de Sitter *S* matrix for the masses

Scott Melville
Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom

Guilherme L. Pimentel
Scuola Normale Superiore and INFN, Piazza dei Cavalieri 7, Pisa, 56126, Italy

(Received 26 September 2023; revised 20 December 2023; accepted 7 October 2024; published 19 November 2024)

We define an *S* matrix for massive scalar fields on a fixed de Sitter spacetime, in the expanding patch coordinates relevant for early Universe cosmology. It enjoys many of the same properties as its Minkowski counterpart, for instance: it is insensitive to total derivatives and field redefinitions in the action; it can be extracted as a particular "on shell" limit of time-ordered correlation functions; and for low-point scattering, kinematics strongly constrains its possible structures. We present explicit formulas relating inflationary observables—namely in-in equal-time correlators and wavefunction coefficients at the conformal boundary—to these *S* matrix elements. This new formalism will allow modern amplitude methods to be applied directly in cosmology and hence provide a wider range of more accurate theoretical predictions for upcoming sky surveys.

DOI: 10.1103/PhysRevD.110.103530

I. INTRODUCTION

The most fundamental observables in cosmology are the spatial correlations present at the "beginning," in the hot big bang of the early Universe. As these initial conditions evolve, they source perturbations in the primordial plasma which are imprinted as temperature and polarization fluctuations in the cosmic microwave background radiation (CMB). The same initial conditions then go on to control the clustering of dark matter and galaxy formation within the large scale structure of the late Universe (LSS). Therefore, almost all cosmological information accessible to observations can ultimately be traced back to these spatial correlations.

The leading paradigm for explaining these initial conditions is cosmic inflation—a period of accelerated spacetime expansion prior to the hot big bang—in which the "initial conditions" are causally sourced by short-distance fluctuations which are stretched to cosmological scales. A useful analogy is that of particle production in a collider experiment, in which an incoming beam of particles gives an energetic environment from which new particles can be produced. In inflationary cosmology, the expansion of the Universe is the source of energy from which new particles can be excited from the vacuum. It is these produced particles which go on to source the initial conditions of our Universe and hence provide an observable output of this cosmological collider.

Pushing the collider analogy further, we can ask how these particles produced by the expansion of the Universe would interact and scatter off each other. In a collider experiment, this scattering is characterized by an observable cross section. In cosmology, particle interactions give rise to an observable deviation of the initial conditions from a Gaussian random process. Measuring and bounding these primordial non-Gaussianities from the CMB and LSS is an important goal of observational cosmology.

From a theoretical standpoint, it is clear that the same observable information can be encoded in more primitive objects. Matrix elements between initial and final states, or scattering amplitudes, are routinely used in collider physics because they are mathematically simple objects which capture all observable cross sections in an efficient way. The main purposes of this paper are to develop the cosmological analog of a scattering amplitude: an *S* matrix for cosmology. These *S* matrix elements should efficiently encode all of the primordial non-Gaussianities produced by the cosmological collider while also enjoying more mathematical structure and a closer relation to collider amplitudes. This will not only expedite modeling of the early inflationary Universe—allowing an efficient exploration of the theory space and their different phenomenological signatures—but will also allow for powerful modern amplitude techniques to be applied directly in a cosmological setting.

Further motivation is that the best understood observables in quantum field theory are asymptotic. As the separation between detectors is taken to be infinitely large in a controllable way, the overlap between different quantum states can be computed at weak coupling using perturbation theory. The S matrix captures this information in a clearer fashion than observables like the cross section. Moreover, as gravity forbids the existence of local operators, asymptotic observables might be the only ones that ultimately make sense in a theory of gravity. Our cosmological S matrix is asymptotic because it relates states in the infinite past to states at the end of the inflationary spacetime, when the hot big bang picture takes over—in that sense the "asymptotic future" for inflation is where the "initial conditions" are formed for the rest of cosmology!

We will focus on the cosmological example of *de Sitter spacetime*. The motivation is twofold: de Sitter is a good approximation of inflation, and being a maximally symmetric spacetime it enjoys many symmetries that make the definition of asymptotic states more amenable, forbidding, e.g., the decay of freely propagating particles. While the S matrix is well understood in Minkowski spacetime [1,2] and, to some extent, in anti–de Sitter (AdS) [3–8], a de Sitter counterpart has received much less attention (though see [9–12]). This is despite much recent activity in the study of cosmological observables in perturbation theory, which typically focuses on the observable non-Gaussianity directly (for reviews see [13,14]).

In this paper we present a concrete definition of the S matrix in de Sitter spacetime that is directly applicable to primordial cosmology. We work throughout in the expanding Poincaré patch, using coordinates $ds^2 = \tau^{-2}(-d\tau^2 + d\mathbf{x}^2)$ where the conformal time $\tau < 0$ (in units where the Hubble rate $H = 1$). Our approach parallels the standard treatment of the Minkowski S matrix and is in the same spirit as the pioneering work of Marolf *et al.* for global de Sitter [15]. An advantage of working in the expanding Poincaré patch is that our S matrix elements explicitly connect to the inflationary wavefunction and primordial non-Gaussianities that characterize the early Universe. These phenomenological connections are one of the main insights which allowed us to make conceptual progress with this problem: since the wavefunction and primordial non-Gaussianities exist and are well understood (at least in perturbation theory), that understanding can now be translated into the language of S matrix elements.

This sheds light on previously observed obstacles to defining an S matrix for light fields on de Sitter. For particular mass values, the wavefunction develops divergences at late times that require holographic renormalization [16,17]. The relation we derive between the wavefunction and S matrix makes it clear that similar divergences must appear in the S matrix elements for light fields with masses $m < d/2$ in d spatial dimensions.

Our first goal is therefore to define and study the finite S-matrices that describe the scattering of *massive scalars* on a fixed de Sitter background—specifically, scalars with masses $m \geq d/2$ and therefore in the principal series of irreducible representations of the de Sitter group. We will then return to light scalars and show that certain S matrix elements remain finite for specific interactions. This closely parallels the development of S matrix theory on Minkowski spacetime, where the assumption of a mass gap is crucial to derive many of the foundational results which we then extend and apply to gapless theories (dealing with any IR divergences as and when they arise). In the context of inflation, the derivative interactions between massless scalars which naturally arise in, e.g., the effective field theory of inflation do not produce any IR divergences in our S matrix and can therefore be treated using the formalism we develop here. Although we study only scalar fields here, we believe that the inclusion of spin is a technical hurdle that can be overcome, at least for massive particles. The case of dynamical gravity is more subtle—see [18] for arguments that the S matrix might not even exist in that case. While these extensions are interesting and deserve further investigation, our results provide a first step toward understanding cosmological observables through their underlying S matrix description.

An important question to address will be how to construct the asymptotic states. Schematically, from the time evolution operator U from conformal time $\tau = -\infty$ to $\tau = 0$, one defines matrix elements $\langle b|U|a\rangle$ that naturally depend on the choice of $|a\rangle$ and $|b\rangle$. Given these elements for a complete basis of $|a\rangle$ and a complete basis of $|b\rangle$, all information about time evolution is specified. Of course one is always free to rotate either the basis of bra's or the basis of ket's, which will produce different representations of U that contain the same physical information. We identify two natural bases to use, and refer to the resulting matrix elements as the "Bunch-Davies" S matrix and the "Unruh-de Witt" S matrix. While these objects encode precisely the same information (as they are matrix elements of the same underlying U), we find that the Bunch-Davies choice enjoys a number of useful features, including

(i) a simple crossing relation that exchanges particles between the in and out states,
(ii) a simple analytic structure in the complex energy plane,
(iii) a close connection to the wavefunction and in-in correlators used in inflationary cosmology,
(iv) a flat space limit in which it coincides with the standard Minkowski S matrix.

With this paper, our aim is to provide the first step and set up a new formalism in which the cosmological collider can provide clear, observer-independent predictions for upcoming sky surveys, in a language closest to the standard of theoretical cosmologists and also phenomenologists, as well as paving the way for scattering amplitude specialists and collider physicists to see that their techniques are also useful in a very different context: namely primordial cosmology. Just as the data from large collider experiments are nowadays used to infer the underlying scattering amplitude (from which we can immediately determine properties of the high-energy quantum fields), in the future we envision large-scale sky surveys being used to determine the underlying S matrix of inflation, which would similarly bridge between what is observed and properties of the high-energy quantum fields in the early Universe.

II. DEFINING AN S MATRIX

We will begin with an abstract discussion of what the different S matrix elements represent, and then provide a concrete definition in terms of field theory correlators.

A. Choice of basis

In order to define S matrix elements, one requires a basis of "in" and "out" states. On Minkowski spacetime, there is a natural choice: using the particle eigenstates $|n\rangle$ of the free theory (i.e., eigenstates of the free Hamiltonian), we can define the "in"/"out" states of the interacting theory to be those which coincide with $|n\rangle$ in the far past/future. However, on de Sitter the number of particles is not conserved due to gravitational particle production: the state $|n,\tau_*\rangle$ which contains n particles at time τ_* is *not* an eigenstate of the free Hamiltonian at later times $\tau \neq \tau_*$, since n particles will generally evolve into a superposition of more/fewer particles due to the expansion of spacetime. This presents a choice in how we define our asymptotic states.

One natural choice is to define the "in"/"out" states of the interacting theory to be those that coincide with $|n,-\infty\rangle$ in the far past/future. We will use $|n,-\infty\rangle_{\text{in}}$ and $|n,-\infty\rangle_{\text{out}}$ to denote these states. They are connected by the S matrix elements

$$\mathcal{S}_{n'\to n} \equiv {}_{\text{out}}\langle n,-\infty|n',-\infty\rangle_{\text{in}}, \tag{1}$$

where we work throughout in the Heisenberg picture. Equivalently, these matrix elements are the coefficients in the expansion

$$|n',-\infty\rangle_{\text{in}} = \sum_n \mathcal{S}_{n'\to n}|n,-\infty\rangle_{\text{out}}, \tag{2}$$

where the sum over n' includes integrals over all momenta and other quantum numbers of the n particles. We will therefore refer to this $\mathcal{S}$ as the *Bunch-Davies S matrix*, since it describes the time evolution of the Bunch-Davies vacuum state $|0,-\infty\rangle$ (and its excitations) in the interacting theory [19]. In the Heisenberg picture, the state itself does not depend on time—as on Minkowksi, when one says that the state $|\alpha\rangle_{\text{in}}$ "coincides" in the far past with the state $|\alpha\rangle$, this really means that physical expectation values will agree: e.g., $\lim_{\tau\to-\infty\text{in}}\langle\alpha|\mathcal{O}(\tau)|\alpha\rangle_{\text{in}} = \lim_{\tau\to-\infty}\langle\alpha|\mathcal{O}(\tau)|\alpha\rangle$.

Another, equally natural, choice is to instead define the "out" states of the interacting theory as those that coincide with $|n,0\rangle$ in the far future. We denote these states by $|n,0\rangle_{\text{out}}$. This produces a different set of S matrix elements,

$$\mathtt{S}_{n'\to n} \equiv {}_{\text{out}}\langle n,0|n',-\infty\rangle_{\text{in}}. \tag{3}$$

We will refer to this $\mathtt{S}$ as the *Unruh-DeWitt S matrix*, since it describes scattering from a state containing n' particles to a state containing n particles, as measured by an Unruh-DeWitt detector in the far past/future.

These two sets of S matrix elements are ultimately related by a Bogoliubov transformation (which maps $|n,-\infty\rangle$ to $|n,0\rangle$ in the free theory). We will initially focus on the Bunch-Davies S matrix because of the properties (i–iv) listed above. It also has no particle production in the free theory, i.e., without interactions, all off diagonal $\mathcal{S}_{n'\to n}$ vanish by construction (unlike $\mathtt{S}_{n'\to n}$, which is nontrivial even in free field theory).

Finally, in practice we typically replace the asymptotic states of the interacting theory with the free theory $|n,-\infty\rangle$ in order to perform perturbative calculations. As on Minkowski, this leads to an expression for the S matrix in terms of the time-evolution operator $U(\tau_1,\tau_2)$ between times τ_1 and τ_2,

$$\mathcal{S}_{n'\to n} = \langle n,-\infty|U^\dagger_{\text{free}}(0,-\infty)U(0,-\infty)|n',-\infty\rangle. \tag{4}$$

The appearance of $U^\dagger_{\text{free}}$ makes it clear that this object reduces to the identity when interactions are turned off.

B. The S matrix from a reduction formula

The S matrix overlap in (1) can be extracted from time-ordered correlation functions by "amputating" the external legs and going "on shell," in analogy with the Lehmann–Symanzik–Zimmermann (LSZ) formula in flat space. Concretely, consider a real scalar field $\phi(\tau,\mathbf{x})$ of mass $m^2 = (d/2)^2 + \mu^2$. We split the action $S = S_{\text{free}} + S_{\text{int}}$, where the free quadratic action is

$$S_{\text{free}} = -\int d\tau d^d\mathbf{x}\sqrt{-g}\frac{Z^2}{2}(g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + m^2\phi^2) \tag{5}$$

and S_{int} contains all nonlinear interactions. Canonical quantization then proceeds as usual: we first quantize the free theory S_{free}, which can be done exactly, and then include the effects of S_{int} as a small perturbation.

Performing a Fourier transform from position $\mathbf{x}$ to momentum $\mathbf{k}$, the free equation of motion for the canonically normalized $\varphi(\tau,\mathbf{k}) \equiv (-\tau)^{-d/2}\phi(\tau,\mathbf{k})$ is [20]

$$\mathcal{E}[k\tau]\varphi(\tau,\mathbf{k}) \equiv [(\tau\partial_\tau)^2 + k^2\tau^2 + \mu^2]\varphi(\tau,\mathbf{k}) = 0. \tag{6}$$

In the Heisenberg picture, the time evolution of the $\hat{\varphi}$ operator can therefore be written as

$$\hat{\varphi}(\tau,\mathbf{k}) = f^-(k\tau)\hat{a}_{-\mathbf{k}} + f^+(k\tau)\hat{a}^\dagger_{\mathbf{k}}, \tag{7}$$

where the mode functions satisfy the free equation of motion $\mathcal{E}[k\tau]f^{\pm}(k\tau) = 0$ with the boundary condition

$$0 = (\tau\partial_\tau \pm i\sqrt{k^2\tau^2 + \mu^2})f^{\pm}(k\tau)|_{\tau=\tau_*}, \tag{8}$$

which ensures that $\hat{a}_{\mathbf{k}}$ diagonalizes the free Hamiltonian at time τ_*. Consequently, $\hat{a}_{\mathbf{k}}|0,\tau_*\rangle = 0$ defines the instantaneous vacuum state $|0,\tau_*\rangle$ (the state with the lowest energy at time τ_*), and $\hat{a}^\dagger_{\mathbf{k}}$ creates a "particle" of momentum $\mathbf{k}$ at time τ_* (an excited eigenstate of the Hamiltonian at time τ_*). A complete basis of states for the Hilbert space is then provided by

$$|n,\tau_*\rangle \equiv \hat{a}^\dagger_n \ldots \hat{a}^\dagger_1|0\rangle, \tag{9}$$

where the label on each $\hat{a}^\dagger$ denotes both the momenta and all other quantum numbers (e.g., mass) of that particle, and $|n\rangle$ denotes the complete list of this n-particle data. For the Bunch-Davies S matrix, we impose the vacuum condition at $\tau_* \to -\infty$, which corresponds to Hankel mode functions,

$$f^+(z) \equiv \frac{\sqrt{\pi}}{2iZ} e^{+\frac{\pi}{2}\mu} H^{(2)}_{i\mu}(-z) = [f^-(z^*)]^*, \tag{10}$$

which have been normalized so that [21]

$$iZ^2 f^-(k\tau)(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}(\tau,\mathbf{k}) = \hat{a}^\dagger_{\mathbf{k}}, \tag{11}$$

where $[\hat{a}_{\mathbf{k}'}, \hat{a}^\dagger_{\mathbf{k}}] = (2\pi)^d \delta^d(\mathbf{k} + \mathbf{k}')$.

In the interacting theory, we now seek to define states $|n,-\infty\rangle_{\text{in}}$ and $|n,-\infty\rangle_{\text{out}}$, which coincide with the state $|n,-\infty\rangle$ as $\tau \to -\infty$ and $\tau \to 0$ respectively. By "coincide," we mean for instance that

$$\lim_{\tau\to-\infty}\langle\alpha|\hat{\mathcal{O}}(\tau)|0,-\infty\rangle_{\text{in}} = \lim_{\tau\to-\infty}\langle\alpha|\hat{\mathcal{O}}(\tau)|0,-\infty\rangle \tag{12}$$

for any operator $\hat{\mathcal{O}}$ and normalizable state $|\alpha\rangle$ in the Heisenberg picture [and strictly speaking the limit on the right-hand side should be $\tau \to -\infty(1 - i\epsilon)$ to ensure convergence].

For brevity, from now on we will denote Bunch-Davies asymptotic states as $|n\rangle_{\text{in}}$ and $|n\rangle_{\text{out}}$.

The idea is then to find an operator which acts on $|0\rangle_{\text{in}}$ to create the one-particle in state $|1\rangle_{\text{in}}$. We claim that

$$\langle\alpha|1\rangle_{\text{in}} = \lim_{\tau\to-\infty}\langle\alpha|iZ^2 f^-(k\tau)(\tau\overleftrightarrow{\partial_\tau})\,\hat{\varphi}(\tau,\mathbf{k})|0\rangle_{\text{in}} \tag{13}$$

for any normalizable state $|\alpha\rangle$. Clearly the right-hand side generates a one-particle state in the free theory thanks to (11), and we argue in the Appendix that in the limit $\tau \to -\infty$ the interactions turn off sufficiently quickly that this operator produces the desired one-particle in state. This is the same "adiabatic hypothesis" used to define LSZ operators on Minkowski. Then any $|n\rangle_{\text{in}}$ can be constructed by repeated application of (13).

Similarly, we claim that for the out states,

$$_{\text{out}}\langle 1|\alpha\rangle = \lim_{\tau\to 0}{}_{\text{out}}\langle 0|iZ^2 f^+(k\tau)(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}^\dagger(\tau,\mathbf{k})|\alpha\rangle \tag{14}$$

for any normalizable state $|\alpha\rangle$. Again this requires that the interactions turn off sufficiently quickly at late times, which is the case for massive fields in the principal series and derivatively coupled fields in the complementary series (see the Appendix for details). We also have the useful corollary that these operators can be used to annihilate $_{\text{out}}\langle 0|$ and $|0\rangle_{\text{in}}$; for instance,

$$\lim_{\tau\to 0}{}_{\text{out}}\langle 0|f^-(k\tau)(\tau\overleftrightarrow{\partial_\tau})\,\hat{\varphi}(\tau,\mathbf{k})|\alpha\rangle = 0. \tag{15}$$

To relate the S matrix elements to a field correlator, we can now follow the analogous steps as in flat space. By applying (13), we see that any particle from the in state can be replaced by a field insertion,

$$\begin{aligned} iZ^{-2}{}_{\text{out}}\langle n'|n\rangle_{\text{in}} &= -\lim_{\tau\to-\infty}{}_{\text{out}}\langle n'|f^-(k_n\tau)(\tau\overleftrightarrow{\partial_\tau})\,\hat{\varphi}(\tau,\mathbf{k}_n)|n-1\rangle_{\text{in}} \\ &= \int_{-\infty}^0 d\tau\partial_\tau[{}_{\text{out}}\langle n'|f^-(k_n\tau)(\tau\overleftrightarrow{\partial_\tau})\,\hat{\varphi}(\tau,\mathbf{k}_n)|n-1\rangle_{\text{in}}] \\ &= \int_{-\infty}^0 \frac{d\tau}{\tau} f^-(k_n\tau)\mathcal{E}[k_n\tau]_{\text{out}}\langle n'|\hat{\varphi}(\tau,\mathbf{k}_n)|n-1\rangle_{\text{in}}. \end{aligned} \tag{16}$$

Note that in going to the penultimate line we have assumed that none of the momenta in $\langle n'|$ coincide with those in $\mathbf{k}_n$, and therefore we can use (15) to discard the $\tau \to 0$ limit of the integral. This amounts to considering the *connected* part of the S matrix element [22]. Proceeding in the same way for each particle in $|n\rangle$ and $\langle n'|$, one can reduce the right-hand side to the vacuum expectation of a (time-ordered) product of field insertions. We therefore define the correlator, the amputated correlator, and the connected part of the Bunch-Davies S matrix element by

$$\begin{aligned} G_{n'\to n} &\equiv {}_{\text{out}}\langle 0|T\prod_{b=1}^{n}\hat{\varphi}^\dagger(\tau_b,\mathbf{k}_b)\prod_{b'=1}^{n'}\hat{\varphi}(\tau'_{b'},\mathbf{k}'_{b'})|0\rangle_{\text{in}}, \\ \mathcal{G}_{n'\to n} &\equiv \left[\prod_{b=1}^{n} iZ^2\mathcal{E}[k_b\tau_b]\right]\left[\prod_{b'=1}^{n'} iZ^2\mathcal{E}[k'_{b'}\tau_{b'}]\right]G_{n\to n'}, \\ \mathcal{S}_{n'\to n} &= \left[\prod_{b=1}^{n}\int_{-\infty}^0 \frac{d\tau_b}{-\tau_b}f^+(k_b\tau_b)\right] \\ &\quad\times\left[\prod_{b'=1}^{n'}\int_{-\infty}^0 \frac{d\tau'_b}{-\tau'_b}f^-(k'_b\tau'_b)\right]\mathcal{G}_{n'\to n}, \end{aligned} \tag{17}$$

where T represents time ordering in τ, and the lower limits of the time integrals are understood to be $\tau \to -\infty(1 \mp i\epsilon)$ for the ingoing/outgoing particles.

Formula (17) is our prescription for the de Sitter S matrix: in words, one should first compute the time-ordered correlation function, then apply the classical equations of motion to each field (this "amputates" its external leg from any Feynman diagram), and finally perform an integral transform using Hankel mode functions (this puts the external legs "on shell").

C. Perturbation theory

To compute $\mathcal{S}_{n'\to n}$ in perturbation theory, one can go to the interaction picture and expand in Feynman diagrams in which

(1) outgoing external lines represent the free mode function $f^+(k\tau)$,
(2) ingoing external lines represent the free mode function $f^-(k\tau)$,
(3) internal lines represent the free theory propagator

$$\langle 0|T\hat{\varphi}(\tau,\mathbf{k})\hat{\varphi}(\tau',\mathbf{k}')|0\rangle \equiv G_2(k\tau,k\tau')(2\pi)^d\delta^d(\mathbf{k}+\mathbf{k}'),$$

(4) n-point vertices represent local interactions involving n powers of φ, and multiply the above propagators by a vertex factor of $i\delta^n S_{\text{int}}/\delta\varphi^n$,
(5) and finally, all internal times and momenta are integrated over.

Regardless of the contention about stability/existence of de Sitter spacetime in a quantum theory of gravity, this $\mathcal{S}_{n'\to n}$ certainly exists perturbatively. For instance, a local interaction $\sqrt{-g}\frac{\lambda_n}{n!}\phi^n$ in S_{int} will produce a calculable "contact" contribution to $\mathcal{S}_{0\to n}$ of

$$\mathcal{S}^{\text{cont}}_{0\to n} = i\lambda_n \int_{-\infty}^{0} \frac{d\tau}{-\tau}(-\tau)^{\frac{d}{2}(n-2)} \prod_{b=1}^{n} f^+(k_b\tau_b), \quad (18)$$

where we have suppressed the overall momentum-conserving δ function. It will also produce "exchange" diagram contributions to higher-order S matrix elements: for instance, a ϕ^3 interaction will give the following contribution to $\mathcal{S}_{0\to 4}$:

$$\mathcal{S}^{\text{exch}}_{0\to 4} = -\lambda_3^2 \int_{-\infty}^{0} \frac{d\tau}{-\tau}(-\tau)^{d/2} \int_{-\infty}^{0} \frac{d\tau'}{-\tau'}(-\tau')^{d/2} f^+(k_1\tau) \times f^+(k_2\tau) G_2(k_s\tau, k_s\tau') f^+(k_3\tau') f^+(k_4\tau') + 2\,\text{perm}, \quad (19)$$

where $k_s \equiv |\mathbf{k}_1 + \mathbf{k}_2|$ and "2 perm." denotes the t and u-channel contributions (again omitting a δ function).

D. The other S matrix

To extract the connected part of the Unruh-DeWitt S matrix (3) from a field correlator, we must make two changes to the LSZ formula (17):

(i) the mode functions for the ingoing and outgoing particles should be replaced by

$$f^-(k\tau) \to \frac{1}{\alpha}\mathtt{f}^-(k\tau), \qquad f^+(k\tau) \to \frac{1}{\alpha} f^+(k\tau), \quad (20)$$

where $\mathtt{f}^\pm$ solves the free equation of motion with the vacuum condition (8) imposed at $\tau_* = 0$, since the operators $\mathtt{f}^-(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}$ and $f^+(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}$ annihilate ${}_{\text{out}}\langle 0,0|$ and $|0,-\infty\rangle_{\text{in}}$ respectively. Concretely, $\mathtt{f}^-$ is a Bessel function,

$$\mathtt{f}^\pm(k\tau) = \frac{\sqrt{\pi}}{Z\sqrt{2\sinh(\mu\pi)}} J_{\mp i\mu}(-k\tau), \quad (21)$$

and is related to the previous Hankel mode function by the Bogoliubov transformation

$$f^+(k\tau) = \alpha\mathtt{f}^+(k\tau) + \beta\mathtt{f}^-(k\tau), \quad (22)$$

where $|\alpha|^2 - |\beta|^2 = 1$ [23]. The factors of α in (20) arise from writing $f^-(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}|n\rangle_{\text{in}} = \frac{1}{\alpha}\mathtt{f}^-(\tau\overleftrightarrow{\partial_\tau})\hat{\varphi}|n\rangle_{\text{in}}$ in the first step of the LSZ reduction (16), where again we focus on the connected component [see (A9) for the disconnected contributions].

(ii) the ${}_{\text{out}}\langle 0,-\infty|$ bra in the time-ordered correlator $G_{n'\to n}$ should be replaced by ${}_{\text{out}}\langle 0,0|$, which changes the boundary condition for internal lines. Concretely, the propagator for the Bunch-Davies S matrix can be written in terms of the Hankel mode functions as

$$G_2(k\tau_1, k\tau_2) = f^-(k\tau_>) f^+(k\tau_<), \quad (23)$$

while for the Unruh-DeWitt S matrix one must instead use the propagator

$$\mathtt{G}_2(k\tau_1, k\tau_2) = \frac{1}{\alpha}\mathtt{f}^-(k\tau_>) f^+(k\tau_<), \quad (24)$$

where $\tau_{>(<)}$ is the greater (lesser) of τ_1 and τ_2.

We stress that neither choice of bases is "more fundamental." Given the (invertible) Bogoliubov transformation (22), in principle one can always expand $\mathcal{S}$ in terms of $\mathtt{S}$, or vice versa. Depending on the question being asked, it may be more convenient to use one basis over the other. Since we have in mind connecting with the well-developed amplitude technology that exists on Minkowski, our guiding principle will be to choose the basis which shares as many properties as possible with the Minkowski S matrix.

Perhaps unsurprisingly, given that the Bunch-Davies state is defined by the condition that it reduces to the Minkowski vacuum on small scales, we find that it is the Bunch-Davies $\mathcal{S}$ elements that bear the most resemblance to the Minkowski S matrix.

III. PROPERTIES OF THE S MATRIX

Before giving explicit examples of these S matrix elements, let us list some model-independent properties.

A. Particle production

Already in free theory, there is an important difference between the Bunch-Davies and Unruh-DeWitt S matrix elements. The only nonzero Bunch-Davies S matrix with two particles is

$$\mathcal{S}_{1\to 1} = (2\pi)^d \delta^d(\mathbf{k} - \mathbf{k}'), \tag{25}$$

and simply reflects the normalization we have chosen for the asymptotic states, namely that $[\hat{a}_{\mathbf{k}'}, \hat{a}^\dagger_{\mathbf{k}}] = \delta^d(\mathbf{k} + \mathbf{k}')$. On the other hand, the Unruh-DeWitt S matrix has a nonzero $\mathsf{S}_{0\to 2}$ since the states $|0, -\infty\rangle$ and $\langle 2, 0|$ are not orthogonal. In fact, given the Bogoliubov transformation (22), their overlap is

$$\mathsf{S}_{0\to 2} = \frac{\beta}{\alpha}(2\pi)^d \delta^d(\mathbf{k}_1 + \mathbf{k}_2). \tag{26}$$

The Bogoliubov coefficient β therefore characterizes the rate of particle production in the free theory, as measured by an Unruh-DeWitt detector (and $\beta/\alpha = e^{-\mu\pi}$ is the characteristic Boltzmann factor which suppresses the production of heavy states). One special feature of the Bunch-Davies basis for the S matrix is that this particle production is accounted for by the choice of asymptotic states: the elements $\mathcal{S}_{n'\to n}$ are the probability that an initial n'-particle state will scatter into the jets of multiparticle "stuff" which would have been created by the free propagation of n particles through the expanding spacetime medium.

For S matrix elements with more than two particles, this free theory particle production shows up as additional contributions to the disconnected parts of the Unruh-DeWitt S matrix: see Figs. 1 and 2 in the Appendix for a concrete example. In the interacting theory, there is additional particle production due to the interactions in S_{int}. These appear explicitly in both the Bunch-Davies and the Unruh-DeWitt bases: for instance both $\mathcal{S}_{0\to n}$ and $\mathsf{S}_{0\to n}$ are generically nonzero when S_{int} contains n-point interactions.

B. Antipodal singularities

Performing the inverse Fourier transform from each momentum $\mathbf{k}$ back to a position $\mathbf{x}$, both the Bunch-Davies propagator $(\tau_1\tau_2)^{d/2} G_2(k\tau_1, k\tau_2)$ and the Unruh-Dewitt propagator $(\tau_1\tau_2)^{d/2}\mathsf{G}_2(k\tau_1, k\tau_2)$ become functions of the invariant chordal separation between $(\tau_1, \mathbf{x}_1)$ and $(\tau_2, \mathbf{x}_2)$,

$$\cosh\sigma \equiv 1 + \frac{(\tau_1 - \tau_2)^2 - |\mathbf{x}_1 - \mathbf{x}_2|^2}{2\tau_1\tau_2}. \tag{27}$$

Explicitly, these functions can be written in terms of the associated Legendre functions [24]

$$\mathsf{G}_2(\cosh\sigma) = \frac{Q^{\frac{d}{2}-\frac{1}{2}}_{i\mu-\frac{1}{2}}(\cosh\sigma)}{(2\pi)^{\frac{d+1}{2}}(\sinh\sigma)^{\frac{d}{2}-\frac{1}{2}}},$$
$$G_2(\cosh\sigma) = \frac{\pi}{2\cosh(\pi\mu)} \frac{P^{\frac{d}{2}-\frac{1}{2}}_{i\mu-\frac{1}{2}}(-\cosh\sigma)}{(2\pi)^{\frac{d+1}{2}}(\sinh\sigma)^{\frac{d}{2}-\frac{1}{2}}}, \tag{28}$$

which are often expressed in terms of either Gegenbauer functions or hypergeometric ${}_2F_1$ functions [25–29].

$\cosh\sigma > 1$ (< 1) corresponds to the two positions being timelike (spacelike) separated. Both propagators have a branch point singularity on the light cone at $\cosh\sigma = 1$, and the branch cut along $\cosh\sigma > 1$ reflects the ambiguity in ordering timelike separated operators. The time ordering relevant for our S matrix is the prescription that these functions are evaluated at $\cosh(\sigma) - i\epsilon$ and the branch cut is approached from below [30,31]. This is precisely analogous to the singularity structure of the Feynman propagator on Minkowski.

One respect in which de Sitter differs qualitatively from Minkowski is the existence of an antipodal map: sending $(\tau_1, \mathbf{x}_1)$ to its antipodal position in the de Sitter spacetime corresponds to sending $\cosh\sigma \to -\cosh\sigma$. The region $\cosh\sigma < -1$ therefore corresponds to $(\tau_2, \mathbf{x}_2)$ being timelike separated from the *antipode* of $(\tau_1, \mathbf{x}_1)$. While the Bunch-Davies propagator G_2 is perfectly regular at $\cosh\sigma = -1$, the Unruh-DeWitt propagator G_2 has an additional branch point singularity there. There is no analog of this antipodal singularity on Minkowski, and the existence of this additional branch cut in G_2 is another important difference between the Bunch-Davies and Unruh-DeWitt boundary conditions.

When Wick rotated to Euclidean AdS, the choice of out state becomes the choice of boundary condition at spatial infinity. G_2 corresponds to the walls of the AdS box being "transparent," while G_2 corresponds to the walls being "reflecting" [30,32]. In that language, it is the reflecting boundary condition that leads to an antipodal image of the coincident singularity in the propagator.

C. Crossing

At the level of the time-ordered correlator, the only difference between an "ingoing" or "outgoing" field is simply our convention for the sign of its momentum, since $\hat{\varphi}^\dagger(\tau, \mathbf{k}) = \hat{\varphi}(\tau, -\mathbf{k})$ for a real scalar field. A physical distinction only arises when we put the fields on shell: we either do this using f^+ or f^-, which is the analog of setting the particle energy $= +\sqrt{k^2+m^2}$ or $= -\sqrt{k^2+m^2}$ on Minkowksi. To relate these, we can make use of the Hankel function identity

$$\lim_{\epsilon\to 0} f^+(-z+i\epsilon) = f^-(z) \tag{29}$$

for real $z > 0$, which is closely related to the invariance of ϕ under CPT transformations. In particular, by replacing each k with a new variable $\tilde{k}$ (which is independent of $\mathbf{k}$) in the final step of the LSZ procedure, we naturally arrive at the object

$$\tilde{\mathcal{S}}_n(\{\tilde{k}\},\{\mathbf{k}\}) \equiv \left[\prod_{b=1}^{n}\int_{-\infty}^{0}\frac{d\tau_b}{-\tau_b}f^+(\tilde{k}_b\tau_b)\right]\mathcal{G}_n(\{\tau\},\{\mathbf{k}\}). \tag{30}$$

This function of $\tilde{k}$ contains the S matrix elements for all $n_1 \to n_2$ processes with $n_1 + n_2 = n$, since the transformation

$$(\tilde{k}_b, \mathbf{k}_b) \to (-\tilde{k}_b, -\mathbf{k}_b) \tag{31}$$

moves a particle from the out state to the in state [33]. For now we restrict our attention to $\tilde{k} = \pm k$ (with an appropriate $i\epsilon$), since these are the values at which $\tilde{\mathcal{S}}_n$ coincides with an $\mathcal{S}_{n_1\to n_2}$ element. We will return to off shell extensions of the S matrix in the Future Directions section below.

Crossing is another important difference between the Bunch-Davies and Unruh-DeWitt S-matrices. The crossing operation that maps a particle from the in to the out state in $\mathbb{S}$ requires the Bogoliubov transformation (22), and as a result there is no longer a simple function like (30) that interpolates between different scattering channels for the Unruh-DeWitt S matrix elements.

D. de Sitter isometries

The $(d+1)(d+2)/2$ isometries of de Sitter spacetime in these coordinates are

(1) d spatial translations, which imply conservation of the total momentum $\mathbf{k}$,

(2) $d(d-1)/2$ spatial rotations, which implies a dependence on $\mathbf{k}_a \cdot \mathbf{k}_b$ only,

(3) one dilation transformation, $(\tau, \mathbf{x}) \to (\gamma\tau, \gamma\mathbf{x})$, which is generated in the momentum domain by

$$D[\tau, \mathbf{k}] = \mathbf{k}\cdot\partial_{\mathbf{k}} - \tau\partial_\tau + d, \tag{32}$$

(4) d "boosts," characterized by a parameter $\mathbf{v}$,

$$\tau \to \gamma\tau, \qquad \mathbf{x} \to \gamma(\mathbf{x} - \mathbf{v}x^2)$$
$$\text{where } \gamma = 1/(1 - 2\mathbf{v}\cdot\mathbf{x} + v^2x^2). \tag{33}$$

This is generated in the momentum domain by

$$\mathbf{K}[\tau, \mathbf{k}] = \mathbf{K}[\mathbf{k}] - \mathbf{k}\tau^2 - 2\tau\partial_\tau\partial_{\mathbf{k}} + 2d\partial_{\mathbf{k}} \tag{34}$$

where $\mathbf{K}[\mathbf{k}] = 2\mathbf{k}\cdot\partial_{\mathbf{k}}\partial_{\mathbf{k}} - \mathbf{k}\partial_{\mathbf{k}}\cdot\partial_{\mathbf{k}}$ is the usual generator of special conformal transformations [34].

Since a scalar field ϕ is invariant under dilations and boosts, the Ward identities for correlators of the rescaled $\varphi = (-\tau)^{-d/2}\phi$ are

$$\sum_{b=1}^{n}\left(D[\tau_b, \mathbf{k}_b] - \frac{d}{2}\right)G_n = 0,$$
$$\sum_{b=1}^{n}(\mathbf{K}[\tau_b, \mathbf{k}_b] - d\partial_{\mathbf{k}_b})G_n = 0. \tag{35}$$

Now applying the LSZ formula, and using the fact that $\mathcal{E}$ represents a quadratic Casimir of the de Sitter algebra and hence commutes with all other generators [35], we find that the S matrix for de Sitter invariant interactions is constrained by the Ward identities

$$\sum_{b=1}^{n}\left(\mathbf{k}_b\cdot\partial_{\mathbf{k}_b} + \frac{d}{2}\right)\tilde{\mathcal{S}}_n = 0,$$
$$\sum_{b=1}^{n}\left(\mathbf{K}[\mathbf{k}_b] + d\partial_{\mathbf{k}_b} + \mu^2\frac{\mathbf{k}_b}{k_b^2}\right)\tilde{\mathcal{S}}_n = 0. \tag{36}$$

For instance, consider the contact contribution (18). Applying the above dilation, the integrand shifts by a total derivative which does not contribute to the S matrix, and so the corresponding Ward identity is satisfied. Applying the above boost, since the mode functions transform as

$$\left(\mathbf{K}[\mathbf{k}] + d\partial_{\mathbf{k}} + \mu^2\frac{\mathbf{k}}{k^2}\right)f^\pm(k\tau) = -\mathbf{k}\tau^2 f^\pm(k\tau) \tag{37}$$

the corresponding Ward identity is automatically satisfied thanks to momentum conservation.

E. Total derivatives and field redefinitions

One main advantage of the S matrix formalism is that, unlike the Lagrangian, there is no ambiguity due to field redefinitions and total derivatives. For instance, consider the following total derivative:

$$\mathcal{L}_{\rm td} = \sqrt{-g}g^{\alpha\beta}\nabla_\alpha(\phi^2\nabla_\beta\phi). \tag{38}$$

It contributes at tree level to the S matrix only via the boundary terms of the form

$$\int_{-\infty}^{0} d\tau\, \partial_\tau (f^+(k_1\tau) f^+(k_2\tau)\tau\partial_\tau[\tau^{d/2} f^+(k_3\tau)]), \quad (39)$$

and both limits separately vanish for principal series fields. It is easy to see that any total covariant derivative of ϕ's will similarly give a vanishing contribution to any S matrix element, simply because $\tau^{d/2} f^{\pm}(k\tau)$ vanishes at both integration boundaries. So while total derivatives can contribute to the correlator, their contribution vanishes once we go on shell.

To show invariance under field redefinitions, it is useful to consider linear and nonlinear redefinitions separately. Linear redefinitions are of the form $\phi' = \gamma\phi$. Since this produces a new $S_{\rm free}$ with $Z^2 \to Z^2/\gamma^2$, the normalization of the mode functions changes in such a way that $f^{\pm} \to \gamma f^{\pm}$. So while the correlator of the new fields is $G'_n = \gamma^n G_n$, it produces the same on shell S matrix defined in (17). This is also explicit in the example (18) given above: since this rescaling produces a new $S_{\rm int}$ with $\lambda_n \to \lambda_n/\gamma^n$, we see that the product of $\lambda_n \times (f^{\pm})^n$ is insensitive to linear field redefinitions. As an example of a nonlinear redefinition, consider $\phi \to \phi + \gamma\phi^2$ applied to the simple Lagrangian $\mathcal{L}_{\rm int} = \sqrt{-g}\frac{\lambda_3}{3!}\phi^3$. This produces a new action,

$$\mathcal{L} \to \mathcal{L} + \gamma Z^2 \mathcal{L}_{\rm td} + \frac{\gamma Z^2}{\tau}(-\tau)^{\frac{d}{2}} \varphi^2 \mathcal{E}\varphi - \frac{\gamma\lambda_3}{2\tau}(-\tau)^d \varphi^4, \quad (40)$$

at leading order in γ. The total derivative does not contribute to the S matrix, but the new cubic and quartic interactions give equal and opposite contributions to any S matrix element. For instance,

$$\tilde{\mathcal{S}}_4^{\rm exch} \to \tilde{\mathcal{S}}_4^{\rm exch} + 4\gamma\lambda_3 \int_{-\infty}^{0} \frac{d\tau}{\tau}(-\tau)^{\frac{d}{2}} \int_{-\infty}^{0} \frac{d\tau'}{\tau'}(-\tau')^{\frac{d}{2}} f^+(\tilde{k}_1\tau) f^+(\tilde{k}_2\tau) Z^2 \mathcal{E}[k_s\tau] G_2(k_s\tau, k_s\tau') f^+(\tilde{k}_3\tau') f^+(\tilde{k}_4\tau') + 2\,\text{perm.}$$
$$\tilde{\mathcal{S}}_4^{\rm cont} \to \tilde{\mathcal{S}}_4^{\rm cont} + 12i\gamma\lambda_3 \int_{-\infty}^{0} \frac{d\tau}{-\tau}(-\tau)^d f^+(\tilde{k}_1\tau) f^+(\tilde{k}_2\tau) f^+(\tilde{k}_3\tau) f^+(\tilde{k}_4\tau) \quad (41)$$

exactly cancel since $Z^2\mathcal{E}[k_s\tau] G_2(k_s\tau, k_s\tau') = i\tau\delta(\tau - \tau')$ and hence collapses one of the time integrals in the exchange diagram. The total S matrix $\tilde{\mathcal{S}}_4^{\rm cont} + \tilde{\mathcal{S}}_4^{\rm exch}$ is therefore unchanged [36]. This shows that while the split into "contact" and "exchange" contributions is ambiguous, the sum is invariant under field redefinitions.

F. Unique structures

This insensitivity to total derivatives and field redefinitions has the important consequence that the three-point S matrix is unique (up to crossing). The argument in perturbation theory is straightforward: since any cubic interaction can be integrated by parts into the form $\phi^2\Box^n\phi$, an arbitrary derivative interaction will contribute to $\tilde{S}_3$ in the same way as $m^{2n}\phi^3$ for some n. The contact integral (18) (and its three crossing images) from ϕ^3 are therefore the unique kinematic structures which can appear for three particles in perturbation theory. To go beyond perturbation theory, notice that these integrals correspond to the four possible solutions to the de Sitter Ward identities [16,37], and therefore any de Sitter invariant set of interactions must produce an $\tilde{S}_3$ of this form. This is the analog of the well-known result on Minkowski that the on shell three-point function is fixed uniquely by the spacetime isometries (to be some, possibly mass dependent, constant).

An interesting corollary of this is that there is a unique four-point exchange structure which describes the interaction of two ϕ's via the exchange of a field σ (again up to crossing). Since the most general cubic vertex is equivalent to a sum of interactions of the form $\phi^2(\Box - m_\sigma^2)^n\sigma$ after integration by parts, for any $n \geq 1$ this can be exchanged via a field redefinition of the form $\sigma \to \sigma + (\Box - m_\sigma^2)^{n-1}\phi^2$, for a quartic interaction $\phi^2(\Box - m_\sigma^2)^n\phi^2$, which corresponds to the contact invariant in (18). This is the analog of $1/(m_\sigma^2 - s)$ on Minkowski: any s-channel exchange diagram for scalars can always be separated into this unique structure plus contact-type contributions.

G. Flat space limit

To take the flat space limit, we will temporarily restore factors of the Hubble rate H. We will also write the conformal time and mass parameter in terms of a new variable t and m using

$$H\tau = -e^{-Ht} \quad \text{and} \quad \mu = \sqrt{\frac{m^2}{H^2} - \frac{d^2}{4}}. \quad (42)$$

In the limit $H \to 0$ at fixed t, k, and m (where we do not assume any further hierarchy, so, e.g., we treat k and m as comparable), the mode functions become [38]

$$f^{\pm}(k\tau) = e^{\mp i\alpha_0/H} \frac{e^{\pm i\Omega_k t}}{\sqrt{2i\Omega_k}}[1 + \mathcal{O}(H)], \quad (43)$$

where we have introduced $\Omega_k = \sqrt{k^2 + m^2}$. Up to an overall phase α_0 (which does not affect physical observables), the leading order term in (43) coincides with the usual Minkowski mode function [39]. Also note that (43) follows from a saddle point approximation of the Hankel function which requires $k > 0$. Assuming instead that

$k < 0$ produces (43) with $\Omega_k = -\sqrt{k^2+m^2}$. The crossing transformation $k \to -k$ therefore implements the usual crossing relation in the Minkowski limit.

Applied to the reduction formula (17), we find that our de Sitter S matrix coincides in the $H \to 0$ limit with the usual Minkowski S matrix, up to an overall phase and with the state normalization (25). Since the Bogoliubov coefficient $\beta \to 0$ in this limit (i.e., the effects of particle production switch off as $H \to 0$), both the Bunch-Davies and the Unruh-DeWitt S matrix elements have the same Minkowski limit. Physically, this reflects the fact that there is no distinction between the vacua $|0,-\infty\rangle$ and $|0,0\rangle$ in the flat space limit (in which the Hamiltonian becomes time independent and there is a unique vacuum state). The same is true of the time-ordered correlation functions: for instance, since the chordal separation (27) becomes $\cosh\sigma = 1 - H^2 s^2 + \mathcal{O}(H^4)$ in the flat space limit, where $s^2 \equiv -(t_1-t_2)^2 + |\mathbf{x}_1 - \mathbf{x}_2|^2$ is the Minkowski geodesic distance, both propagators have the same limiting behavior at fixed μ, namely [40]

$$\lim_{H\to 0} G_2(\cosh\sigma) = \lim_{H\to 0} \mathrm{G}_2(\cosh\sigma) = \frac{\Gamma\left(\frac{d-1}{2}\right)}{4\pi^{\frac{d+1}{2}}}\left(\frac{1}{s^2+i\epsilon}\right)^{\frac{d-1}{2}}, \tag{44}$$

which coincides with the massless Minkowski propagator. The massive propagator is obtained by taking $H \to 0$ at fixed m.

IV. SOME EXAMPLES

To illustrate some of these features, we now list some simple S matrix elements in particular models.

Consider a scalar field σ with mass $m^2 = (d^2-1)/4$ [i.e., conformal weight $\Delta = (d-1)/2$]. This complementary series field has arguably the simplest mode function, since the Hankel function at $i\mu = 1/2$ reduces to a plane wave,

$$Z_\sigma f^\pm(k\tau) = \frac{e^{\pm ik\tau}}{\sqrt{\mp 2ik\tau}} \quad \text{when } i\mu = \frac{1}{2}. \tag{45}$$

From the interaction Lagrangian $\mathcal{L}_{\text{int}} = \sqrt{-g}\frac{\lambda_n}{n!}\sigma^n$, the n-point Bunch-Davies S matrix elements are given by

$$\mathcal{S}_{0\to n} = \frac{i\lambda_n}{Z_\sigma^n}\frac{\Gamma(j_n)}{(ik_T)^{j_n}}\frac{(2\pi)^d\delta^d(\sum_{b=1}^4 \mathbf{k}_b)}{\prod_{c=1}^n\sqrt{2ik_c}} \tag{46}$$

and its various crossing images, where $k_T = \sum_{b=1}^n k_b$ is the "total energy" flowing into this vertex, and the power $j_n \equiv n(\frac{d-1}{2}) - d$ uniquely satisfies the dilation Ward identity. Note that this result is formally infinite whenever the total conformal weight $n(\frac{d-1}{2})$ coincides with $d-N$ for any integer N, since then $j_n = -N$ and the $\Gamma(j_n)$ factor diverges. This is a general feature: while interactions of principal series fields always have a total conformal weight with $\mathrm{Re}(\Delta_T) > d$ and are free of such divergences, for complementary series fields our adiabatic hypothesis can break down whenever $\mathrm{Re}(\Delta_T) = d - N$. However, that is not to say that every such interaction of light fields leads to problematic divergences in every S matrix element. For instance, even though σ^3 gives a divergent contribution to $\mathcal{S}_{0\to 3}$, it gives a finite exchange contribution to $\mathcal{S}_{0\to 4}$ since a nonzero k_s (or k_t or k_u) effectively regulates the divergence. Explicitly, we find that for s-channel scattering in $d = 3$ it is given by

$$\mathcal{S}_{0\to 4}^{\text{exch}} = \frac{\lambda_3^2}{Z_\sigma^6}\frac{\mathrm{Li}_2\left(\frac{k_1+k_2-k_s}{k_T}\right) + \mathrm{Li}_2\left(\frac{k_3+k_4-k_s}{k_T}\right) - \frac{\pi^2}{6}}{2ik_s\sqrt{2ik_1}\sqrt{2ik_2}\sqrt{2ik_3}\sqrt{2ik_4}}, \tag{47}$$

where we now omit the total-momentum δ function.

Now consider a massive field ϕ coupled to $n-1$ of these σ fields, namely $\mathcal{L}_{\text{int}} = \sqrt{-g}\frac{\lambda_n'}{(n-1)!}\sigma^{n-1}\phi$. The n-point Bunch-Davies S matrix is

$$\mathcal{S}_{0\to n} = \frac{i\lambda_n'}{Z_\sigma^{n-1}Z_\phi}\frac{\left|\Gamma\left(\frac{2j_n+1}{2} - i\mu\right)\right|^2}{(k_\sigma^2 - k_\phi^2)^{j_n}\sqrt{2ik_\phi}}\frac{P^{-j_n}_{i\mu-\frac{1}{2}}\left(\frac{k_\sigma}{k_\phi}\right)}{\prod_{b=1}^n\sqrt{2ik_b}}, \tag{48}$$

where $k_\sigma \equiv k_1 + \cdots + k_{n-1}$ is the total energy of the σ fields, $k_\phi \equiv k_n$ is the energy of the heavy ϕ field, and $P^{-j_n}_{i\mu-\frac{1}{2}}$ is the associated Legendre polynomial [41].

In all of the above examples, when $d = 3$ any of the σ fields may be replaced by the time derivative of a massless "pion" field $\dot{\pi}$ by simply multiplying the corresponding $\mathcal{S}_{0\to n}$ by a factor of ik for that external leg.

Note that if we had instead normalized our asymptotic states by a factor of $\sqrt{2ik}$ per particle, the analytic structure of each of these S matrix elements would be very simple. In particular, the only singularities of these $\tilde{\mathcal{S}}_n$ in the complex $\tilde{k}$ plane would be at $\tilde{k}_T = 0$ when the total energy flowing into the diagram vanishes and, in the case of the exchange diagram (47), also when the energy flowing into either vertex vanishes (namely when $\tilde{k}_L = \tilde{k}_1 + \tilde{k}_2 + k_s$ or $\tilde{k}_R = \tilde{k}_3 + \tilde{k}_4 + k_s$ vanishes). Depending on whether each particle is in the initial or final state, these branch points will correspond to the vanishing of a particular linear combination of the k_b (e.g., for $\mathcal{S}_{\mathbf{k}'\to\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}$, this would be $k_1 + k_2 + k_3 - k' = 0$).

In contrast, the Unruh-DeWitt S matrix elements can have several singularities from each vertex. For instance, from the σ^n interaction considered above,

$$\mathsf{S}_{1\to n-1} = \frac{i\lambda_n}{Z_\sigma^n \alpha^n} \prod_{b=1}^{n-1} \frac{1}{\sqrt{2ik_b}} \times \left[\frac{\alpha\Gamma(j_n)}{(ik_T - ik')^{j_n}\sqrt{-2ik'}} - \frac{\beta\Gamma(j_n)}{(ik_T + ik')^{j_n}\sqrt{2ik'}} \right], \tag{49}$$

where k' is the ingoing momentum and k_T the total outgoing energy. From the $\sigma^{n-1}\phi$ interaction,

$$\mathsf{S}_{1\to n-1} \propto \frac{Q^{-j_n}_{i\mu-\frac{1}{2}}\left(\frac{k_\sigma}{k_\phi}\right)}{(k_\sigma^2 - k_\phi^2)^{j_n}\sqrt{2ik_\phi}} \prod_{b=1}^{n} \frac{1}{\sqrt{2ik_b}}. \tag{50}$$

The former manifestly has singularities at both $k_T = \pm k'$, while the latter has singularities at both $k_\sigma = \pm k_\phi$ (cf. the discussion of G_2 above). This illustrates that choosing different bases for the de Sitter S matrix can result in very different singularities.

As a final example, consider the unique three-particle S matrix. For three general masses, this unique structure can be written explicitly in terms of the Appell F_4 function. For example

$$\mathsf{S}_{\mathbf{k}_1\mathbf{k}_2\to\mathbf{k}_3} \propto \frac{1}{(ik_3)^{d/2}} \left(\frac{k_2}{k_3}\right)^{i\mu_1} \left(\frac{k_1}{k_3}\right)^{i\mu_2} \times F_4\left(a_+, a_-; 1+i\mu_1, 1+i\mu_2; \frac{k_1^2}{k_3^2}, \frac{k_2^2}{k_3^2}\right), \tag{51}$$

where the indices are $a_\pm = \frac{1}{2}(\frac{d}{2} + i\mu_1 + i\mu_2 \pm i\mu_3)$. For particular mass values (e.g., if the masses coincide or take the special value $i\mu = 1/2$) this general Appell function often simplifies into ${}_2F_1$ functions. This example illustrates that the special functions routinely encountered when computing observables in the expanding Poincaré patch are an unavoidable consequence of these coordinates (in particular labeling the fields by $\mathbf{k}$) and are not rendered simpler by considering a "better" observable such as the S matrix.

V. COSMOLOGICAL OBSERVABLES FROM THE S MATRIX

Finally, we turn to the question of how to extract inflationary observables from the de Sitter S matrix. From an observational standpoint, we are metaobservers who live outside of the de Sitter spacetime in which the scattering is taking place: the usual objections to measuring an asymptotic S matrix at future infinity (at which all spatial points are causally disconnected) therefore do not apply, since once the inflationary perturbations reenter the horizon and come back into causal contact we *do* have observational access to the asymptotic out-state via the CMB and LSS, as we shall now show.

A. Wavefunction of the Universe

The de Sitter S matrix elements are closely related to non-Gaussianities of the late-time Bunch-Davies wavefunction, which can in fact be constructed from the S matrix once a particular field basis has been chosen. Specifically, the wavefunction of a state at time τ is defined by projecting the state onto a basis of field eigenstates, $|\phi(\tau)\rangle$, which are defined by $\hat{\phi}(\tau,\mathbf{k})|\phi(\tau)\rangle = \phi(\tau,\mathbf{k})|\phi(\tau)\rangle$. The wavefunction of the Bunch-Davies vacuum $|0,-\infty\rangle$ is of particular importance in early Universe cosmology since it describes the statistics of primordial perturbations, which seed inhomogeneities in the cosmic microwave background, as well as density perturbations of the large-scale structure of the Universe. This wavefunction can be characterized by a set of "wavefunction coefficients," which are most conveniently written as the connected part of the following matrix element [42]:

$$\psi_n(\tau) \equiv \langle \phi(\tau) = 0| \left[\prod_{b=1}^{n} \frac{i\hat{\Pi}(\tau,\mathbf{k}_b)}{Z^2} \right] |0,-\infty\rangle_{\text{in}}, \tag{52}$$

where $\hat{\Pi}$ is the momentum conjugate to $\hat{\phi}$ [43].

To relate these coefficients to the S matrix elements, we expand the bra in terms of our asymptotic out states,

$$\lim_{\tau\to 0} \langle \phi = 0| \frac{i\hat{\Pi}_1}{Z^2} \dots \frac{i\hat{\Pi}_n}{Z^2} = \sum_j {}_{\text{out}}\langle j, -\infty| B_n^j, \tag{53}$$

since then we can write the wavefunction coefficient as a sum over $0 \to j$ matrix elements,

$$\lim_{\tau\to 0} \psi_n(\tau) = \sum_j B_n^j \mathcal{S}_{0\to j}. \tag{54}$$

The B_n^j coefficients can be evaluated in the free theory, since by the adiabatic hypothesis

$$|j,-\infty\rangle_{\text{out}} \to \hat{a}_1^\dagger \dots \hat{a}_j^\dagger |0,-\infty\rangle$$
$$(-\tau)^d \hat{\Pi}(\tau,\mathbf{k}) \to Z^2 \tau\partial_\tau \phi(\tau,\mathbf{k}) \tag{55}$$

as $\tau \to 0$ and the interactions turn off. This gives

$$\lim_{\tau\to 0} \psi_{\mathbf{k}_1\dots\mathbf{k}_n}(\tau) = \sum_{j=0}^{\infty} \left[\prod_{\ell=1}^{j} \int_{\mathbf{q}_\ell \mathbf{q}'_\ell} P_{\mathbf{q}_\ell \mathbf{q}'_\ell}(\tau) \right] \frac{(-1)^j \mathcal{S}_{0\to \mathbf{k}_1\dots\mathbf{k}_n \mathbf{q}_1 \mathbf{q}'_1 \dots \mathbf{q}_j \mathbf{q}'_j}}{[\prod_{b=1}^n (-\tau)^{d/2} f^+(k_b\tau)][\prod_{c=1}^j (-\tau)^d f^+(q_c\tau) f^+(q'_c\tau)]}, \tag{56}$$

where $P_{\mathbf{q}\mathbf{q}'}(\tau) = \langle 0, -\infty|\hat{\phi}(\tau, \mathbf{q})\hat{\phi}(\tau, \mathbf{q}')|0, -\infty\rangle$ is the free-theory power spectrum and $\int_{\mathbf{q}\mathbf{q}'} \equiv \int \frac{d^d\mathbf{q}}{(2\pi)^d}\frac{d^d\mathbf{q}'}{(2\pi)^d}$. Equation (56) allows one to explicitly construct any wavefunction coefficient from the S matrix elements and the mode functions of the fields. In practice, this infinite sum will always truncate at a given order in perturbation theory, so often only the first few S matrix elements $(\mathcal{S}_{0\to n}, \mathcal{S}_{0\to n+2}, \ldots)$ are needed.

This is best illustrated with an example. Consider the quartic coefficient ψ_4 generated by the interactions $\mathcal{L}_{\text{int}} = \sqrt{-g}(\frac{\lambda_3}{3!}\phi^3 + \frac{\lambda_4}{4!}\phi^4)$ at tree level. From the ϕ^4 interaction, the contact Feynman-Witten diagram gives

$$\psi^{\text{cont}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}(\tau') = i\lambda_4\int_{-\infty}^{\tau'}\frac{d\tau}{-\tau}(-\tau)^d\prod_{b=1}^{4}K_{k_b}(\tau;\tau'), \tag{57}$$

where the "bulk-to-boundary" propagator is given in terms of the Hankel mode functions by

$$K_k(\tau;\tau') = \frac{(-\tau)^{d/2}f^+(k\tau)}{(-\tau')^{d/2}f^+(k\tau')}. \tag{58}$$

Comparing with (18), we see that this contribution to the wavefunction can indeed be written as

$$\psi^{\text{cont}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}(\tau) = \frac{\mathcal{S}^{\text{cont}}_{0\to\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}}{\prod_{b=1}^4(-\tau)^{d/2}f^+(k_b\tau)} \tag{59}$$

at late times [which corresponds to the general formula (56) with $n = 4$ and $j = 0$]. The exchange contribution $\psi_4^{\text{exch}}(\tau)$ is almost identical to (19), but with $G_2(k_s\tau, k_s\tau')$ replaced by the "bulk-to-bulk propagator,"

$$\begin{aligned}-iG_k^{\text{bulk}}(\tau_1, \tau_2; \tau) = G_2(k\tau_1, k\tau_2) \\ -\frac{f^-(k\tau)}{f^+(k\tau)}(-\tau_1)^{d/2} \\ \times f^+(k\tau_1)(-\tau_2)^{d/2}f^+(k\tau_2).\end{aligned}$$

The $\phi^3 \times \phi^3$ contribution to the wavefunction can therefore be written as

$$\begin{aligned}\psi^{\text{exch}}_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}(\tau) = \frac{1}{\prod_{b=1}^4(-\tau)^{d/2}f^+(k_b\tau)}\Bigg[\mathcal{S}^{\text{exch}}_{0\to\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} \\ -\int_{\mathbf{q}\mathbf{q}'}\frac{P_{\mathbf{q}\mathbf{q}'}(\tau)\mathcal{S}_{0\to\mathbf{k}_1\mathbf{k}_2\mathbf{q}}\mathcal{S}_{0\to\mathbf{k}_3\mathbf{k}_4\mathbf{q}'}}{(-\tau)^d f^+(q\tau)f^+(q'\tau)} - 2\,\text{perm}\Bigg]\end{aligned} \tag{60}$$

at late times. So the full $\psi_4^{\text{cont}} + \psi_4^{\text{exch}}$ is indeed given by (56), since the only nonzero $\mathcal{S}_{0\to n}$ at this order in perturbation theory are $\mathcal{S}_{0\to 4} = \mathcal{S}^{\text{cont}}_{0\to 4} + \mathcal{S}^{\text{exch}}_{0\to 4}$ and the disconnected part of $\mathcal{S}_{0\to 6}$ $(= \mathcal{S}_{0\to 3}\mathcal{S}_{0\to 3})$, which becomes connected once integrated over $\mathbf{q}$ and $\mathbf{q}'$.

Note that it is also straightforward to invert (56) and express a given S matrix element in terms of the wavefunction. For example

$$\begin{aligned}\mathcal{S}_{0\to 4} = \lim_{\tau\to 0}\left[\prod_{b=1}^{4}(-\tau)^{\frac{d}{2}}f^+(k_b\tau)\right]\Bigg(\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4}(\tau) \\ + \int_{\mathbf{q}\mathbf{q}'}P_{\mathbf{q}\mathbf{q}'}(\tau)\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}}(\tau)\psi_{\mathbf{k}_3\mathbf{k}_4\mathbf{q}'}(\tau) + 2\,\text{perm}\Bigg).\end{aligned} \tag{61}$$

Previous results for the wavefunction coefficients can then be readily translated into S matrix elements.

B. In-in correlators

We can similarly extract from the S matrix any desired equal-time correlator at late times. These objects are the closest to what one would observe in primordial non-Gaussianity, and are defined by [44]

$$\langle\phi_1\ldots\phi_n\rangle \equiv \lim_{\tau\to 0}{}_{\text{in}}\langle 0|\hat{\phi}(\tau, \mathbf{k}_1)\ldots\hat{\phi}(\tau, \mathbf{k}_n)|0\rangle_{\text{in}}. \tag{62}$$

If we therefore decompose the product

$$\lim_{\tau\to 0}\hat{\phi}_1\ldots\hat{\phi}_n = \sum_{j,j'}|j, -\infty\rangle_{\text{out}}C_n^{jj'}{}_{\text{out}}\langle j', -\infty| \tag{63}$$

and use the definition of the S matrix (2), we have that

$$\langle\phi_1\ldots\phi_n\rangle = \sum_{j,j'}\mathcal{S}^*_{0\to j}C_n^{jj'}\mathcal{S}_{0\to j'}. \tag{64}$$

Since the field operators and out states in (63) coincide with those of the free theory at $\tau\to 0$, we can immediately evaluate the $C_n^{jj'}$ in terms of the late-time limit of $(-\tau)^{d/2}f^\pm(k\tau)$, which we denote by $f_k^\pm$. This produces a very explicit relation between the observable correlator and the de Sitter S matrix:

$$\begin{aligned}\langle\phi_1\ldots\phi_n\rangle = 2\text{Re}\Bigg[\sum_{j'=0}^{n/2}\left[\prod_{b=1}^{n-j'}f^-_{k_b}\right]\left[\prod_{b=n-j'+1}^{n}f^+_{k_b}\right]\sum_{j=0}^{\infty}\int_{\substack{\mathbf{q}_1\ldots\mathbf{q}_j\\ \mathbf{q}'_1\ldots\mathbf{q}'_j}} \\ \mathcal{S}_{0\to\mathbf{k}_1\ldots\mathbf{k}_{n-j'}\mathbf{q}_1\ldots\mathbf{q}_j}\mathcal{S}^*_{0\to\mathbf{k}_{n-j'+1}\ldots\mathbf{k}_n\mathbf{q}'_1\ldots\mathbf{q}'_j} + \text{perm}\Bigg],\end{aligned} \tag{65}$$

where $\int_{\substack{\mathbf{q}_1\ldots\mathbf{q}_j\\ \mathbf{q}'_1\ldots\mathbf{q}'_j}}$ is an integral over the pairs $(\mathbf{q}_\ell, \mathbf{q}'_\ell)$ subject to the condition $\mathbf{q}_\ell + \mathbf{q}'_\ell = 0$. The "+perm" denotes a symmetrization over all possible permutations of the external $\mathbf{k}$'s. We can immediately notice some differences with the wavefunction relation (56): the correlators

(1) depend quadratically on the S matrix elements, while the wavefunction coefficients are linear,

(2) are manifestly real, since ϕ is real Hermitian and we have assumed spatial parity (so $\langle\phi_{\mathbf{k}_1}\ldots\phi_{\mathbf{k}_n}\rangle = \langle\phi_{-\mathbf{k}_1}\ldots\phi_{-\mathbf{k}_n}\rangle$), and
(3) are insensitive to phase information in the S matrix (which would determine the late-time mixed correlators of $\hat{\phi}$ and its conjugate momentum $\hat{\Pi}$).

As a concrete example, consider the four-point correlator induced by the same Lagrangian $\mathcal{L}_{\rm int} = \sqrt{-g}(\frac{\lambda_3}{3!}\phi^3 + \frac{\lambda_4}{4!}\phi^4)$. This correlator can be determined from the wavefunction coefficients studied above, and can be separated into three different contributions,

$$\langle\phi_1\ldots\phi_4\rangle^{\rm cont} = 2{\rm Re}[\psi_4^{\rm cont}]\prod_{b=1}^4 |f_{k_b}^+|^2,$$
$$\langle\phi_1\ldots\phi_4\rangle^{\rm exch} = 2{\rm Re}[\psi_4^{\rm exch}]\prod_{b=1}^4 |f_{k_b}^+|^2,$$
$$\langle\phi_1\ldots\phi_4\rangle^{\rm quad} = \prod_{b=1}^4 |f_{k_b}^+|^2\left(\int_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}2{\rm Re}[\psi_{\mathbf{k}_1\mathbf{k}_2\mathbf{q}}]2{\rm Re}[\psi_{\mathbf{k}_3\mathbf{k}_4\mathbf{q}'}] + 2\,{\rm perm}\right). \tag{66}$$

Let us now confirm that this is correctly reproduced by (65). At this order in perturbation theory, there are only three S matrix elements that contribute, namely, $\mathcal{S}_{0\to3}$, $\mathcal{S}_{0\to4}$ and $\mathcal{S}_{0\to6} \supset \mathcal{S}_{0\to3}\mathcal{S}_{0\to3}$. Comparing their integral representations with (65), we see that

$$\langle\phi_1\ldots\phi_4\rangle^{\rm cont} = 2{\rm Re}\left[\prod_{b=1}^4 f_{k_b}^- \mathcal{S}^{\rm cont}_{0\to\mathbf{k}_1\ldots\mathbf{k}_4}\right]$$
$$\langle\phi_1\ldots\phi_4\rangle^{\rm exch} = 2{\rm Re}\left[\prod_{b=1}^4 f_{k_b}^- \left(\mathcal{S}^{\rm exch}_{0\to\mathbf{k}_1\ldots\mathbf{k}_4} - \int_{\mathbf{q}\mathbf{q}'}\frac{P_{\mathbf{q}\mathbf{q}'}}{f_q^+ f_{q'}^+}\mathcal{S}_{0\to\mathbf{k}_1\ldots\mathbf{k}_4\mathbf{q}\mathbf{q}'}\right)\right]$$
$$\langle\phi_1\ldots\phi_4\rangle^{\rm quad} = 2{\rm Re}\left[\int_{\mathbf{q}\mathbf{q}'}\frac{P_{\mathbf{q}\mathbf{q}'}}{f_q^+ f_{q'}^+}\mathcal{S}_{0\to\mathbf{k}_1\ldots\mathbf{k}_4\mathbf{q}\mathbf{q}'}\prod_{b=1}^4 f_{k_b}^- + f_{k_1}^- f_{k_2}^- f_{k_3}^+ f_{k_4}^+ \int_{\mathbf{q}\mathbf{q}'}\mathcal{S}_{0\to\mathbf{k}_1\mathbf{k}_2\mathbf{q}}\mathcal{S}^*_{0\to\mathbf{k}_3\mathbf{k}_4\mathbf{q}'} + {\rm perm},\right] \tag{67}$$

and so the sum of all three contributions to the correlator is indeed given by the general relation (65).

VI. FUTURE DIRECTIONS

In summary, we have defined a perturbative S matrix for scalar fields in the expanding patch of a fixed de Sitter spacetime background, and demonstrated that it enjoys many of the useful properties of the Minkowski S matrix. We believe that fundamental properties like unitarity, causality, and locality will be simpler to express in terms of these S matrix elements (as opposed to, say, the wavefunction or in-in correlators). This expectation stems from the fact that this S matrix describes the time evolution in a field-independent way, and is the natural extrapolation of the Minkowski S matrix to nonzero values of the Hubble rate. They are also in many cases easier to compute and analyze than their wavefunction counterparts. In an upcoming companion paper [45], we will describe in more detail how to efficiently compute these S matrix elements.

A. Unitarity

Interestingly, the particular combination of wavefunction coefficients that corresponds to an S matrix element was previously constructed in [46] via an independent argument. There, this combination [which we shall denote at finite times by $\tilde{\psi}_n(\tau)$], was engineered as the unique combination of the wavefunction coefficients which remains invariant under the free evolution for any initial condition. Constraints from unitarity were therefore formulated most simply in terms of $\tilde{\psi}_n(\tau)$ because its time dependence stems only from the interactions, whereas the time dependence of the original $\psi_n(\tau)$ is a convolution of both the interactions and the initial condition. Here we have uncovered a deeper reason for the simplicity of $\tilde{\psi}_n(\tau)$: it is a finite-time counterpart to the S matrix (and coincides with the S matrix as $\tau \to 0$). Furthermore, recent cutting rules for wavefunction coefficients have found a proliferation of terms not present in the usual Cutkosky rules on Minkowski due to the presence of the boundary term in $G_k^{\rm bulk}$ [46–52]. Since the de Sitter S matrix uses the Feynman propagator for internal lines, it obeys simpler cutting rules than the wavefunction of the Universe. They are essentially identical to the usual Cutkosky rules. Finally, the analytic continuation to negative values of $|\mathbf{k}|$ which has played a central role in cosmological cutting rules can now be understood as a crossing transformation which relates the $0 \to n$ matrix element to a conjugate channel such as $n \to 0$, which therefore recovers the usual form of the optical theorem.

B. Analyticity

Although we have focused on scattering particles with an on shell energy $\tilde{k} = \pm|\mathbf{k}|$, our definition of $\tilde{\mathcal{S}}_n$ in (30) can be evaluated at any value of $\tilde{k}$. In particular, since $f^+(\tilde{k}\tau) \sim e^{+i\tilde{k}\tau}$ and the integration domain restricts $\tau < 0$, the off shell $\tilde{\mathcal{S}}_n$ must be *analytic* in the lower half of the complex $\tilde{k}$ plane (for any time-ordered correlator that is exponentially bounded at large τ). This is the precise analog of the Kramers-Kronig analyticity that underpins nonrelativistic dispersion relations. $\tilde{S}_n$ is therefore a natural extension of the off shell wavefunction of [53] to de Sitter spacetime.

C. Locality

One crucial consequence of locality (together with unitarity and analyticity) on Minkowski is the "Froissart bound," which limits the growth of scattering amplitudes at large center-of-mass energies. Since our S matrix reduces to the Minkowski S matrix in the high-energy limit $k \to \infty$, we expect that a similar bound will apply to the growth of the de Sitter S matrix. A rigorous proof of this is left for the future.

D. Late-time divergences

The divergences which appear at the conformal boundary as $\tau \to 0$ and complicate the S matrix for light fields are similar to the ones encountered near the conformal boundary of AdS. For the latter, there is a well-understood procedure of holographic renormalization. Perhaps in most cases the S matrix for light fields can be safely defined (or at least reliably computed in perturbation theory) by applying an analogous renormalization procedure on de Sitter. Progress in that direction would extend the S matrix construction described here to fields of any stable mass (in principal or complementary series). Phenomenologically, since these late-time divergences do not arise when computing the two-, three-, or four-point correlators in a large class of inflationary models (including all single-field models via the effective field theory of inflation), the S matrix developed above can already be compared with astrophysical observations. However, given that IR divergences are still raising new and interesting questions for the Minkowski S matrix, we anticipate that future studies of this aspect of the de Sitter S matrix will reveal surprisingly rich and subtle structures.

E. Other off shell extensions

Finally, we note that the extension $k \to \tilde{k}$ is not the only way to define an "off shell" S matrix. In particular, another option is to replace the mass parameter $\mu \to \tilde{\mu}$ (which is now independent of the mass m^2), and interpret the LSZ reduction formula as a Kontorovich-Lebedev integral transform from τ to $\tilde{\mu}$. This alternative procedure for going off shell has a closer connection to the Källén-Lehmann spectral representation of flat space, and we aim to discuss it further in [45].

ACKNOWLEDGMENTS

It is a pleasure to thank Dionysios Anninos, Santiago Agüí Salcedo, Tarek Anous, Nima Arkani-Hamed, Daniel Baumann, Paolo Benincasa, Carlos Duaso Pueyo, Austin Joyce, Hayden Lee, Juan Maldacena, Enrico Pajer, Sasha Polyakov, and Dong-Gang Wang for many enlightening discussions on de Sitter space over the years. S. M. is supported by a UKRI Stephen Hawking Fellowship (No. EP/T017481/1). G. L. P. is supported by Scuola Normale, by a Rita-Levi Montalcini fellowship from the Italian Ministry of Universities and Research (MUR), and by INFN (IS GSS-Pi).

APPENDIX

1. Adiabatic hypothesis

Here we give a technical account of our "adiabatic hypothesis," which is the assumption that the interactions turn off sufficiently quickly in the far past/future so that asymptotic states in the interacting theory are reliably captured by the corresponding states in the free theory. In particular, we wish to highlight that while adiabaticity in the far past ($\tau \to -\infty$) follows from essentially the same argument as in Minkowski space, the far future of de Sitter ($\tau \to 0$) is qualitatively different. In particular, the adiabatic hypothesis only strictly applies for sufficiently massive fields. We will also focus on modes with $\mathbf{k} \neq 0$ which regulates possible IR divergences.

The basic idea is that $\hat{\varphi}(\tau, \mathbf{k})$ acting on the vacuum should produce a new state which contains (with some nonzero probability) a single particle of momentum $\mathbf{k}$. We write this probability amplitude as

$$ {}_{\text{out}}\langle \mathbf{q}, -\infty|\hat{\varphi}(\tau,\mathbf{k})|0,-\infty\rangle_{\text{out}} \equiv f^+_{\text{out}}(k\tau)(2\pi)^d\delta^d(\mathbf{k}+\mathbf{q}) \tag{A1} $$

for the out states, and

$$ {}_{\text{in}}\langle \mathbf{q}, -\infty|\hat{\varphi}(\tau,\mathbf{k})|0,-\infty\rangle_{\text{in}} \equiv f^+_{\text{in}}(k\tau)(2\pi)^d\delta^d(\mathbf{k}+\mathbf{q}) \tag{A2} $$

for the in states. If $\hat{\varphi}$ is an operator in the principal series, then f^+_{out} and f^+_{in} take the same form as f^+ in (10) but with a possibly renormalized Z and μ.

The technical issue is that, in the interacting theory, $\hat{\varphi}$ can also create multiple particles. In particular, while

$Z^2 f^-(\tau\overleftrightarrow{\partial}_\tau)\hat{\varphi}$ creates a normalized one-particle state in the free theory, in the interacting theory it creates

$$\begin{aligned} & iZ^2_{\text{out}} f^-_{\text{out}}(\tau\overleftrightarrow{\partial}_\tau)\hat{\varphi}(\tau,\mathbf{k})|0,-\infty\rangle_{\text{out}} \\ &\quad = |\mathbf{k},-\infty\rangle_{\text{out}} + \sum_{n=2}^{\infty} c_{\text{out}}(k\tau;n)|n,-\infty\rangle_{\text{out}}, \end{aligned} \tag{A3}$$

where the c_{out} are the probability amplitudes for creating n particles from the vacuum. There is an analogous set of c_{in} defined by $Z^2_{\text{in}} f^-_{\text{in}}(\tau\overleftrightarrow{\partial}_\tau)\hat{\varphi}$ acting on the in vacuum. The adiabatic hypothesis that interactions "switch off" at early and late times is formally the requirement that

$$\lim_{\tau\to 0} c_{\text{out}}(k\tau;n) = 0, \qquad \lim_{\tau\to-\infty} c_{\text{in}}(k\tau;n) = 0. \tag{A4}$$

The vanishing of c_{in} at early times is ultimately the same assumption that is made to define the Minkowski S matrix, and since de Sitter is indistinguishable from Minkowski as $k\tau \to -\infty$ the usual arguments can be used to justify this weak limit [54] (although see [54–58] and more recently [59] for subtleties related to composite or unstable particles). The only qualitative difference is that any finite mass parameter μ will blueshift away in the far past, and all such fields behave essentially as if massless—this can lead to the same kind of IR divergences which appear for massless particles on Minkowski, but these do not affect the S matrix for scattering hard modes with $k \neq 0$ [60].

The vanishing of c_{out} at late times is more subtle, and our argument for this is essentially perturbative. Since in the Heisenberg picture the time evolution of $\hat{\varphi}$ is determined by the equation of motion,

$$Z^2 \mathcal{E}[k\tau]\varphi(\tau,\mathbf{k}) = \tau \frac{\delta S_{\text{int}}}{\delta\varphi(\tau,\mathbf{k})}, \tag{A5}$$

we can write the general solution as

$$\begin{aligned} \varphi(\tau,\mathbf{k}) &= f^+_{\text{out}}(k\tau)a^\dagger_{\text{out}}(\mathbf{k}) + f^-_{\text{out}}(k\tau)a_{\text{out}}(-\mathbf{k}) \\ &\quad + \int_\tau^0 d\tau' G^{\text{ret}}_{\text{out}}(k\tau,k\tau')\frac{\delta S_{\text{int}}}{\delta\varphi(\tau',\mathbf{k})}, \end{aligned} \tag{A6}$$

where S_{int} is the (suitably renormalized) nonlinear part of the action, $\hat{a}_{\text{out}}$ annihilates $|0,-\infty\rangle_{\text{out}}$, and $G^{\text{ret}}_{\text{out}}$ is the retarded propagator built from f_{out} mode functions. One can then verify that de Sitter invariant interactions for massive fields in S_{int} will give contributions to c_{out} that vanish at late times. For example, the interaction $S_{\text{int}} = \lambda\sqrt{-g}\phi^3$ appears at first order in λ as the coefficient

$$\begin{aligned} & c_{\text{out}}(k\tau;q_1\tau,q_2\tau) \\ &\quad = \lambda\int_\tau^0 \frac{d\tau'}{\tau'}(-\tau')^{\frac{d}{2}} f^-_{\text{out}}(k\tau') f^+_{\text{out}}(q_1\tau') f^+_{\text{out}}(q_2\tau'), \end{aligned} \tag{A7}$$

which describes the overlap with the two-particle state $|\mathbf{q}_1\mathbf{q}_2,-\infty\rangle_{\text{out}}$ [and we have suppressed a factor of $\delta^d(\mathbf{q}_1+\mathbf{q}_2-\mathbf{k})$]. For fields in the principal series, this integral is finite for all τ and vanishes as $\tau\to 0$.

In general, for an n-point interaction involving both light and heavy fields, the integrand which appears in c_{out} behaves like $\sim\tau^{\alpha-1}$ at small τ, with

$$\alpha = \frac{d}{2}(n-2) - \sum_{b=1}^{n} |\text{Im}\mu_b|. \tag{A8}$$

The corresponding integral therefore diverges if the total $|\text{Im}\mu_T|$ exceeds $\frac{d}{2}(n-2)$, and in that case c_{out} is no longer guaranteed to vanish at late times [63].

However, there are nonetheless some interactions involving light fields which switch off sufficiently fast to avoid any issues at late times. An example would be the cubic interaction $\phi^2\pi$ between two heavy fields ϕ and a light field π, for which c_{out} vanishes for every nonzero π mass. More generally, if the late-time divergence in c_{out} is suitably regularized, a renormalized theory of the boundary degrees of freedom may have well-defined S matrix elements for an even wider range of interactions and mass values.

A physical picture of this divergence is the following. Suppose two wave packets are prepared in the far past, sufficiently localized so that their overlap vanishes, and the two particles therefore do not interact. Since the expansion of de Sitter spreads out these wave packets at late times, no matter how localized the two particles were initially they will inevitably overlap to some extent in the far future. This effect competes with the diluting strength of the interaction between the particles [e.g., $\sqrt{-g}\phi^n$ scales as $\tau^{d(n-2)}$ at small τ for principal series fields]. For sufficiently heavy fields, the interaction strength falls off fast enough that the two particles do not interact at late times. For sufficiently light fields, the spreading of their wave packets at late times gives rise to a non-negligible interaction between the two particles. In the latter case, the particles do not decouple into their free theory eigenstates, and a careful renormalization of their long-lived interactions is required.

Finally, we should distinguish between the late-time divergences which can appear due to the conformal boundary at $\tau\to 0$ and the soft/colinear IR divergences which can appear as on Minkowski. For instance, the interaction σ^4 of conformally coupled fields produces an $\mathcal{S}_{\mathbf{k}_1\mathbf{k}_2\to\mathbf{k}_3\mathbf{k}_4}$ which is free of any late-time divergence, but contains a kinematic singularity at $k_1+k_2=k_3+k_4$. Since such singularities also appear in the Minkowski S

matrix for light fields, we believe they should be renormalized in the usual way. Such singularities do not affect generic (noncolinear) kinematics, and so we leave further exploration of these exceptional points for the future.

2. Disconnected components

In the main text we focused on the connected contributions to the S matrix. The complete LSZ formula also contains disconnected contributions. These arise from the nonzero commutators between $\hat{a}_{\mathbf{k}}$, $\hat{a}^{\dagger}_{\mathbf{k}}$, $\hat{b}_{\mathbf{k}}$, and $\hat{b}^{\dagger}_{\mathbf{k}}$, where the $\hat{b}$ operators are the Unruh-DeWitt analog of the $\hat{a}$ operators, i.e., they create $|1,0\rangle$ from $|0,0\rangle$ in the free theory. Specifically, the LSZ derivation in (16) also produces disconnected terms such as

$$\hat{a}_{\mathbf{k}}|n\rangle_{\mathrm{in}} = \sum_{b=1}^{n}(2\pi)^d\delta^d(\mathbf{k}-\mathbf{k}_b)|n-1\rangle_{\mathrm{in}},$$
$$\alpha|n\rangle_{\mathrm{in}} = \hat{b}^{\dagger}_{\mathbf{k}_n}|n-1\rangle_{\mathrm{in}} - \beta\sum_{b=1}^{n-1}(2\pi)^d\delta^d(\mathbf{k}_n-\mathbf{k}_b)|n-1\rangle_{\mathrm{in}}. \tag{A9}$$

For instance, the full expression for the $2\to 2$ S matrix elements is shown diagrammatically in Figs. 1 and 2.

3. Quantum mechanics analogy

In nonrelativistic quantum mechanics, the goal of scattering theory is to find the wavefunction $\psi(t,x) = e^{-iEt/\hbar}\psi(x)$ that solves the Schrödinger equation

$$E\psi(x) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x) \tag{A10}$$

subject to particular boundary conditions. Typically, we consider potentials $V(x)$ that go to zero at $x\to\pm\infty$. In that case, when waves with energy $E=\frac{\hbar^2k^2}{2m}$ are sent in from the left (from $x=-\infty$) we look for wavefunctions with the asymptotic form

$$\psi_-(x) \sim \begin{cases} e^{+ikx} + r_- e^{-ikx} & \text{as } x\to-\infty \\ t_- e^{+ikx} & \text{as } x\to+\infty. \end{cases} \tag{A11}$$

The rationale is that e^{ikx} is the right-moving solution for $\psi(x)$ when $V=0$, and so when the potential is introduced its effect is to reflect some of this wave (captured by the reflection coefficient r_-) and transmit the remainder (captured by the transmission coefficient t_-). If waves of the same energy are sent in from the right (from $x=+\infty$), then we look for solutions with

$$\psi_+(x) \sim \begin{cases} t_+ e^{-ikx} & \text{as } x\to-\infty \\ e^{-ikx} + r_+ e^{+ikx} & \text{as } x\to+\infty. \end{cases} \tag{A12}$$

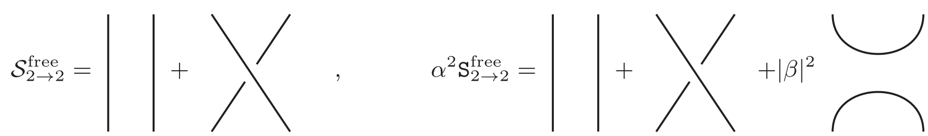

FIG. 1. The free theory S matrix element for "$2\to2$ scattering" in both the Bunch-Davies and Unruh-DeWitt basis. The latter contains an additional contribution from particle production. A line joining two external points represents a momentum conserving δ function.

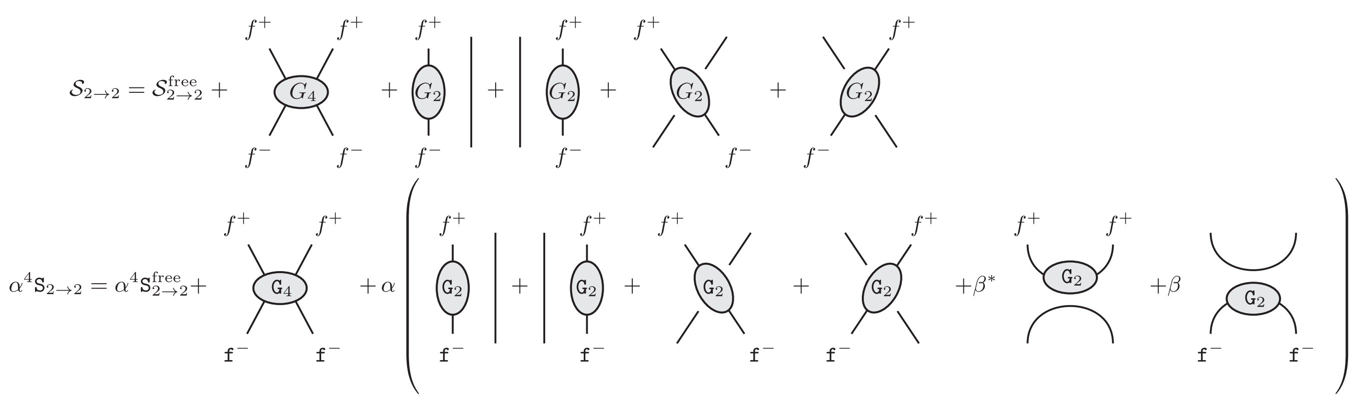

FIG. 2. In the interacting theory, the S matrix element for "$2\to2$ scattering" is given by the free contribution shown in Fig. 1 plus the diagrams shown above. The Unruh-DeWitt basis again contains additional disconnected diagrams due to the free theory particle production. A line joining two external points represents a momentum-conserving δ function, and the gray blobs represent the amputated Green's functions shown (which are then put on shell using the mode functions shown).

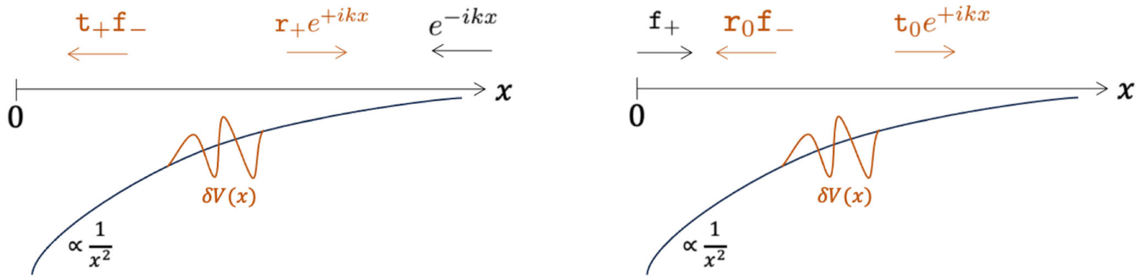

FIG. 3. Cartoon of the scattering from the potential (A15). Left: the ψ_+ boundary condition in (A16). Right: the ψ_0 boundary condition in (A17).

The S matrix is then conventionally defined as

$$S = \begin{pmatrix} t_- & r_- \\ r_+ & t_+ \end{pmatrix} \tag{A13}$$

and conveniently encodes how plane waves scatter off $V(x)$ (i.e., how they are reflected or transmitted), as shown in Fig 3. The S matrix enjoys many useful properties, including for instance $S^\dagger S = 1$ for any real potential—an immediate consequence of $|r_\pm|^2 + |t_\pm|^2 = 1$ together with the relation

$$t_+ = t_-, \qquad r_+ = -r_-^* t_-/t_-^* \tag{A14}$$

between right- and left-scattering coefficients (which follows from comparing the solution $\psi_-^* - r^*\psi_-$ with ψ_+).

The story changes somewhat if the potential does *not* vanish at the boundaries. This is the situation of interest for cosmology, where at late times particles continue to feel the effects of the expanding spacetime since they are redshifted to ever larger areas. To give an analog in the simple setting of nonrelativistic quantum mechanics, suppose we scatter particles between $x = 0$ to $x = +\infty$ subject to a potential of the form

$$V(x) = -\frac{\hbar^2}{2m}\frac{\nu^2 + \frac{1}{4}}{x^2} + \delta V(x), \tag{A15}$$

where now $\delta V(x)$ represents some localized target and vanishes at $x \to 0$ and $x \to +\infty$, but there is a background potential that $\sim 1/x^2$ and dominates the physics near the $x = 0$ boundary. Looking at (A10), we see that when $\delta V = 0$ the left- and right-moving solutions near $x = 0$ are $\mathtt{f}_- = (kx/\nu)^{1/2 - i\nu}$ and $\mathtt{f}_+ = (kx/\nu)^{1/2 + i\nu}$ [64]. On the other hand, as $x \to +\infty$ the potential becomes unimportant, and we recover the $e^{\pm ikx}$ plane waves from before. So one natural analog of scattering from this potential would be to look for solutions of the form

$$\psi_+(x) \sim \begin{cases} \mathtt{t}_+\mathtt{f}_-(kx) & \text{as } x \to 0 \\ e^{-ikx} + \mathtt{r}_+ e^{+ikx} & \text{as } x \to +\infty \end{cases} \tag{A16}$$

to describe waves being sent in from $x = +\infty$, and also solutions of the form

$$\psi_0(x) \sim \begin{cases} \mathtt{f}_+(kx) + \mathtt{r}_0\mathtt{f}_-(kx) & \text{as } x \to 0 \\ \mathtt{t}_0 e^{+ikx} & \text{as } x \to +\infty \end{cases} \tag{A17}$$

to describe waves being sent in from $x = 0$. The resulting S matrix,

$$\mathtt{S} = \begin{pmatrix} \mathtt{t}_0 & \mathtt{r}_0 \\ \mathtt{r}_+ & \mathtt{t}_+ \end{pmatrix}, \tag{A18}$$

is the analog of the Unruh-de Witt S matrix defined in the main text. Notice that as long as δV is real and $|\mathtt{r}_+|^2 + |\mathtt{t}_+|^2 = 1$, then $\mathtt{S}^\dagger \mathtt{S} = 1$ as before since

$$\mathtt{t}_+ = \mathtt{t}_0, \qquad \mathtt{r}_+ = -\mathtt{r}_0^* \mathtt{t}_0/\mathtt{t}_0^*. \tag{A19}$$

However, *unlike* the previous S matrix, when we remove the localized scatterer (i.e., send $\delta V \to 0$), the Unruh-de Witt S matrix does not reduce to the identity. In fact, since we can solve (A10) exactly when $\delta V = 0$ and find a general solution,

$$\psi(x) = \sqrt{kx}[AJ_{i\nu}(kx) + BJ_{-i\nu}(kx)], \tag{A20}$$

we can determine the S matrix explicitly:

$$\mathtt{S} = \begin{pmatrix} e^{i\chi}\sqrt{1 - e^{-2\nu\pi}} & -ie^{-\nu\pi + 2i\chi} \\ -ie^{-\nu\pi} & e^{i\chi}\sqrt{1 - e^{-2\nu\pi}} \end{pmatrix}, \tag{A21}$$

where $e^{i\chi} = \sqrt{\frac{\nu\pi}{i\sinh(\nu\pi)}}(\frac{2}{\nu})^{i\nu}/\Gamma(1 - i\nu)$ is a pure phase that could have been absorbed into $\mathtt{f}_\pm$. This nontrivial S matrix arises because, even when the scatterer's potential vanishes, the background $1/x^2$ part of the potential can still reflect some of the ingoing waves.

For some questions, $\mathtt{S}$ is the most convenient answer. For instance if we actually can send in $\mathtt{f}_+$ waves from $x = 0$ and measure the output at $x = +\infty$, then the matrix $\mathtt{S}$ is useful because its values describe the actual transmission and reflection we would measure in that experiment (which is a combination of the reflections from δV plus the scattering from the background $1/x^2$ potential). In the cosmological context, the nontrivial (A21) reflects the physical fact that, even in the free theory, particles can be produced from the vacuum thanks to the expanding spacetime. If our goal is to study this free theory particle production, then $\mathtt{S}$ is a natural object to compute.

However, in practice the physical questions we want to answer in inflationary cosmology are somewhat different. In that context, we *cannot* prepare and send in waves from $x = 0$. Assuming the Universe was in the Bunch-Davies vacuum state in the far past corresponds in this analogy to having e^{-ikx} boundary conditions at $x = +\infty$ (in the far past), and the natural question we would like to answer is how some small δV interactions during inflation affect the evolution of this state. Since we can quantize the background $1/x^2$ potential exactly, this amounts to asking how the "Bunch-Davies" solution $f_- \propto \sqrt{kx}H^{(2)}_{i\nu}(kx)$ and its conjugate $f_+ \propto \sqrt{kx}H^{(1)}_{i\nu}(kx)$ are affected by the presence of δV as $x \to 0$ (in the far future). The corresponding scattering problem is then to find wavefunction solutions with the asymptotic behavior

$$\psi_+(x) \sim \begin{cases} t_+ f_-(kx) & \text{as } x \to 0 \\ e^{-ikx} + r_+ e^{+ikx} & \text{as } x \to +\infty \end{cases} \tag{A22}$$

to describe waves being sent in from $x = +\infty$, and also solutions of the conjugate problem

$$\psi_0(x) \sim \begin{cases} f_+(kx) + r_0 f_-(kx) & \text{as } x \to 0 \\ t_0 e^{+ikx} & \text{as } x \to +\infty. \end{cases} \tag{A23}$$

Notice that now the $f_\pm$ behavior at $x \to 0$ does not correspond to left moving/right moving, but rather corresponds to what we would expect to see if δV were absent and there was no scatterer. The "Bunch-Davies" S matrix,

$$\mathcal{S} = \begin{pmatrix} t_0 & r_0 \\ r_+ & t_+ \end{pmatrix}, \tag{A24}$$

is therefore an alternative description of δV which has the useful property that it reduces to the identity when $\delta V \to 0$. Furthermore, for real potentials and $|r_+|^2 + |t_+|^2 = 1$, we retain the usual properties such as $\mathcal{S}^\dagger \mathcal{S} = 1$.

To sum up, when $V(x)$ vanishes at the boundaries of our scattering domain, there is essentially a unique definition of the S matrix (up to unimportant phases) which describes how plane waves are reflected/transmitted from $V(x)$ when incident from left or right. However, when $V(x)$ can be separated into an exactly solvable background (which *does not* vanish at the boundary) with a weak perturbation δV on top (which *does* vanish at the boundaries), then there is a choice to be made about what reflection/transmission coefficients to consider. Physically, the distinction between the "Unruh-de Witt" and "Bunch-Davies" choices is ultimately whether to compute the reflection/transmission from the total $V(x)$, or from only the perturbation $\delta V(x)$. They are not independent objects, and in practice can always be related to one another: for instance using (22) to replace $f^\pm$ with $\mathtt{f}^\pm$ in ψ_0 gives the identification

$$\mathtt{r}_0 = \frac{\beta + r_0\alpha^*}{\alpha + r_0\beta^*}, \qquad \mathtt{t}_0 = \frac{t_0}{\alpha + r_0\beta^*}. \tag{A25}$$

Notice that this Bogoliubov transformation has preserved $|r|^2 + |t|^2 = 1$, and that the other $\mathtt{r}_+$ and $\mathtt{t}_+$ coefficients can be found immediately from (A19). As we have shown in the main text, the "Bunch-Davies" S matrix seems better suited to practical applications in cosmological collider physics, since it is often easier to compute in perturbation theory and has a more direct connection with the in-in correlators that ultimately seed the CMB fluctuations that we observe.

[1] R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix* (Cambridge University Press, Cambridge, England, 1966).
[2] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity* (Cambridge University Press, Cambridge, England, 2015).
[3] S. B. Giddings, Phys. Rev. Lett. **83**, 2707 (1999).
[4] I. Heemskerk, J. Penedones, J. Polchinski, and J. Sully, J. High Energy Phys. 10 (2009) 079.
[5] J. Penedones, J. High Energy Phys. 03 (2011) 025.
[6] A. L. Fitzpatrick and J. Kaplan, J. High Energy Phys. 10 (2012) 127.
[7] S. Raju, Phys. Rev. D **85**, 126009 (2012).
[8] M. F. Paulos, J. Penedones, J. Toledo, B. C. van Rees, and P. Vieira, J. High Energy Phys. 11 (2017) 133.
[9] M. Spradlin and A. Volovich, Phys. Rev. D **65**, 104037 (2002).
[10] R. Bousso, Phys. Rev. D **71**, 064024 (2005).

[11] G. Dvali, C. Gomez, and S. Zell, J. Cosmol. Astropart. Phys. 06 (2017) 028.

[12] C. Cheung, J. Parra-Martinez, and A. Sivaramakrishnan, J. High Energy Phys. 05 (2022) 027.

[13] D. Baumann, D. Green, A. Joyce, E. Pajer, G. L. Pimentel, C. Sleight, and M. Taronna, arXiv:2203.08121.

[14] P. Benincasa, 10.1142/S0217751X22300101 (2022).

[15] D. Marolf, I. A. Morrison, and M. Srednicki, Classical Quantum Gravity **30**, 155023 (2013).

[16] A. Bzowski, P. McFadden, and K. Skenderis, J. High Energy Phys. 03 (2016) 066.

[17] A. Bzowski, P. McFadden, and K. Skenderis, J. High Energy Phys. 11 (2018) 159.

[18] G. Dvali, Symmetry **13**, 3 (2020).

[19] Note that we work throughout in the Heisenberg picture. In the Schrödinger picture, (2) corresponds to expanding $\hat{U}(0,-\infty)\hat{a}^\dagger_{n'}\ldots\hat{a}^\dagger_{1'}|\Omega\rangle$ in terms of the states $\hat{U}_{\rm free}(0,-\infty)\hat{a}^\dagger_n\ldots\hat{a}^\dagger_1|\Omega\rangle$, where $|\Omega\rangle$ is the Bunch-Davies vacuum state in the far past and $\hat{U}(\tau_2,\tau_1)$ is the time evolution operator from τ_1 to τ_2.

[20] The Fourier transform is performed using the flat metric, so $\phi(\tau,\mathbf{k}) = \int d^d\mathbf{x} e^{i\mathbf{k}\cdot\mathbf{x}}\phi(\tau,\mathbf{x})$ where $\mathbf{k}\cdot\mathbf{x} = k^i\delta_{ij}x^i$ is τ-independent. We also abuse notation and use the same symbol to denote functions in different representations: for instance $\phi(\tau,\mathbf{x})$ and $\phi(\tau,\mathbf{k})$ are of course different functions (one is the field in position space, and the other is the field in momentum space).

[21] The real constant Z describes the power spectrum at early times,

$$\lim_{\tau\to-\infty}\langle\hat{\varphi}(\tau,\mathbf{k}')\hat{\varphi}(\tau,\mathbf{k})\rangle = \frac{1}{2kZ^2}\delta^d(\mathbf{k}+\mathbf{k}'),$$

and is fixed by the overall normalization of the ϕ kinetic term in the Lagrangian. The phase of $f^\pm$ has been fixed so that the crossing relation (29) has a trivial phase.

[22] One exception is the trivial $1\to1$ scattering amplitude, for which translation invariance requires that $\mathbf{k}=\mathbf{k}'$, and an additional boundary term must be included in (16)—see the Appendix for more details about these disconnected boundary contributions.

[23] The Bogoliubov coefficients are given explicitly by $\alpha = e^{+\mu\pi/2}/\sqrt{2\sinh(\mu\pi)}$ and $\beta = e^{-\mu\pi/2}/\sqrt{2\sinh(\mu\pi)}$.

[24] Our conventions for the Legendre functions are

$$P^n_{i\mu-\frac{1}{2}}(z) = \frac{1}{\Gamma(1-n)}\left(\frac{z+1}{z-1}\right)^{n/2}{}_2F_1\left(\frac{1}{2}-i\mu,\frac{1}{2}+i\mu;1-n;\frac{1-z}{2}\right)$$

$$\frac{Q^n_{i\mu-\frac{1}{2}}(z)}{(z^2-1)^{n/2}} = \frac{e^{i\pi n}}{2^{n+1}}\frac{\Gamma\left(i\mu+n+\frac{1}{2}\right)\Gamma\left(i\mu+\frac{1}{2}\right)}{\Gamma(2i\mu+1)}\left(\frac{z-1}{2}\right)^{-i\mu-\frac{1}{2}-n}$$
$$\times\,{}_2F_1\left(i\mu+n+\frac{1}{2},i\mu+\frac{1}{2},2i\mu+1;\frac{2}{1-z}\right).$$

[25] N. A. Chernikov and E. A. Tagirov, Ann. Inst. Henri Poincare Phys. Theor. A **9**, 109 (1968).

[26] T. S. Bunch and P. C. W. Davies, Proc. R. Soc. A **360**, 117 (1978).

[27] B. Allen, Phys. Rev. D **32**, 3136 (1985).

[28] C. Schomblond and J. Géhéniau, Bulletins de l'Académie Royale de Belgique **54**, 1147 (1968).

[29] A. M. Polyakov, Nucl. Phys. **B797**, 199 (2008).

[30] B. Allen and T. Jacobson, Commun. Math. Phys. **103**, 669 (1986).

[31] M. Fukuma, S. Sugishita, and Y. Sakatani, Phys. Rev. D **88**, 024041 (2013).

[32] S. J. Avis, C. J. Isham, and D. Storey, Phys. Rev. D **18**, 3565 (1978).

[33] For S-matrix elements with only two particles, crossing must be applied carefully since momentum conservation fixes $k_1 = k_2$. The naive procedure of "flipping the sign of k" simply maps $\mathcal{S}_{0\to2}$ to $\mathcal{S}_{2\to0}$, both of which are zero for the Bunch-Davies S-matrix. In terms of $\tilde{\mathcal{S}}_2$, the boundary term responsible for the nonzero $\mathcal{S}_{1\to1}$ element is essentially the Wronskian $f^+(\tilde{k}_1\tau)(\tau\overleftrightarrow{\partial_\tau})f^+(\tilde{k}_2\tau)$, which indeed vanishes when both $\tilde{k}$ have the same sign but is nonzero when both $\tilde{k}$ have different signs.

[34] Note that when acting on a function of only $k=|\mathbf{k}|$, the generator can be written as $\mathbf{K}[\mathbf{k}] + d\partial_\mathbf{k} = \frac{\mathbf{k}}{k^2}(k\partial_k)^2$. This is in line with $\mathbf{K}$ being the "momentum" associated with translations of the inverted position $\mathbf{x}/x^2$, since the special conformal transformation is equivalent to an inversion-translation-inversion.

[35] Given the canonical normalization of φ, it is the extended generators $D[\tau,\mathbf{k}] - d/2$ and $\mathbf{K}[\tau,\mathbf{k}] - d\partial_\mathbf{k}$ that commute with $\mathcal{E}[k\tau]$.

[36] Here it is important that we set each $\tilde{k} = \pm k$, since otherwise $\mathcal{E}[k\tau]f^+(\tilde{k}\tau)\neq0$ and there would be additional contributions to $\tilde{\mathcal{S}}_4^{\rm exch}$, i.e., the off shell extension of the S-matrix need not be invariant under field redefinitions.

[37] A. Bzowski, P. McFadden, and K. Skenderis, J. High Energy Phys. 03 (2014) 111.

[38] G. Boerner and H. P. Duerr, Nuovo Cimento A **64**, 669 (1969).

[39] Explicitly, $\alpha_0 = \Omega_k - m\,\mathrm{arcsinh}(\frac{m}{k})$.

[40] Equation (44) also demonstrates that the short-distance singularity at $\cosh\sigma\to1$ in both de Sitter propagators matches that of Minkowski [30], as expected since dS is locally flat.

[41] Note that since

$$P_0^{-j}(z) = \frac{1}{\Gamma(1+j)}\left(\frac{z-1}{z+1}\right)^{j/2},$$

(48) indeed reduces to (46) when $i\mu$ is continued to the value $1/2$.

[42] Here, "connected" corresponds to keeping only contributions which are proportional to a single momentum-conserving delta function—perturbatively, this corresponds to keeping only connected Feynman-Witten diagrams.

[43] This is often written in terms of field eigenstates as

$$\langle\phi(\tau)|0,-\infty\rangle_{\rm in} = \exp\left(\sum_n^\infty\left[\prod_{b=1}^n\int\frac{d^d\mathbf{k}_b}{(2\pi)^d}\phi(\tau,\mathbf{k}_b)\right]\frac{\psi_n(\tau)}{n!}\right)$$

since in that basis each $i\hat{\Pi}$ operator in (52) implements the field derivative $Z^2\delta/\delta\phi$.

[44] The correlator can also be defined in terms of the wavefunction coefficient using the functional integral

$$\langle\phi_1\ldots\phi_n\rangle = \frac{\int\mathcal{D}\phi\phi_1\ldots\phi_n|\langle\phi|0,-\infty\rangle_{\rm in}|^2}{\int\mathcal{D}\phi|\langle\phi|0,-\infty\rangle_{\rm in}|^2}.$$

[45] S. Melville and G. L. Pimentel, J. High Energy Phys. 08 (2024) 211.
[46] S. Céspedes, A.-C. Davis, and S. Melville, J. High Energy Phys. 02 (2021) 012.
[47] H. Goodhew, S. Jazayeri, and E. Pajer, J. Cosmol. Astropart. Phys. 04 (2021) 021.
[48] S. Melville and E. Pajer, J. High Energy Phys. 05 (2021) 249.
[49] H. Goodhew, S. Jazayeri, M. H. Gordon Lee, and E. Pajer, J. Cosmol. Astropart. Phys. 08 (2021) 003.
[50] D. Baumann, W.-M. Chen, C. Duaso Pueyo, A. Joyce, H. Lee, and G. L. Pimentel, J. High Energy Phys. 09 (2022) 010.
[51] S. Albayrak, P. Benincasa, and C. D. Pueyo, SciPost Phys. **16**, 157 (2024).
[52] S. Agui-Salcedo and S. Melville, J. High Energy Phys. 12 (2023) 076.
[53] S. A. Salcedo, M. H. G. Lee, S. Melville, and E. Pajer, J. High Energy Phys. 06 (2023) 020.
[54] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).
[55] R. Haag, Phys. Rev. **112**, 669 (1958).
[56] R. Haag, Nuovo Cimento **14**, 131 (1959).
[57] D. Ruelle, *On the Asymptotic Condition in Quantum Field Theory* (Verlag nicht ermittelbar, 1962).
[58] K. Hepp, Commun. Math. Phys. **1**, 95 (1965).
[59] J. Collins, arXiv:1904.10923.

[60] Notice that although the “comoving energy” $k\tau$ would seem to diverge at early times, this is an artifact of the conformal coordinates: in particular, since $\tau\partial_\tau$ is the scale-free derivative that defines the conjugate momentum, the offending factors of τ cancel and the $\tau\to-\infty$ limit recovers the usual Minkowski energy. This is related to recent discussions of the trans-Planckian problem [61,62].

[61] J. Martin and R. H. Brandenberger, Phys. Rev. D **63**, 123501 (2001).
[62] I. Komissarov, A. Nicolis, and J. Staunton, J. High Energy Phys. 05 (2023) 216.

[63] Though note that even when $\alpha<0$ the integral may still converge and $c_{\rm out}$ vanish. An example of this would be the $\tilde{S}_3$ from the cubic interaction σ^3 given in (46): this particular integral is perfectly finite $d=2$ ($\alpha=-1/2$). So while $\alpha>0$ is sufficient for the adiabatic hypothesis, it is not necessary for particular *S*-matrix elements.

[64] To justify why $(kx)^{1/2+i\nu}$ is right moving, consider the phase of $\psi(t,x)$, namely $\phi(t,x)=\nu\log(kx)-Et/\hbar$. The effective frequency and wave number are $\tilde{\omega}=-\partial_t\phi=E/\hbar$ and $\tilde{k}=\partial_x\phi=\nu/x$, which gives a positive phase velocity of $\tilde{\omega}/\tilde{k}=Ex/(\hbar\nu)$. So while the velocity vanishes at exactly $x=0$, at a small displacement from the boundary there is a meaningful separation of $(kx)^{1/2\pm i\nu}$ into left moving and right moving. The overall normalization of $1/\nu$ ensures that $|\mathtt{r}_+|^2+|\mathtt{t}_+|^2=1$.