpole" diagrams. In practice,

JHEP12(2023)076

these should be renormalised away by redefining $\phi \to \phi - v$, which introduces an extra step into the renormalisation process (which can mix e.g. the bispectrum into the power spectrum). Happily, the Cosmological Tree Theorem proves that this is not necessary: at least for massless fields at one loop order.

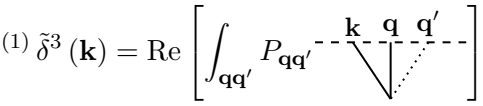
To see this explicitly, consider the loop expansion of (3.6),

$$v^{(1)}\,\tilde{\delta}^3\left(\mathbf{k}\right) = \mathrm{Re}\left[\right. \quad + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \quad \left. \right] \,. \tag{3.20}$$



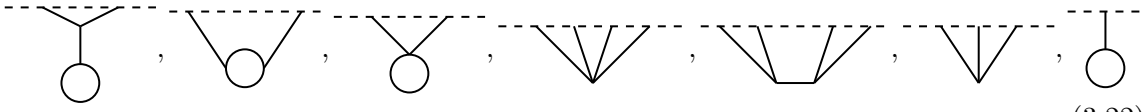
Applying the tree theorem produces simply,

$$v^{(1)}\,\tilde{\delta}^3\left(\mathbf{k}\right) = \mathrm{Re}\left[\int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \right] \,. \tag{3.21}$$



Finally, we invoke the "energy conservation condition" to determine the location of any singular points in this diagram. Since the total energy flowing into this vertex is independent of the loop momenta q, we conclude that the only singular points in the integrand come from P_q. Since these singular points are independent of k, this loop integral does not introduce any kinematic branch points beyond those already present in ψ_3. For instance, for a massless scalar on both Minkowski and de Sitter (1.7), since P_q is simply a power of q this integral *vanishes* in dim reg.

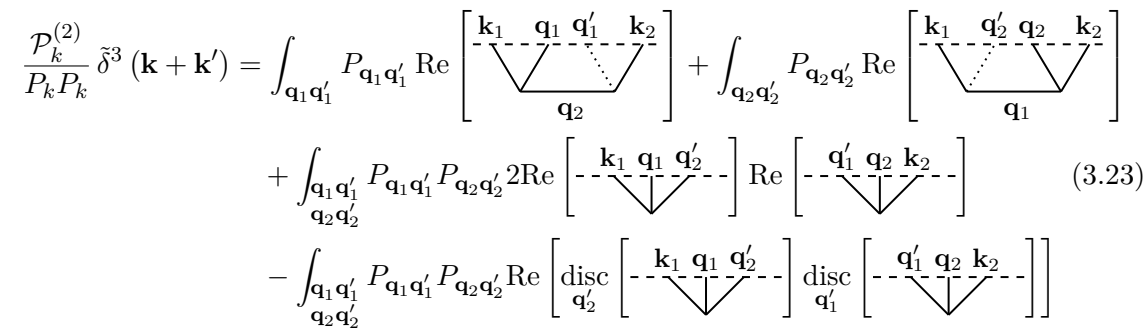
One-loop power spectrum. The next simplest example would be the one-loop correction to the power spectrum, given by expanding (3.4) to $\mathcal{O}\left(g_*^2\right)$. Doing so produces terms built from the following 7 diagrams,



(3.22)

These organise into the combinations (3.12) and (3.14). Applying the tree theorem, we find that all of the (3.12) pairs of terms will vanish in dim reg for a massless scalar, and the (3.14) terms combine to give,

$$\begin{aligned}\frac{\mathcal{P}_k^{(2)}}{P_k P_k}\,\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right) = &\int_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_1\mathbf{q}_1'}\,\mathrm{Re}\left[\right] + \int_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_2\mathbf{q}_2'}\,\mathrm{Re}\left[\right] \\ &+ \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'}\, 2\mathrm{Re}\left[\right]\mathrm{Re}\left[\right] \\ &- \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'}\,\mathrm{Re}\left[\mathrm{disc}_{\mathbf{q}_2'}\left[\right]\mathrm{disc}_{\mathbf{q}_1'}\left[\right]\right]\end{aligned} \tag{3.23}$$



The fact that the one-loop integrand from the power spectrum can be constructed from just two diagrams — the exchange four-point and the contact three-point — is a great

JHEP12(2023)076

simplification compared with the seven diagrams in (3.22). Furthermore, since we have argued that the exchange diagram can be effectively bootstrapped from the contact three-point diagram, in practice the only time integral one needs to compute is that of $\psi_3^{\rm tree}$: from that input alone, one can construct both the tree-level four-point exchange and the entire loop integrand above.

For example, using the results (2.10) and (2.11) for these tree-level diagrams for a ϕ^3 interaction on Minkowski, our simplified equation (3.23) produces,

$$\frac{\mathcal{P}_k^{(2)}}{P_k^2}\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right) = \frac{\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)}{\omega_k+\omega_{k'}}\int_{\mathbf{q}}\frac{\omega_k+\omega_{k'}+\omega_{q_1}+\omega_{q_2}}{2\omega_{q_1}\omega_{q_2}(\omega_k+\omega_{q_1}+\omega_{q_2})(\omega_{k'}+\omega_{q_1}+\omega_{q_2})}\,, \tag{3.24}$$

for a massless scalar, where $\mathbf{q}_2-\mathbf{q}_1=\mathbf{k}$ and $\mathbf{q}$ denotes the remaining loop momentum. We have checked that this agrees with performing a traditional in-in calculation (details can be found in appendix C, cf. (C.12)). Notice that, as anticipated in section 3.1, there are no singularities in this integrand in $\omega_T=\omega_k+\omega_{k'}$ thanks to the analytic continuation of the external lines (beyond the ω_T pole already present in the tree-level wavefunction). However, note also that while the individual exchange terms in (3.23) also introduce singularities at $\omega_{k'}+\omega_{q_2}-\omega_{q_1}$ and $\omega_k-\omega_{q_2}+\omega_{q_1}$, these singular points exactly cancel once combined with the disc$[\psi_3^{\rm tree}]$ terms — there are therefore further cancellations taking place and leading to an even simpler correlator than (3.23) would suggest.[37]

In this simple example, we can see these cancellations explicitly by performing the loop momentum integrals in both (3.24) and its wavefunction counterpart (2.43). These are given explicitly by,

$$\begin{aligned}\frac{\mathcal{P}_k^{(2)}}{P_k^2}\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right) &= \frac{\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)}{8\pi^2(\omega_k+\omega_{k'})}\left[\frac{\omega_{k'}\log(k+\omega_k)-\omega_k\log(k+\omega_{k'})}{\omega_k-\omega_{k'}}+\text{analytic}\right]\,,\\ \psi^{1-\text{loop}}_{\mathbf{k}_1\mathbf{k}_2} &= \frac{\tilde{\delta}^3\left(\mathbf{k}+\mathbf{k}'\right)}{8\pi^2(\omega_k+\omega_{k'})}\left[\frac{\omega_{k'}\log(k+\omega_k)-\omega_k\log(k+\omega_{k'})}{\omega_k-\omega_{k'}}+\text{analytic}\right]\\ &\quad-\frac{1}{16\pi^2 k}\left(\frac{1}{2}\log^2\left(\frac{\omega_k+k}{\omega_{k'}+k}\right)+\text{Li}_2\left(\frac{k-\omega_k}{k+\omega_{k'}}\right)+\text{Li}_2\left(\frac{k-\omega_{k'}}{k+\omega_k}\right)\right).\end{aligned} \tag{3.25}$$

where "analytic" denotes terms which are analytic in both ω_k and $\omega_{k'}$, which we treat as independent variables. As described in [127], the singular points in $\psi_2^{1-\text{loop}}$ in this example are at,

(i) $\omega_k+\omega_{k'}=0$,

(ii) $\omega_k=-k$,

(iii) $\omega_{k'}=-k$.

We see that while the power spectrum has the same three singular points, it has a simpler kind of singularity at each. In particular, (i) is a branch point in ψ_2, but the cancellation

[37] In fact, if this observation were promoted to an assumption about the loop integrands of the Bunch-Davis wavefunction, then one could proceed as in section 2.1 and fix the analytically continued exchange diagrams appearing in (3.23) so that they cancelled the undesired singularities in the disc$[\psi_3^{\rm tree}]$ terms.

JHEP12(2023)076

between real and virtual contributions highlighted above has reduced this to a simple pole in the power spectrum. Furthermore, the branch points at (ii) and (iii) "overlap" in ψ_2, but appear separately in the power spectrum. By overlap, we mean that when we simultaneously approach both singular points, i.e. $\omega_k = -k + \epsilon$ and $\omega_{k'} = -k + \epsilon'$, the power spectrum and wavefunction diverge like,

$$\mathcal{P}^{(2)} \sim \log(\epsilon) + \log(\epsilon')\,, \qquad \psi_2^{1-\text{loop}} \sim \log(\epsilon)\log(\epsilon')\ . \tag{3.26}$$

The separation of overlapping singularities when passing from the wavefunction to the equal-time correlation function also seems quite generic, and suggests that our KLN theorem could be extended in future to also encompass a simplification of other singularities (beyond the total energy branch point).

Finally, once momentum conservation is imposed, the two-point function is special in that $\omega_k = \omega_{k'} = \omega_T/2$ and therefore branch points in either partial energy become indistinguishable from branch points in the total energy. However, one virtue of writing the wavefunction/correlators in terms of general ω_k functions is that it is straightforward to consider the analogous contribution to the 3-point function, in which $\omega_k = \omega_{k_1} + \omega_{k_2}$ and $\omega_{k'} = \omega_{k_3}$ are no longer tied together by momentum conservation. So while our KLN theorem above should be applied carefully when considering the on-shell two-point function, it applies unambiguously to all higher-point correlators.

EFT of inflation. As our last example, we consider the effective field theory of inflation. The leading correction to the power spectrum comes from a loop of two $\dot{\pi}^3$ interactions, and the tree-level wavefunction coefficients from this interaction on a quasi-de Sitter inflationary spacetime were given in (2.19) and (2.22). From these, our simplified expression (3.23) immediately gives the loop integrand,

$$\frac{\mathcal{P}_k^{(2)}}{P_k^2} = \int_{\mathbf{q}} \frac{k_1^2 k_2^2}{k_T^5}\left[-12q_T + q_1 q_2\left(\frac{2k_T^2(k_T+q_T)(k_T^2+2k_Tq_T+4q_T^2)}{e_L^3 e_R^3} + \frac{6k_Tq_T(k_T+2q_T)}{e_L^2 e_R^2} + \frac{12q_T}{e_L e_R}\right)\right]$$

where we have adopted the shorthands,

$$e_L = k + q_1 + q_2\ , \qquad e_R = k' + q_1 + q_2\ , \qquad k_T = k + k'\ , \qquad q_T = q_1 + q_2\ , \tag{3.27}$$

Again we see that the only singular points depend on k and k' individually, and that the apparent singularities at $k' + q_2 - q_1$ and $k - q_2 + q_1$ in (3.23) have cancelled out. Performing the remaining momentum integral (using e.g. the identities in [127]) reproduces the result of [13] from a traditional in-in calculation. In fact, we find that complete off-shell expression (treating $\omega_{k_1}, \omega_{k_2}$ and $k = |\mathbf{k}_1| = |\mathbf{k}_2|$ as independent) is given by,

$$\frac{\mathcal{P}_k^{(2)}}{P_k^2} = -\frac{\pi\text{Poly}_1(k,\omega_{k_1},\omega_{k_2})}{40(\omega_{k_1}-\omega_{k_2})^4(\omega_{k_1}+\omega_{k_2})^3} + \frac{\pi\text{Poly}_2(k,\omega_{k_1},\omega_{k_2})\log\left(\frac{\omega_{k_1}}{\Lambda}\right) - (\omega_{k_1} \leftrightarrow \omega_{k_2})}{120(\omega_{k_1}^2-\omega_{k_2}^2)^5} \tag{3.28}$$

JHEP12(2023)076

where:

$$\begin{aligned}\text{Poly}_1(k,\omega_{k_1},\omega_{k_2}) &= 10k^2\omega_{k_1}^2\omega_{k_2}^2(\omega_{k_1}+\omega_{k_2})^4 - 5\omega_{k_1}^3\omega_{k_2}^3(3\omega_{k_1}^4+4\omega_{k_1}^3\omega_{k_2}+34\omega_{k_1}^2\omega_{k_2}^2+4\omega_{k_1}\omega_{k_2}^3+3\omega_{k_2}^4)\\ &\quad + k^4(\omega_{k_1}^6-\omega_{k_1}^5\omega_{k_2}-21\omega_{k_1}^4\omega_{k_2}^2-6\omega_{k_1}^3\omega_{k_2}^3-21\omega_{k_1}^2\omega_{k_2}^4-\omega_{k_1}\omega_{k_2}^5+\omega_{k_2}^6) \qquad (3.29)\\ \text{Poly}_2(k,\omega_{k_1},\omega_{k_2}) &= 4\omega_{k_1}^2\omega_{k_2}^3(9k^4(5\omega_{k_1}^4+10\omega_{k_1}^2\omega_{k_2}^2+\omega_{k_2}^4)+45\omega_{k_1}^2\omega_{k_2}^2(5\omega_{k_1}^4+10\omega_{k_1}^2\omega_{k_2}^2+\omega_{k_2}^4)\\ &\quad - 5k^2(9\omega_{k_1}^6+55\omega_{k_1}^4\omega_{k_2}^2+31\omega_{k_1}^2\omega_{k_2}^4+\omega_{k_2}^6)) \qquad (3.30)\end{aligned}$$

Indeed, we find that the power spectrum contains only partial energy branch cuts in ω_{k_1} and ω_{k_2}, unlike the wavefunction coefficients $\psi_2^{1-\text{loop}}$, which contains a dilogarithmic branch cut in $\omega_{k_1}+\omega_{k_2}$.

Overall, the Cosmological Tree Theorem (2.60) for wavefunction coefficients can be used to replace "virtual" (loop) contributions with additional "real" (tree) contributions when computing equal-time correlators, and this makes manifest the various cancellations which can take place in a scheme such as dim reg. This provides a simpler way to compute the loop corrections to observable cosmological correlators from the wavefunction of the Universe.

4 Discussion

To sum up, we have shown how non-relativistic causality — that free theory propagation can be described by a retarded Green's function — can be used to place perturbative constraints on both the cosmological wavefunction and cosmological correlators. When combined with recent cutting rule from unitarity and an analytic structure mandated by the Bunch-Davies initial condition, this provides a new way to bootstrap tree-level exchange diagrams from their simpler contact building blocks. At loop-level, causality and unitarity naturally lead to a cosmological analogue of Feynman's tree theorem which can replace any closed loop in a Feynman-Witten diagram with a sum over cut diagrams. This Cosmological Tree Theorem fixes the whole loop integrand of wavefunction coefficient, complementing the recent unitarity cutting rules which fix only the discontinuity, and can be applied to fields of any mass or spin with any unitarity interaction on any time-dependent spacetime background. We have given several explicit examples of these constraints at both tree- and loop-level for both Minkowski and de Sitter backgrounds. Applying the Cosmological Tree Theorem to cosmological correlators leads to various cancellations between real and virtual contributions which closely parallels the KLN theorem from scattering amplitudes. We have therefore named this phenomenon the Cosmological KLN theorem, and have shown that it leads to a greatly simplified expression for the one-loop power spectrum in terms of just two tree-level Feynman-Witten diagrams. In particular, the loop integration may not introduce any additional total energy singularities, and so the loop-level correlators have the same analytic structure in ω_T as the tree-level wavefunction.

Thanks to these results, *any* Feynman diagram (with an arbitrary number of edges and loops) can now be expressed in terms of the tree-level single-vertex Feynman diagrams of the theory. Once these have been determined (e.g. by performing a single time integral), there is no need for any further time integration. For Minkowski scattering amplitudes, energy conservation removes the need to do any time integration — here, we are showing that

JHEP12(2023)076

for the Bunch-Davies wavefunction on an arbitrary time-dependent background, unitarity and causality remove the need to do all but one time integration. It would be interesting to combine these results with other bootstrap techniques for the single-vertex diagrams, which would remove the need for any time integration whatsoever.

There are a number of interesting directions to be explored in the future:

(i) *Perturbative vs. non-perturbative.* Wavefunction identities can be organised according to whether they hold:

   (a) for the full ψ_n coefficients of the interacting theory (i.e. sum over all Feynman diagrams with n external legs),

   (b) for the $\psi_n^{L\text{-loop}}$ coefficients at a fixed order in the interactions (i.e. sum over all Feynman diagrams with n external legs and L loops)

   (c) for individual Feynman diagrams.

   At present, very few relations of type (a) are known. The tree theorem presented in this work, as well as the earlier cutting rules from perturbative unitarity in [106], are of type (b). Further cutting rules, which involve any analytic continuation of the internal line energies, are of type (c). It is always natural to ask, therefore, which identities might be promoted to a more general (non-perturbative) type. One way in which the Cosmological Tree Theorem might be extended to type (a) would be to consider the causality constraints on the propagator of the full interacting theory: for instance using an expansion like (A.2) developed in the appendix.

(ii) *Landau analysis.* As mentioned briefly in section 3.1, the Cosmological Tree Theorem can provide new insights into the analytic structure of wavefunction coefficients in curved spacetime. In essence, the analytic structure of loop diagrams should now be determinable by the analytic structure of the tree-level diagrams (which is comparatively much easier to determine). In particular, this will mean that the branch points of loop diagram are determined by the poles of its tree-level cut diagrams, analogous to what was recently found for Minkowski [127].

(iii) *UV/IR sum rules.* This tree theorem is therefore an important step towards UV/IR sum rules for cosmological spacetimes. At present, such relations are mostly limited to subhorizon scattering amplitudes [124, 125, 169], which share many properties of the Minkowski amplitude but are fundamentally disconnected from the horizon-scale physics that we ultimately observe in the CMB (although see [126] for important recent progress towards positivity bounds directly on correlation functions). Building on these recent subhorizon applications of unitarity and causality [170–176], it would be interesting to see the wavefunction emerge as a new object which shares enough similarly with the Minkowski amplitude that it admits usable UV/IR relations and yet remains firmly connected to the cosmological correlators that we actually measure.

(iv) *Stronger causality conditions.* On Minkowksi, the free propagators are also constrained by the future light-cone, i.e. $t_1 - t_2 > |\mathbf{x}_1 - \mathbf{x}_2|$, which is stronger than the

JHEP12(2023)076

$t_1 - t_2 > 0$ condition used in this work. For amplitudes, this ultimately corresponds to analyticity in the Lorentz-invariant p^2 rather than the energy, and this is what underpins the Kallen-Lehmann spectral representation and other relativistic dispersion relations. It would be interesting to explore whether such a stronger condition, and corresponding spectral representation, can exists for wavefunction coefficients on a cosmological background.

Acknowledgments

We thank Harry Goodhew, Mang Hei Gordon Lee, Prahar Mitra, Enrico Pajer and David Stefanyszyn for useful discussions. S.M. is supported by a UKRI Stephen Hawking Fellowship (EP/T017481/1). S.A.S. is supported by a Harding Distinguished Postgraduate Scholarship. This work has been partially supported by STFC consolidated grant ST/T000694/1.

Note added. In the final stages of preparing this manuscript, [115] appeared on the arXiv, which discusses cutting rules from a different (polytope) perspective. It would be interesting to investigate how the causality conditions and loop-level cutting rules presented here could be recovered from the optical polytope of [115].

A Comparison with previous cutting rules

In this appendix, we give an short overview of the various cutting rules which follow from unitarity (along the lines of [106, 107]) and from causality (this work).

General procedure. Throughout this work, we have introduced a number of different "cutting rules" for the wavefunction coefficients: equations which relate diagrams with different internal and external lines. All of these rules can be derived following the same three-step procedure,

(i) Identify a combination of the propagators G_p and K_k (and their complex conjugates) which vanishes thanks to unitarity or causality properties of the free theory,

(ii) Multiply this combination by any further function of G_p, K_k and vertex factors, and then integrate over all time arguments,

(iii) Relate each term in the resulting identity to a Feynman-Witten diagram, invoking the unitarity of the interacting theory (hermiticity of $\mathcal{H}_{\rm int}$) so that the discontinuity operations may be used to convert between G_p, K_k and G_p^*, K_k^* thanks to (1.13).

Free propagation. Step (i) of this procedure requires finding combinations of the propagators that vanish identically. To achieve this, we introduce a useful representation for $G_p(t_1, t_2)$,

$$G_p(t_1, t_2) = \frac{i}{2}\left(\Delta_p^H(t_1, t_2) + i\Delta_p^S(t_1, t_2)\text{sign}(t_1 - t_2)\right) \tag{A.1}$$

JHEP12(2023)076

where we have split the two-point function into its symmetric/antisymmetric parts,

$$\Delta_k^H(t_1,t_2) = \langle \phi = 0 | \{\phi_{\mathbf{k}}(t_1), \phi_{\mathbf{k}'}(t_2)\} | \Omega \rangle' \,, \quad \Delta_k^S(t_1,t_2) = \langle \phi = 0 | [\phi_{\mathbf{k}}(t_1), \phi_{\mathbf{k}'}(t_2)] | \Omega \rangle' \,. \tag{A.2}$$

where the prime denotes that a momentum-conserving $\tilde{\delta}$ function has been removed. These are the wavefunction analogues of the Hadamard and Schwinger functions. In particular, the Schwinger function Δ_k^S is closely related to the classical response of the field to applied sources (more on this below), and in fact for the free theory it is insensitive to the choice of in- and out- state thanks to the canonical commutation relations.[38] The Hadamard function Δ_k^H, on the other hand, characterises quantum aspects of the propagation and depends on the in- and out-state.[39] Unitarity and causality of the free theory then imply relations between Δ_k^H, Δ_k^S and the bulk-to-boundary propagator K_k, which lead to various cutting rules.

Tree-level cutting rules. For example, the tree-level cutting rules of [106, 107] follow from two particular properties of these Hadamard/Schwinger functions,

$$\mathrm{Im}\, \Delta_k^S(t_1,t_2) = 0, \tag{U_1}$$
$$\mathrm{Re}\, \Delta_k^H(t_1,t_2) = 4P_k \mathrm{Im} K_k(t_1) \mathrm{Im} K_k(t_2) \tag{U_2}$$

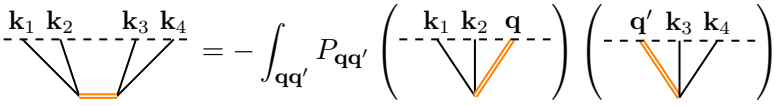
Eq. (U_1) guarantees that the imaginary part of any product of N bulk-to-bulk propagators will contain at most $N-1$ factors of $\mathrm{sign}(t_1 - t_2)$, and can therefore be re-expressed in terms of products of at most $N-1$ bulk-to-bulk propagators. Eq. (U_2) then allows us to rewrite these reduced products in terms of G_p and K_k only. For instance, the combination,

$$\mathcal{I}_p(t_1,t_2) \equiv \mathrm{Im}\, G_p(t_1,t_2) - 2P_p \mathrm{Im}\, K_p(t_1) \mathrm{Im}\, K_p(t_2) \tag{A.4}$$

vanishes identically once (U_1) and (U_2) are imposed. Integrating this over time leads to a cutting rule for any diagram containing at least one internal line: for instance,

$$\int_{t_L,t_R} K_{k_1}(t_L) K_{k_2}(t_L) K_{k_3}(t_R) K_{k_4}(t_R) \mathcal{I}_{p_s}(t_L,t_R) = 0 \tag{A.5}$$

corresponds to the diagrammatic identity,[40]



(A.6)

[38] This Δ_k^S is therefore equal to the usual Schwinger function for the vacuum state, $\tilde{\Delta}_k^S(t_1,t_2) = \langle \Omega | [\phi_{\mathbf{k}}(t_1), \phi_{\mathbf{k}'}(t_2)] | \Omega \rangle'$.

[39] For instance, the usual Hadamard function for the vacuum state,

$$\tilde{\Delta}_k^H(t_1,t_2) = \langle \Omega | \{\phi_{\mathbf{k}}(t_1), \phi_{\mathbf{k}'}(t_2)\} | \Omega \rangle' = \Delta_k^H(t_1,t_2) + 2P_k K_k(t_1) K_k(t_2) \,, \tag{A.3}$$

differs from (A.2) by a boundary term, and would give the usual Feynman propagator if used in (A.1) in place of Δ_p^H.

[40] Note that the momentum integral is trivially performed using the δ function in $P_{\mathbf{qq}'}$. Introducing separate $\mathbf{q}$ and $\mathbf{q}'$ labels for the cut line is simply a useful book-keeping device: particularly for later identities in which we will perform further Disc operations on either the left- or right-hand side of the cut.

JHEP12(2023)076

As a second example, the combination

$$\begin{aligned}\mathcal{I}_{p_1p_2}(t_1,t_2,t_3) \equiv & \,\mathrm{Im}\left(G_{p_1}(t_1,t_2)G_{p_2}(t_2,t_3)\right)+4P_{p_1}P_{p_2}\mathrm{Im}(K_{p_1}(t_1))\mathrm{Im}(K_{p_1}(t_2)K_{p_2}(t_2))\mathrm{Im}(K_{p_2}(t_3)) \\ & -2P_{p_1}\mathrm{Im}(K_{p_1}(t_1))\mathrm{Im}(K_{p_1}(t_2)G_{p_2}(t_2,t_3))-2P_{p_2}\mathrm{Im}(G_{p_1}(t_1,t_2)K_{p_2}(t_2))\mathrm{Im}(K_{p_2}(t_3))\end{aligned}$$

also vanishes thanks to (U_1) and (U_2), and leads to cutting rules for any diagram containing at least two internal lines. Similar cutting rules for an arbitrary tree-level diagram were systematically developed in [106, 107], and also follow from (U_1) and (U_2).

Loop-level cutting rules. By contrast, the loop-level cutting rules of [106] follow instead from exploiting the full two-point function,

$$\langle\phi=0|\,\hat{\phi}_{\mathbf{k}}(t_1)\hat{\phi}_{\mathbf{k}'}(t_2)|\Omega\rangle' = 2P_kK_k(t_2)\mathrm{Im}K_k(t_1) \tag{U}$$

as well as the identity,

$$\sum_{\{n_j\}}\prod_{j=1}^{N}\mathrm{sign}(t_j-t_{j+1})^{n_j}=0 \qquad \left(\text{sum over all } n_j=0 \text{ or } 1 \text{ such that } N-\sum_j n_j \text{ is even}\right), \tag{C_1}$$

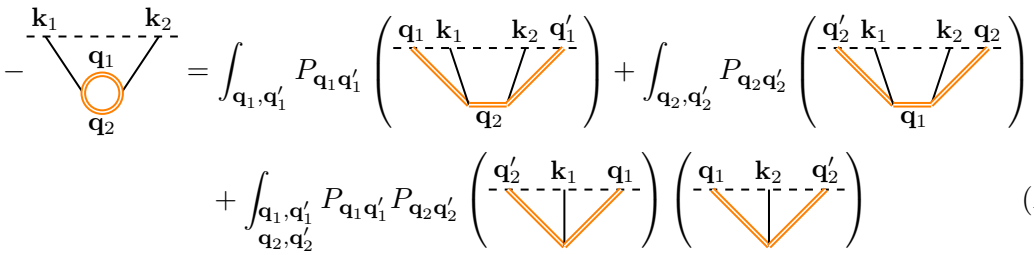
which holds whenever $t_{N+1}=t_1$. Note that (U) fully specifies both Δ_k^H and Δ_k^S (and hence implies both (U_1) and (U_2)). The identity (C_1) is crucial because it guarantees that a product of N bulk-to-bulk propagators whose arguments form a closed loop will contain at most $N-1$ factors of $\mathrm{sign}(t_1-t_2)$, and hence can be re-expressed in terms of products of at most $N-1$ bulk-to-bulk propagators. For instance, the combination,

$$\begin{aligned}\mathcal{R}_{p_1p_2}(t_1,t_2) \equiv & \,2\mathrm{Re}\left(G_p(t_1,t_2)G_{p_2}(t_2,t_1)\right)+4P_{p_1}P_{p_2}\mathrm{Im}(K_{p_1}(t_1)K_{p_2}(t_1))\mathrm{Im}(K_{p_2}(t_2)K_{p_1}(t_2)) \\ & -2P_{p_1}\mathrm{Im}(K_{p_1}(t_2)G_{p_2}(t_2,t_1)K_{p_1}(t_1))-2P_{p_2}\mathrm{Im}(K_{p_2}(t_1)G_{p_1}(t_1,t_2)K_{p_2}(t_2))\end{aligned} \tag{A.7}$$

vanishes identically once (U) and (C_1) are imposed (the latter in this case is $\mathrm{sign}(t_1-t_2)\mathrm{sign}(t_2-t_1)=-1$). Integrating this over time leads to a cutting rule for any diagram containing a loop with two internal lines: for instance,

$$\int_{t_1,t_2} K_{k_1}(t_1)K_{k_2}(t_2)\mathcal{R}_{p_1p_2}(t_1,t_2)=0 \tag{A.8}$$

corresponds to the diagrammatic identity,



(A.9)

Analogous cutting rules for the Disc of an arbitrary loop diagram were developed in [106], and all follow from (U) and (C_1).[41]

[41] Stated this way, it is clear why additional cutting rules can exist for loop-diagrams but not for trees: because the identity (C_1) only holds when the time arguments form a closed loop.

JHEP12(2023)076

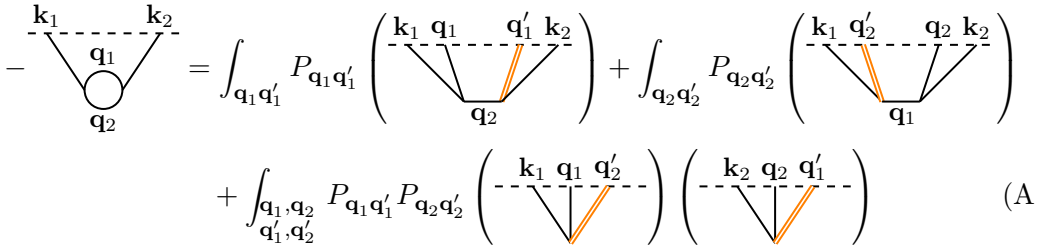
New cutting rules. Since both the Im and Re part of the bulk-to-bulk propagators appearing in a loop can be reduced to terms with fewer bulk-to-bulk propagators, the entire complex loop diagram (without any discontinuity) must be expressible in terms of cut tree-level diagrams. This could be achieved by straightforwardly applying the above identities: for instance $\mathcal{R}_{p_1p_2}(t_1,t_2)+\mathcal{I}_{p_1p_2}(t_1,t_2,t_1)=0$ could be used to cut any $G_{p_1}(t_1,t_2)G_{p_2}(t_2,t_1)$ loop. However, we will show that a more elegant formulation of the loop-level cutting rules follows from the identity,

$$\prod_{j=1}^{N}\Theta(t_j-t_{j+1})=0\,, \tag{C}$$

which holds whenever $t_{N+1}=t_1$. For instance, the combination,

$$\begin{aligned}\mathcal{L}_{p_1p_2}(t_1,t_2)\equiv{}&G_{p_1}(t_1,t_2)G_{p_2}(t_2,t_1)+4P_{p_1}P_{p_2}K_{p_1}(t_1)\mathrm{Im}(K_{p_2}(t_1))K_{p_2}(t_2)\mathrm{Im}(K_{p_1}(t_2))\\&-2P_{p_1}\mathrm{Im}(K_{p_1}(t_2))G_{p_2}(t_2,t_1)K_{p_1}(t_1)-2P_{p_2}\mathrm{Im}(K_{p_2}(t_1))G_{p_1}(t_1,t_2)K_{p_2}(t_2)\,,\end{aligned} \tag{A.10}$$

vanishes identically thanks to (U) and (C). This leads to a diagrammatic cutting rule,



$$-\quad=\int_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_1\mathbf{q}_1'}\quad+\int_{\mathbf{q}_2\mathbf{q}_2'}P_{\mathbf{q}_2\mathbf{q}_2'}\quad+\int_{\substack{\mathbf{q}_1,\mathbf{q}_2\\ \mathbf{q}_1',\mathbf{q}_2'}}P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}\quad \tag{A.11}$$

which goes beyond (A.9) because it captures both the real and imaginary part of the loop diagram.

Note that (C) has assumed a particular orientation for the loop. The other orientation (again with $t_{N+1}=t_1$),

$$\prod_{j=1}^{N}\Theta(t_{j+1}-t_j)=0\,, \tag{C'}$$

gives rise to a second identity ($\mathcal{L}_{p_2p_1}(t_1,t_2)=0$ in the preceding example), and hence a second cutting rule, which differs from the first only in which momenta are held fixed in the disc's.

In fact, since $\Theta(t_1-t_2)=\frac{1}{2}\left(1+\mathrm{sign}(t_1-t_2)\right)$, the identities ($C$) and ($C'$) are equivalent to ($C_1$) and a further identity,

$$\sum_{\{n_j\}}\prod_{j=1}^{N}\mathrm{sign}(t_j-t_{j+1})^{n_j}=0\qquad\left(\text{sum over all } n_j=0 \text{ or } 1 \text{ such that } N-\sum_j n_j \text{ is odd}\right)\,. \tag{C_2}$$

This leads to one further set of cutting rules, which are quite distinct from all previous examples in that they combine the same diagram with different collinear kinematics in order to achieve a cut. For instance, the combination,

$$\begin{aligned}&\mathcal{C}_{p_1p_2}(t_1,t_2) \qquad\qquad (A.12)\\ &\equiv \mathrm{Re}G_{p_1}(t_1,t_2)\mathrm{Im}\Big(K_{p_2}(t_2)K^*_{p_2}(t_1)\Big)-2\mathrm{Im}K_{p_1}(t_1)\mathrm{Im}K_{p_2}(t_2)\mathrm{Re}K_{p_1}(t_2)\mathrm{Re}K_{p_2}(t_1)-(p_1\leftrightarrow p_2).\end{aligned}$$

vanishes once (U) and (C_2) are imposed (the latter in this case is $\mathrm{sign}(t_1-t_2)+\mathrm{sign}(t_2-t_1)=0$). When integrated against two bulk-to-boundary propagators as in (A.5), this leads to the diagrammatic relation shown in (2.30). So while unitarity can fix only the Disc of tree-level diagrams, causality imposes additional constraints on their Disc-less part.

Unitarity and causality. Property (U) (and hence also (U_1) and (U_2)) can be viewed as *unitarity* of the free theory, since Hermiticity of the free Hamiltonian allows us to resolve the identity using a complete basis of n-particle states,

$$\sum_n \langle\phi=0|\hat{\phi}_{\mathbf{k}}(t)|n\rangle\langle n|\hat{\phi}_{\mathbf{k}'}(t')|\Omega\rangle = \int_{\mathbf{p}}\langle\phi=0|\hat{\phi}_{\mathbf{k}}(t)\hat{a}^\dagger_{\mathbf{p}}|\Omega\rangle\langle\Omega|\hat{a}_{\mathbf{p}}\hat{\phi}_{\mathbf{k}'}(t')|\Omega\rangle\ , \qquad (A.13)$$

since in the free theory $\hat{\phi}_{\mathbf{k}}|\Omega\rangle$ overlaps only with the 1-particle states. The matrix elements on the right-hand-side evaluate to $\sqrt{P_k}\mathrm{Im}K_k(t)$ and $\sqrt{P_{k'}}K_{k'}(t')$ (up to an overall unimportant phase), hence producing (U).[42]

The identities (C) and (C'), on the other hand, are related to *causality*. Causality in the classical sense: signals may not precede their sources. This amounts to the existence of a retarded Green's function,

$$G^R_p(t_1,t_2) = G_p(t_1,t_2) - \langle\phi=0|\hat{\phi}_{\mathbf{p}}(t_1)\hat{\phi}_{\mathbf{p}_2}(t_2)|\Omega\rangle' = \Delta^S_p(t_1,t_2)\Theta(t_1-t_2)\ . \qquad (A.14)$$

which would describe the classical response of the field to an applied source, and for which identity (C) implies,

$$\prod_{j=1}^N G^R_p(t_j,t_{j+1}) = 0\ , \qquad (A.15)$$

whenever $t_{N+1}=t_1$ and the arguments form a closed loop. Put another way: a closed loop is forbidden by classical causality, since it requires the signal to propagate backwards in time. This is ultimately the reason why loop-level Feynman diagrams are so tightly constrained, and can be expressed in terms of tree-level diagrams. The identity (C') similarly implies that the product of advanced propagators is zero for a closed loop. Since the identities (C_1) and (C_2) are equivalent to (C) and (C'), they too can be viewed as causality conditions.

Finally, beyond unitarity of the free theory, all of the above cutting rules make use of unitarity of the interactions. This is ultimately the condition that all vertex factors commute with the discontinuity operations, so that we can extract the Im part of propagators inside the diagram without changing the structure of any interaction vertex.

[42] There is another sense in which (U_1) is related to unitarity: it follows from the canonical commutation relation for $[\phi_{\mathbf{k}},\hat{\Pi}_{\mathbf{k}}]$, which (by the Stone-von Neumann theorem) is what guarantees that the Hilbert spaces at different times are unitarily related, i.e. that time evolution is implemented by a unitary operator.

JHEP12(2023)076

Identity	follows from	used to fix	leads to cutting rules for
$\mathcal{I}$=0	(U_1) + (U_2)	$\text{Im}(G_1...G_N)$	Disc of any tree diagram, see [106, 107]
$\mathcal{R}$=0	(U) + (C_1)	$\text{Re}(G_1...G_N)_{\text{loop}}$	Disc of any loop diagram, see [106]
$\mathcal{L}$=0	(U) + (C)	$(G_1...G_N)_{\text{loop}}$	any loop diagram, see (2.60)
$\mathcal{L}'$=0	(U) + (C')	$(G_1...G_N)_{\text{loop}}$	any loop diagram, see (2.61)
$\mathcal{C}$=0	(U) + (C_2)	$\sum^N_{\text{perm.}}\Delta_1^S G_2...G_N$	Disc of collinear tree diagrams, see (2.62)

Table 2. Summary of the various cutting rules considered in this section, and the different unitarity ($(U) \Rightarrow (U_1), (U_2)$) and causality ($(C_1), (C_2) \Leftrightarrow (C), (C')$) properties from which they folow.

Cosmological tree theorem. Altogether, the different propagator identities, their corresponding cutting rules and their underlying assumptions are listed in table 2. The cutting rules stemming from $(C)/(C')$ will turn out to be particularly useful, and so we have given them a name: the *Cosmological Tree Theorem*. In particular, when applied to equal-time correlation functions, these "tree theorem" identities immediately explain the cancellation of certain singularities (see section 3). Furthermore, singling out the Cosmological Tree Theorem seems natural because it is a complex relation (whereas the other cutting rules reviewed here are purely real), and so it is a more efficient way of encoding the cutting rules. Of course, since all of these rules follow in some way from the same (U) identity, they are ultimately related to one another. For instance, in the examples given above, the two complex relations $\mathcal{L}_{p_1p_2}(t_1,t_2)=0$ and $\mathcal{L}_{p_2p_1}(t_1,t_2)=0$ are actually related to the four real relations $\{\mathcal{R}_{p_1p_2}(t_1,t_2),\mathcal{I}_{p_1p_2}(t_1,t_2,t_1),\mathcal{C}_{p_1p_2}(t_1,t_2),\mathcal{I}_p(t)\}$, since,

$$\begin{aligned}
\text{Re}\left[\mathcal{L}_{p_1p_2}(t_1,t_2)+\mathcal{L}_{p_2p_1}(t_1,t_2)\right] &= \mathcal{R}_{p_1p_2}(t_1,t_2)+2\text{Re}\left(K_{p_1}(t_1)K^*_{p_1}(t_2)\right)\mathcal{I}_{p_2}(t_2,t_1)\\
&\quad +2\text{Re}\left(K_{p_2}(t_2)K^*_{p_2}(t_1)\right)\mathcal{I}_{p_1}(t_1,t_2)\\
\text{Im}\left[\mathcal{L}_{p_1p_2}(t_1,t_2)+\mathcal{L}_{p_2p_1}(t_1,t_2)\right] &= \mathcal{I}_{p_1p_2}(t_1,t_2,t_1)+\text{Re}G^R_{p_1}(t_1,t_2)\mathcal{I}_{p_2}(t_2,t_1)+\text{Re}G^R_{p_2}(t_1,t_2)\mathcal{I}_{p_1}(t_1,t_2)\\
\text{Re}\left[\mathcal{L}_{p_1p_2}(t_1,t_2)-\mathcal{L}_{p_2p_1}(t_1,t_2)\right] &= 2\mathcal{C}_{p_1p_2}(t_1,t_2)\\
\text{Im}\left[\mathcal{L}_{p_1p_2}(t_1,t_2)-\mathcal{L}_{p_2p_1}(t_1,t_2)\right] &= 2\text{Im}\left(K_{p_1}(t_1)K^*_{p_1}(t_2)\right)\mathcal{I}_{p_2}(t_2,t_1)+2\text{Im}\left(K_{p_2}(t_2)K^*_{p_2}(t_1)\right)\mathcal{I}_{p_1}(t_1,t_2)
\end{aligned} \tag{A.16}$$

irrespective of any connection[43] between G_p and K_k.

Disc of the tree theorem. We remarked above (2.37) that existing loop-level cutting rules immediately follow from taking discontinuities of the Cosmological Tree Theorem. In addition to that one-vertex example, it is straightforward to show this for loops with any number of vertices. For instance, consider the tree theorems (2.40) and (2.42) for the two-vertex loop. Since the underlying propagator identity can be written as,

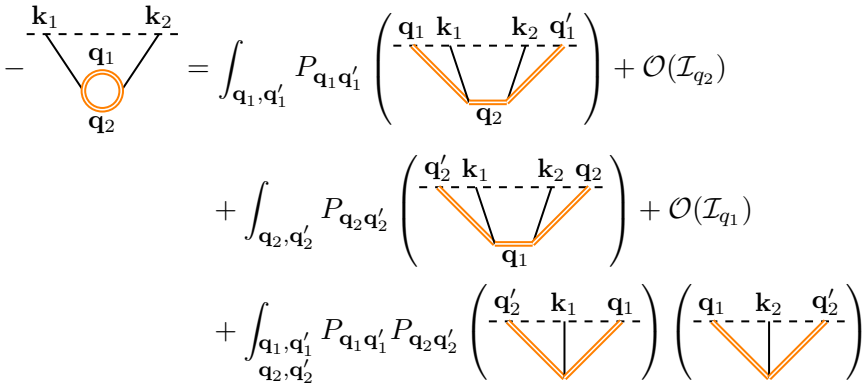
$$\text{Re}\left(\mathcal{L}_{p_1p_2}(t_1,t_2)+\mathcal{L}_{p_2p_1}(t_1,t_2)\right)=\mathcal{R}_{p_1p_2}(t_1,t_2)+\Delta^H_{p_1}(t_1,t_2)\mathcal{I}_{p_2}(t_2,t_1)+\Delta^H_{p_2}(t_2,t_1)\mathcal{I}_{p_1}(t_1,t_2)\,,$$

the cutting rule which follows from $\mathcal{L}_2=0$, together with the unitarity condition $\mathcal{I}_1=0$, is enough to reproduce the $\mathcal{R}_2=0$ cutting rule. In diagrammatic form, this amounts to

[43] In this equation we have used the shorthand, $G^R_p(t_1,t_2)=G_p(t_1,t_2)-2K_p(t_1)\text{Im}K_{p_1}(t_2)$.

JHEP12(2023)076

taking a Disc of (2.40) or (2.42):

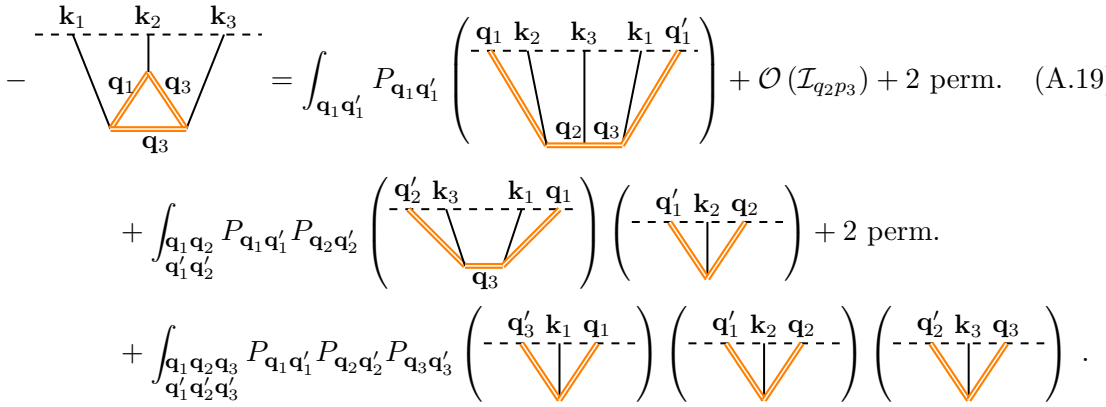


(A.17)

which indeed reproduces (A.9). Similarly for the three-vertex loop, since the propagator identity can be written as,

$$\begin{aligned}&\mathrm{Re}(\mathcal{L}_{p_1p_2p_3}(t_1,t_2,t_3)+\mathcal{L}_{p_3p_2p_1}(t_1,t_3,t_2))\\&\qquad=2\mathcal{R}_{p_1p_2p_3}(t_1,t_2,t_3)+\Delta^{H}_{p_1}(t_1,t_2)\mathcal{I}_{p_2p_3}(t_2,t_3,t_1)+2\text{ Perm}\end{aligned}\tag{A.18}$$

which corresponds to the diagrammatic relation,



(A.19)

This again reproduces the cutting rule for the Disc of this loop which were given in [106]. Since these rules require a Disc, they must invoke unitarity of the interacting Hamitonian (to ensure that coupling constant are real, or at least commute with taking the Disc).

One might wonder whether the tree-level identities like (A.6) which follow from perturbative unitarity can be reconstructed by subtracting the $\mathcal{R}$-type identities from the Disc of the tree theorem, for instance by subtracting (A.9) from (A.17)). This leads to an identity which differs from (A.6) by an analytic continuation of the external energies. This turns out to be a general feature: the difference between the $\mathcal{R}$-type cutting rules of [106] and the Disc of the Cosmological Tree Theorem is always proportional to the $\mathcal{I}$-type tree-level identities of [104, 105] with some of the external lines continued to negative energies.

B Non-Gaussianity at next-to-next-to-leading order

In this appendix we briefly describe primordial non-Gaussianity at next-to-next-to-leading order (NNLO). By "order", we are referring to the power of the small coupling g_* which

JHEP12(2023)076

suppresses field insertions in the Lagrangian. For instance, the "leading-order" non-Gaussianity is the tree-level bispectrum, which is $\mathcal{O}(g_*)$. In the main text we discuss the next-to-leading order (NLO) effects at $\mathcal{O}(g_*^2)$, which includes the one-loop correction to the power spectrum. Here, we discuss NNLO effects at $\mathcal{O}(g_*^3)$, namely the one-loop bispectrum and the two-loop vev.

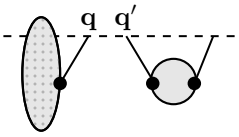
B.1 Bispectrum at one loop

Beyond the power spectrum (3.3), we can analogously extract higher-point correlation functions from the wavefunction. The first of these is the bispectrum,[44]

$$\frac{\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\hat{\phi}_{\mathbf{k}_3}\rangle}{\mathcal{P}_{k_1}\mathcal{P}_{k_2}\mathcal{P}_{k_3}} = \mathcal{B}_{k_1k_2k_3}\tilde{\delta}^3\left(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3\right) \ . \tag{B.1}$$

Note that we normalise the bispectrum using the power spectrum $\mathcal{P}_q$ of the full interacting theory. This normalisation leads to a convenient cancellation of all diagrams in which the propagation of a single external leg to the boundary is modified, i.e. it removes all terms of the form,

$$\int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \tag{B.2}$$



where the left blob can have arbitrarily many external legs but the right blob may only have the two external legs shown.

The bispectrum is related to the non-Gaussian wavefunction coefficients by the Born rule (3.1). Expanding this perturbatively in g_* gives,

$$\begin{aligned}\mathcal{B}_{k_1k_2k_3}\tilde{\delta}^3\left(\sum_{a=1}^{3}\mathbf{k}_a\right) &= 2\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\ \left(\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{q}\mathbf{q}'} + 4\,\mathrm{Re}\,\psi_{\mathbf{q}'}\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{q}}\right)\\ &+ 2\int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}\left(\mathrm{Re}\,\psi_{\mathbf{q}_1'\mathbf{q}_2\mathbf{q}_2'}\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{q}_1} + \sum_{\text{perm.}}^{3}\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}_2'}\mathrm{Re}\,\psi_{\mathbf{k}_2\mathbf{k}_3\mathbf{q}_2\mathbf{q}_1'}\right)\\ &+ 4\int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'\\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_1\mathbf{q}_1'}P_{\mathbf{q}_2\mathbf{q}_2'}P_{\mathbf{q}_3\mathbf{q}_3'}\,\mathrm{Re}\,\psi_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}_2'}\mathrm{Re}\,\psi_{\mathbf{k}_2\mathbf{q}_2\mathbf{q}_3'}\mathrm{Re}\,\psi_{\mathbf{k}_3\mathbf{q}_3\mathbf{q}_1'} + \mathcal{O}\left(g_*^4\right) \ .\end{aligned} \tag{B.3}$$

Each wavefunction coefficient can then be expanded in terms of Feynman-Witten diagrams, and as a result one arrives at the expansion,

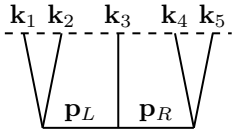
$$\mathcal{B}_{k_1k_2k_3} = \mathcal{B}^{(1)}_{k_1k_2k_3} + \mathcal{B}^{(3)}_{k_1k_2k_3} + \mathcal{O}\left(g_*^5\right) \tag{B.4}$$

[44] Strictly speaking, the full $\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\hat{\phi}_{\mathbf{k}_3}\rangle$ also contains disconnected contributions which $\sim \tilde{\delta}(\mathbf{k}_1+\mathbf{k}_2)\tilde{\delta}^3(\mathbf{k}_3)$ and its permutations. One should therefore interpret the bispectrum (B.1) as the *connected* part of the three-point function.

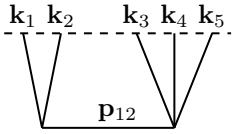
JHEP12(2023)076

where $\mathcal{B}^{(1)} = 2\mathrm{Re}\,\psi_3^{\mathrm{tree}}$ is the tree-level part, and $\mathcal{B}^{(3)}$ is determined by $\psi_3^{1-\mathrm{loop}}, \psi_5^{\mathrm{tree}}, \psi_4^{\mathrm{tree}}$ and ψ_3^{tree}. The tree-level diagrams for ψ_3 and ψ_4 are given in (1.12), while for ψ_5 they are,

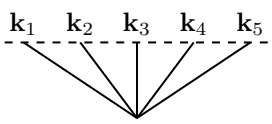
$$\psi^{\mathrm{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4\mathbf{k}_5} = \sum_{\mathrm{perm.}}^{15}$$



$$+ \sum_{\mathrm{perm.}}^{10}$$



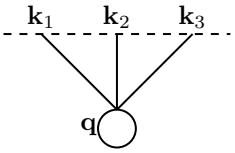
$+$



(B.5)

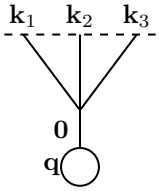
and for the 1-loop cubic coefficient they are,

$$\psi^{1-\mathrm{loop}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} = \frac{1}{2}$$

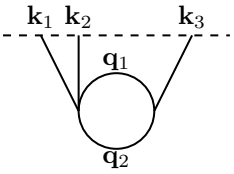


(B.6)

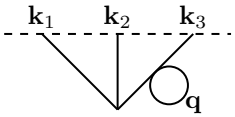
$$+\frac{1}{2}$$



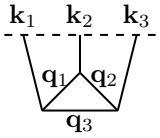
$$+\frac{1}{2}\sum_{\mathrm{perm.}}^{3}$$



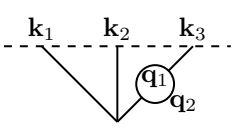
$$+\frac{1}{2}\sum_{\mathrm{perm.}}^{3}$$



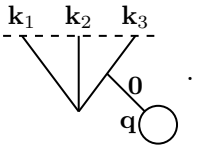
$+$



$$+\frac{1}{2}\sum_{\mathrm{perm.}}^{3}$$



$$+\frac{1}{2}\sum_{\mathrm{perm.}}^{3}$$



.

Notice that the one- and two-vertex loops come with symmetry factors of 1/2 (more on this in appendix C), while the three-vertex loop has a symmetry factor of 1.

Once the $\mathcal{B}^{(3)}$ correction is written in terms of Feynman-Witten diagrams, we find that they collect precisely into the groups identified in (3.12), (3.14) and (3.16). To be precise, the diagrams can be organised according to which vertices they contain and whether the external lines end on the same or different vertices. This leads to,

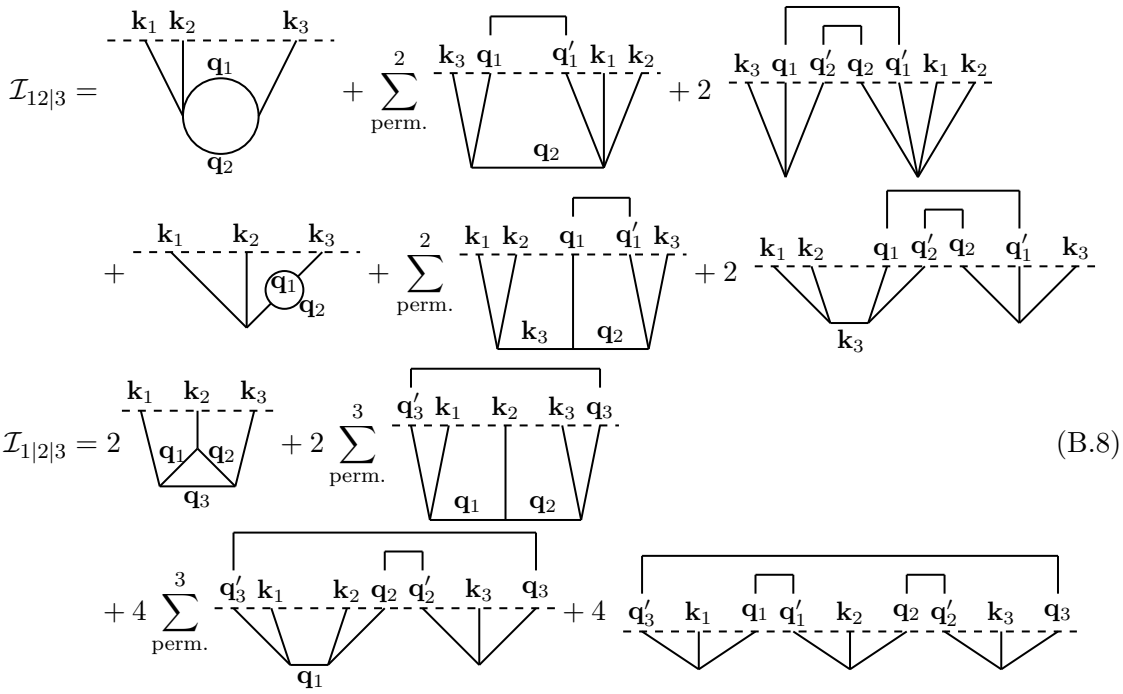
$$\mathcal{B}^{(3)}_{k_1k_2k_3}\,\tilde{\delta}^3\left(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3\right) = \mathcal{I}_{123} + \mathcal{I}_{12|3} + \mathcal{I}_{13|2} + \mathcal{I}_{23|1} + \mathcal{I}_{1|2|3} \tag{B.7}$$

where the diagrams in $\mathcal{I}_{123}$ are all of the form (3.12), those in $\mathcal{I}_{12|3}$ and its permutations are of the form (3.14), and those in $\mathcal{I}_{1|2|3}$ are of the form (3.16). Using our tree theorem to replace all loop diagrams with tree-level diagrams then leads to the precise KLN cancellations between virtual and real emission described in section 3.1, and as a result there are no new total energy branch points introduced by the integration over loop momenta.

In spite of these cancellations, there are still many tree-level diagrams required to compute the bispectrum for general massive fields. However, if we focus on massless internal lines, then many of these diagrams become scale-free integrals and hence vanish in dimensional regularisation. In that case, the only terms in (B.7) which give a non-zero

JHEP12(2023)076

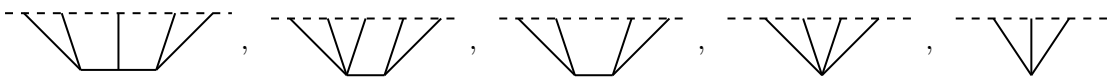
contribution are,



(B.8)

where we have omitted the Re which should be taken of every connected component, and replaced each $\int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}$ factor by a "contraction" of the corresponding $\mathbf{q}\mathbf{q}'$ pair in the diagram.

Applying the tree theorem, we find that the one-loop correction to the bispectrum from massless fields can be written in terms of just 5 tree-level diagrams,



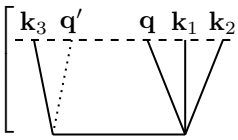
(B.9)

which is just 2 more than the power spectrum. By contrast, the original (B.3) requires 13 separate Feynman-Witten diagrams (6 more than the power spectrum). In fact, applying a bootstrap argument to fix the exchange diagrams in terms of contact diagrams, we can actually construct the integrand for $\mathcal{P}^{(2)}$ entirely from the single ψ_3^{tree} contact diagram and $\mathcal{B}^{(3)}$ from the two contact diagrams shown in (B.9).

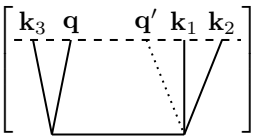
JHEP12(2023)076

The explicit relation between (B.9) and the one-loop bispectrum (B.7) is,

$$\mathcal{I}_{12|3} = \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \left(\mathrm{Re} \left[\right. \right. \tag{B.10}$$

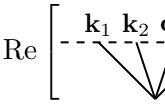


$$\left. \left. \right] + \mathrm{Re} \left[\right. \right.$$

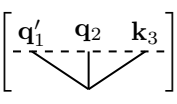


$$\left. \left. \right] \right)$$

$$+ 2 \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \, \mathrm{Re} \left[\right.$$

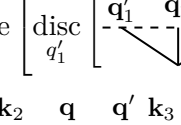


$$\left. \right] \mathrm{Re} \left[\right.$$

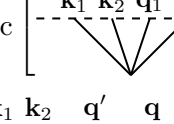


$$\left. \right]$$

$$- \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \, \mathrm{Re} \left[\mathrm{disc}_{q_1'} \left[\right. \right.$$

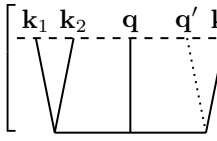


$$\left. \left. \right] \mathrm{disc}_{q_2'} \left[\right. \right.$$

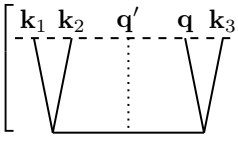


$$\left. \left. \right] \right]$$

$$+ \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \left(\mathrm{Re} \left[\right. \right.$$

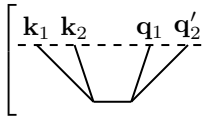


$$\left. \left. \right] + \mathrm{Re} \left[\right. \right.$$

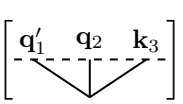


$$\left. \left. \right] \right)$$

$$+ 2 \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \mathrm{Re} \left[\right.$$

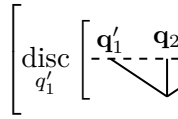


$$\left. \right] \mathrm{Re} \left[\right.$$

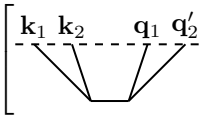


$$\left. \right]$$

$$- \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \mathrm{Re} \left[\mathrm{disc}_{q_1'} \left[\right. \right.$$



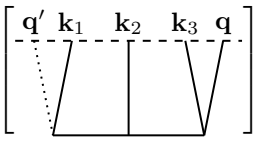
$$\left. \left. \right] \mathrm{disc}_{q_2'} \left[\right. \right.$$



$$\left. \left. \right] \right]$$

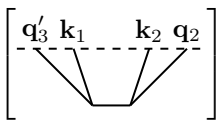
and its permutations, together with,

$$\mathcal{I}_{1|2|3} = \sum_{\mathrm{perm.}}^{3} \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \mathrm{Re} \left[\right. \tag{B.11}$$

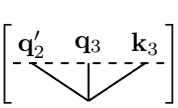


$$\left. \right]$$

$$+ \sum_{\mathrm{perm.}}^{3} 2 \int_{\substack{\mathbf{q}_2\mathbf{q}_2' \\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_3\mathbf{q}_3'} \mathrm{Re} \left[\right.$$

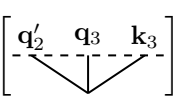


$$\left. \right] \mathrm{Re} \left[\right.$$

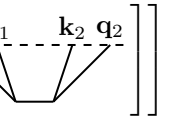


$$\left. \right]$$

$$- \sum_{\mathrm{perm.}}^{3} \int_{\substack{\mathbf{q}_2\mathbf{q}_2' \\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_3\mathbf{q}_3'} \mathrm{Re} \left[\mathrm{disc}_{q_2'} \left[\right. \right.$$

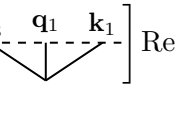


$$\left. \left. \right] \mathrm{disc}_{q_3'} \left[\right. \right.$$

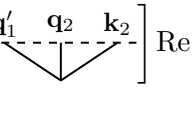


$$\left. \left. \right] \right]$$

$$+ 4 \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2' \\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_3\mathbf{q}_3'} \mathrm{Re} \left[\right.$$



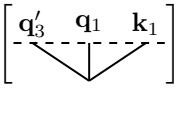
$$\left. \right] \mathrm{Re} \left[\right.$$



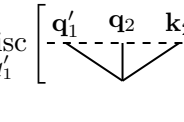
$$\left. \right] \mathrm{Re} \left[\right.$$

$$\left. \right]$$

$$- \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2' \\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_3\mathbf{q}_3'} \mathrm{Re} \left[\mathrm{disc}_{q_3'} \left[\right. \right.$$



$$\left. \left. \right] \mathrm{disc}_{q_1'} \left[\right. \right.$$



$$\left. \left. \right] \mathrm{disc}_{q_2'} \left[\right. \right.$$

$$\left. \left. \right] \right] .$$

JHEP12(2023)076

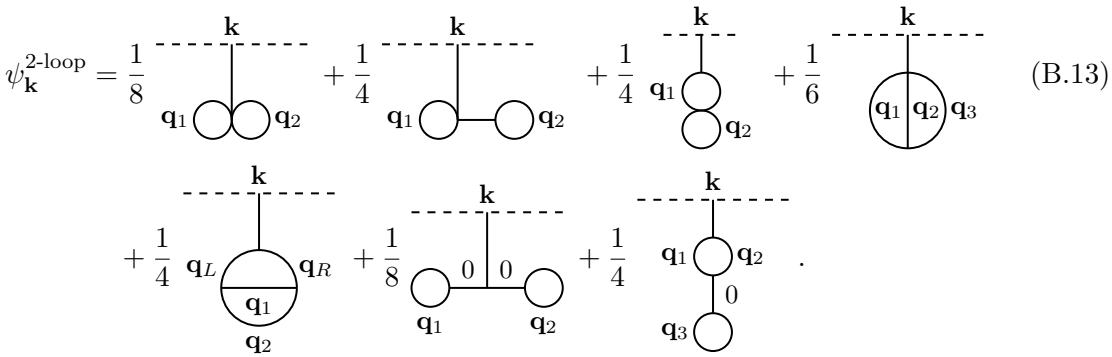
B.2 Field vev at two loops

To compute the vev (3.5) at next-to-next-to-leading order, we first expand the Born rule up to $\mathcal{O}(g_*^3)$,

$$v\,\tilde{\delta}^3\left(\mathbf{k}\right) = 2\mathrm{Re}\,\psi_{\mathbf{k}} + \int_{\mathbf{q}\mathbf{q}'} \mathcal{P}_{\mathbf{q}\mathbf{q}'}\,\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{q}\mathbf{q}'} + \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'}\,\frac{1}{4}\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{q}_1\mathbf{q}_1'\mathbf{q}_2\mathbf{q}_2'} + \int_{\substack{\mathbf{q}_1\mathbf{q}_1'\\ \mathbf{q}_2\mathbf{q}_2'\\ \mathbf{q}_3\mathbf{q}_3'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} P_{\mathbf{q}_3\mathbf{q}_3'}\,\frac{2}{3}\mathrm{Re}\,\psi_{\mathbf{k}\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3}\mathrm{Re}\,\psi_{\mathbf{q}_1'\mathbf{q}_2'\mathbf{q}_3'} + \mathcal{O}\left(g_*^4\right) \tag{B.12}$$

where note that the $\mathcal{P}_{\mathbf{q}\mathbf{q}'}$ appearing in the next-to-leading order term is the full power spectrum and should be expanded as in (3.4). Next, we expand each wavefunction coefficient as a series in the number of loops. This gives an expression for $v^{(3)}$ which depends on $\psi_1^{2-\text{loop}}$, $\psi_2^{1-\text{loop}}$, $\psi_3^{1-\text{loop}}$, ψ_5^{tree}, ψ_4^{tree} and ψ_3^{tree}. The corresponding tree-level diagrams are given in (1.12) and (B.5), while the 1-loop diagrams are given in (B.6). The necessary two-loop diagrams are,

$\psi_{\mathbf{k}}^{2\text{-loop}} =$



(B.13)

We find that the terms in $v^{(3)}$ organise into the combinations (3.12), (3.14) and (3.16) for the one-loop diagrams, as well as (3.19) for the two-loop diagrams. Expanding each loop using the Cosmological Tree Theorem therefore leads to very many cancellations, and in particular the only connected terms which remain are found to have pairs of energies which are analytically continued so that there is no set of vertices whose total energy depends on both the total external energy and the loop momenta. As a result, we find that while $v^{(3)}$ could be non-zero even for a massless field, it does not contain any additional branch cuts in k. Of course, once momentum conservation is imposed, in this example k is fixed to be zero regardless. But when computing higher-point correlators, e.g. the $\mathcal{P}_k^{(4)}$ correction to the power spectrum, there will be a series of diagrams which resemble those above but with the $\mathbf{k}$ leg split into a pair of legs carrying momentum $\mathbf{k}$ and $\mathbf{k}'$. Since the KLN cancellation takes place at finite k, it means that this subset of diagrams in such higher-point correlators (for which the energy $k + k'$ is no longer set to zero) also do not introduce any additional branch points.

C Wavefunction symmetry factors

In the literature, often single diagrams are considered so the overall normalisation is arbitrary. Here, when we compute in-in correlators in section 3, the relative coefficient between

JHEP12(2023)076

loop and tree diagrams matters (in particular, the cancellation of singularities in subsection 3.1 depends crucially on using the correct relative normalisation).

In this short appendix, we give three separate proofs that the correct symmetry factor for the sunset diagram in (1.12) is 1/2: computing the Wick contractions directly, solving the Schrödinger equation and comparing the Born rule with an explicit in-in calculation. In each case, it is clear how the counting of symmetry factors should work for more general diagrams.

From Wick contractions. The way symmetry factors are often introduced in the context of amplitudes is through the Dyson series expansion of $\hat{U}$. The sunset diagram appears naturally in the self energy for a scalar field at one loop with $\frac{g}{3!}\phi^3$ interaction:

$$\frac{1}{2}\frac{g^2}{3!^2}\langle\hat{\phi}(x)\hat{\phi}(y)\phi^3(z_1)\phi^3(z_2)\rangle = \frac{g^2}{2}\int_{z_1 z_2}\Delta(x-z_1)\Delta^2(z_1-z_2)\Delta(z_2-y), \tag{C.1}$$

where $\Delta(x_1-x_2)$ is the Feynman propagator. The symmetry factor of 1/2 arises from there being three different ways to contract the internal line with each vertex. There are then two ways to contract the remaining internal lines. There is an overall factor of 2 to account for the vertex permutation:

$$\frac{3^2\times 2\times 2}{3!^2 2} = \frac{1}{2} \tag{C.2}$$

In the case of the wavefunction, the analogous series expansion of $\hat{U}$ inside (1.9) leads to the same set of Wick contractions. Consequently, the symmetry factors appearing in the wavefunctions coefficients are identical to those appearing in the usual amplitude calculation.

From Schrodinger equation. We now turn to solving the Schrodinger equation,

$$i\partial_t\Psi[\phi] = \hat{\mathcal{H}}\Psi[\phi], \tag{C.3}$$

where t is the time to which we have evolved the Bunch-Davies state. For the free Hamiltonian (1.8),[45] the Schrodinger equation in momentum space becomes [104],

$$-\partial_t\Gamma[\phi] = \frac{1}{2a^{d-1}(t)}\int_{\mathbf{p}_1\mathbf{p}_2}\delta(\mathbf{p}_1+\mathbf{p}_2)\left(\frac{\delta\Gamma[\phi]}{\delta\phi_{\mathbf{p}_1}}\frac{\delta\Gamma[\phi]}{\delta\phi_{\mathbf{p}_2}} - i\frac{\delta^2\Gamma[\phi]}{\delta\phi_{\mathbf{p}_1}\delta\phi_{\mathbf{p}_2}}\right) \tag{C.4}$$

In order to discuss the symmetry factor of the sunset diagram we expand the phase (the on-shell action) to quadratic order in the coupling g:

$$\begin{aligned}\Gamma[\phi] = &\frac{1}{2}\int_{\mathbf{k}_1\mathbf{k}_2}\psi_2(\mathbf{k}_1,\mathbf{k}_2)\phi_{\mathbf{k}_1}\phi_{\mathbf{k}_2}\tilde{\delta}(\mathbf{k}_T) + \frac{1}{6}\int_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_2}\psi_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\phi_{\mathbf{k}_1}\phi_{\mathbf{k}_2}\phi_{\mathbf{k}_3}\tilde{\delta}(\mathbf{k}_T)\\ &+\frac{1}{24}\int_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_2\mathbf{k}_4}\psi_4(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)\phi_{\mathbf{k}_1}\phi_{\mathbf{k}_2}\phi_{\mathbf{k}_3}\phi_{\mathbf{k}_4}\tilde{\delta}(\mathbf{k}_T)\end{aligned} \tag{C.5}$$

[45] Including the interaction Hamiltonian $\mathcal{H}_{\rm int}$ will only source the contact Feynman-Witten diagrams, but does not affect the exchange or loop-type diagrams for which we wish to count the symmetry factor.

JHEP12(2023)076

Doing an expansion $\psi_2(\mathbf{k}_1,\mathbf{k}_2)=\psi_2^{\text{free}}(\mathbf{k}_1,\mathbf{k}_2)+\psi_2^{1-\text{loop}}(\mathbf{k}_1,\mathbf{k}_2)+\dots$ leaves an Schrödinger equation of the form:[46]

$$\frac{\partial_t(i\psi_2^{1-\text{loop}}(\mathbf{k}_1,\mathbf{k}_2)f_{k_1}^*(t)f_{k_2}^*(t))}{f_{k_1}^*(t)f_{k_2}^*(t)}=-\frac{1}{2a^{d-1}(t)}\int_{\mathbf{q}}\psi_4(\mathbf{k}_1,\mathbf{k}_2,\mathbf{q},-\mathbf{q}). \tag{C.6}$$

where we have also used $\psi_2'^{\text{free}}(\mathbf{k}_1,\mathbf{k}_2)=a^{d-1}\partial_t\log(f_{k_1}(t))$. This equation is solved by:

$$\begin{aligned}\psi_2^{1-\text{loop}}(\mathbf{k}_1,\mathbf{k}_2)&=\frac{g^2}{2}\int_{\mathbf{q}_1\mathbf{q}_2}\int_{t_1t_2}\frac{\tilde{\delta}(\mathbf{q}_1+\mathbf{q}_2+\mathbf{k}_1)}{a^{d+1}(t_1)a^{d+1}(t_2)}K_{k_1}(t_1)K_{k_2}(t_2)G_{q_1}(t_1,t_2)G_{q_2}(t_1,t_2)\\ \psi_4(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)&=ig^2\int_{t_1t_2}\frac{1}{a^{d+1}(t_1)a^{d+1}(t_2)}K_{k_1}(t_1)K_{k_2}(t_1)G_{p_s}(t_1,t_2)K_{k_3}(t_2)K_{k_4}(t_2)\\ &\quad+2\text{ perm.}\end{aligned} \tag{C.7}$$

since we have the useful relation [75, 104],

$$\partial_tG_q(t_1,t_2)=-a^{1-d}(t)K_q(t_1)K_q(t_2) \tag{C.8}$$

for the derivative of the bulk-to-bulk propagator with respect to the boundary time. The relative factor of 1/2 in (C.7) is the same symmetry factor we find in amplitudes for similar diagrams.

From canonical quantisation. Finally, in the in-in calculation of cosmological correlators one finds that the second order correction towards the expectation value of an observable is:

$$\langle\hat{\mathcal{O}}\rangle=-\int_{-\infty}^{\tau}d\tau_2\int_{-\infty}^{\tau_2}d\tau_1\langle[\hat{\mathcal{H}}(\tau_1),[\hat{\mathcal{H}}(\tau_2),\hat{\mathcal{O}}]]\rangle, \tag{C.9}$$

which can be rewritten as the sum of two terms:

$$\langle\hat{\mathcal{O}}\rangle=-2\int_{-\infty}^{\tau}d\tau_2\int_{-\infty}^{\tau_2}d\tau_1\left(\text{Re}\left(\langle\hat{\mathcal{H}}(\tau_1)\hat{\mathcal{H}}(\tau_2)\hat{\mathcal{O}}\rangle\right)-\text{Re}\left(\langle\hat{\mathcal{H}}(\tau_1)\hat{\mathcal{O}}\hat{\mathcal{H}}(\tau_2)\rangle\right)\right). \tag{C.10}$$

In order to fix the symmetry factor for the sunset diagram, we will focus on the two point function $\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\rangle$ and the interaction $\frac{g}{3!}\phi^3$. We do the calculation in Minkowski spacetime, as the symmetry factor does not depend on the spacetime. From equation (C.10) we find two terms contributing towards $\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\rangle$:

$$\begin{aligned}\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\rangle&=\delta(\mathbf{k}_1+\mathbf{k}_2)g^2\int_{\mathbf{q}_1\mathbf{q}_2}\delta(\mathbf{q}_1+\mathbf{q}_2+\mathbf{k}_1)\left(\mathcal{I}_1-\mathcal{I}_2\right)\\ \mathcal{I}_1&=\frac{k_1+k_2+2(q_1+q_2)}{16k_1k_2q_1q_2(k_1+k_2)(k_1+q_1+q_2)(k_2+q_1+q_2)}\\ \mathcal{I}_2&=\frac{k_2-k_1}{16k_1k_2q_1q_2(k_1-k_2)(k_1+q_1+q_2)(k_2+q_1+q_2)},\end{aligned} \tag{C.11}$$

which leaves:

$$\langle\hat{\phi}_{\mathbf{k}_1}\hat{\phi}_{\mathbf{k}_2}\rangle=g^2\delta(-\mathbf{k}_1-\mathbf{k}_2)\int_{\mathbf{q}_1\mathbf{q}_2}\frac{2k+q_1+q_2}{16k^3q_1q_2(k+q_1+q_2)^2}. \tag{C.12}$$

This coincides with the calculation done in section 3.1 using the 1/2 symmetry factor for the sunset diagram.

[46] This agrees with [106] once we account for the factor of 1/2 which is missing from the right-hand-side of their equation (A.15).

JHEP12(2023)076

D Further identities

In this appendix we list a number of further identities that stem from the same conditions of non-relativistic causality and unitarity of the free theory.

D.1 Symmetric version of the cutting rules and tree theorem

In the main text, we mostly focus on replacing the bulk-to-bulk propagators with retarded propagators using the identity,

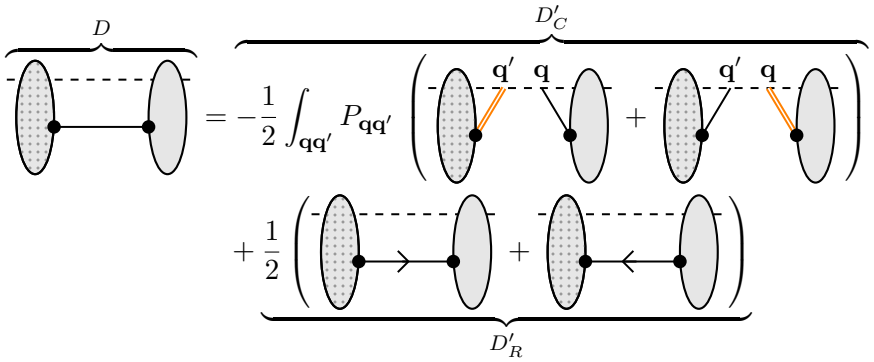
$$G_p(t_1,t_2) = G_p^R(t_1,t_2) + 2P_pK_p(t_1)\mathrm{Im}\,K_p(t_2)\ . \tag{D.1}$$

In a few places, we also made use of the advanced form of this identity,

$$G_p(t_1,t_2) = G_p^R(t_2,t_1) + 2P_pK_p(t_2)\mathrm{Im}\,K_p(t_1) \tag{D.2}$$

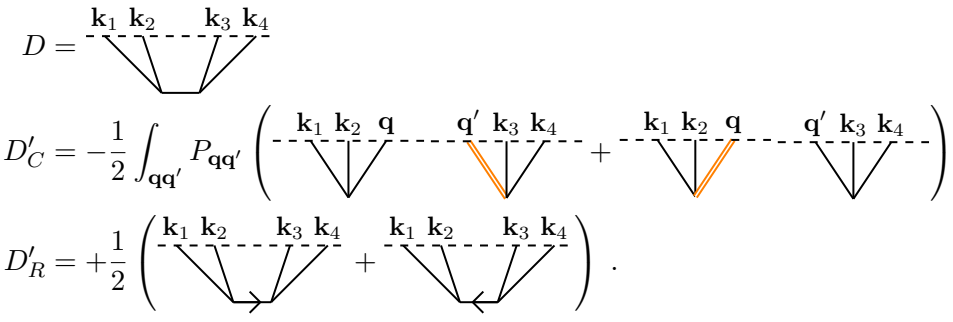
In each of these expressions, the $t_1 \leftrightarrow t_2$ symmetry on the left-hand-side is no longer manifest once we introduce the retarded propagator. An alternative replacement is to use the sum of (D.1) and (D.2). This will lead to diagrammatic identities with more terms, but which preserve the exchange symmetries of the original diagram.

Tree-level. Applying this to a general tree-level exchange diagram produces,



(D.3)

where this cut contribution D'_C and retarded contribution D'_R now have twice as many terms as in the main text, but are manifestly symmetric under the exchange of the two blobs. The bootstrap procedure we described in section 2.1 can be applied in the same way to determine to D'_R. For instance, let us repeat the first example from section 2.1,



(D.4)

D'_C is determined by $\psi_3^{\rm tree}$ and its discontinuity (assumed known), while D'_R corresponds to,

$$D'_R = P_{p_s} \underset{p_s}{\rm disc} \left[\int_{t_1 t_2} K_{k_1}(t_1) K_{k_2}(t_1) K_{k_3}(t_2) K_{k_4}(t_2)\, K_{p_s}(t_1) K^*_{p_s}(t_2) {\rm sign}\,(t_1 - t_2) \right] . \quad \text{(D.5)}$$

Rather than compute this explicitly, we can notice that since it takes the form $P_{p_s} \underset{p_s}{\rm disc}\,[f(p_s)]$, it must be an even function of p_s. Demanding that it also cancels the unphysical singularities in D'_C (and does not introduce any further unphysical singularities) then fixes D'_R up to an analytic remainder. In the case of Minkowski, where $\psi_3^{\rm tree}$ is given by (2.10), this procedure leads to,

$$\begin{aligned} D'_C &= \frac{\lambda^2 \tilde{\delta}^3 \left(\sum_{b=1}^4 \mathbf{k}_b\right)}{2 E_L E_R} \left(\frac{1}{k_3 + k_4 - p_s} + \frac{1}{k_1 + k_2 - p_s} \right) \\ \Rightarrow \quad D'_R &= \frac{\lambda^2 \tilde{\delta}^3 \left(\sum_{b=1}^4 \mathbf{k}_b\right)}{k_T} \left(\frac{1}{p_s^2 - (k_3 + k_4)^2} + \frac{1}{p_s^2 - (k_1 + k_2)^2} \right) , \end{aligned} \quad \text{(D.6)}$$

which indeed sum to the simple result (2.13).

Loop-level. In the main text we gave two versions of the Cosmological Tree Theorem, which correspond to cutting the loop either clockwise or anticlockwise. Using the sum of these two identities is sometimes useful, for instance to retain a manifest symmetry in all external lines. For instance, instead of using the identity $\mathcal{L}_{p_1 p_2}(t_1, t_2) = 0$ given in (A.10), a symmetric version of the tree theorem follows from $\mathcal{L}_{p_1 p_2}(t_1, t_2) + \mathcal{L}_{p_2 p_1}(t_1, t_2)$: this has twice as many terms as (2.40), but would be manifestly symmetric in $k_1 \leftrightarrow k_2$ if the two external lines are indistinguishable.

We have noticed an intriguing feature of this symmetric form of the tree theorem. It turns out that the various terms can always be combined in such a way that each term contains exactly one imaginary part (and otherwise depends only on $K_k(t)$ and $G_p(t_1, t_2)$ without any complex conjugation), and it therefore expresses each one-loop diagram as a sum of *single* discontinuities of tree-level diagrams. For instance, the $\mathcal{L}_{p_1 p_2}(t_1, t_2) + \mathcal{L}_{p_2 p_1}(t_1, t_2)$ identity can be written as,

$$\begin{aligned} & G_{p_1}(t_1, t_2) G_{p_2}(t_1, t_2) \quad \text{(D.7)} \\ & = P_1 {\rm Im}(K_{p_1}(t_1)) K_{p_1}(t_2) G_{p_2}(t_1, t_2) + P_1 {\rm Im}(K_{p_1}(t_2)) K_{p_1}(t_1) G_{p_2}(t_1, t_2) \\ & \quad + P_2 {\rm Im}(K_{p_2}(t_1)) K_{p_1}(t_2) G_{p_1}(t_1, t_2) + P_2 {\rm Im}(K_{p_2}(t_2)) K_{p_1}(t_1) G_{p_1}(t_1, t_2) \\ & \quad + i P_1 P_2 {\rm Im}(K_{p_1}(t_1)) K_{p_1}(t_2) K_{p_2}(t_1) K_{p_2}(t_2) + i P_1 P_2 {\rm Im}(K_{p_1}(t_2)) K_{p_1}(t_1) K_{p_2}(t_1) K_{p_2}(t_2) \\ & \quad + i P_1 P_2 {\rm Im}(K_{p_2}(t_1)) K_{p_1}(t_2) K_{p_2}(t_1) K_{p_1}(t_2) + i P_1 P_2 {\rm Im}(K_{p_2}(t_2)) K_{p_1}(t_1) K_{p_2}(t_1) K_{p_1}(t_2) \\ & \quad - i P_1 P_2 {\rm Im}(K_{p_1}(t_1) K_{p_2}(t_2)) K_{p_2}(t_1) K_{p_1}(t_2) - i P_1 P_2 {\rm Im}(K_{p_1}(t_2) K_{p_2}(t_1)) K_{p_2}(t_2) K_{p_1}(t_1), \end{aligned}$$

JHEP12(2023)076

which implies relations between wavefunction coefficients such as,

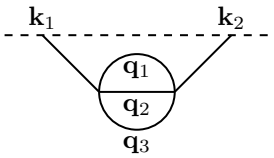
$$
\begin{aligned}
-4\,\psi^{1-\text{loop}}_{\mathbf{k}_1\mathbf{k}_2} = &\int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}\left(\operatorname*{disc}_{q'}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}\mathbf{q}'}\right] + 3 \text{ perm.}\right) \\
&+ \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'}\left(\operatorname*{disc}_{q_1'}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{q}_1'\mathbf{q}_2}\psi^{\text{tree}}_{\mathbf{k}_2\mathbf{q}_1\mathbf{q}_2'}\right] + 3 \text{ perm.}\right. \\
&\qquad \left. - \operatorname*{disc}_{q_1'q_2'}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{q}_1'\mathbf{q}_2}\psi^{\text{tree}}_{\mathbf{k}_2\mathbf{q}_1\mathbf{q}_2'}\right] + 1 \text{ perm.}\right).
\end{aligned}
\tag{D.8}
$$

In practice, this way of writing the cutting rules is less useful because regrouping into single discontinuities leads to a proliferation of terms. While the Cosmological Tree Theorem (2.60) for a graph with E edges produces a sum of $2^E - 1$ terms, symmetrising and then regrouping produces a single-Disc identity with $2(3^E - 2^E)$ terms ($= 10$ when $E = 2$ in the example above). This is always larger than both OFPT and TOPT (see table 1).

D.2 Higher loops

While most of our examples so far have been for one-loop diagrams, the Cosmological Tree Theorem can be applied straightforwardly to higher loop graphs. The key observation is that we can apply the tree theorem (2.60) to any closed loop within a diagram, even when that diagram contains other loops. By recursively removing one loop at a time, (2.60) can be used to reduce any arbitrary loop diagram into a sum of (momentum integrals of) tree-level diagrams. While removing the loops in different orders can produce different identities for any given diagram, these differ only by combinations which vanish thanks to lower-order identities.

A two-loop example. Let us illustrate this with an example. Consider the following two-loop correction to the power spectrum,



. (D.9)

This diagram contains three different closed cycles which can be cut open using (2.40), namely:

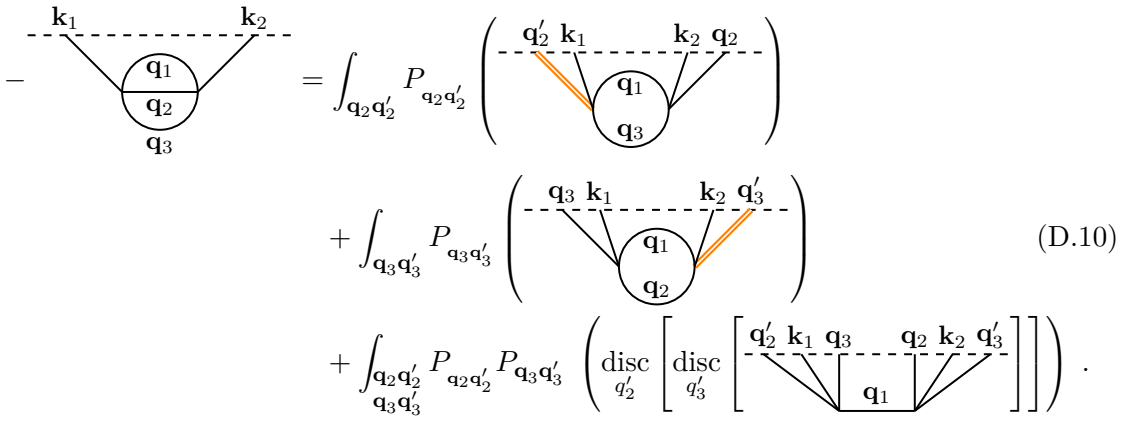
(a) the upper one-loop subgraph, $G_{q_1}G_{q_2}$,

(b) the lower one-loop subgraph, $G_{q_2}G_{q_3}$,

(c) the outer one-loop subgraph, $G_{q_3}G_{q_1}$.

JHEP12(2023)076

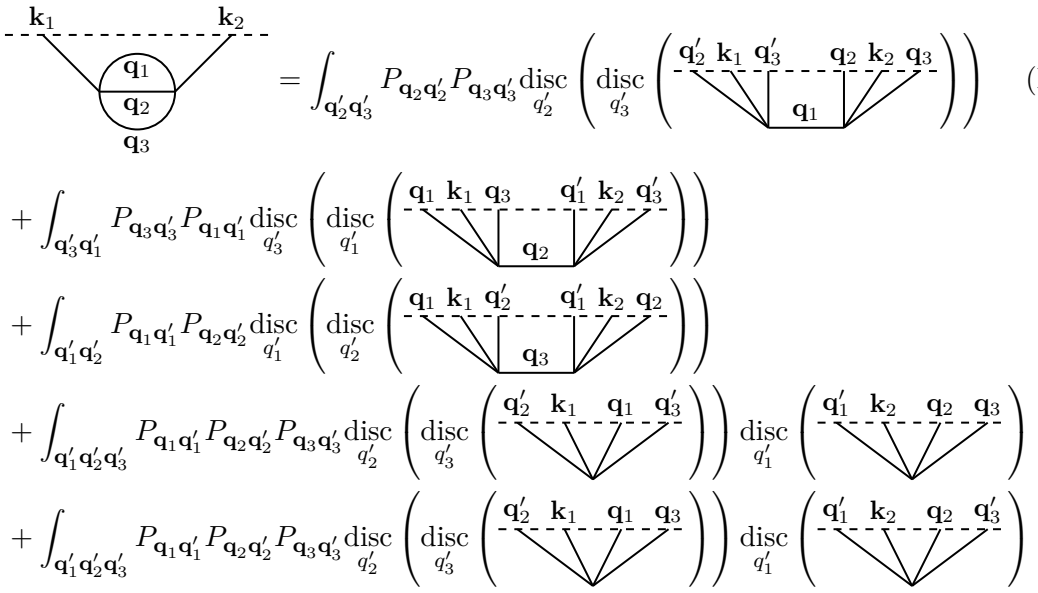
In all of these cases the resulting product expansion decomposes the integrand into one-loop and tree level contributions. For instance, cutting the lower loop (b) produces,



(D.10)

Notice that we must take two consecutive discontinuities on the same wavefunction coefficient in order to extract the required imaginary parts. In one-loop diagrams, cutting any pair of lines in the loop separates the diagram into two disconnected pieces and produces a product of single discontinuities. For higher loops, if a diagram remains connected affecting cutting two or more internal lines, then multiple discontinuities will appear in its tree theorem representation. In general there can be up to C consecutive discontinuities for each term with C cut lines.

To fully reduce this diagram to a sum of trees, we can now apply the tree theorem (2.40) again to the remaining loops in (D.10). This will produce twice-cut diagrams (i.e. a six-point exchange diagram) with two discontinuities and thrice-cut diagrams (i.e. a product of four-point contact diagrams) with three discontinuities. Explicitly,



(D.11)

Note that this tree-level representation is not unique, since we could have started by cutting any of the closed loops and used either orientation. Eq. (D.11) corresponds to

JHEP12(2023)076

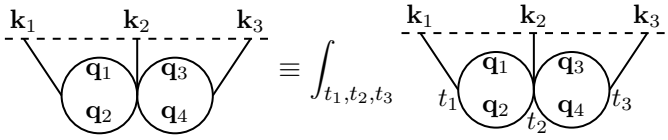
first cutting (b) clockwise and then cutting the remaining (c) anticlockwise and (a) clockwise. But there are 23 other possibilities. The consistency of these different expressions is guaranteed by the one-loop tree theorem and the tree-level relation (2.30). Explicitly, the difference between initially cutting loop (b) or loop (a) corresponds to,

$$\begin{aligned} &G_{q_1}(t_1,t_2)\mathcal{L}_{q_2q_3}(t_1,t_2) - \mathcal{L}_{q_1q_2}(t_1,t_2)G_{q_3}(t_1,t_2) \\ &\quad = 2K_{q_1}(t_1)\mathrm{Im}(K_{q_1}(t_2))\mathcal{L}_{q_2q_3}(t_1,t_2) + 2\mathrm{Im}(K_{q_2}(t_1)K^*_{q_2}(t_2))\mathcal{L}_{q_3q_1}(t_2,t_1) \\ &\qquad - 2K_{q_3}(t_2)\mathrm{Im}(K_{q_3}(t_1))\mathcal{L}_{q_1q_2}(t_1,t_2), \end{aligned} \tag{D.12}$$

and vanishes thanks to $\mathcal{L}_2 = 0$ for loop (c). Similarly, the difference between initially cutting loop (b) or loop (a) with opposite orientation also vanishes thanks to $\mathcal{L}_2 = 0$.

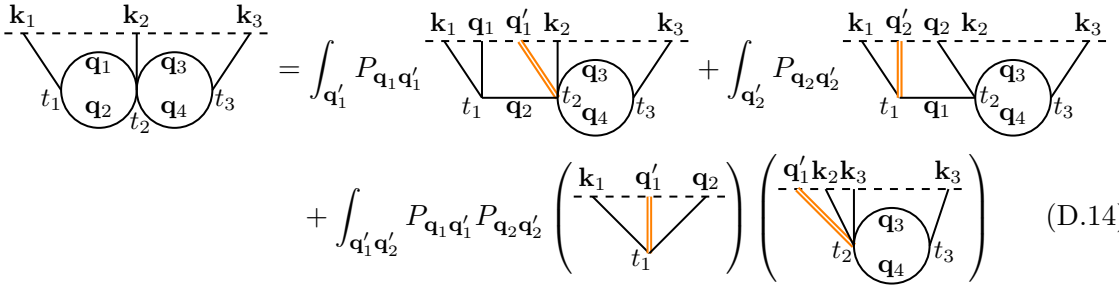
Coincident times. Finally, we point out that many higher-loop diagrams can be viewed as one-loop diagrams in which two or more vertices are identified. The consistency of cutting such higher-loop diagrams in different ways follows automatically follows from the consistency of cutting a one-loop diagram in either orientation (guaranteed by the tree-level identity (2.30)).

To see this, it is useful to introduce a diagrammatic representation for the time integrand. For instance, consider the following two-loop correction to the bispectrum,



(D.13)

We can apply the tree theorem (2.42) to either the left or right loop, and with either orientation. For instance, expanding the left loop using (2.40) gives,



(D.14)

The remaining loop can now be expanded, again with either orientation. There are therefore four possible representations of this diagram in terms of only tree-level diagrams, which we denote by:

(CC) Cut both loops clockwise, using $\mathcal{L}_{q_1q_2}(t_1,t_2)\mathcal{L}_{q_3q_4}(t_2,t_3) = 0$,

(CA) Cut the left (right) loop clockwise (anticlockwise), using $\mathcal{L}_{q_1q_2}(t_1,t_2)\mathcal{L}_{q_4q_3}(t_3,t_2) = 0$,

(AC) Cut the left (right) loop anticlockwise (clockwise), using $\mathcal{L}_{q_2q_1}(t_2,t_1)\mathcal{L}_{q_3q_4}(t_2,t_3) = 0$

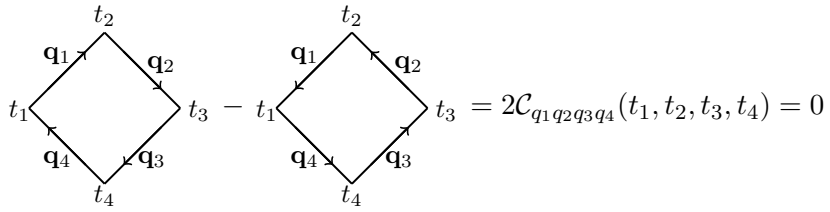
(AA) Cut both loops anticlockwise, using $\mathcal{L}_{q_2q_1}(t_2,t_1)\mathcal{L}_{q_4q_3}(t_3,t_2) = 0$

JHEP12(2023)076

To show that these different representations must be equivalent, it suffices to notice that this diagram can be constructed from a single four-vertex loop by identifying two opposite vertices. We have already shown that the difference between cutting such a one-loop diagram either clockwise or anticlockwise is given by,

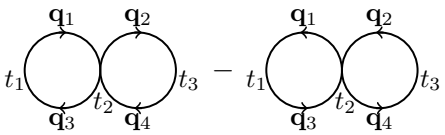
$$\mathcal{L}_{q_1q_2q_3q_4}(t_1,t_2,t_3,t_4) - \mathcal{L}_{q_1q_2q_3q_4}(t_2,t_1,t_4,t_3) = 2\mathcal{C}_{q_1q_3q_4q_2}(t_1,t_2,t_3,t_4) = 0 \tag{D.15}$$

and vanishes thanks to (2.30). Diagrammatically, we can represent this as,



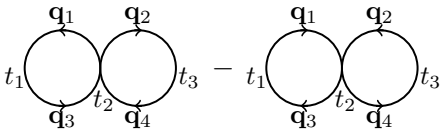
$$= 2\mathcal{C}_{q_1q_2q_3q_4}(t_1,t_2,t_3,t_4) = 0 \tag{D.16}$$

where we do not include the external lines since this is a propagator identity which may be used for any external lines. We therefore see that identifying $t_2 = t_4$ produces,



$$= 2\mathcal{C}_{q_1q_3q_4q_2}(t_1,t_2,t_3,t_2) = 0 \tag{D.17}$$

and guarantees that the (CC) and (AA) representations of the diagram (D.13) above are equivalent. A relabelling of the momenta in (D.16) leads analogously to,



$$= 2\mathcal{C}_{q_1q_3q_2q_4}(t_1,t_2,t_3,t_2) = 0 \tag{D.18}$$

which guarantees that (AC) and (CA) are equivalent. Finally, (CC) and (AC) are equivalent thanks to consistency of cutting a single two-vertex loop in either orientation, namely,

$$(\mathcal{L}_{q_1p_4}(t_1,t_2) - \mathcal{L}_{q_4p_1}(t_2,t_1))\mathcal{L}_{q_2q_3}(t_2,t_3) = 2\mathcal{C}_{q_1q_2}(t_1,t_2)\mathcal{L}_{q_2q_3}(t_2,t_3) = 0\,. \tag{D.19}$$

D.3 Finite time and restricted integration regions

Finally, we discuss how the various propagator identities can be used to constrain other objects, such as the wavefunction evaluated at a finite time or with local contributions removed. This is straightforward mathematically: rather than integrate $\mathcal{L}_n = 0$ over all $-\infty < t < 0$ as in (2.59), we may instead integrate over any desired region R,

$$\left[\prod_{a=1}^{N}\int_R dt_a K_{k_a}(t_a)\right]\mathcal{L}_{p_1\ldots p_N}(t_1,\ldots,t_N) = 0\,. \tag{D.20}$$

We will now describe three particularly interesting choices of R.

JHEP12(2023)076

Bulk wavefunction. Consider $\int_R dt_a = \int_{-\infty}^{t} dt_a$ where t is a fixed time in the bulk. In that case, the causal identity (2.59) implies a tree theorem for the bulk wavefunction coefficients, $\psi_n(t)$, which are the coefficients appearing in the Bunch-Davies wavefunction at time t. The derivation from the main text is essentially unchanged, and for instance the identity (1.28) for the quadratic wavefunction coefficient becomes,

$$-2\psi^{\text{1-loop}}_{\mathbf{k}_1\mathbf{k}_2}(t) = \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \operatorname*{disc}_{q'} \left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}\mathbf{q}'}(t)\right] + \int_{\substack{\mathbf{q}_1\mathbf{q}_1' \\ \mathbf{q}_2\mathbf{q}_2'}} P_{\mathbf{q}_1\mathbf{q}_1'} P_{\mathbf{q}_2\mathbf{q}_2'} \operatorname*{disc}_{q_2'} \left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}_2'}(t)\right] \operatorname*{disc}_{q_1'} \left[\psi^{\text{tree}}_{\mathbf{k}_2\mathbf{q}_2\mathbf{q}_1'}(t)\right] . \tag{D.21}$$

The cutting rules for $\text{Disc}\psi_n^{1-\text{loop}}(t)$ at a finite time follow from the conservation of the total probability under this interacting time evolution [104, 106]. The finite-time tree theorem, on the other hand, follows simply from the perturbative form of the Schrodinger equation (C.4), in which both exchange and loop corrections source the time evolution of the wavefunction phase equally (and so one can be traded for the other).

Boundary divergences. Now consider $\int_R dt_a = \lim_{t_0\to 0}\int_{-\infty}^{t_0} dt_a$, where t_0 is an infinitesimal cut-off which can be used to regulate otherwise divergent interactions. Since the Cosmological Tree Theorem can be applied at any finite t, it also applies order-by-order in t_0: in particular, it can be used to express the late-time divergence of certain loop diagrams in terms of the late-time divergence of their cut tree graphs. As an example, consider a π^4 interaction for a massless scalar field on de Sitter. The two- and four-point wavefunction coefficients are given at late times by,

$$\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}(t_0) = -\frac{(k_1^3+k_2^3+k_3^3+k_4^3)\log(-it_0(k_1+k_2+k_3+k_4))}{3H^4} + \mathcal{O}(1) \tag{D.22}$$

$$\psi^{1-\text{loop}}_{\mathbf{k}_1\mathbf{k}_2}(t_0) = \int_{\mathbf{q}} \frac{(k_1^3+k_2^3+2q^3)\log(-it_0(k_1+k_2+2q)) + (-k_1^3-k_2^3)\log(-it_0(k_1+k_2))}{6H^2q^3} + \mathcal{O}(1) \tag{D.23}$$

where we have focused on the logarithmic divergence as $t_0 \to 0$ (dropping local power-law divergences which are trivially renormalised) and suppressed the overall momentum-conserving δ function. One can verify that (D.22) indeed satisfies the tree theorem (D.21), i.e. the one-loop integrand diverges at late-times in precisely the same way as the tree-level four-point function.

In general, such a boundary divergence appears on de Sitter at particular "extremal" values of the external masses such as $\sum_{b=1}^{n}\Delta_b^{\pm} = d+2\ell$, where $\Delta_b^{\pm} = \frac{d}{2}\pm\sqrt{\frac{d^2}{4}-m^2}$, d is the number of spatial dimensions, and ℓ is an integer which counts the number of derivatives in the interaction. For massless scalar fields, this condition is met for all n when $\ell = 0$, signalling that all π^n interactions have such a logarithmic divergence as $t_0 \to 0$. Such divergences can be renormalised using a suitable choice of boundary operators [71], related to the original bulk fields via a Boundary Operator Expansion as described in [104]. Our Cosmological Tree Theorem expresses this logarithmic divergence in terms of (a momentum integral of) tree-level data, and it would be interesting to further explore the implications of this for the BOE or for dynamical RG.

JHEP12(2023)076

Local contributions and residuals. Finally, consider the integration region in which coincident times have been removed. This allows us to apply the tree theorem to wavefunction coefficients in which certain local-type interactions have been removed. To be precise, if a particular diagram D corresponds to the following time integral,

$$D \equiv \int_{-\infty}^{0} dt_1 \int_{-\infty}^{0} dt_2 \; \mathcal{D}(t_1, t_2) \tag{D.24}$$

then we define $\hat{D}$ to be the same integral but with coincident points removed,

$$\hat{D} \equiv \int_{\substack{t_a<0 \\ t_1 \neq t_2}} dt_1 dt_2 \; \mathcal{D}(t_1, t_2) \; . \tag{D.25}$$

This definition is useful because the difference, which we refer to as the "residual" of the diagram,

$$\mathcal{R}\left[D\right] \equiv D - \hat{D} \tag{D.26}$$

can often be written in terms of a simpler diagram in which internal lines have been contracted.

As an example, consider the interaction $\frac{a(t)}{3!}\lambda\dot{\pi}^3$ for a massless scalar field on de Sitter. This contributes to ψ_4^{tree} via the exchange diagram (1.10), which in this case is given by,

$$D \equiv i\lambda^2 \int_{t_1 t_2} \frac{K'_{k_1}(t_1) K'_{k_2}(t_1) K'_{k_3}(t_2) K'_{k_4}(t_2) \partial_{t_1}\partial_{t_2} G_{p_s}(t_1, t_2)}{H^2 t_1 t_2} \, . \tag{D.27}$$

One common method for performing this integral first assumes an ordering such as $t_1 > t_2$ and then symmetrises the result. This will neglect the equal-time contributions from the $\delta(t_1 - t_2)$ functions inside $\partial_{t_1}\partial_{t_2} G_p(t_1, t_2)$ (from the derivatives acting on the $\Theta(t_1 - t_2)$), and corresponds to evaluating this time integral over the restricted domain R, namely:

$$\hat{D} = i\lambda^2 \int_{t_1 > t_2} \frac{K'_{k_1}(t_1) K'_{k_2}(t_1) K'_{k_3}(t_2) K'_{k_4}(t_2) \partial_{t_1}\partial_{t_2} G_{p_s}(t_1, t_2)}{H^2 t_1 t_2} + (t_1 \leftrightarrow t_2) \; . \tag{D.28}$$

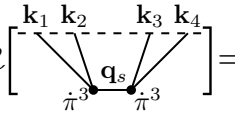
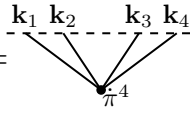
Another method is to integrate by parts (using that $G_p(0, t) = 0$),

$$D = i\lambda^2 \int_{t_1 t_2} \partial_{t_1} \left(K'_{k_1}(t_1) K'_{k_2}(t_1) \frac{1}{H t_1} \right) G_{p_s}(t_1, t_2) \partial_{t_2} \left(K'_{k_1}(t_2) K'_{k_2}(t_2) \frac{1}{H t_2} \right) ,$$

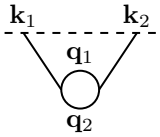
so that now there are no δ function contributions in the integrand. The difference between these two methods is a simple total energy pole,

$$D - \hat{D} = -\frac{24\lambda^2 k_1^2 k_2^2 k_3^2 k_4^2}{(k_1 + k_2 + k_3 + k_4)^5} \tag{D.29}$$

which can be absorbed into a local interaction like $\dot{\pi}^4$. Diagrammatically, the residual of this exchange diagram is therefore the contact diagram given by collapsing the internal line,

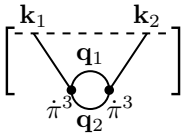
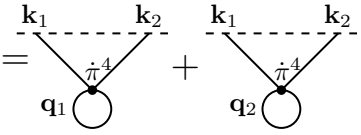
$\mathcal{R}\Big[$  $\Big] =$  , (D.30)

where the $\dot\pi^4$ vertex factor is $\lambda^2/4!$. Similarly, the one-loop correction to ψ_2 from this interaction



$$= \lambda^2 \int_{\mathbf{q}_1\mathbf{q}_2} \int_{t_1 t_2} \frac{K'_{k_1}(t_1)\partial_{t_1}\partial_{t_2} G_{q_1}(t_1,t_2)\partial_{t_1}\partial_{t_2} G_{q_2}(t_1,t_2) K'_{k_2}(t_2)}{H^2 t_1 t_2} \,, \tag{D.31}$$

contains coincident $\delta(t_1 - t_2)$ singularities, and whether or not these are included in the integration region corresponds to the freedom to add a local $\dot\pi^4$ interaction (again with coefficient $\lambda^2/4!$),

$\mathcal{R}\Big[$  $\Big] =$ 

$$= \frac{3\lambda^2 H^2}{4}\left[\int_{\mathbf{q}_1}\left(\frac{k^4 q_1}{(k+q_1)^5} - \frac{q_1}{k}\right) + \int_{\mathbf{q}_2}\left(\frac{k^4 q_2}{(k+q_2)^5} - \frac{q_2}{k}\right)\right] , \tag{D.32}$$

where in the second line we have written $|\mathbf{k}_1| = |\mathbf{k}_2| = k$ and removed the overall momentum-conserving δ-function.

Since we can apply the identity (D.20) on both integration domains (D.24) and (D.25), we can apply the tree theorem to both D and $\hat D$ separately. In particular, if we define $\hat\psi_n$ as the sum over all $\hat D$ with n external lines (and which therefore differs from ψ_n only by the inclusion of coincident points in the time integrals), we have an analogous tree theorem for $\hat\psi_2$,

$$-2\hat\psi^{\text{1-loop}}_{\mathbf{k}_1\mathbf{k}_2} = \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \operatorname*{disc}_{q'}\left[\hat\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}\mathbf{q}'}\right] + \int_{\substack{\mathbf{q}_1\mathbf{q}'_1\\ \mathbf{q}_2\mathbf{q}'_2}} P_{\mathbf{q}_1\mathbf{q}'_1} P_{\mathbf{q}_2\mathbf{q}'_2} \operatorname*{disc}_{q'_2}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{q}_1\mathbf{q}'_2}\right] \operatorname*{disc}_{q'_1}\left[\psi^{\text{tree}}_{\mathbf{k}_2\mathbf{q}_2\mathbf{q}'_1}\right] \,, \tag{D.33}$$

where ψ_3^{tree} does not require a hat since it only contains a single time integral (so there is no coincident point to remove). If we subtract this from the tree theorem (1.28) for ψ_2, we find that the residuals obey a simple relation,

$$\mathcal{R}\left[2\psi^{\text{1-loop}}_{\mathbf{k}_1\mathbf{k}_2} + \int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'} \operatorname*{disc}_{q'}\left[\psi^{\text{tree}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}\mathbf{q}'}\right]\right] = 0 \tag{D.34}$$

where we have used that the $\mathcal{R}$ and disc operations commute. Comparing (D.30) and (D.32), we see that the wavefunction of our previous $\dot\pi^3$ example indeed satisfies this relation. Causality guarantees that this should hold more generally, for any interaction on any time-dependent background.

Finally, note that in appendix C of [106] the above one-loop diagram from $\dot\pi^3 \times \dot\pi^3$ was computed without including the local contact contributions at $t_1 = t_2$, and hence differs from the result (2.44) presented here. It was not necessary for [106] to include the residual terms since they have vanishing discontinuity: for this particular diagram, Disc D = Disc $\hat D$. This is most easily seen from (D.32), which expresses the residual terms as simpler one-loop diagrams, followed by our tree theorem, which expresses these one-vertex loops as contact diagrams (which have vanishing Disc).

JHEP12(2023)076

E Cosmological loop-tree duality

Finally, we establish an explicit connection with the "loop-tree duality" (LTD), a modern descendent of Feynman's tree theorem. The main idea, due to Catani et al. [149], is that any one-loop diagram can be written as a sum over single-cut (tree) diagrams, where a diagram in which the line i is cut uses the modified propagator,

$$\tilde{G}^i_j = G_j + \omega^i_j H_j \,. \tag{E.1}$$

For amplitudes, $H_j \sim K_j K_j^*$, and for wavefunction coefficients $H_j \sim K_j \mathrm{Im}\, K_j$ (where the time arguments are distinguished by going either clockwise or anticlockwise around the loop).

The key to this construction is that the ω^i_j coefficients obey the relations,

$$\sum_i^N \prod_{j\neq i} \omega^i_j = 1 \,, \tag{E.2}$$

where N is any subset of the edges in the loop. For instance, [149] use $\omega^i_j = \Theta(q^0_j - q^0_i)$, where q_j is the 4-momentum of the j^{th} internal line, since it is not hard to show that energy conservation then implies (E.2). For instance when $N = 2$,

$$\omega^2_1 + \omega^1_2 = \Theta(k^0) + \Theta(-k^0) = 1 \tag{E.3}$$

where $\{+k, -k\}$ are the momenta of the external lines. When $N = 3$,

$$\omega^3_1\omega^3_2 + \omega^2_1\omega^2_3 + \omega^1_2\omega^1_3 = \Theta(k^0_1)\Theta(-k^0_3) + \Theta(-k^0_2)\Theta(k^0_3) + \Theta(k^0_2)\Theta(-k^0_1) = 1 \tag{E.4}$$

since $k^0_1 + k^0_2 + k^0_3 = 0$.

For instance, the LTD for the triangle graph follows from expanding,

$$\begin{aligned}\tilde{G}^3_1\tilde{G}^3_2 H_3 + 2 \text{ perm.} = {} & G_1 H_2 H_3 \left(\omega^3_2 + \omega^2_3\right) + 2 \text{ perm.} \\ & + H_1 H_2 H_3 \left(\omega^3_1\omega^3_2 + \omega^2_1\omega^2_3 + \omega^1_2\omega^1_3\right)\end{aligned} \tag{E.5}$$

and then using (E.2) to set each ω^i_j sum to unity.

So in order to use the LTD for the wavefunction, one needs to introduce a set of ω^i_j which obey (E.2). Since energy is no longer conserved, the Catani choice of $\Theta(q^0_j - q^0_i)$ will not satisfy (E.2). However, the spatial momentum in any fixed direction is conserved, so an alternative modification of the propagator would be to introduce a fixed reference vector $\hat{n}$ and use $\Theta\left(\hat{n}\cdot(\mathbf{q}_j - \mathbf{q}_i)\right)$ instead. We have not explored this particularly thoroughly, and it would be interesting to develop this direction further in future, particularly in light of the recent progress which has been made in implementing the LTD both numerically [156–160] and analytically for diagrams with up to five loops [177].

JHEP12(2023)076

References

[1] S. Weinberg, *Quantum contributions to cosmological correlations*, *Phys. Rev. D* **72** (2005) 043514 [hep-th/0506236] [INSPIRE].

[2] D. Boyanovsky, H.J. de Vega and N.G. Sanchez, *Quantum corrections to the inflaton potential and the power spectra from superhorizon modes and trace anomalies*, *Phys. Rev. D* **72** (2005) 103006 [astro-ph/0507596] [INSPIRE].

[3] D. Boyanovsky, H.J. de Vega and N.G. Sanchez, *Quantum corrections to slow roll inflation and new scaling of superhorizon fluctuations*, *Nucl. Phys. B* **747** (2006) 25 [astro-ph/0503669] [INSPIRE].

[4] S. Weinberg, *Quantum contributions to cosmological correlations. II. Can these corrections become large?*, *Phys. Rev. D* **74** (2006) 023508 [hep-th/0605244] [INSPIRE].

[5] M.S. Sloth, *On the one loop corrections to inflation and the CMB anisotropies*, *Nucl. Phys. B* **748** (2006) 149 [astro-ph/0604488] [INSPIRE].

[6] M.S. Sloth, *On the one loop corrections to inflation. II. The consistency relation*, *Nucl. Phys. B* **775** (2007) 78 [hep-th/0612138] [INSPIRE].

[7] A. Bilandzic and T. Prokopec, *Quantum radiative corrections to slow-roll inflation*, *Phys. Rev. D* **76** (2007) 103507 [arXiv:0704.1905] [INSPIRE].

[8] D. Seery, *One-loop corrections to the curvature perturbation from inflation*, *JCAP* **02** (2008) 006 [arXiv:0707.3378] [INSPIRE].

[9] D. Seery, *One-loop corrections to a scalar field during inflation*, *JCAP* **11** (2007) 025 [arXiv:0707.3377] [INSPIRE].

[10] M. van der Meulen and J. Smit, *Classical approximation to quantum cosmological correlations*, *JCAP* **11** (2007) 023 [arXiv:0707.0842] [INSPIRE].

[11] P. Adshead, R. Easther and E.A. Lim, *Cosmology with many light scalar fields: stochastic inflation and loop corrections*, *Phys. Rev. D* **79** (2009) 063504 [arXiv:0809.4008] [INSPIRE].

[12] P. Adshead, R. Easther and E.A. Lim, *The 'in-in' formalism and cosmological perturbations*, *Phys. Rev. D* **80** (2009) 083521 [arXiv:0904.4207] [INSPIRE].

[13] L. Senatore and M. Zaldarriaga, *On loops in inflation*, *JHEP* **12** (2010) 008 [arXiv:0912.2734] [INSPIRE].

[14] X. Chen, Y. Wang and Z.-Z. Xianyu, *Loop corrections to standard model fields in inflation*, *JHEP* **08** (2016) 051 [arXiv:1604.07841] [INSPIRE].

[15] A. Ota, M. Sasaki and Y. Wang, *One-loop tensor power spectrum from an excited scalar field during inflation*, *Phys. Rev. D* **108** (2023) 043542 [arXiv:2211.12766] [INSPIRE].

[16] K. Inomata, M. Braglia, X. Chen and S. Renaux-Petel, *Questions on calculation of primordial power spectrum with large spikes: the resonance model case*, *JCAP* **04** (2023) 011 [*Erratum ibid.* **09** (2023) E01] [arXiv:2211.02586] [INSPIRE].

[17] H. Firouzjahi, *One-loop corrections in power spectrum in single field inflation*, *JCAP* **10** (2023) 006 [arXiv:2303.12025] [INSPIRE].

[18] J. Fumagalli et al., *One-loop infrared rescattering by enhanced scalar fluctuations during inflation*, arXiv:2307.08358 [INSPIRE].

JHEP12(2023)076

[19] A. Riotto and M.S. Sloth, *On resumming inflationary perturbations beyond one-loop*, *JCAP* **04** (2008) 030 [arXiv:0801.1845] [INSPIRE].

[20] V. Assassi, D. Baumann and D. Green, *Symmetries and loops in inflation*, *JHEP* **02** (2013) 151 [arXiv:1210.7792] [INSPIRE].

[21] V. Gorbenko and L. Senatore, $\lambda\phi^4$ *in dS*, arXiv:1911.00022 [INSPIRE].

[22] D. Green and A. Premkumar, *Dynamical RG and critical phenomena in de Sitter space*, *JHEP* **04** (2020) 064 [arXiv:2001.05974] [INSPIRE].

[23] T. Cohen, D. Green, A. Premkumar and A. Ridgway, *Stochastic inflation at NNLO*, *JHEP* **09** (2021) 159 [arXiv:2106.09728] [INSPIRE].

[24] A. Premkumar, *Regulating loops in dS*, arXiv:2110.12504 [INSPIRE].

[25] T. Cohen, D. Green and A. Premkumar, *A tail of eternal inflation*, *SciPost Phys.* **14** (2023) 109 [arXiv:2111.09332] [INSPIRE].

[26] L.-T. Wang, Z.-Z. Xianyu and Y.-M. Zhong, *Precision calculation of inflation correlators at one loop*, *JHEP* **02** (2022) 085 [arXiv:2109.14635] [INSPIRE].

[27] Z.-Z. Xianyu and H. Zhang, *Bootstrapping one-loop inflation correlators with the spectral decomposition*, *JHEP* **04** (2023) 103 [arXiv:2211.03810] [INSPIRE].

[28] T. Heckelbacher, I. Sachs, E. Skvortsov and P. Vanhove, *Analytical evaluation of cosmological correlation functions*, *JHEP* **08** (2022) 139 [arXiv:2204.07217] [INSPIRE].

[29] H. Motohashi and Y. Tada, *Squeezed bispectrum and one-loop corrections in transient constant-roll inflation*, *JCAP* **08** (2023) 069 [arXiv:2303.16035] [INSPIRE].

[30] L. Iacconi and D.J. Mulryne, *Multi-field inflation with large scalar fluctuations: non-Gaussianity and perturbativity*, *JCAP* **09** (2023) 033 [arXiv:2304.14260] [INSPIRE].

[31] M.H.G. Lee, C. McCulloch and E. Pajer, *Leading loops in cosmological correlators*, *JHEP* **11** (2023) 038 [arXiv:2305.11228] [INSPIRE].

[32] M. Sasaki, T. Suyama, T. Tanaka and S. Yokoyama, *Primordial black holes — perspectives in gravitational wave astronomy*, *Class. Quant. Grav.* **35** (2018) 063001 [arXiv:1801.05235] [INSPIRE].

[33] J. Kristiano and J. Yokoyama, *Ruling out primordial black hole formation from single-field inflation*, arXiv:2211.03395 [INSPIRE].

[34] A. Riotto, *The primordial black hole formation from single-field inflation is not ruled out*, arXiv:2301.00599 [INSPIRE].

[35] J. Kristiano and J. Yokoyama, *Response to criticism on "Ruling out primordial black hole formation from single-field inflation": a note on bispectrum and one-loop correction in single-field inflation with primordial black hole formation*, arXiv:2303.00341 [INSPIRE].

[36] A. Riotto, *The primordial black hole formation from single-field inflation is still not ruled out*, arXiv:2303.01727 [INSPIRE].

[37] S. Choudhury, M.R. Gangopadhyay and M. Sami, *No-go for the formation of heavy mass primordial black holes in single field inflation*, arXiv:2301.10000 [INSPIRE].

[38] S. Choudhury, S. Panda and M. Sami, *PBH formation in EFT of single field inflation with sharp transition*, *Phys. Lett. B* **845** (2023) 138123 [arXiv:2302.05655] [INSPIRE].

JHEP12(2023)076

[39] G. Franciolini, J.A. Iovino, M. Taoso and A. Urbano, *One loop to rule them all: perturbativity in the presence of ultra slow-roll dynamics*, arXiv:2305.03491 [INSPIRE].

[40] S. Choudhury, S. Panda and M. Sami, *Quantum loop effects on the power spectrum and constraints on primordial black holes*, *JCAP* **11** (2023) 066 [arXiv:2303.06066] [INSPIRE].

[41] S. Choudhury, S. Panda and M. Sami, *Galileon inflation evades the no-go for PBH formation in the single-field framework*, *JCAP* **08** (2023) 078 [arXiv:2304.04065] [INSPIRE].

[42] K.N. Ananda, C. Clarkson and D. Wands, *The cosmological gravitational wave background from primordial density perturbations*, *Phys. Rev. D* **75** (2007) 123518 [gr-qc/0612013] [INSPIRE].

[43] D. Baumann, P.J. Steinhardt, K. Takahashi and K. Ichiki, *Gravitational wave spectrum induced by primordial scalar perturbations*, *Phys. Rev. D* **76** (2007) 084019 [hep-th/0703290] [INSPIRE].

[44] J. Fumagalli, S. Renaux-Petel and L.T. Witkowski, *Oscillations in the stochastic gravitational wave background from sharp features and particle production during inflation*, *JCAP* **08** (2021) 030 [arXiv:2012.02761] [INSPIRE].

[45] J. Fumagalli et al., *Primordial gravitational waves from excited states*, *JHEP* **03** (2022) 196 [arXiv:2111.14664] [INSPIRE].

[46] V. Aragam, S. Paban and R. Rosati, *Primordial stochastic gravitational wave backgrounds from a sharp feature in three-field inflation. Part I. The radiation era*, *JCAP* **11** (2023) 014 [arXiv:2304.00065] [INSPIRE].

[47] A. Strominger, *Inflation and the dS/CFT correspondence*, *JHEP* **11** (2001) 049 [hep-th/0110087] [INSPIRE].

[48] T. Heckelbacher and I. Sachs, *Loops in dS/CFT*, *JHEP* **02** (2021) 151 [arXiv:2009.06511] [INSPIRE].

[49] J.M. Penín, K. Skenderis and B. Withers, *Massive holographic QFTs in de Sitter*, *SciPost Phys.* **12** (2022) 182 [arXiv:2112.14639] [INSPIRE].

[50] S. Fichet, *On holography in general background and the boundary effective action from AdS to dS*, *JHEP* **07** (2022) 113 [arXiv:2112.00746] [INSPIRE].

[51] L.H. Ford, *Quantum instability of de Sitter space-time*, *Phys. Rev. D* **31** (1985) 710 [INSPIRE].

[52] I. Antoniadis, J. Iliopoulos and T.N. Tomaras, *Quantum instability of de Sitter space*, *Phys. Rev. Lett.* **56** (1986) 1319 [INSPIRE].

[53] A.A. Starobinsky and J. Yokoyama, *Equilibrium state of a selfinteracting scalar field in the de Sitter background*, *Phys. Rev. D* **50** (1994) 6357 [astro-ph/9407016] [INSPIRE].

[54] A. Dolgov and D.N. Pelliccia, *Scalar field instability in de Sitter space-time*, *Nucl. Phys. B* **734** (2006) 208 [hep-th/0502197] [INSPIRE].

[55] D. Marolf and I.A. Morrison, *The IR stability of de Sitter: loop corrections to scalar propagators*, *Phys. Rev. D* **82** (2010) 105032 [arXiv:1006.0035] [INSPIRE].

[56] C.P. Burgess, R. Holman, L. Leblond and S. Shandera, *Breakdown of semiclassical methods in de Sitter space*, *JCAP* **10** (2010) 017 [arXiv:1005.3551] [INSPIRE].

JHEP12(2023)076

[57] J.B. Hartle and S.W. Hawking, *Wave function of the universe*, *Phys. Rev. D* **28** (1983) 2960 [INSPIRE].

[58] N. Arkani-Hamed and J. Maldacena, *Cosmological collider physics*, arXiv:1503.08043 [INSPIRE].

[59] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne, *The analytic S-matrix*, Cambridge University Press, Cambridge, U.K. (1966).

[60] M. Kruczenski, J. Penedones and B.C. van Rees, *Snowmass white paper: S-matrix bootstrap*, arXiv:2203.02421 [INSPIRE].

[61] Z. Bern and J. Trnka, *Snowmass TF04 report: scattering amplitudes and their applications*, arXiv:2210.03146 [INSPIRE].

[62] M. Baumgart et al., *Snowmass theory frontier: effective field theory*, in the proceedings of the *Snowmass* 2021, (2022) [arXiv:2210.03199] [INSPIRE].

[63] C. de Rham et al., *Snowmass white paper: UV constraints on IR physics*, in the proceedings of the *Snowmass* 2021, (2022) [arXiv:2203.06805] [INSPIRE].

[64] N. Arkani-Hamed, D. Baumann, H. Lee and G.L. Pimentel, *The cosmological bootstrap: inflationary correlators from symmetries and singularities*, *JHEP* **04** (2020) 105 [arXiv:1811.00024] [INSPIRE].

[65] D. Baumann et al., *The cosmological bootstrap: weight-shifting operators and scalar seeds*, *JHEP* **12** (2020) 204 [arXiv:1910.14051] [INSPIRE].

[66] D. Baumann et al., *The cosmological bootstrap: spinning correlators from symmetries and factorization*, *SciPost Phys.* **11** (2021) 071 [arXiv:2005.04234] [INSPIRE].

[67] D. Baumann et al., *Linking the singularities of cosmological correlators*, *JHEP* **09** (2022) 010 [arXiv:2106.05294] [INSPIRE].

[68] P. McFadden and K. Skenderis, *Holography for cosmology*, *Phys. Rev. D* **81** (2010) 021301 [arXiv:0907.5542] [INSPIRE].

[69] P. McFadden and K. Skenderis, *Holographic non-Gaussianity*, *JCAP* **05** (2011) 013 [arXiv:1011.0452] [INSPIRE].

[70] A. Bzowski, P. McFadden and K. Skenderis, *Implications of conformal invariance in momentum space*, *JHEP* **03** (2014) 111 [arXiv:1304.7760] [INSPIRE].

[71] A. Bzowski, P. McFadden and K. Skenderis, *Scalar 3-point functions in CFT: renormalisation, beta functions and anomalies*, *JHEP* **03** (2016) 066 [arXiv:1510.08442] [INSPIRE].

[72] D. Baumann et al., *Snowmass white paper: the cosmological bootstrap*, in the proceedings of the *Snowmass* 2021, (2022) [arXiv:2203.08121] [INSPIRE].

[73] H. Gomez, R.L. Jusinskas and A. Lipstein, *Cosmological scattering equations*, *Phys. Rev. Lett.* **127** (2021) 251604 [arXiv:2106.11903] [INSPIRE].

[74] H. Gomez, R. Lipinski Jusinskas and A. Lipstein, *Cosmological scattering equations at tree-level and one-loop*, *JHEP* **07** (2022) 004 [arXiv:2112.12695] [INSPIRE].

[75] D. Anninos, T. Anous, D.Z. Freedman and G. Konstantinidis, *Late-time structure of the Bunch-Davies de Sitter wavefunction*, *JCAP* **11** (2015) 048 [arXiv:1406.5490] [INSPIRE].

JHEP12(2023)076

[76] G. Konstantinidis, R. Mahajan and E. Shaghoulian, *Late-time structure of the Bunch-Davies FRW wavefunction*, *JHEP* **10** (2016) 103 [arXiv:1608.06163] [INSPIRE].

[77] D. Baumann, D. Green and T. Hartman, *Dynamical constraints on RG flows and cosmology*, *JHEP* **12** (2019) 134 [arXiv:1906.10226] [INSPIRE].

[78] C. Sleight, *A Mellin space approach to cosmological correlators*, *JHEP* **01** (2020) 090 [arXiv:1906.12302] [INSPIRE].

[79] C. Sleight and M. Taronna, *Bootstrapping inflationary correlators in Mellin space*, *JHEP* **02** (2020) 098 [arXiv:1907.01143] [INSPIRE].

[80] C. Sleight and M. Taronna, *From AdS to dS exchanges: spectral representation, Mellin amplitudes, and crossing*, *Phys. Rev. D* **104** (2021) L081902 [arXiv:2007.09993] [INSPIRE].

[81] C. Sleight and M. Taronna, *From dS to AdS and back*, *JHEP* **12** (2021) 074 [arXiv:2109.02725] [INSPIRE].

[82] L. Di Pietro, V. Gorbenko and S. Komatsu, *Analyticity and unitarity for cosmological correlators*, *JHEP* **03** (2022) 023 [arXiv:2108.01695] [INSPIRE].

[83] M. Hogervorst, J. Penedones and K.S. Vaziri, *Towards the non-perturbative cosmological bootstrap*, *JHEP* **02** (2023) 162 [arXiv:2107.13871] [INSPIRE].

[84] J. Penedones, K. Salehi Vaziri and Z. Sun, *Hilbert space of quantum field theory in de Sitter spacetime*, arXiv:2301.04146 [INSPIRE].

[85] M. Loparco, J. Penedones, K. Salehi Vaziri and Z. Sun, *The Källén-Lehmann representation in de Sitter spacetime*, arXiv:2306.00090 [INSPIRE].

[86] E. Pajer, D. Stefanyszyn and J. Supeł, *The boostless bootstrap: amplitudes without Lorentz boosts*, *JHEP* **12** (2020) 198 [*Erratum ibid.* **04** (2022) 023] [arXiv:2007.00027] [INSPIRE].

[87] E. Pajer, *Building a boostless bootstrap for the bispectrum*, *JCAP* **01** (2021) 023 [arXiv:2010.12818] [INSPIRE].

[88] S. Jazayeri, E. Pajer and D. Stefanyszyn, *From locality and unitarity to cosmological correlators*, *JHEP* **10** (2021) 065 [arXiv:2103.08649] [INSPIRE].

[89] J. Bonifacio, E. Pajer and D.-G. Wang, *From amplitudes to contact cosmological correlators*, *JHEP* **10** (2021) 001 [arXiv:2106.15468] [INSPIRE].

[90] D. Meltzer, *The inflationary wavefunction from analyticity and factorization*, *JCAP* **12** (2021) 018 [arXiv:2107.10266] [INSPIRE].

[91] A. Hillman and E. Pajer, *A differential representation of cosmological wavefunctions*, *JHEP* **04** (2022) 012 [arXiv:2112.01619] [INSPIRE].

[92] Z. Qin and Z.-Z. Xianyu, *Helical inflation correlators: partial Mellin-Barnes and bootstrap equations*, *JHEP* **04** (2023) 059 [arXiv:2208.13790] [INSPIRE].

[93] X. Tong, Y. Wang and Y. Zhu, *Cutting rule for cosmological collider signals: a bulk evolution perspective*, *JHEP* **03** (2022) 181 [arXiv:2112.03448] [INSPIRE].

[94] Z. Qin and Z.-Z. Xianyu, *Inflation correlators at the one-loop order: nonanalyticity, factorization, cutting rule, and OPE*, *JHEP* **09** (2023) 116 [arXiv:2304.13295] [INSPIRE].

[95] G. Cabass, E. Pajer, D. Stefanyszyn and J. Supeł, *Bootstrapping large graviton non-Gaussianities*, *JHEP* **05** (2022) 077 [arXiv:2109.10189] [INSPIRE].

JHEP12(2023)076

[96] S. Jazayeri and S. Renaux-Petel, *Cosmological bootstrap in slow motion*, *JHEP* **12** (2022) 137 [arXiv:2205.10340] [INSPIRE].

[97] G. Cabass, S. Jazayeri, E. Pajer and D. Stefanyszyn, *Parity violation in the scalar trispectrum: no-go theorems and yes-go examples*, *JHEP* **02** (2023) 021 [arXiv:2210.02907] [INSPIRE].

[98] G.L. Pimentel and D.-G. Wang, *Boostless cosmological collider bootstrap*, *JHEP* **10** (2022) 177 [arXiv:2205.00013] [INSPIRE].

[99] D.-G. Wang, G.L. Pimentel and A. Achúcarro, *Bootstrapping multi-field inflation: non-Gaussianities from light scalars revisited*, *JCAP* **05** (2023) 043 [arXiv:2212.14035] [INSPIRE].

[100] J. Bonifacio et al., *The graviton four-point function in de Sitter space*, *JHEP* **06** (2023) 212 [arXiv:2212.07370] [INSPIRE].

[101] D. Ghosh, K. Panchal and F. Ullah, *Mixed graviton and scalar bispectra in the EFT of inflation: soft limits and boostless bootstrap*, *JHEP* **07** (2023) 233 [arXiv:2303.16929] [INSPIRE].

[102] S. Jazayeri, S. Renaux-Petel and D. Werth, *Shapes of the cosmological low-speed collider*, arXiv:2307.01751 [INSPIRE].

[103] C. Duhr, *QCD at NNLO and beyond*, *Nucl. Part. Phys. Proc.* **273-275** (2016) 2128 [INSPIRE].

[104] S. Céspedes, A.-C. Davis and S. Melville, *On the time evolution of cosmological correlators*, *JHEP* **02** (2021) 012 [arXiv:2009.07874] [INSPIRE].

[105] H. Goodhew, S. Jazayeri and E. Pajer, *The cosmological optical theorem*, *JCAP* **04** (2021) 021 [arXiv:2009.02898] [INSPIRE].

[106] S. Melville and E. Pajer, *Cosmological cutting rules*, *JHEP* **05** (2021) 249 [arXiv:2103.09832] [INSPIRE].

[107] H. Goodhew, S. Jazayeri, M.H. Gordon Lee and E. Pajer, *Cutting cosmological correlators*, *JCAP* **08** (2021) 003 [arXiv:2104.06587] [INSPIRE].

[108] D. Meltzer and A. Sivaramakrishnan, *CFT unitarity and the AdS Cutkosky rules*, *JHEP* **11** (2020) 073 [arXiv:2008.11730] [INSPIRE].

[109] N. Arkani-Hamed, P. Benincasa and A. Postnikov, *Cosmological polytopes and the wavefunction of the universe*, arXiv:1709.02813 [INSPIRE].

[110] N. Arkani-Hamed and P. Benincasa, *On the emergence of Lorentz invariance and unitarity from the scattering facet of cosmological polytopes*, arXiv:1811.01125 [INSPIRE].

[111] P. Benincasa, *From the flat-space S-matrix to the wavefunction of the universe*, arXiv:1811.02515 [INSPIRE].

[112] P. Benincasa, *Cosmological polytopes and the wavefuncton of the universe for light states*, arXiv:1909.02517 [INSPIRE].

[113] P. Benincasa, A.J. McLeod and C. Vergu, *Steinmann relations and the wavefunction of the universe*, *Phys. Rev. D* **102** (2020) 125004 [arXiv:2009.03047] [INSPIRE].

[114] P. Benincasa and W.J.T. Bobadilla, *Physical representations for scattering amplitudes and the wavefunction of the universe*, *SciPost Phys.* **12** (2022) 192 [arXiv:2112.09028] [INSPIRE].

JHEP12(2023)076

[115] S. Albayrak, P. Benincasa and C.D. Pueyo, *Perturbative unitarity and the wavefunction of the universe*, arXiv:2305.19686 [INSPIRE].

[116] A. Adams et al., *Causality, analyticity and an IR obstruction to UV completion*, *JHEP* **10** (2006) 014 [hep-th/0602178] [INSPIRE].

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

[118] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Positivity bounds for scalar field theories*, *Phys. Rev. D* **96** (2017) 081702 [arXiv:1702.06134] [INSPIRE].

[119] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *UV complete me: positivity bounds for particles with spin*, *JHEP* **03** (2018) 011 [arXiv:1706.02712] [INSPIRE].

[120] A.J. Tolley, Z.-Y. Wang and S.-Y. Zhou, *New positivity bounds from full crossing symmetry*, *JHEP* **05** (2021) 255 [arXiv:2011.02400] [INSPIRE].

[121] B. Bellazzini et al., *Positive moments for scattering amplitudes*, *Phys. Rev. D* **104** (2021) 036006 [arXiv:2011.00037] [INSPIRE].

[122] J. Davighi, S. Melville and T. You, *Natural selection rules: new positivity bounds for massive spinning particles*, *JHEP* **02** (2022) 167 [arXiv:2108.06334] [INSPIRE].

[123] X. Li et al., *Positivity in multifield effective field theories*, *Phys. Rev. Lett.* **127** (2021) 121601 [arXiv:2101.01191] [INSPIRE].

[124] D. Baumann, D. Green, H. Lee and R.A. Porto, *Signs of analyticity in single-field inflation*, *Phys. Rev. D* **93** (2016) 023523 [arXiv:1502.07304] [INSPIRE].

[125] T. Grall and S. Melville, *Positivity bounds without boosts: new constraints on low energy effective field theories from the UV*, *Phys. Rev. D* **105** (2022) L121301 [arXiv:2102.05683] [INSPIRE].

[126] P. Creminelli, O. Janssen and L. Senatore, *Positivity bounds on effective field theories with spontaneously broken Lorentz invariance*, *JHEP* **09** (2022) 201 [arXiv:2207.14224] [INSPIRE].

[127] S.A. Salcedo, M.H.G. Lee, S. Melville and E. Pajer, *The analytic wavefunction*, *JHEP* **06** (2023) 020 [arXiv:2212.08009] [INSPIRE].

[128] C.Y.-R. Chen, C. de Rham, A. Margalit and A.J. Tolley, *A cautionary case of casual causality*, *JHEP* **03** (2022) 025 [arXiv:2112.05031] [INSPIRE].

[129] C. de Rham, A.J. Tolley and J. Zhang, *Causality constraints on gravitational effective field theories*, *Phys. Rev. Lett.* **128** (2022) 131102 [arXiv:2112.05054] [INSPIRE].

[130] N. Arkani-Hamed, Y.-T. Huang, J.-Y. Liu and G.N. Remmen, *Causality, unitarity, and the weak gravity conjecture*, *JHEP* **03** (2022) 083 [arXiv:2109.13937] [INSPIRE].

[131] M. Carrillo Gonzalez, C. de Rham, V. Pozsgay and A.J. Tolley, *Causal effective field theories*, *Phys. Rev. D* **106** (2022) 105018 [arXiv:2207.03491] [INSPIRE].

[132] B. Bellazzini, G. Isabella, M. Lewandowski and F. Sgarlata, *Gravitational causality and the self-stress of photons*, *JHEP* **05** (2022) 154 [arXiv:2108.05896] [INSPIRE].

[133] B. Bellazzini, G. Isabella and M.M. Riva, *Classical vs quantum eikonal scattering and its causal structure*, *JHEP* **04** (2023) 023 [arXiv:2211.00085] [INSPIRE].

JHEP12(2023)076

[134] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez and D. Simmons-Duffin, *Causality constraints on corrections to Einstein gravity*, *JHEP* **05** (2023) 122 [arXiv:2201.06602] [INSPIRE].

[135] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez and D. Simmons-Duffin, *Graviton partial waves and causality in higher dimensions*, *Phys. Rev. D* **108** (2023) 026007 [arXiv:2205.01495] [INSPIRE].

[136] R.P. Feynman, *Quantum theory of gravitation*, *Acta Phys. Polon.* **24** (1963) 697 [INSPIRE].

[137] B.S. DeWitt, *Quantum theory of gravity. 2. The manifestly covariant theory*, *Phys. Rev.* **162** (1967) 1195 [INSPIRE].

[138] D. Vaman and Y.-P. Yao, *QCD recursion relations from the largest time equation*, *JHEP* **04** (2006) 030 [hep-th/0512031] [INSPIRE].

[139] A. Brandhuber, B. Spence and G. Travaglini, *From trees to loops and back*, *JHEP* **01** (2006) 142 [hep-th/0510253] [INSPIRE].

[140] S. Caron-Huot, *Loops and trees*, *JHEP* **05** (2011) 080 [arXiv:1007.3224] [INSPIRE].

[141] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys. B* **425** (1994) 217 [hep-ph/9403226] [INSPIRE].

[142] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl. Phys. B* **435** (1995) 59 [hep-ph/9409265] [INSPIRE].

[143] R. Britto, F. Cachazo and B. Feng, *Generalized unitarity and one-loop amplitudes in $N = 4$ super-Yang-Mills*, *Nucl. Phys. B* **725** (2005) 275 [hep-th/0412103] [INSPIRE].

[144] C. Anastasiou et al., *Unitarity cuts and reduction to master integrals in d dimensions for one-loop amplitudes*, *JHEP* **03** (2007) 111 [hep-ph/0612277] [INSPIRE].

[145] D. Forde, *Direct extraction of one-loop integral coefficients*, *Phys. Rev. D* **75** (2007) 125019 [arXiv:0704.1835] [INSPIRE].

[146] Z. Bern, L.J. Dixon and D.A. Kosower, *On-shell methods in perturbative QCD*, *Annals Phys.* **322** (2007) 1587 [arXiv:0704.2798] [INSPIRE].

[147] C.F. Berger and D. Forde, *Multi-parton scattering amplitudes via on-shell methods*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 181 [arXiv:0912.3534] [INSPIRE].

[148] S.D. Badger, *Direct extraction of one loop rational terms*, *JHEP* **01** (2009) 049 [arXiv:0806.4600] [INSPIRE].

[149] S. Catani et al., *From loops to trees by-passing Feynman's theorem*, *JHEP* **09** (2008) 065 [arXiv:0804.3170] [INSPIRE].

[150] G. Rodrigo et al., *From multileg loops to trees (by-passing Feynman's tree theorem)*, *Nucl. Phys. B Proc. Suppl.* **183** (2008) 262 [arXiv:0807.0531] [INSPIRE].

[151] I. Bierenbaum, S. Catani, P. Draggiotis and G. Rodrigo, *A tree-loop duality relation at two loops and beyond*, *JHEP* **10** (2010) 073 [arXiv:1007.0194] [INSPIRE].

[152] I. Bierenbaum, S. Catani, P. Draggiotis and G. Rodrigo, *Feynman's tree theorem and loop-tree dualities*, *PoS* **LC2010** (2010) 034 [arXiv:1011.0585] [INSPIRE].

[153] I. Bierenbaum et al., *Tree-loop duality relation beyond simple poles*, *JHEP* **03** (2013) 025 [arXiv:1211.5048] [INSPIRE].

JHEP12(2023)076

[154] S. Buchta et al., *On the singular behaviour of scattering amplitudes in quantum field theory*, *JHEP* **11** (2014) 014 [arXiv:1405.7850] [INSPIRE].

[155] S. Buchta, *Theoretical foundations and applications of the loop-tree duality in quantum field theories*, Ph.D. thesis, IFIC, Valencia U., Valencia, Spain (2015) [arXiv:1509.07167] [INSPIRE].

[156] S. Buchta, G. Chachamis, P. Draggiotis and G. Rodrigo, *Numerical implementation of the loop-tree duality method*, *Eur. Phys. J. C* **77** (2017) 274 [arXiv:1510.00187] [INSPIRE].

[157] R.J. Hernandez-Pinto, G.F.R. Sborlini and G. Rodrigo, *Towards gauge theories in four dimensions*, *JHEP* **02** (2016) 044 [arXiv:1506.04617] [INSPIRE].

[158] G.F.R. Sborlini, F. Driencourt-Mangin, R. Hernandez-Pinto and G. Rodrigo, *Four-dimensional unsubtraction from the loop-tree duality*, *JHEP* **08** (2016) 160 [arXiv:1604.06699] [INSPIRE].

[159] G.F.R. Sborlini, F. Driencourt-Mangin and G. Rodrigo, *Four-dimensional unsubtraction with massive particles*, *JHEP* **10** (2016) 162 [arXiv:1608.01584] [INSPIRE].

[160] F. Driencourt-Mangin, G. Rodrigo and G.F.R. Sborlini, *Universal dual amplitudes and asymptotic expansions for $gg \to H$ and $H \to \gamma\gamma$ in four dimensions*, *Eur. Phys. J. C* **78** (2018) 231 [arXiv:1702.07581] [INSPIRE].

[161] N.C. Tsamis and R.P. Woodard, *Quantum gravity slows inflation*, *Nucl. Phys. B* **474** (1996) 235 [hep-ph/9602315] [INSPIRE].

[162] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650 [INSPIRE].

[163] T.D. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, *Phys. Rev.* **133** (1964) B1549 [INSPIRE].

[164] C. Frye et al., *Infrared finiteness and forward scattering*, *Phys. Rev. D* **99** (2019) 056015 [arXiv:1810.10022] [INSPIRE].

[165] H. Hannesdottir and M.D. Schwartz, *S-matrix for massless particles*, *Phys. Rev. D* **101** (2020) 105001 [arXiv:1911.06821] [INSPIRE].

[166] S. Melville and G.L. Pimentel, *A de Sitter S-matrix for the masses*, to appear.

[167] C. Cheung et al., *The effective field theory of inflation*, *JHEP* **03** (2008) 014 [arXiv:0709.0293] [INSPIRE].

[168] K. Skenderis, *Lecture notes on holographic renormalization*, *Class. Quant. Grav.* **19** (2002) 5849 [hep-th/0209067] [INSPIRE].

[169] T. Grall and S. Melville, *Inflation in motion: unitarity constraints in effective field theories with (spontaneously) broken Lorentz symmetry*, *JCAP* **09** (2020) 017 [arXiv:2005.02366] [INSPIRE].

[170] S. Melville and J. Noller, *Positivity in the sky: constraining dark energy and modified gravity from the UV*, *Phys. Rev. D* **101** (2020) 021502 [*Erratum ibid.* **102** (2020) 049902] [arXiv:1904.05874] [INSPIRE].

[171] S. Kim, T. Noumi, K. Takeuchi and S. Zhou, *Heavy spinning particles from signs of primordial non-Gaussianities: beyond the positivity bounds*, *JHEP* **12** (2019) 107 [arXiv:1906.11840] [INSPIRE].

JHEP12(2023)076

[172] G. Ye and Y.-S. Piao, *Positivity in the effective field theory of cosmological perturbations*, *Eur. Phys. J. C* **80** (2020) 421 [arXiv:1908.08644] [INSPIRE].

[173] A.-C. Davis and S. Melville, *Scalar fields near compact objects: resummation versus UV completion*, *JCAP* **11** (2021) 012 [arXiv:2107.00010] [INSPIRE].

[174] S. Melville and J. Noller, *Positivity bounds from multiple vacua and their cosmological consequences*, *JCAP* **06** (2022) 031 [arXiv:2202.01222] [INSPIRE].

[175] M. Freytsis, S. Kumar, G.N. Remmen and N.L. Rodd, *Multifield positivity bounds for inflation*, *JHEP* **09** (2023) 041 [arXiv:2210.10791] [INSPIRE].

[176] K. Aoki, S. Mukohyama and R. Namba, *Positivity vs. Lorentz-violation: an explicit example*, *JCAP* **10** (2021) 079 [arXiv:2107.01755] [INSPIRE].

[177] S. Ramírez-Uribe, R.J. Hernández-Pinto, G. Rodrigo and G.F.R. Sborlini, *From five-loop scattering amplitudes to open trees with the loop-tree duality*, *Symmetry* **14** (2022) 2571 [arXiv:2211.03163] [INSPIRE].

JHEP12(2023)076