One-particle-irreducible consistency relations for cosmological perturbations

Walter D. Goldberger[*]

Physics Department, Yale University, New Haven, Connecticut 06520, USA

Lam Hui[†] and Alberto Nicolis[‡]

Physics Department and Institute for Strings, Cosmology and Astroparticle Physics, Columbia University,
New York, New York 10027, USA

We derive consistency relations for correlators of scalar cosmological perturbations that hold in the "squeezed limit" in which one or more of the external momenta become soft. Our results are formulated as relations between suitably defined one-particle-irreducible N-point and $(N-1)$-point functions that follow from residual spatial conformal diffeomorphisms of the unitary gauge Lagrangian. As such, some of these relations are *exact* to all orders in perturbation theory and do not rely on approximate de Sitter invariance or other dynamical assumptions (e.g., properties of the operator product expansion or the behavior of modes at the horizon crossing). The consistency relations apply model-independently to cosmological scenarios in which the time evolution is driven by a single scalar field. Besides reproducing the known results for single-field inflation in the slow-roll limit, we verify that our consistency relations hold more generally, for instance, in ghost condensate models *in flat space*. We comment on possible extensions of our results to multifield models.

I. INTRODUCTION

While the observed temperature anisotropies of the cosmic microwave background (CMB) are consistent with the possibility that the early Universe underwent a period of exponential inflation, there is as of yet no definitive test of the dynamics of inflation. In the near future, experiments such as Planck will provide further constraints from studies of the CMB B modes, which would probe the spectrum of gravitational waves produced during inflation, and possibly of non-Gaussian correlations of the temperature anisotropies.

The possibility of observing non-Gaussian features in the CMB and large scale structure has motivated much theoretical work in the last decade aimed at understanding the predictions of inflation for three-point and higher correlation functions of density perturbations. In the context of single-field, slow-roll inflation, a systematic treatment of scalar and tensor non-Gaussianity was first given in Ref. [1]. The explicit results of Ref. [1] indicate the existence of a set of *consistency relations* between the two-point and three-point correlation functions, which hold in the soft (or "squeezed") limit in which one of the momenta approaches zero. For the usual mode ζ that describes adiabatic density perturbations, this relation takes the form

$$\langle \zeta_{\vec{k}\to 0} \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \rangle = -(n_s - 1)\delta(\vec{k}_1 + \vec{k}_2)P(\vec{k}\to 0)P(\vec{k}_1), \quad (1)$$

where $P(\vec{k}) \sim k^{-3+(n_s-1)}$ is the power spectrum of adiabatic modes. Heuristically, this relation can be understood

to follow from the behavior of long-wavelength adiabatic modes [1]. Such modes freeze out at horizon crossing, becoming indistinguishable from a redefinition of the background scale factor. Therefore, an equal-time correlator with one insertion of a long wavelength ζ is equivalent to a correlator for the remaining fields, evaluated at rescaled coordinates. This line of reasoning was subsequently used in Ref. [2] to argue that the relation Eq. (1) is in fact a general statement about single-field inflation, independent of the slow-roll approximation.

Since the work of Refs. [1,2], much recent effort has been dedicated to further understanding and generalizing the original consistency relations such as Eq. (1). Reference [3] considered the behavior of N-point correlators in the limit of soft *internal* momenta in which linear combinations approach zero. New single-field relations relating gradients of $(N + 1)$-point functions and N-point correlators were derived in Ref. [4]. These results were obtained both on the basis of approximate de Sitter invariance (in the spectator field limit) and, more generally, at tree level but all orders in slow roll, by exploiting a relation [5] between long-wavelength adiabatic modes and "large" spatial diffeomorphisms, i.e., dilations and special conformal transformations, acting on a homogeneous background. See Ref. [6] for a generalization of this approach.

These large diffeomorphisms also play a role in the work of Ref. [7], which approaches the dilation consistency relations from the point of view of Ward identities associated with nonlinearly realized dilation invariance, with ζ playing the role of the Goldstone mode. In addition to contributions from single ζ poles, which reproduce [4], Ref. [7] finds additional contributions from higher

[*]walter.goldberger@yale.edu
[†]lhui@astro.columbia.edu
[‡]nicolis@phys.columbia.edu

Fock states, although these are found to be subleading for single-field models with Bunch-Davies initial vacuum state. Reference [8] proves that the superhorizon constancy of ζ holds as an operator statement in single-field scenarios with Bunch-Davies initial conditions, closing a possible loophole in the derivation of Ref. [4]. Relations between correlators in multifield models, obtained via approximate de Sitter invariance and the operator product expansion, were obtained in Ref. [9], while Ref. [10] used approximate $SO(4, 1)$ invariance of the background to reproduce the relation between two- and three-point functions.

It should be pointed out, however, that despite these recent developments, there seems to be no universal consensus on the status of the consistency relations, even in single-field theories. A proposed counterexample was given in Ref. [11], while the analysis presented in Ref. [12] indicates the relation in Eq. (1) is violated for generic initial states. Reference [13] argues that although Eq. (1) is correct, it does not reflect measurements made by physical observers. Rather, physical (gauge invariant) non-Gaussianities should have softer behavior in the squeezed limit than what is suggested by Eq. (1).

In this paper, we present consistency relations that follow from residual diffeomorphism invariance of "ζ-gauge" scalar cosmological perturbations. The main assumptions that go into our derivation are gauge invariance of the action and of the path integral measure. Because the set of starting assumptions is relatively small, our results are expected to be quantum mechanically exact and to have a wider scope of validity than just single-field, slow-roll scenarios. In particular, our relations hold independently of dynamical assumptions (approximate isometries, superhorizon behavior, etc.) or of the details of the initial state (but see Sec. IV and Appendix A on subtleties). We find it most natural to express our results in terms of in-in, one-particle irreducible (1PI) correlators of ζ at equal times rather than the more conventional connected Green's functions used to describe non-Gaussianity. However, besides implying new, nontrivial, squeezed relations between the two- and three-point function, our 1PI relations are completely consistent with those obtained in the literature under more restrictive assumptions.

In Sec. II we introduce the effective action $\Gamma[\zeta]$ that generates in-in 1PI correlators of ζ at equal times and derive formal constraints due to invariance under spatial diffeomorphisms. Explicit results for the special cases of residual spatial dilation and special conformal invariance are given in Secs. II A and II B, respectively. These consist of relations between N- and $(N + 1)$-point correlators with one zero-momentum mode (Sec. II A) and between derivatives at zero momentum (Sec. II B). When reexpressed in terms of connected correlators, our dilation relations agree with the standard results in the literature. For the case of special conformal transformations, which do not preserve the closure of the momentum space polygon, we find that a

procedure for computing derivatives of correlators must be specified in order to obtain well-defined relations. For any such procedure, our results are similar in form to those of Ref. [4] and also of Ref. [14] involving pure de Sitter space correlators.

Explicit examples are give in Sec. III. Section III A applies our relations to slow-roll inflation. In particular, by working in terms of 1PI correlators, we are able to formulate and verify a new special conformal relation between the two-point and three-point functions. In Sec. III B, we check that the consistency relations also hold for scalar perturbations of the ghost condensate [15] for a choice of model parameters in which the background spacetime is tuned to be exactly Minkowski. This example serves to illustrate that the consistency relations do not require the presence of a nearly de Sitter cosmological horizon and the freezing-out of long-wavelength modes in order to hold.

However, we stress that our results are not completely universal. In Sec. III C we present a simple model that *violates* the consistency relations, namely, a free massless scalar field in flat spacetime, decoupled from gravity. This theory admits a spatially homogeneous background solution $\phi \propto t$ with constant energy density and equation of state $p = \rho$. It has, in addition, adiabatic perturbations ζ that start at quadratic order in the underlying scalar field fluctuations π. The violation of the consistency relations in this example can be traced to singularities in the Jacobian relating the path integral measure for π and ζ fluctuations, as further explained in that section.

Conclusions, as well as a discussion of possible extensions of this work to tensors and to multifield models, are given in Sec. IV.

II. FORMALISM AND RESULTS

We work in the context of single-field models, in which case there is a gauge—to which we refer as "unitary gauge"—in which the physical scalar perturbations $\zeta(x)$ and traceless tensor modes $\gamma_{ij}(x)$ are encoded in the metric [1]. Written in Arnowitt-Deser-Misner (ADM) form, this is

$$ds^2 = -N^2 dt^2 + a^2(t) h_{ij} (dx^i + N^i dt)(dx^j + N^j dt), \quad (2)$$

where

$$h_{ij} = e^{2\zeta}(e^\gamma)_{ij} = e^{2\zeta}(\delta_{ij} + \gamma_{ij} + \cdots), \qquad \gamma_{ii} = 0. \quad (3)$$

The lapse N and shift N^i can be integrated out through the ADM Hamiltonian and momentum constraints (see Ref. [1]). The physical tensor mode is then the transverse projection of γ_{ij}.

In this paper we focus on momentum-space scalar correlators at some fixed time t_*, $\langle \zeta(t_*)_{\vec{k}_1} \ldots \zeta(t_*)_{\vec{k}_n} \rangle$, with $\zeta_{\vec{k}}(t) = \int d^3x \, e^{-i\vec{k}\cdot\vec{x}} \zeta(\vec{x}, t)$. For these, it is particularly convenient to introduce a three-dimensional Euclidean generalization of the quantum effective action formalism of the

quantum field theory (QFT) (see, e.g., Ref. [16]). We define a generating functional

$$Z[J] = \int D\zeta(t_*, \vec{x}) P[\zeta, t_*] e^{\int d^3 x \zeta(\vec{x}, t_*) J(\vec{x})}, \qquad (4)$$

where the probability measure $P[\zeta, t_*]$ on scalar modes at fixed time t_* is an integral over the vacuum wave functional $\Psi[\zeta, \gamma, t_*]$ at time t_*:

$$P[\zeta, t_*] = \int D\gamma_{ij}(t_*, \vec{x}) |\Psi[\zeta, \gamma, t_*]|^2. \qquad (5)$$

The wave functional at time t_* has the path integral representation

$$\Psi[\zeta, \gamma, t_*] = \int^{\zeta(t_*), \gamma(t_*)} D\zeta(x) D\gamma_{ij} e^{iS[\zeta, \gamma]} \Psi_0[\zeta, \gamma]. \qquad (6)$$

Here, the initial wave function in the far past is $\Psi_0 = \langle \zeta, \gamma; t \to -\infty | 0 \rangle$. Although the initial state $|0\rangle$ is usually taken to be the Bunch-Davies adiabatic vacuum, the specific form of the wave function will not matter for our results, as long as it is diffeomorphism-invariant. Note that we only integrate over $\zeta(x)$ and $\gamma_{ij}(x)$ as we assume that the nondynamical fields N, N^i have been integrated out via the ADM constraints. Note also that the fact that Ψ appears quadratically in Eq. (5) means the corresponding path integral has a doubling of fields—this is precisely the in-in formalism [17].

Equal-time correlators can be retrieved from $Z[J]$ in the usual way, e.g.,

$$\langle \zeta(t_*)_{\vec{k}_1} \zeta(t_*)_{\vec{k}_2} \rangle = \frac{1}{Z[0]} \frac{\delta^2 Z[J]}{\delta J_{-\vec{k}_1} \delta J_{-\vec{k}_2}} \bigg|_{J=0}. \qquad (7)$$

In addition to the generating functional $Z[J]$, it is useful to introduce $W[J] = \ln Z[J]$, which generates connected correlators, as well as its Legendre transform

$$\Gamma[\bar{\zeta}] = W[J] - \int d^3 x J(\vec{x}) \bar{\zeta}(\vec{x}), \qquad (8)$$

which generates 1PI correlators.[1] It follows from the definitions that

$$\bar{\zeta}(\vec{x}) = \frac{\delta W[J]}{\delta J(\vec{x})}, \qquad J(\vec{x}) = -\frac{\delta \Gamma[\bar{\zeta}]}{\delta \bar{\zeta}(\vec{x})}. \qquad (9)$$

In particular, $\bar{\zeta}(\vec{x})$ is the one-point function of $\zeta(\vec{x})$ for given external source $J(\vec{x})$:

$$\bar{\zeta}(\vec{x}) = \langle \zeta(\vec{x}) \rangle_J. \qquad (10)$$

The equal-time 1PI correlation functions are[2]

$$\frac{\delta}{\delta \bar{\zeta}_{-\vec{k}_1}} \cdots \frac{\delta}{\delta \bar{\zeta}_{-\vec{k}_n}} \Gamma[\bar{\zeta}] \bigg|_{\bar{\zeta}=0} = \delta\left(\sum_a \vec{k}_a\right) \Gamma^{(n)}(\vec{k}_1, \ldots \vec{k}_n, t_*).$$
$$(11)$$

The connected Green's functions generated by $W[J]$ can then be expressed in terms of the 1PI correlators $\Gamma^{(n)}(\vec{k}_1, \ldots, \vec{k}_n, t_*)$. For instance, the connected two-point function (or power spectrum), $\langle \zeta(t_*)_{\vec{k}_1} \zeta(t_*)_{\vec{k}_2} \rangle_c = \delta(\vec{k}_1 + \vec{k}_2) G_c^{(2)}(\vec{k}_1, \vec{k}_2, t_*)$, obtained by differentiation of $W[J]$, is

$$G_c^{(2)}(\vec{k}, -\vec{k}, t_*) = P(\vec{k}, t_*) = -\frac{1}{\Gamma^{(2)}(\vec{k}, -\vec{k}, t_*)}, \qquad (12)$$

while the three-point function can be expressed as

$$G_c^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3, t_*) = \left[\prod_{a=1}^3 P(\vec{k}_a, t_*)\right] \Gamma^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3, t_*). \qquad (13)$$

In general, the $\Gamma^{(n)}$ should be interpreted as quantum-corrected vertices that, combined with the equal-time propagators $P(\vec{k}, t_*)$, yield the fully quantum-mechanical equal-time connected correlators via standard *tree-level* diagrammatics. We will refer to $\Gamma[\bar{\zeta}]$ as the "3D effective action," and to its Taylor coefficients $\Gamma^{(n)}(\vec{k}_1, \ldots, \vec{k}_n)$ as the "3D effective vertices," to remind ourselves that they contain information only about equal-time correlators. From now on, we drop the explicit dependence on time as it is understood that all correlators are at time $t = t_*$.

Our strategy then is quite simple: we will show that the 3D effective action is invariant under the residual diffeomorphisms of Eqs. (2) and (3). These act nonlinearly on $\bar{\zeta}$. Upon expanding Γ in powers of $\bar{\zeta}$, these nonlinear symmetries yield relations between successive 3D effective vertices, $\Gamma^{(n)}$ and $\Gamma^{(n+1)}$, for all n's—and, therefore, between n- and $(n + 1)$-point equal-time correlation functions. Since the parameterization (2) and (3) admits residual diffeomorphisms only at zero momentum, these relations between $\Gamma^{(n)}$ and $\Gamma^{(n+1)}$ arise only when, in the latter, one of the momenta is "soft," $\vec{q} \to 0$.

To show the invariance of $\Gamma[\bar{\zeta}]$ under residual diffeomorphisms, we follow closely the standard manipulations that yield the symmetries of the *standard* four-dimensional quantum effective action; see, e.g., Ref. [16]. We assume that the integration measure—including the wave function Ψ_0—in Eq. (6) is invariant under spatial diffeomorphisms at the fixed time t_*. This is actually a subtle assumption for nonlinearly realized symmetries, like our diffeomorphisms, or the global symmetries acting on Goldstone bosons. We briefly touch upon this subtlety in the

---

[1] More precisely, these are correlators with diagrams that are 1PI with respect to internal ζ lines but not 1PI for the other fields.

[2] We employ the momentum-space notations $\delta(\vec{k}) \equiv (2\pi)^3 \delta^3(\vec{k})$ and $\int_{\vec{k}} \equiv \int \frac{d^3 k}{(2\pi)^3}$.

appendix. Then, since the action $S[\zeta, \gamma]$ is diff-invariant, the measure $D\zeta(t_*)P[\zeta, t_*]$ in Eq. (4) is also invariant. For an infinitesimal diffeomorphism $\zeta \to \zeta + \Delta\zeta$, we have, after changing variables in the definition of $Z[J]$ and expanding to linear order in $\Delta\zeta$,

$$0 = \int d^3x J(\vec{x}) \int D\zeta(t_*, \vec{x}) P[\zeta, t_*] e^{\int d^3x J\zeta} \Delta\zeta(t_*, \vec{x}). \quad (14)$$

Upon dividing by $Z[J]$, this result then implies

$$\int d^3x \langle \Delta\zeta(t_*, \vec{x}) \rangle_J \frac{\delta\Gamma[\bar{\zeta}]}{\delta\bar{\zeta}(t_*, \vec{x})} = 0, \quad (15)$$

where we have used Eq. (9). The above equation is a symmetry statement: it states that $\Gamma[\bar{\zeta}]$ is invariant under the (infinitesimal) transformation

$$\bar{\zeta} \to \bar{\zeta} + \langle \Delta\zeta(t_*, \vec{x}) \rangle_J, \quad (16)$$

where J and $\bar{\zeta}$ are related via Eq. (9). In Sec. II A we consider the constraints on the form of correlators $\Gamma^{(n)}$ that follow from Eq. (15) as a result of invariance under residual spatial dilations. In Sec. II B, we consider constraints from spatial conformal transformations. This will require a discussion of possible modifications to Eq. (15) due to shifts in tensor modes.

A. Dilations

It is simplest to first consider scale transformations of the spatial coordinates. Under the infinitesimal dilation $\vec{x} \to (1 - \lambda)\vec{x}$, the transformation of the spatial metric $g_{ij} = e^{2\zeta} h_{ij}$ on constant time hypersurfaces implies that

$$\Delta\zeta(\vec{x}) = \lambda(1 + \vec{x} \cdot \nabla\zeta(\vec{x})), \quad (17)$$

or in momentum space

$$\Delta\zeta_{\vec{k}} = \lambda\delta(\vec{k}) - \lambda\nabla_{\vec{k}} \cdot (\vec{k}\zeta_{\vec{k}}). \quad (18)$$

Since this transformation is at most linear in ζ, it follows that $\langle \Delta\zeta \rangle_J = \Delta\bar{\zeta}$ [recall Eq. (10)], and thus from Eq (15),

$$\frac{\delta\Gamma[\bar{\zeta}]}{\delta\bar{\zeta}_{\vec{q}=0}} = -\int_{\vec{k}} \bar{\zeta}_{\vec{k}} (\vec{k} \cdot \nabla_{\vec{k}}) \frac{\delta\Gamma}{\delta\zeta_{-\vec{k}}}. \quad (19)$$

This holds for a generic configuration $\bar{\zeta}_{\vec{k}}$. We can Taylor expand about $\bar{\zeta} = 0$. Applying n partial derivatives $\delta^n/\delta\bar{\zeta}_{-\vec{k}_1} \ldots \delta\bar{\zeta}_{-\vec{k}_n}$ to both sides, and evaluating them at $\bar{\zeta} = 0$, we obtain the relations

$$\frac{\delta^{n+1}\Gamma}{\delta\bar{\zeta}_{\vec{q}=0}\delta\bar{\zeta}_{\vec{k}_1} \ldots \delta\bar{\zeta}_{\vec{k}_n}} \Bigg|_{\bar{\zeta}=0}$$
$$= -\left(\sum_{a=1}^n \vec{k}_a \cdot \nabla_{\vec{k}_a}\right) \frac{\delta^n\Gamma}{\delta\bar{\zeta}_{\vec{k}_1} \ldots \delta\bar{\zeta}_{\vec{k}_n}} \Bigg|_{\bar{\zeta}=0}. \quad (20)$$

Defining a dilation derivative $\mathcal{D}_n = \sum_{a=1}^n \vec{k}_a \cdot \nabla_{\vec{k}_a}$, it follows that $\mathcal{D}_n\delta(\sum_{a=1}^n \vec{k}_a) = -3$, and thus we can write the above equation as

$$\Gamma^{(n+1)}(0, \vec{k}_1, \ldots, \vec{k}_n) = (3 - \mathcal{D}_n)\Gamma^{(n)}(\vec{k}_1, \ldots, \vec{k}_n). \quad (21)$$

Note that although the momenta on the rhs of this equation are constrained to add up to zero, the derivative operator $\mathcal{D}_n$ acts to generate an overall rescaling that preserves the surface $\sum_a \vec{k}_a = 0$. Consequently, $\mathcal{D}_n$ is well-defined when acting on $\Gamma^{(n)}$. One can also think of $\Gamma^{(n)}$ as dependent on only $n - 1$ momenta, e.g., $\vec{k}_1, \ldots, \vec{k}_{n-1}$, and thus $\mathcal{D}_n\Gamma^{(n)} = \mathcal{D}_{n-1}\Gamma^{(n)}$.

We find that Eq. (21) is the most transparent way to state the dilation consistency relation. Starting from this equation, it is straightforward to derive constraints on the connected Green's functions:

$$\langle \zeta_{\vec{k}_1} \ldots \zeta_{\vec{k}_n} \rangle_c = \delta\left(\sum_a \vec{k}_a\right) G_c^{(n)}(\vec{k}_1, \ldots, \vec{k}_n). \quad (22)$$

For example, from Eq. (21) with $n = 2$ and Eq. (13), we deduce that

$$G_c^{(3)}(0, \vec{k}_1, \vec{k}_2) = -P(0)(3 + \mathcal{D}_1)P(k_1)$$
$$= -P(0)P(k_1)\frac{\partial \ln (k_1^3 P(k_1))}{\partial \ln k_1}, \quad (23)$$

reproducing Maldacena's consistency relation for the bispectrum, obtained here without making assumptions regarding the behavior of ζ at and after horizon crossing or even the existence of a Hubble horizon.

As another example, we may consider constraints on the four-point connected correlator. Schematically, this can be expressed in terms of 1PI correlators as

$$G_c^{(4)} = \left[\prod_a P(\vec{k}_a)\right]\left[\Gamma^{(4)} + \sum_{s,t,u} \Gamma^{(3)}P(\vec{k})\Gamma^{(3)}\right], \quad (24)$$

where the sum on the rhs is over all three "crossings" of the external momenta. Taking the limit $\vec{k}_4 \to 0$ on the lhs, it follows immediately from Eq. (21) that

$$G_c^{(4)}(0, \vec{k}_1, \vec{k}_2, \vec{k}_3) = -P(0)(6 + \mathcal{D}_2)G_c^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3), \quad (25)$$

in agreement with the results of Ref. [4]. It is not difficult to show, by induction on n, that Eq. (21) is in fact consistent with the results of Ref. [4] for all possible connected Green's functions $G_c^{(n)}$: $G_c^{(n+1)}(0, \vec{k}_1, \ldots) = -P(0)[3(n - 1) + \mathcal{D}_{n-1}]G_c^{(n)}(\vec{k}_1, \ldots)$.

B. Special conformal transformations

We can repeat essentially the same analysis for infinitesimal special conformal transformations, $\delta\vec{x} = \vec{x}^2\vec{b} - 2(\vec{b} \cdot \vec{x})\vec{x}$, under which ζ shifts by

$$\Delta\zeta(\vec{x}) = 2\vec{b} \cdot \vec{x} - (\vec{x}^2\vec{b} - 2(\vec{x} \cdot \vec{b})\vec{x}) \cdot \vec{\nabla}\zeta(\vec{x}), \quad (26)$$

or in momentum space

$$\Delta\zeta_{\vec{k}} = 2i\vec{b}\cdot\vec{\nabla}_{\vec{k}}\delta(\vec{k}) - i\nabla^2_{\vec{k}}(\vec{b}\cdot\vec{k}\zeta_{\vec{k}}) + 2i\vec{b}\cdot\nabla_{\vec{k}}(\nabla_{\vec{k}}\cdot(\vec{k}\zeta_{\vec{k}})).$$
(27)

It should be noted, however, that, in the presence of the tensor mode γ_{ij}, this is *not* a residual gauge freedom of Eqs. (2) and (3). More precisely, under a special conformal diffeomorphism, on top of the ζ shift just alluded to, one gets a transformation of γ_{ij} that does not preserve its transversality. It should be possible to modify the following analysis to explicitly keep track of this. In this paper we just ignore this subtlety and leave addressing it for future work. It suffices to note that the 1PI correlation functions involving ζ only should not be affected by this subtlety. The reason is the following: under a special conformal transformation, one gets a nontransverse contribution to γ of order $\delta\gamma = \mathcal{O}(b\gamma)$. This contains a traceless scalar piece of the form $(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2)\chi$, which can be removed via a further spatial diff, thus modifying Eq. (26) by an extra contribution of order $\chi = \mathcal{O}(b\gamma)$. It is important to notice that this extra diffeomorphism is not simply undoing Eq. (26). Then, if we generalized our formalism to include tensors in Γ, we would get a new symmetry statement replacing Eq. (16):

$$\int d^3x \frac{\delta\Gamma}{\delta\bar{\zeta}}\Delta\bar{\zeta} + \frac{\delta\Gamma}{\delta\bar{\gamma}_{ij}}\Delta\bar{\gamma}_{ij} = 0.$$
(28)

By taking functional derivatives with respect to $\bar{\zeta}$, and setting $\bar{\zeta} = \bar{\gamma}_{ij} = 0$—which is all we need for 1PI vertices of $\bar{\zeta}$ only—we see that these $\mathcal{O}(b\gamma)$ corrections to the transformation laws simply do not contribute. Since the 1PI ζ vertices are not affected, consistency relations at the level of connected ζ Green's functions should thus remain valid as long as diagrams for which 1PI vertices are connected by internal graviton legs are subdominant.

Inserting Eq. (27) into Eq. (15)—for the same reason as before, we have $\langle\Delta\zeta\rangle_J = \Delta\bar{\zeta}$—and differentiating n times with respect to ζ then yields

$$2\nabla_{\vec{q}\to 0}\left[\delta\left(\vec{q} + \sum_a\vec{k}_a\right)\Gamma^{(n+1)}\left(\vec{q}, \vec{k}_1, \ldots, \vec{k}_n\right)\right]$$
$$= \sum_a\vec{\mathcal{S}}_a\left[\delta\left(\sum_a\vec{k}_a\right)\Gamma^{(n)}\left(\vec{k}_1, \ldots, \vec{k}_n\right)\right],$$
(29)

where we have introduced the (vector) differential operator

$$\vec{\mathcal{S}}_a \equiv \vec{k}_a\nabla^2_{\vec{k}_a} - 2(\vec{k}_a\cdot\vec{\nabla}_{\vec{k}_a})\vec{\nabla}_{\vec{k}_a}.$$
(30)

In order to convert Eq. (29) into a relation involving only gradients acting on $\Gamma^{(n+1)}$ and $\Gamma^{(n)}$, a prescription for evaluating derivatives $\nabla_{\vec{k}_a}$ acting on functions on the surface $\sum_a\vec{k}_a = 0$ must be specified. In other words, by construction the G's and Γ's are defined only on the surface $\sum_a\vec{k}_a = 0$; one can, in principle, continue them to arbitrary, unconstrained momenta, but there is no unique

prescription for doing so. One could impose total symmetry under generic permutations of the momenta, but that is not enough to give an unambiguous continuation. For instance, away from $\sum_a\vec{k}_a = 0$, one can continue the inflationary spectrum $G^{(2)} \sim 1/k^3$ into a symmetric function of two independent momenta, but there are many inequivalent choices that reduce to the desired spectrum when these momenta are taken to be equal in magnitude and opposite in direction, e.g., $1/(k_1 + k_2)^3$ and $1/(k_1k_2)^{3/2}$. Generic derivatives with respect to the momenta take us (infinitesimally) away from the $\sum_a\vec{k}_a = 0$ surface and are thus sensitive to which choice we adopt. As we show in Appendix B, it is straightforward to check that in the end these ambiguities cancel between the left-hand side and the right-hand side of the final consistency relation, provided one preserves the *dilation* consistency relations in moving off the momentum-conserving surface. However, instead of having a consistency relation with equally ambiguous sides, we prefer to avoid the ambiguity altogether. As we will see in the next section, our pickiness about this will be rewarded.

One possible choice is to eliminate, e.g., $\vec{k}_n$ using the delta functions so that $\Gamma^{(n)}$ is only a function of the $n - 1$ momenta $\vec{k}_{a=1,\ldots,n-1}$, and $\Gamma^{(n+1)}$ of those momenta as well as of $\vec{k}$. This has the advantage of expressing the consistency conditions directly in terms of observationally relevant quantities, such as the spectrum, the bispectrum, etc. With this prescription in mind, the lhs of Eq. (29) can then be expressed as

$$2\delta\left(\sum_a\vec{k}_a\right)\nabla_{\vec{q}\to 0}\Gamma^{(n+1)}$$
$$+ 2\left[\nabla_{\vec{q}\to 0}\delta\left(\vec{q} + \sum_a\vec{k}_a\right)\right](3 - \mathcal{D}_n)\Gamma^{(n)},$$
(31)

where we have made use of the dilation consistency relation, Eq. (21).

The rhs of Eq. (29) is a sum of terms with up to two derivatives acting on $\delta(\sum_a\vec{k}_a)$. The term with two gradients on the delta function is

$$\Gamma^{(n)}\sum_a\vec{\mathcal{S}}_a\delta\left(\sum_b\vec{k}_b\right) = 6\left[\nabla_{\vec{q}\to 0}\delta\left(\vec{q} + \sum_a\vec{k}_a\right)\right]\Gamma^{(n)},$$
(32)

where we have made use of the identity[3]

$$\sum_a\vec{k}_a^i\nabla^j_{\vec{k}_a}\nabla^k_{\vec{k}_a}\delta\left(\sum_b\vec{k}_b\right)$$
$$= -(\delta^{ij}\nabla^k_{\vec{q}\to 0} + \delta^{ik}\nabla^j_{\vec{q}\to 0})\delta\left(\vec{q} + \sum_a\vec{k}_a\right).$$
(33)

---

[3]Such an identity can be proven via standard manipulations upon rewriting the delta function as $\delta(\sum_b\vec{k}_b) = \int d^3x\, e^{-i(\sum_b\vec{k}_b)\cdot\vec{x}}$.

The term involving a single gradient acting on the delta function is

$$2[(L^{ij} - \delta^{ij}\mathcal{D}_n)\Gamma^{(n)}]\left[\nabla^j_{\vec{q}\to 0}\delta\left(\vec{q} + \sum_a \vec{k}_a\right)\right], \quad (34)$$

where $L^{ij} \equiv \sum_a (\vec{k}_a^i \nabla^j_{\vec{k}_a} - \vec{k}_a^j \nabla^i_{\vec{k}_a})$ is the angular momentum generator. By rotational invariance, we have $L^{ij}\Gamma^{(n)} = 0$, so Eq. (34) reduces to

$$-2\left[\nabla_{\vec{q}\to 0}\delta\left(\vec{q} + \sum_a \vec{k}_a\right)\right]\mathcal{D}_n\Gamma^{(n)}. \quad (35)$$

Finally, the term with no gradients acting on the delta functions is simply

$$\delta\left(\sum_a \vec{k}_a\right)\sum_{a=1}^{n-1}\vec{\mathcal{S}}_a\Gamma^{(n)}, \quad (36)$$

where we have interpreted $\Gamma^{(n)} = \Gamma^{(n)}(\vec{k}_1, \ldots, \vec{k}_{n-1}, -\sum_{a=1}^{n-1}\vec{k}_a)$ as discussed previously.

Comparing the various terms in Eqs. (32), (34), and (36), with Eq. (31), we find that all terms involving the gradient $\nabla_{\vec{q}\to 0}\delta(\vec{q} + \sum_i \vec{k}_i)$ cancel so that Eq. (29) finally becomes

$$2\nabla_{\vec{q}\to 0}\Gamma^{(n+1)}\left(\vec{q}, \vec{k}_1, \ldots, -\vec{q} - \sum_{a=1}^{n-1}\vec{k}_a\right)$$
$$= \sum_{a=1}^{n-1}\vec{\mathcal{S}}_a\Gamma^{(n)}\left(\vec{k}_1, \ldots, \vec{k}_{n-1}, -\sum_{a=1}^{n-1}\vec{k}_a\right). \quad (37)$$

This result can be combined with its counterpart for the dilation, Eq. (21), into a Taylor expansion in $\vec{q}$:

$$\Gamma^{(n+1)}(\vec{q}, \vec{k}_1, \ldots) = \left(3 - \mathcal{D}_n + \frac{1}{2}\sum_{a=1}^{n-1}\vec{q}\cdot\vec{\mathcal{S}}_a\right)\Gamma^{(n)}(\vec{k}_1, \ldots), \quad (38)$$

with corrections coming at $O(q^2)$.

It is straightforward to convert Eq. (37) into a statement regarding the connected Green's functions $G_c^{(n)}$. For example, for $n = 2$, we obtain

$$\nabla_{\vec{q}\to 0}G^{(3)}(\vec{q}, \vec{k}_1, -\vec{q} - \vec{k}_1)$$
$$= -\nabla_{\vec{q}\to 0}P(\vec{q})(3 + \mathcal{D}_1)P(\vec{k}_1)$$
$$- \frac{1}{2}P(0)(6\nabla_{\vec{k}_1} - \vec{\mathcal{S}}_1)P(\vec{k}_1). \quad (39)$$

In deriving this, it is important to keep in mind $G^{(3)}(\vec{q}, \vec{k}_1, -\vec{q} - \vec{k}_1) = \Gamma^{(3)}(\vec{q}, \vec{k}_1, -\vec{q} - \vec{k}_1)P(\vec{q})P(\vec{k}_1)\times P(\vec{q} + \vec{k}_1)$, and thus its derivative with respect to $\vec{q}$ has several terms, including the derivative of $P(\vec{q} + \vec{k}_1)$, which is easy to miss. Equation (39)—at first sight—looks the same as the consistency relation derived in Ref. [4]. There is one important difference, however: our $\vec{k}_n$ momentum has been expressed in terms of the other momenta, as enforced by the delta-functions, *before* taking the derivatives. As the

example in Sec. III A shows, this apparently minor technicality has crucial consequences.

III. EXAMPLES

A. Slow-roll inflation

It is well-known that the three-point correlator in single-field slow-roll inflation obeys the relation in Eq. (1), which in our language is associated with residual dilations. What has been overlooked so far is that it also obeys a nontrivial relation associated with special conformal transformations. More precisely, the special conformal relation between two-point and three-point correlators has been claimed to be trivially obeyed [4], with the left- and right-hand sides both being zero in the squeezed limit. In fact, we claim that this is an artifact of insisting on using (somewhat ambiguous) n-point functions that depend on n unconstrained momenta. If this ambiguity is resolved by a prescription such as the one discussed above, and one expresses the n-point correlators in terms of $n - 1$ independent momenta, there is a nontrivial check to perform already at the level of the three-point function.

The three-point function has been computed by Maldacena [1]. At lowest order in the slow roll expansion, it reads

$$G^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3) \simeq \frac{H^4}{4\epsilon^2 M_{\text{Pl}}^4}\frac{1}{\Pi_i(2k_i^3)}$$
$$\times\left[(3\epsilon - 2\eta)\sum_i k_i^3 + \epsilon\sum_{i\neq j}k_i k_j^2\right.$$
$$\left. + 8\epsilon\frac{1}{\sum_i k_i}\sum_{i<j}k_i^2 k_j^2\right], \quad (40)$$

while the spectrum is

$$P(k) = -\frac{1}{\Gamma^{(2)}(k)} \simeq \frac{H^2}{4\epsilon M_{\text{Pl}}^2}\frac{1}{k^{3-2(\eta-3\epsilon)}}. \quad (41)$$

By setting $\vec{k}_2 = -(\vec{k}_1 + \vec{k}_3)$ and taking the soft limit $\vec{k}_3 \to 0$, after straightforward manipulations we get

$$\Gamma^{(3)} = \prod_a (-\Gamma^{(2)}(k_a))\cdot G^{(3)}$$
$$= (\eta - 3\epsilon)\Gamma^{(2)}(k_1)\left[2 + 3\frac{\vec{k}_1\cdot\vec{k}_3}{k_1^2}\right] + \mathcal{O}(k_3^2). \quad (42)$$

The zeroth-order term obeys the dilation consistency condition:

$$\Gamma^{(3)}|_{\vec{k}_3\to 0} = 2(\eta - 3\epsilon)\Gamma^{(2)}(k_1) = (3 - \mathcal{D}_1)\Gamma^{(2)}(k_1), \quad (43)$$

in agreement with the standard results. What is new here is that the linear term in $\vec{k}_3 \to 0$ obeys the special-conformal consistency condition:

$$2\nabla_{\vec{k}_3\to 0}\Gamma^{(3)} = 6(\eta - 3\epsilon)\frac{\vec{k}_1}{k_1^2}\cdot\Gamma^{(2)}(k_1) = \vec{\mathcal{S}}_{\vec{k}_1}\Gamma^{(2)}(k_1), \quad (44)$$

as can be checked straightforwardly from the explicit form of the spectrum, Eq. (41).

B. Flat-space limit of ghost condensate

A less traditional example is the ghost condensate in Minkowski spacetime, coupled to dynamical gravity. This is the theory of a peculiar derivatively coupled scalar field ϕ. At lowest order in the derivative expansion, in addition to the gravitational (Einstein-Hilbert) term, the theory is defined by the Lagrangian [15]

$$\mathcal{L} = \sqrt{-g} M^4 P(X), \qquad (45)$$

with $X = -g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$, M as some mass scale, and P as a fairly generic function. If such a function has a minimum at some X—say $X = 1$—and if the value of P at that minimum is adjusted to be zero (by tuning the cosmological constant), then

$$\phi(x) = t, \qquad g_{\mu\nu}(x) = \eta_{\mu\nu} \qquad (46)$$

is a solution of the equations of motion *in the presence of gravity.* This model thus shares with single-field inflationary scenarios the presence of a "rolling" physical scalar that can serve as a clock—thus allowing us to choose the unitary gauge that we have introduced in Sec. II—but does not feature a cosmological expansion nor a Hubble horizon. In particular, modes do not "freeze out" at late times, nor do they become classical.

Before proceeding, it should be mentioned that higher-derivative terms in the action are crucial in order to stabilize the solution $\phi = t$ against small perturbations; from the $P(X)$ Lagrangian by itself, perturbations do not have gradient energy, which results—neglecting gravity—in a trivial dispersion relation, $\omega = 0$. When higher-derivative terms are included, the leading gradient energy gives the perturbations a quadratic dispersion relation at short distances,

$$\omega_{\vec{k}} \simeq \frac{\bar{M}}{M^2} k^2, \qquad (47)$$

where $\bar{M}$ is a mass parameter associated with the higher-derivative terms. In the far infrared limit, at momenta lower than

$$k_J = \frac{M^2}{M_{\rm Pl}}, \qquad (48)$$

gravitational effects become important, and a slow, Jeans-type instability sets in [15]. To be safe, we will work at shorter distances, $k \gg k_J$, where the solution (46) is free of instabilities.[4]

The perturbations about cosmological solutions of this theory were analyzed in Ref. [18], working in a unitary gauge where all fluctuations appear in the metric, written in the ADM parametrization employed in Eq. (2). The action for the scalar perturbation ζ is quite involved; see Ref. [18] for details. We will use those results (and extend them to cubic order) for the case in which the background metric is Minkowskian. At high momenta, $k \gg k_J$, the quadratic Lagrangian reads

$$\mathcal{L}_2 \simeq \frac{2M_{\rm Pl}^4}{\bar{M}^2}\dot{\zeta}^2 - \frac{2M_{\rm Pl}^4}{M^4}(\nabla^2\zeta)^2, \qquad (49)$$

while the cubic interactions relevant for checking the three-point consistency relations (i.e., those that survive when one of the three momenta becomes soft) are given by[5]

$$\mathcal{L}_3 \simeq \frac{6M_{\rm Pl}^4}{\bar{M}^2}\zeta\dot{\zeta}^2 + \frac{2M_{\rm Pl}^4}{M^4}\zeta(\nabla^2\zeta)^2. \qquad (51)$$

The time-ordered two-point function of ζ, from Eq. (49), is

$$\langle T\zeta_{\vec{k}}(t)\zeta_{\vec{q}}(t')\rangle = \delta(\vec{k}+\vec{q})\frac{\bar{M}^2}{8M_{\rm Pl}^4\omega_{\vec{k}}}\Big[\theta(t-t')e^{-i\omega_{\vec{k}}(t-t')}$$
$$+ \theta(t'-t)e^{-i\omega_{\vec{k}}(t'-t)}\Big], \qquad (52)$$

with $\omega_{\vec{k}}$ given in Eq. (47), so that

$$\Gamma^{(2)}(\vec{k},-\vec{k}) = -\frac{8M_{\rm Pl}^4}{\bar{M}^2}\omega_{\vec{k}}, \qquad (53)$$

while the cubic ζ self-interaction with incoming momenta $\vec{k}_{1,2,3}$ can be written as

$$-\frac{8iM_{\rm Pl}^4}{\bar{M}^2}(\omega_{\vec{k}_1}\omega_{\vec{k}_2} + \omega_{\vec{k}_1}\omega_{\vec{k}_3} + \omega_{\vec{k}_2}\omega_{\vec{k}_3}), \qquad (54)$$

up to contact terms that do not contribute to long-distance correlations [an overall delta function $\delta(\sum_a\vec{k}_a)$ has been suppressed]. The connected three-point function at equal time t is therefore

$$G_c^{(3)}(\vec{k}_1,\vec{k}_2,\vec{k}_3,t) = -\frac{8iM_{\rm Pl}^4}{\bar{M}^2}\left(\frac{\bar{M}^2}{8M_{\rm Pl}^4}\right)^3\left[\sum_a\omega_{\vec{k}_a}^{-1}\right]\mathcal{I}(\omega_a,t),$$
$$(55)$$

where $\mathcal{I}(\omega_a,t)$ is the integral,

---

[4]We work in the weak gravity limit, $M_{\rm Pl} \gg \bar{M}, M$. For simplicity we can also assume $\bar{M} \sim M$. These conditions automatically ensure that the Jeans scale k_J is much below the strong-coupling scale of the theory, which is some combination of M and $\bar{M}$.

[5]In the notation of Ref. [18], we have used the following unitary-gauge Lagrangian for the ghost condensate:

$$S = \int d^4x\sqrt{h}\left[\frac{1}{8}M^4(1/N^2 - 1)^2 - \frac{1}{2}\bar{M}^2(E^i_{\ i})^2\right], \qquad (50)$$

which is not the most general one. However, for the purposes of our check, we need not be completely general—just consistent.

$$\mathcal{I}(\omega_a, t) = \int_{-\infty}^{\infty} dt' \left[\theta(t - t') e^{-i\sum_a \omega_{\vec{k}_a}(t-t')} \right.$$
$$\left. + \theta(t' - t) e^{-i\sum_a \omega_{\vec{k}_a}(t'-t)} \right]$$
$$= \frac{2i}{-\sum_a \omega_{\vec{k}_a} + i\epsilon}. \quad (56)$$

The connected Green's function is then

$$G_c^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3, t) = -\frac{16M_{\rm Pl}^4}{\bar{M}^2} \left(\frac{\bar{M}^2}{8M_{\rm Pl}^4}\right)^3 \frac{\sum_a \omega_{\vec{k}_a}^{-1}}{\sum_a \omega_{\vec{k}_a}}, \quad (57)$$

while the 1PI correlator at time t is, from Eq. (13),

$$\Gamma^{(3)}(\vec{k}_1, \vec{k}_2, \vec{k}_3) = -\frac{16M_{\rm Pl}^4}{\bar{M}^2} \frac{\omega_{\vec{k}_1}\omega_{\vec{k}_2} + \omega_{\vec{k}_1}\omega_{\vec{k}_3} + \omega_{\vec{k}_2}\omega_{\vec{k}_3}}{\omega_{\vec{k}_1} + \omega_{\vec{k}_2} + \omega_{\vec{k}_3}}. \quad (58)$$

In the limit $\vec{k}_3 \rightarrow 0$, this becomes

$$\Gamma^{(3)}(\vec{k}, -\vec{k}, 0) = -\frac{8M_{\rm Pl}^4}{\bar{M}^2} \omega_{\vec{k}} = \Gamma^{(2)}(\vec{k}, -\vec{k}). \quad (59)$$

Given that $(3 - \mathcal{D}_1)\omega_{\vec{k}} = \omega_{\vec{k}}$, this verifies the dilation consistency relation for the three-point function. If, before taking the soft limit $\vec{k}_3 \rightarrow 0$, we differentiate with respect to $\vec{k}_3$, we can check the special conformal consistency condition as well:

$$2\nabla_{\vec{k}_3 \rightarrow 0}\Gamma^{(3)}(\vec{k}, -(\vec{k} + \vec{k}_3), \vec{k}_3) = -16 \frac{M_{\rm Pl}^4}{\bar{M}M^2} \vec{k}$$
$$= \tilde{\mathcal{S}}_{\vec{k}} \Gamma^{(2)}(\vec{k}, -\vec{k}), \quad (60)$$

in agreement with Eq. (37).

Because the background is flat, this example serves to illustrate that the consistency relations hold not as a consequence of the superhorizon freeze-out of modes but rather because of the residual diffeomorphisms associated with scale and special conformal invariance.

C. Nonexample: flat-space scalar field

Despite the relatively few assumptions that go into our derivation of the consistency relations, our results are not completely model-independent. A simple example that violates the consistency relations is the theory of a free massless scalar field $\phi(x)$ in flat spacetime—decoupled from gravity—with a time-dependent background. In a coordinate system $(x^0, \vec{x})$ in which the metric is Minkowskian, the theory has a solution with $\phi(x) = M^2 x^0$ for some arbitrary scale M, and fluctuations about this background are parametrized as $\phi(x) = M^2(x^0 + \pi(x))$.

It is also possible to describe the physics in unitary gauge with coordinates $(t, \vec{X})$ in which ϕ is spatially homogeneous on surfaces of constant time t,

$$t(x^0, \vec{x}) = x^0 + \pi(x^0, \vec{x}), \quad (61)$$

$$\vec{X}(x^0, \vec{x}) = \vec{x}. \quad (62)$$

We will need the following metric components in these coordinates:

$$g^{tt}(t, \vec{X}) = -\frac{1}{(1 - \dot{\pi})^2}[1 - (\nabla\pi)^2], \quad (63)$$

$$g_{ij}(t, \vec{X}) = \delta_{ij} + \nabla_i \pi \nabla_j \pi, \quad (64)$$

where π is now regarded as a function of $(t, \vec{X})$ via $\pi(t, \vec{X}) \equiv \pi(x^0(t, \vec{X}), \vec{x}(t, \vec{X}))$ and $\dot{} = \partial/\partial t$, $\nabla_i = \partial/\partial X^i$. Thus, in these coordinates the action for the fluctuation π becomes

$$S[\pi] = -\frac{M^4}{2} \int dt d^3\vec{X}(1 - \dot{\pi})g^{tt}(t, \vec{X})$$
$$= \frac{M^4}{2} \int dt d^3\vec{X}[\dot{\pi}^2 - (\nabla\pi)^2] + \mathcal{O}(\pi^3). \quad (65)$$

To extract ζ from Eq. (64), we decompose $g_{ij} = e^{2\zeta} h_{ij}$, with h_{ij} unimodular, and perform a further diffeomorphism $\vec{X} \rightarrow \vec{X} + \vec{\xi}(t, \vec{X})$ such that $\gamma_{ij} = h_{ij} - \delta_{ij}$ is transverse, $\nabla_i \gamma_{ij} = 0$, in the new coordinates. To quadratic order in π, this requires

$$\frac{4}{3}\nabla \cdot \vec{\xi} = \frac{1}{\nabla^2}\nabla_i\nabla_j(\nabla_i\pi\nabla_j\pi) - \frac{1}{3}(\nabla\pi)^2 + \mathcal{O}(\pi^3), \quad (66)$$

yielding

$$4\zeta = (\nabla\pi)^2 - \frac{1}{\nabla^2}\nabla_i\nabla_j(\nabla_i\pi\nabla_j\pi) + \mathcal{O}(\pi^3). \quad (67)$$

Seeing that the consistency relations do not hold in this example is now a matter of simple power counting in the scale M. From Eq. (65) the connected two-point function is $\langle\zeta\zeta\rangle \sim M^{-8}$, and thus $\Gamma^{(2)}(\vec{k}, -\vec{k}) \sim M^8$. On the other hand, the connected three-point function scales as M^{-12} so that $\Gamma^{(3)} \sim M^{12}$ up to terms involving more powers of momenta divided by M. Barring accidental cancellations, it follows that no simple relation between $\Gamma^{(2)}$ and $\Gamma^{(3)}$ is possible in this model.

In this example, there is a subtlety in applying the formal path integral manipulations of Sec. II. Although we can choose the standard unitary gauge of Eq. (3), the mapping between the scalar perturbation π that appears in the original Lagrangian and the ζ variable defined in unitary gauge starts at quadratic order in the π field. This is quite an unusual change of field variables for a perturbative field theory—quite different from the standard field redefinitions we are used to in quantum field theory, which, even if nonlinear, always start with linear terms. Usually, we require such linear plus higher-order field redefinitions because they do not affect S-matrix elements. In our case we are not computing S-matrix elements, and such a

requirement should not be relevant. Perhaps more to the point, our $\zeta = \mathcal{O}(\pi^2)$ change of variables is not invertible in any small neighborhood of $\pi = \zeta = 0$, and it has a singular $\delta\pi/\delta\zeta$ Jacobian about that point. Since our theory is perturbative in π (free, in fact) about $\pi = 0$, such a change of integration variables in the path integral is clearly pathological. We believe this to be the cause of the violation of the consistency relations in this example.

IV. DISCUSSION AND OUTLOOK

We derived fully quantum-mechanical consistency relations—or, in the language of quantum field theory, soft-ζ theorems—for 1PI vertex functions of scalar perturbations in cosmology. Our results follow from the residual gauge invariance of an equal-time analog of the quantum effective action for ζ, which generates equal-time correlation functions. The near-de Sitter isometries of inflation play no role in our derivation, nor do the existence of a cosmological horizon, the freezing-out of long-wavelength perturbations, or their becoming approximately classical. As a result, our consistency relations hold as exact quantum mechanical statements and in more general situations than the standard inflationary cosmology scenarios, like the flat-space ghost-condensate example of Sec. III B shows.

The residual gauge invariance of which we make use is the three-dimensional conformal group, under which ζ transforms as a dilaton. For dilations, the presence of tensor modes does not affect our arguments, and our consistency relations can be generalized straightforwardly to include tensor modes on the external legs. On the other hand, special conformal transformations do not preserve the transversality of tensor modes and must be dealt with more carefully. For 1PI correlation functions that only involve ζ on the external legs, this does not constitute a problem, for the reasons outlined in Sec. II B. For 1PI correlation functions that also have external tensors, this subtlety has to be addressed directly, perhaps via transverse projectors acting on the external tensors.[6] We leave investigating this subtlety for future work.

Dilations and special-conformal transformations have also been used recently in Ref. [4] to derive, via more traditional methods, consistency relations, which—not surprisingly—are in agreement with ours. However, we feel that the novelty of our work lies in clarifying the wide scope of validity of such consistency relations and the minimality of the assumptions that go into deriving them. Moreover, we identified (and removed) an implicit ambiguity in the special-conformal consistency relations. Our unambiguous version can be tested nontrivially already at the level of the slow-roll inflation three-point function.

Given how few assumptions we have made in deriving the consistency relations, it is not completely obvious to

see what goes wrong in situations where they are known to fail, like, for example, multifield inflation. There, we can still choose a gauge in which one of the time-dependent scalars has zero fluctuations. In that case, one can then parameterize the scalar modes in terms of the ζ field appearing in the metric in this gauge and of the remaining scalar fields. All these modes will transform in some definite way under our residual gauge transformations, and, once fields other than ζ have been integrated out, our $\Gamma[\bar\zeta]$ should have the same (gauge) symmetries as we have used and should thus obey the same consistency relations as we have derived.

The catch in this argument is that this definition of ζ does not coincide with the observationally relevant late-time curvature perturbation; all multifield examples that violate the consistency relations do so thanks to a substantial conversion of ''isocurvature'' fluctuations into adiabatic ones. One might try to work around this by choosing a ''late-time unitary gauge,'' where the late-time curvature perturbations are encoded in the spatial metric as in Eq. (3), and extrapolating this gauge choice backwards in time. It is not clear how this could work in general though; the most natural and universal definition of a post-reheating unitary gauge is to define time via the $\rho = $ const hypersurfaces, but for our purposes this is not quite the same as the unitary gauge. This is evident if one employs a $P(X)$ parameterization of the post-inflationary (vorticity-free) cosmic fluid, where $X = (\partial\phi)^2$ for some scalar ϕ, and P is a generic function (related to the fluid's equation of state.) In this parameterization, $\rho = $ const is the same as $X = $ const, which is not the same as $\phi = $ const. In other words, in the $\delta\rho = 0$ gauge, $\delta\phi$ is not zero, and scalar perturbations are parameterized by ζ *and* $\delta\phi$. The real unitary gauge choice would be to choose $\delta\phi = 0$ and define scalar perturbations in terms of the ζ variable appearing in the metric in that gauge, but it is not obvious how to extrapolate the $\delta\phi = 0$ gauge condition to earlier times, i.e., to inflationary times, when this scalar field ϕ has no meaning in general.

A subtle paradox remains, though, at least for models in which all isocurvature perturbations eventually disappear. One could choose the *initial* time in our effective action manipulations to be, in fact, quite late in the evolution of the Universe, say *after* the isocurvature perturbations are gone. Then, one would be left with adiabatic scalar perturbations in the post-inflationary cosmic fluid, which can be parameterized in the $\delta\phi = 0$ unitary gauge above by ζ only. Now there is no need to extrapolate this gauge back in time. It seems to us that in this case $\Gamma[\zeta]$ should obey all our symmetries and should thus respect the consistency relations. We are actually unclear about why it does not.

Similar considerations apply to the case of ''solid inflation,'' for which the three-point function also violates the consistency relations [20]. There, there is no standard unitary gauge choice; once the fluctuations in the matter

---

[6]A different approach will appear in Ref. [19].

fields are set to zero via a gauge choice, the metric cannot be put into the form (3). In fact, the most convenient version of the unitary gauge in that case would be to *remove* ζ from the metric via a time redefinition and parameterize scalar perturbations in terms of a scalar component of the traceless γ_{ij}, which in this gauge cannot be taken as transverse [20]. Like in the multifield case, here, too, there is no gauge in which the curvature perturbation that is relevant for observations at late times appears in the spatial metric as ζ does in Eq. (3). Finally, there is the counterexample of Sec. III C, which, despite the fact that it admits a standard unitary gauge, violates the consistency relations. In this case the violation can be traced to the specific form of the relation between ζ and the underlying scalar fluctuation, as discussed in Sec. III C.

ACKNOWLEDGMENTS

APPENDIX A: NONLINEAR SYMMETRIES OF Γ

The standard manipulations that are normally used to prove the symmetries of the effective action Γ are not entirely correct when applied to nonlinearly realized symmetries, like, for instance, spontaneously broken global symmetries. The reason is that the action of such symmetries on the field variables in the path integral changes the boundary conditions for those fields, e.g., from zero to something nonzero for Goldstone fields. Or, in the language of Sec. II, the vacuum wave functional Ψ_0 is not invariant under nonlinearly realized symmetries; for spontaneously broken global symmetries, this is true *by definition*. Since our residual diffeomorphisms (17) and (20) formally act nonlinearly on ζ (even though they do not correspond to spontaneous breaking in any standard sense), we have to take this subtlety into account.

For nonlinearly realized symmetries, to see any invariance at the level of the vacuum wave functional, we have to take into account that at any fixed time t, the wave functional depends not only on the field configuration $\zeta(\vec{x})$, but secretly also on a background configuration $\zeta_b(\vec{x})$, which can be thought of as the boundary condition for $\zeta(\vec{x})$ at spatial infinity, or more in general as the vacuum expectation violation $\langle\zeta(\vec{x})\rangle$ in the absence of external sources (the necessity for an $\vec{x}$ argument for ζ_b will be clear in a moment). The standard choice for the vacuum would be $\zeta_b = 0$, but a nonlinear symmetry transformation can

change that, so that the general symmetry statement for the vacuum wave functional is[7]

$$\Psi_0[\zeta, \zeta_b = 0] = \Psi_0[\zeta + \Delta\zeta, \zeta_b = \langle\Delta\zeta\rangle]. \quad (A1)$$

For notational simplicity, from now on we will be ignoring the tensor modes, and we will omit the "observation time" t_*. The symmetry transformations in which we will be interested are the scale transformations (17), which generate a constant background,

$$\zeta_b = \lambda, \quad (A2)$$

and the special conformal transformations (26), which generate a linear one:

$$\zeta_b(\vec{x}) = \vec{b} \cdot \vec{x}. \quad (A3)$$

One can now go through the same steps as in Sec. II, keeping the background dependence of the various quantities explicit:

$$\Psi[\zeta] \rightarrow \Psi[\zeta, \zeta_b], \qquad P[\zeta] \rightarrow P[\zeta, \zeta_b], \quad (A4)$$

$$W[J] \rightarrow W[J, \zeta_b], \qquad \Gamma[\bar{\zeta}] \rightarrow \Gamma[\bar{\zeta}, \zeta_b], \quad (A5)$$

where Γ is still defined as the Legendre transform of W with respect to J only. In particular, standard properties of the Legendre transform imply

$$\left.\frac{\delta\Gamma}{\delta\zeta_b}\right|_{\bar{\zeta}} = \left.\frac{\delta W}{\delta\zeta_b}\right|_J. \quad (A6)$$

Using the symmetry statement (A1) and (15) now gets generalized to

$$\int d^3x \langle\Delta\zeta(\vec{x})\rangle_J \frac{\delta\Gamma[\bar{\zeta}, \zeta_b = 0]}{\delta\bar{\zeta}(\vec{x})} + \langle\Delta\zeta(\vec{x})\rangle_0 \frac{\delta\Gamma[\bar{\zeta}, \zeta_b = 0]}{\delta\bar{\zeta}_b(\vec{x})} = 0, \quad (A7)$$

which is just the statement that $\Gamma[\bar{\zeta}, \zeta_b = 0]$ is invariant under a simultaneous transformation on $\bar{\zeta}$ and ζ_b:

$$\Gamma[\bar{\zeta}, \zeta_b = 0] = \Gamma[\bar{\zeta} + \langle\Delta\zeta\rangle_J, \zeta_b = \langle\Delta\zeta\rangle_0]. \quad (A8)$$

From the viewpoint of our mathematical manipulations so far, $\bar{\zeta}$ and ζ_b are independent variables. However, the generating functional provides a nontrivial relation between them—the expected one, $\zeta_b(\vec{x})$, is the configuration $\bar{\zeta}(\vec{x})$ reduces to in the absence of external sources:

---

[7]For spontaneously broken symmetries, this property of the wave functional is just the statement that there are many physically equivalent vacua, all related by symmetry transformations, and the wave functionals of the fields in all these vacua are the same, provided one transforms the fields properly.

$$\bar{\zeta}(\vec{x})|_{J=0} = \delta W/\delta J(x)|_{J=0} = \langle \zeta(\vec{x}) \rangle_{J=0} = \zeta_b(\vec{x}). \quad \text{(A9)}$$

Therefore, if we restrict the effective action Γ to configurations that are localized perturbations—which we can turn on via localized sources—of some background field configuration, we can identify $\zeta_b(\vec{x})$ with such a background configuration, which can be inferred from the behavior of $\bar{\zeta}(\vec{x})$ at spatial infinity.[8] In such a case, the effective action just becomes a functional of $\bar{\zeta}$,

$$\Gamma[\bar{\zeta}] \equiv \Gamma[\bar{\zeta}; \zeta_b = \bar{\zeta}(\infty)], \quad \text{(A10)}$$

and now Eq. (A8) reduces to the expected symmetry statement:

$$\Gamma[\bar{\zeta}] = \Gamma[\bar{\zeta} + \langle \zeta \rangle_J], \quad \text{(A11)}$$

where it is understood that $\bar{\zeta}(\vec{x})$ vanishes at spatial infinity [and, as a consequence, so does $J(\vec{x})$].

To summarize, in Sec. II all statements based on symmetry prior to Eq. (16) need qualifying, but the final symmetry statement on $\Gamma[\bar{\zeta}]$ is correct. Note that the same qualifications—and the same conclusion—apply to the case of spontaneously broken global symmetries.

APPENDIX B: AMBIGUITY CANCELLATION

Suppose we extrapolate our 1PI vertices $\Gamma^{(n+1)}$, $\Gamma^{(n)}$ out of the momentum-conserving surface by making them functions respectively of $n + 1$ and n *unconstrained* momenta. Since Eq. (29) involves first derivatives on the lhs and second derivatives on the rhs, as well as momentum-conserving delta functions, we only need to consider displacements from the momentum-conserving surface that are of first order for the lhs and up to second order for the rhs:

$$\Gamma^{(n+1)} \to \Gamma^{(n+1)} + \gamma^{(n+1)} \equiv \Gamma^{(n+1)} + \vec{u}_{n+1} \cdot \vec{g}_{n+1} \quad \text{(B1)}$$

$$\Gamma^{(n)} \to \Gamma^{(n)} + \gamma^{(n)} \equiv \Gamma^{(n)} + \vec{u}_n \cdot \vec{g}_n + u_n^i u_n^j g_n^{ij}, \quad \text{(B2)}$$

where the $\vec{u}$'s are the relevant total momenta,

$$\vec{u}_{n+1} = \vec{q} + \sum_b \vec{k}_b, \qquad \vec{u}_n = \sum_b \vec{k}_b, \quad \text{(B3)}$$

and the g_{n+1}, g_n vectors and tensors are generic functions of all the $n + 1$ and n momenta involved. If we now expand the derivatives of Eq. (29) like we did in Sec. II B, we end up with the following additional contributions:

For the left hand side:

$$2\vec{\nabla}_{\vec{q}\to0}\delta(\vec{u}_{n+1})\gamma^{(n+1)}|_{\vec{q}\to0} + 2\delta(\vec{u}_n)\vec{g}_{n+1}|_{\vec{q}\to0}. \quad \text{(B4)}$$

For the right hand side:

$$2\vec{\nabla}_{\vec{q}\to0}\delta(\vec{u}_{n+1})(3 - \mathcal{D}_n)\gamma^{(n)} + \delta(\vec{u}_n)\bigg[2(2 - \mathcal{D}_n)\vec{g}_n$$
$$+ g_n^{ij}\sum_b \vec{\mathcal{S}}_b(u_n^i u_n^j)\bigg], \quad \text{(B5)}$$

where we have used the same identities for derivatives of the delta function as in Sec. II B. We have also used the vanishing of $L^{ij}\gamma^{(n)}$ due to rotational invariance as well as an analogous property of $\vec{g}_n$.[9] It is a matter of simple algebra to show that

$$\sum_b \mathcal{S}_b^k(u_n^i u_n^j) = 2(\delta^{ij} u_n^k - \delta^{ik} u_n^j - \delta^{jk} u_n^i), \quad \text{(B6)}$$

which thus vanishes when multiplied by the delta function above.

We are thus left with these new contributions to our special-conformal consistency relation:

$$\text{lhs: } 2\vec{\nabla}_{\vec{q}\to0}\delta(\vec{u}_{n+1})\gamma^{(n+1)}|_{\vec{q}\to0} + 2\delta(\vec{u}_n)\vec{g}_{n+1}|_{\vec{q}\to0} \quad \text{(B7)}$$

$$\text{rhs: } 2\vec{\nabla}_{\vec{q}\to0}\delta(\vec{u}_{n+1})(3 - \mathcal{D}_n)\gamma^{(n)} + 2\delta(\vec{u}_n)(2 - \mathcal{D}_n)\vec{g}_n. \quad \text{(B8)}$$

The terms involving the derivative of the delta function cancel between the lhs and rhs *only if* we assume that $\gamma^{(n+1)}$ and $\gamma^{(n)}$ obey the dilation consistency relation, that is, if we make sure that the extrapolations (B1) and (B2) preserve the dilation consistency relation, now between 1PI vertices evaluated at momenta that do not add up to zero. This is an in-principle nontrivial check one should perform on the candidate "off-shell" 1PI vertices. Once they pass it, one is then left with the terms involving the nondifferentiated delta functions, which also vanish, for exactly the same reason: given the relation between the γ's and the $\vec{g}$'s [see Eqs. (B1) and (B2)], the $\vec{g}$'s obey the dilation consistency relation

$$\vec{g}_{n+1}|_{\vec{q}\to0} = (2 - \mathcal{D}_n)\vec{g}_n \quad \text{(B9)}$$

as long as the γ's obey the usual one.

---

[1] J. M. Maldacena, J. High Energy Phys. 05 (2003) 013.

[2] P. Creminelli and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 10 (2004) 006.

[3] L. Senatore and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 08 (2012) 001.

[4] P. Creminelli, J. Norena, and M. Simonovic, J. Cosmol. Astropart. Phys. 07 (2012) 052.

[5] S. Weinberg, Phys. Rev. D **67**, 123504 (2003).

[6] K. Hinterbichler, L. Hui, and J. Khoury, J. Cosmol. Astropart. Phys. 08 (2012) 017.

[7] V. Assassi, D. Baumann, and D. Green, J. Cosmol. Astropart. Phys. 11 (2012) 047.

[8] V. Assassi, D. Baumann, and D. Green, J. High Energy Phys. 02 (2013) 151.

[9] A. Kehagias and A. Riotto, Nucl. Phys. **B864**, 492 (2012); **B868**, 577 (2013).

[10] K. Schalm, G. Shiu, and T. van der Aalst, J. Cosmol. Astropart. Phys. 03 (2013) 005.

[11] X. Chen, H. Firouzjahi, M. H. Namjoo, and M. Sasaki, arXiv:1301.5699.

[12] N. Agarwal, R. Holman, A. J. Tolley, and J. Lin, arXiv:1212.1172.

[13] T. Tanaka and Y. Urakawa, J. Cosmol. Astropart. Phys. 05 (2011) 014.

[14] J. M. Maldacena and G. L. Pimentel, J. High Energy Phys. 09 (2011) 045.

[15] N. Arkani-Hamed, H.-C. Cheng, M. A. Luty, and S. Mukohyama, J. High Energy Phys. 05 (2004) 074.

[16] S. Weinberg, *Modern Applications*, The Quantum Theory of Fields Vol. 2 (Cambridge University Press, Cambridge, England, 1996), p. 489.

[17] J. S. Schwinger, J. Math. Phys. (N.Y.) **2**, 407 (1961); L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)]; K. T. Mahanthappa, Phys. Rev. **126**, 329 (1962); P. M. Bakshi and K. T. Mahanthappa, J. Math. Phys. (N.Y.) **4**, 12 (1963); R. D. Jordan, Phys. Rev. D **33**, 444 (1986); E. Calzetta and B. L. Hu, Phys. Rev. D **35**, 495 (1987).

[18] P. Creminelli, M. A. Luty, A. Nicolis, and L. Senatore, J. High Energy Phys. 12 (2006) 080.

[19] K. Hinterbichler, L. Hui, and J. Khoury (unpublished).

[20] S. Endlich, A. Nicolis, and J. Wang, arXiv:1210.0569.