

Energy's and amplitudes' positivity

Alberto Nicolis^a, Riccardo Rattazzi^b, Enrico Trincherini^{c,d}

^a *Department of Physics and ISCAP,
Columbia University, New York, NY 10027, USA*

^b *Institut de Théorie des Phénomènes Physiques, EPFL,
CH1015 Lausanne, Switzerland*

^c *Scuola Normale Superiore,
Piazza dei Cavalieri 7, 56126 Pisa, Italy*

^d *Scuola Internazionale di Studi Superiori Avanzati,
Via Beirut 2-4, I-34151 Trieste, Italy*

Abstract

In QFT, the null energy condition (NEC) for a classical field configuration is usually associated with that configuration's stability against small perturbations, and with the sub-luminality of these. Here, we exhibit an effective field theory that allows for stable NEC-violating solutions with exactly luminal excitations only. The model is the recently introduced 'galileon', or more precisely its conformally invariant version. We show that the theory's low-energy S -matrix obeys standard positivity as implied by dispersion relations. However we also show that if the relevant NEC-violating solution is inside the effective theory, then *other* (generic) solutions allow for superluminal signal propagation. While the usual association between sub-luminality and positivity is not obeyed by our example, that between NEC and sub-luminality is, albeit in a less direct way than usual.

1 Introduction, and relation to previous work

Of all energy conditions of general relativity (GR), the 'null' one stands out as the best defined—thus the most likely to follow from fundamental properties of well-behaved physical systems. The reason is that the null energy condition (NEC),

$$T_{\mu\nu} n^\mu n^\nu \geq 0 \quad \text{for all null } n^\mu\text{'s}, \quad (1)$$

is saturated by the cosmological constant contribution's to $T_{\mu\nu}$. It is therefore impossible to violate the NEC by adding a suitably negative c.c. to an healthy system, or conversely, to obey it by adding a large positive c.c. to an otherwise negative-energy system. This is not the case for other energy conditions. One could postulate that the latter be obeyed by the "non-cosmological

constant part” of $T_{\mu\nu}$, but the splitting between the cosmological constant and everything else is at best non-local in field space and thus in real space—in the sense that it cannot be performed on the basis of local observations—and more likely ill-defined.

1.1 A no-go theorem...

Since the value of the c.c. does not affect the behavior of a theory at distances smaller than the curvature radius—and at all if we decouple gravity—it is on the NEC that we should focus to look for a generic relation with good properties of healthy physical systems. In fact it is notoriously difficult to violate the NEC while retaining stability. Indeed, it has been shown in ref. [1] that for a generic effective theory of scalar fields ϕ^I ¹, the NEC follows from stability and isotropy, or alternatively from stability and sub-luminality. Explicitly, given the low-energy effective Lagrangian as a derivative expansion,

$$\mathcal{L} = \mathcal{L}(\phi^I, \partial\phi^I, \partial\partial\phi^I, \dots) , \quad (2)$$

at low energies one can restrict to the lowest-derivative order—at most one derivative per field—and ignore higher derivative terms,

$$\mathcal{L} \rightarrow \mathcal{L}(\phi^I, B^{IJ}) , \quad B^{IJ} \equiv \partial_\mu\phi^I \partial^\mu\phi^J . \quad (3)$$

Then one can relate the derivatives of the Lagrangian w.r.t. the metric to the derivatives of the Lagrangian w.r.t. the scalars’ derivatives. That is, there is a very direct relation between a generic solution’s stress-energy tensor, and the kinetic action for small perturbations of the same solution. One can thus show that if a solution violates the NEC, either *(i)* it is unstable, or *(ii)* if it is stable, both its stress-energy tensor is anisotropic, and some of its perturbations are super-luminal. Conversely, if a solution is stable and has an isotropic stress-energy tensor, or if it stable and admits sub-luminal excitations only, then it obeys the NEC [1].

The instabilities we are concerned about are those associated with negative kinetic or gradient energy for the fluctuations, rather than with negative mass terms². This is because such instabilities plague all short-wavelength fluctuations, down to the UV cutoff of the theory. The decay rate is thus extremely fast, dominated by the shortest scales in the theory, and cannot be reliably computed within the effective theory. Conversely, a standard tachyon-like instability, like Jeans’s or that associated with the negative mode of a true-vacuum bubble, is dominated by modes with momenta of order of the tachyonic mass scale, which can be parameterically smaller than the UV cutoff.

Now, the theorem summarized above crucially relies on neglecting higher-derivative contributions to the Lagrangian. Is this justified? Usually it is, for higher-derivative terms are irrelevant at low energies and, perhaps more to the point, if they do become important the derivative expansion is expected to break down—with an infinite tower of even higher-order terms all becoming

¹Such a system is quite general. For instance it describes the low-energy dynamics of perfect fluids and solids [1].

²For background solutions that break Lorentz-invariance spontaneously, the definition of stability may be subtler. See ref. [1] for a discussion about this point.

important at the same time. At the level of the classical e.o.m., higher-derivative terms signal the presence of massive ghost-like degrees of freedom, thus *demanding* that the effective theory break down at energies below their masses and effectively making the higher-derivative terms unimportant within the regime of validity of the effective theory.

1.2 ... with exceptions

However there can be exceptions to this general logic. The first exception is provided by the ghost-condensate [2]. There, the fluctuations' Lagrangian is degenerate (or nearly so) at the lowest-derivative level, and as a consequence certain higher-derivative terms can be important at low energies, below the cutoff of the theory. This is consistent with the derivative expansion because the background solution—the ‘condensate’—spontaneously breaks Lorentz invariance. It is then higher *spatial* derivative terms that become important for the fluctuations' dynamics. They do not introduce any new ghost-like degrees of freedom at the classical level, and being the leading source of gradient energy, they change the scaling dimensions of higher-order operators thus keeping the derivative expansion under control [2]. Indeed the ghost-condensate was used in ref. [3] to construct a cosmological solution that violates the NEC while being stable and isotropic, and having (highly) sub-luminal excitations only.

Here we consider a second exception—that provided by the galileon [4]. The model involves a scalar π enjoying a shift-symmetry on itself as well as on its first derivatives,

$$\pi \rightarrow \pi + c + b_\mu x^\mu, \quad (4)$$

with constant c and b_μ . This is reminiscent of the invariance of non-relativistic mechanics under galilean transformations $x(t) \rightarrow x(t) + x_0 + v_0 t$ —hence the name of the model. Now, galilean invariance (4) forces each π in the e.o.m. to be acted upon by *at least* two derivatives. Absence of ghosts requires instead that there be *no more* than two derivatives per field. There is therefore a preferred set of higher-derivative terms in the Lagrangian—those that yield *purely* two-derivative equations of motion for π . Such terms have been classified in ref. [4]. In four spacetime dimensions there are four: the standard kinetic term

$$\mathcal{L}_2 = -\frac{1}{2}(\partial\pi)^2, \quad (5)$$

a cubic interaction

$$\mathcal{L}_3 = -\frac{1}{2}(\partial\pi)^2 \square\pi, \quad (6)$$

and quartic and quintic ones

$$\mathcal{L}_4 \sim \partial\pi\partial\pi(\partial^2\pi)^2, \quad \mathcal{L}_5 \sim \partial\pi\partial\pi(\partial^2\pi)^3, \quad (7)$$

with suitable Lorentz contractions³.

Now, if we stick to this particular set of interactions, at the classical level no new ghost-like degree of freedom shows up in the non-linear regime—by construction. This is a necessary

³Formally, the tadpole π should also be added to this class [4]. However in its presence the trivial configuration $\pi = 0$ is not a solution. We will therefore set it to zero.

condition for the quantum effective theory not to break down for sizable classical non-linearities. In fact the model exhibits remarkable properties at the quantum level as well. First, thanks to galilean invariance, the Lagrangian terms we discussed do not get renormalized upon inclusion of quantum corrections, nor do the latter generate terms with fewer derivatives, like e.g. $(\partial\pi)^4$ [5]. Second, with a consistent, if not generic choice of counter-terms, our effective theory can be extrapolated deep into the non-linear regime of classical solutions. Essentially, galilean invariance allows for a robust disentanglement of large classical non-linearities from large quantum effects. See ref. [6] for details and discussions about this point.

In summary, the galileon provides us with a classically as well as quantum-mechanically consistent higher-derivative Lagrangian, which can potentially evade the no-go theorem reviewed above. Analyzing this possibility is the goal of our paper.

2 The NEC-violating solution

For practical purposes and for reasons that will become clear below, it is more convenient to work with a slight deformation of the theory. We can promote the Poincaré group supplemented by the galilean transformations (4) (a fifteen-dimensional symmetry group overall) to the conformal group $SO(4, 2)$ [4]. Indeed, if we think of the conformal group as the subgroup of diffeomorphisms that leaves the fictitious metric

$$g_{\mu\nu}^{(\pi)}(x) \equiv e^{2\pi(x)} \eta_{\mu\nu} \quad (8)$$

conformally flat, then we recognize in eq. (4) the small- π limit of infinitesimal dilations

$$\pi(x) \rightarrow \pi(\lambda x) + \log \lambda , \quad (9)$$

and of infinitesimal special conformal transformations

$$\pi(x) \rightarrow \pi(x + (b x^2 - 2(b \cdot x)x)) - 2b_\mu x^\mu . \quad (10)$$

Then one can show that for each galilean invariant term, there is a corresponding conformally invariant one that reduces to that for small π . For instance the kinetic and cubic terms (5, 6) get promoted to

$$\mathcal{L}_2 \rightarrow -\frac{1}{2} e^{2\pi} (\partial\pi)^2 \quad (11)$$

$$\mathcal{L}_3 \rightarrow -\frac{1}{2} (\partial\pi)^2 \square\pi - \frac{1}{4} (\partial\pi)^4 . \quad (12)$$

Such conformally invariant ‘completion’ is most easily constructed by taking suitable curvature invariants for the metric (8), as described in ref. [4]. Working with the model’s conformal version makes finding interesting highly symmetric solutions and studying their properties particularly easy. For instance, there is a deSitter-like solution where the fictitious metric (8) describes (part of) deSitter space:

$$e^{\pi_{\text{dS}}} = -\frac{1}{H_0 t} , \quad -\infty < t < 0 . \quad (13)$$

Such a solution spontaneously breaks the conformal group $SO(4, 2)$ down to one of its maximal subgroups, the deSitter group $SO(4, 1)$. The ‘curvature’ scale H_0 depends on the model’s parameters. Starting from a generic linear combination of our Lagrangian terms

$$\mathcal{L} = c_2 \mathcal{L}_2 + \cdots + c_5 \mathcal{L}_5 , \quad (14)$$

where with an abuse of notation by \mathcal{L}_n we refer to the conformal versions of the galilean invariants (5–7), H_0 is a solution to the algebraic equation [4]

$$H_0^2 \left(-2c_2 + 3c_3 H_0^2 - 3c_4 H_0^4 + \frac{3}{2}c_5 H_0^6 \right) = 0 . \quad (15)$$

Notice that, since we work with the dimensionless field π , the lagrangian coefficients are dimensionful: $c_n \propto E^{6-2n}$. To make conformal invariance explicit, one could conveniently define a dimensionful dilaton field $\phi \equiv M e^\pi$, where M is an arbitrary mass scale, and write the lagrangian in terms of ϕ and dimensionless couplings $\bar{c}_n = c_n M^{2n-6}$. The action parameterized this way would not depend explicitly on M , thanks to scale invariance. In section 3 we shall conveniently pick $M \equiv H_0$. Here we prefer to use the somewhat redundant notation with dimensionful c_n for basically two reasons. First we want to make contact with the non-conformal galileon limit. Second, the role of conformal invariance in our discussion is mostly technical: to simplify computations and to deliver a simple globally defined solution around which to expand. The lesson about NEC violation is more general.

Now, let us consider fluctuations about eq. (13). They are exactly luminal, because of the high degree of symmetry: by construction the full action only involves the fictitious metric (8) and its curvature invariants. For the background solution we are expanding about, the latter are proportional to the former, since the background is maximally symmetric. Therefore, although the solution (13) spontaneously breaks Lorentz invariance, the only background tensor structure appearing in the fluctuations’ action is $\eta_{\mu\nu}$. In particular, it is easy to see that the quadratic Lagrangian for the fluctuation $\varphi \equiv \pi - \pi_{\text{dS}}$ is

$$\mathcal{L}_\varphi = d_2 \frac{1}{H_0^4 t^4} \left[-\frac{1}{2} H_0^2 t^2 (\partial\varphi)^2 - \frac{1}{2} (-4H_0^2) \varphi^2 \right] , \quad (16)$$

where the overall normalization factor d_2 is [4]

$$d_2 \equiv c_2 - 3c_3 H_0^2 + \frac{9}{2}c_4 H_0^4 - 3c_5 H_0^6 . \quad (17)$$

The kinetic term is fixed by the above symmetry considerations, and its normalization by the local analysis of ref. [4]. The mass term is also enforced by symmetry. The background π configuration spontaneously break time-translational invariance, $t \rightarrow t + \epsilon$. The fluctuation φ is the associated Goldstone boson, which has to non-linearly represent the broken symmetry:

$$\varphi \rightarrow \varphi + \pi'_{\text{dS}}(t) \cdot \epsilon = \varphi - \frac{\epsilon}{t} . \quad (18)$$

The variation of the kinetic term under an infinitesimal transformation is non-zero:

$$\delta \left(-\frac{1}{2} \frac{1}{H_0^2 t^2} (\partial\varphi)^2 \right) = \frac{1}{H_0^2 t^2} \partial_\mu \varphi \partial^\mu \frac{\epsilon}{t} = \epsilon \cdot 4\varphi \frac{1}{H_0^2 t^5} , \quad (19)$$

where in the last step we integrated by parts. Then the kinetic term has to be accompanied by a mass term that compensates for this variation,

$$\delta\left(-\frac{1}{2}\frac{1}{H_0^4 t^4}m^2\varphi^2\right) = \epsilon \cdot m^2\varphi \frac{1}{H_0^4 t^5}, \quad (20)$$

which happens precisely for $m^2 = -4H_0^2$. This is obviously the same logic that for spontaneously broken *internal* symmetries ensures the masslessness of the associated Goldstone bosons.

Notice that while the fluctuation's kinetic energy is positive definite for

$$d_2 > 0, \quad (21)$$

the mass term always has the wrong sign. However its size is of order of the typical time variation rate of the background solution. As a consequence ascertaining the solution's stability is subtler than for a time-independent one. In fact the very concept of stability is somewhat ambiguous for a time-dependent solution (see a related discussion in ref. [6]). As to short wave-length perturbations, the solution is certainly stable, for the fluctuations' kinetic energy is positive, and the mass term is negligible at high frequencies. For longer wavelengths, we can tentatively define stability in the following way. For given initial conditions, the fluctuation's equation of motion yields two solutions. At zero momentum, one of these will always be the Goldstone solution:

$$\varphi_{\text{Goldstone}} \equiv \pi'_{\text{dS}}(t) \cdot \epsilon = -\frac{\epsilon}{t}, \quad (22)$$

with constant ϵ . If this is the leading solution at late times ($t \rightarrow 0^-$, in our case), we say that the background solution π_{dS} is stable. Indeed in such a case any long-wavelength small departures from the background solution will be diluted away at sufficiently late times, and we will be left with the background solution itself, slightly translated in time. This is hardly an instability. It is straightforward to check that for the Lagrangian (16) the Goldstone solution (22) is indeed the leading one at late times.

In summary, the deSitter solution is stable against small perturbations if d_2 defined in eq. (17) is positive. Also perturbations are exactly luminal, thanks to the solution's high degree of symmetry.

2.1 The stress-tensor

It remains to be determined whether such a solution can violate the NEC. To answer this question, we have to compute the stress-energy tensor,

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S_\pi}{\delta g_{\mu\nu}}. \quad (23)$$

For a generic π configuration the computation is straightforward, but painful due to the higher-derivative structure of the Lagrangian. However for our deSitter solution (13) things drastically simplify, as we now show. The reader not interested in the details of the computation can jump directly to eq. (52), which expresses the condition on the coefficients c_i for the solution to violate the NEC.

All conformal-invariant operators are of the form [4]

$$e^{(4-2n)\pi} (\partial^2 \pi)^m (\partial \pi)^{2n-2m} \sim \frac{1}{t^4} \quad (24)$$

This is just a consequence of scale-invariance, which is unbroken by the solution:

$$e^{2\pi_{\text{as}}(t)} \rightarrow e^{2\pi_{\text{as}}(\lambda t) + 2 \log \lambda} = e^{2\pi_{\text{as}}(t)} \quad (25)$$

Therefore

$$T^{\mu\nu} = \tau^{\mu\nu} \frac{1}{t^4}, \quad (26)$$

with a constant and isotropic $\tau^{\mu\nu}$. Now, conservation of $T^{\mu\nu}$ for an homogeneous solution in flat space is simply

$$\frac{d\rho}{dt} = 0, \quad (27)$$

which implies $\tau^{00} = 0$. We are left with the pressure $T^{ij} = p \delta^{ij}$, which we can get from the trace

$$3p = T^\mu{}_\mu, \quad (28)$$

which in turn we can compute by varying the action just w.r.t. the conformal mode of the metric:

$$T^\mu{}_\mu = \eta_{\mu\nu} 2 \left. \frac{\delta S_\pi}{\delta g_{\mu\nu}} \right|_{\eta_{\mu\nu}} = \left. \frac{\delta S_\pi}{\delta \varphi} \right|_{\varphi=0}, \quad (29)$$

where

$$g_{\mu\nu}(x) = e^{2\varphi(x)} \eta_{\mu\nu} \equiv g_{\mu\nu}^{(\varphi)}(x). \quad (30)$$

Now we can use the fact that S_π is built out of curvature invariants involving the fictitious metric (8), where π itself appears as the conformal mode. More precisely, suppose we compute the Ricci tensor for the metric $e^{2\pi} g_{\mu\nu}$, with $g_{\mu\nu}$ generic. We get

$$R_{\mu\nu}[e^{2\pi} g] = R_{\mu\nu}[g] + R_{\mu\nu}[e^{2\pi} \eta, g], \quad (31)$$

where by the last term we denote the Ricci tensor for the metric (8), ‘‘covariantized’’ with the metric $g_{\mu\nu}$ —precisely what makes up S_π upon minimal coupling. Actually, to perform the contractions appearing in S_π we have to raise one index with the metric $e^{2\pi} g_{\mu\nu}$:

$$R^\mu{}_\nu[e^{2\pi} g] = e^{-2\pi} R^\mu{}_\nu[g] + e^{-2\pi} g^{\mu\alpha} R_{\alpha\nu}[e^{2\pi} \eta, g], \quad (32)$$

where by $R^\mu{}_\nu[G]$ we are denoting $G^{\mu\alpha} R_{\alpha\nu}[G]$. As we mentioned, minimal coupling corresponds to the replacement

$$\begin{aligned} S_\pi &= \int d^4x e^{4\pi} F(R^\mu{}_\nu[e^{2\pi} \eta]) \rightarrow \int d^4x \sqrt{-g} e^{4\pi} F(e^{-2\pi} g^{\mu\alpha} R_{\alpha\nu}[e^{2\pi} \eta, g]) \\ &= \int d^4x \sqrt{-g} e^{4\pi} F(R^\mu{}_\nu[e^{2\pi} g] - e^{-2\pi} R^\mu{}_\nu[g]) \end{aligned} \quad (33)$$

which for the conformally flat metric (30), at linear order in φ becomes

$$\begin{aligned}
S_\pi[\pi, \varphi] &= \int d^4x e^{4(\pi+\varphi)} F(R^\mu{}_\nu[e^{2(\pi+\varphi)}\eta] - e^{-2\pi}R^\mu{}_\nu[e^{2\varphi}\eta]) \\
&\simeq S_\pi[\pi + \varphi, 0] - \int d^4x e^{2\pi} \frac{\partial F}{\partial R^\mu{}_\nu} R^\mu{}_\nu[e^{2\varphi}\eta] \\
&\simeq S_\pi[\pi + \varphi, 0] + \int d^4x \varphi(x) (2\partial^\mu\partial_\nu + \delta_\nu^\mu\Box) \left(e^{2\pi} \frac{\partial F}{\partial R^\mu{}_\nu} \right)
\end{aligned} \tag{34}$$

On a solution of the π eom, the first term has vanishing variation w.r.t. φ , and we are left just with the second:

$$T^\mu{}_\mu = (2\partial^\mu\partial_\nu + \delta_\nu^\mu\Box) \left(e^{2\pi} \frac{\partial F}{\partial R^\mu{}_\nu} \right). \tag{35}$$

The above expression holds for any solution. However in our deSitter case we can simplify this expression even further. For the deSitter solution,

$$R_{\text{dS}}{}^\mu{}_\nu = 3H_0^2\delta_\nu^\mu. \tag{36}$$

It is straightforward to convince oneself that

$$\left. \frac{\partial F}{\partial R^\mu{}_\nu} \right|_{\text{dS}} = \frac{1}{12} \delta_\nu^\mu \frac{dF(R_{\text{dS}}{}^\mu{}_\nu)}{dH_0^2} \tag{37}$$

Then derivatives in eq. (35) only act on the $e^{2\pi}$ factor, and we finally get

$$T^\mu{}_\mu = -\frac{3}{H_0^2 t^4} \frac{dF(R_{\text{dS}}{}^\mu{}_\nu)}{dH_0^2} \tag{38}$$

In summary, to compute the pressure, we just have to express the action of π in terms of curvature invariants of the metric (8), plug in the deSitter solution, and then derive w.r.t. to H_0^2 .

2.2 The problem with \mathcal{L}_3

If we apply eq. (38) to \mathcal{L}_3 , we run into a problem. In order to express \mathcal{L}_3 in terms of curvature invariants, on top of R^2 and $(R^\mu{}_\nu)^2$ we need the operator [4]

$$\int d^d x \sqrt{-g_d} \frac{1}{d-4} \left[\frac{(R^\mu{}_\nu)^2}{d-1} - \frac{R^2}{(d-1)^2} \right], \tag{39}$$

which diverges when $d \rightarrow 4$ for generic metric, but is regular for conformally flat metrics $g_{\mu\nu} = e^{2\pi}\eta_{\mu\nu}$, for which it reduces to

$$\int d^d x e^{(d-4)\pi} \left[\frac{(d-2)(3d-4)}{2(d-1)} \Box\pi(\partial\pi)^2 + \frac{(d-2)^3}{2(d-1)} (\partial\pi)^4 + (\Box\pi)^2 \right]. \tag{40}$$

Now, the problem is that to get from eq. (39) to eq. (40), we integrated by parts. This amounts to neglecting a boundary term, whose coefficient diverges for $d \rightarrow 4$. This is fine for asymptotically

trivial π 's, but for our deSitter solution we have to keep track of the boundary term. Indeed eq. (39) diverges when computed on a deSitter solution, even though the latter *is* conformally flat. The bottom line is that if we want to apply eq. (38) to an action containing \mathcal{L}_3 , we must be more careful about boundary terms.

We can bypass this problem by treating \mathcal{L}_3 separately, since the computation of its stress-energy tensor is straightforward anyway:

$$\int d^4x \mathcal{L}_3 \quad \rightarrow \quad T_{\mu\nu}^{(3)} = \frac{1}{2} [2 \partial_\mu \pi \partial_\nu \pi \square \pi - (\partial_\mu \pi \partial_\nu (\partial\pi)^2 + \partial_\nu \pi \partial_\mu (\partial\pi)^2) + \eta_{\mu\nu} \partial_\alpha \pi \partial^\alpha (\partial\pi)^2] \\ + \frac{1}{4} [4(\partial\pi)^2 \partial_\mu \pi \partial_\nu \pi - \eta_{\mu\nu} (\partial\pi)^4]. \quad (41)$$

[In computing the first line, which comes from eq. (6), care has to be taken in varying the \square , because it contains a covariant derivative. However this is straightforward to do if we rewrite $\square\pi$ as $\frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} g^{\mu\nu} \partial_\nu \pi)$.]

In particular, the trace $T^{(3)\mu}{}_\mu$ computed on the deSitter solution $e^{2\pi_{\text{dS}}} = 1/H_0^2 t^2$ is

$$T^{(3)\mu}{}_\mu = \frac{3}{t^4}. \quad (42)$$

Now we can subtract the \mathcal{L}_3 contribution from the action, and apply eq. (34) to the remainder. The first term in eq. (34), which would vanish on-shell for the full action, now can be rewritten as minus the e.o.m. contribution coming from \mathcal{L}_3 . Combining everything we thus get

$$T^\mu{}_\mu = c_3 T^{(3)\mu}{}_\mu - c_3 \frac{\delta \mathcal{L}_3}{\delta \pi} - \frac{3}{H_0^2 t^4} \frac{dF^{(-3)}(R_{\text{dS}}{}^\mu{}_\nu)}{dH_0^2}, \quad (43)$$

where the last term is the same as eq. (38), applied to the non- \mathcal{L}_3 part of the action. As to second term, for the deSitter solution it is

$$-c_3 \frac{\delta \mathcal{L}_3}{\delta \pi} = -c_3 [(\square\pi)^2 - \partial_\mu \partial_\nu \pi \partial^\mu \partial^\nu \pi + \partial^\mu (\partial_\mu \pi (\partial\pi)^2)] = -c_3 \frac{3}{t^4}, \quad (44)$$

which exactly cancels the first term. In conclusion, eq. (38) holds on-shell if we restrict it to the non- \mathcal{L}_3 part of the action:

$$T^\mu{}_\mu = -\frac{3}{H_0^2 t^4} \frac{dF^{(-3)}(R_{\text{dS}}{}^\mu{}_\nu)}{dH_0^2} \quad (45)$$

2.3 The recipe

We are now in a position to check whether there exists a choice for the Lagrangian coefficients c_2, \dots, c_5 that makes our deSitter solution (13) violate the NEC without giving up stability. (As we stressed, absence of superluminality is automatic for this particular solution.) The various contributions to

$$F^{(-3)} = c_2 \mathcal{L}_2 + c_4 \mathcal{L}_4 + c_5 \mathcal{L}_5 \quad (46)$$

computed on the deSitter solution are [4]

$$\mathcal{L}_2 = -\frac{1}{12}R = -H_0^2 \quad (47)$$

$$\mathcal{L}_4 = \frac{1}{24} \left[\frac{7}{36} R^3 + R(R^\mu{}_\nu)^2 - (R^\mu{}_\nu)^3 \right] = \frac{55}{2} H_0^6 \quad (48)$$

$$\mathcal{L}_5 = \frac{1}{16} \left[\frac{93}{2 \cdot 6^4} R^4 - \frac{39}{4 \cdot 6^2} R^2 (R^\mu{}_\nu)^2 + \frac{5}{12} R (R^\mu{}_\nu)^3 + \frac{3}{16} ((R^\mu{}_\nu)^2)^2 - \frac{3}{8} (R^\mu{}_\nu)^4 \right] = \frac{3}{32} H_0^8 . \quad (49)$$

We thus get

$$p = \frac{1}{3} T^\mu{}_\mu = -\frac{1}{H^2 t^4} \left[-c_2 + \frac{3 \cdot 55}{2} c_4 H_0^4 + \frac{3}{8} c_5 H_0^6 \right] \quad (50)$$

$$= -\frac{1}{H_0^2 t^4} \left[-\frac{1}{2} c_2 - \frac{3}{4} c_3 H_0^2 + \frac{333}{4} c_4 H_0^4 \right] , \quad (51)$$

where in the second step we used the e.o.m. (15). To violate the NEC, since the energy density vanishes on the de Sitter solution, it is enough to have negative pressure, that is

$$-\frac{1}{2} c_2 - \frac{3}{4} c_3 H_0^2 + \frac{333}{4} c_4 H_0^4 > 0 . \quad (52)$$

On the other hand, for the de Sitter solution to be stable, we need (21), that is

$$-3c_2 + 3c_3 H_0^2 - \frac{3}{2} c_4 H_0^4 > 0 , \quad (53)$$

where we used the e.o.m. (15). Notice that the e.o.m. (15) is the only one of our conditions where c_5 appears. We can thus think of it as determining c_5 for given H_0^2 and c_2, c_3, c_4 . Furthermore, we can eliminate H_0^2 by absorbing it into a redefinition of c_2, c_3, c_4 . That is, for the moment we can work in $H_0^2 = 1$ units and reintroduce H_0^2 later by dimensional analysis. Then, combining eqs. (52) and (53) we get a condition on c_4 :

$$\frac{1}{333} (2c_2 + 3c_3) < c_4 < 2(c_3 - c_2) . \quad (54)$$

This interval exists if and only if

$$c_3 > \frac{2}{3} \cdot \frac{334}{221} c_2 . \quad (55)$$

So, summarizing, here is the recipe for a system that has a stable NEC-violating, de Sitter-like solution:

1. Pick $H_0^2 > 0$ for the deSitter-like solution;
2. Pick c_2 ;
3. Pick $c_3 > \frac{2}{3} \cdot \frac{334}{221} c_2 / H_0^2$;
4. Pick c_4 in the interval $\frac{1}{333} (2c_2 + 3c_3 H_0^2) < c_4 H_0^4 < 2(c_3 H_0^2 - c_2)$;
5. c_5 is fixed by the e.o.m.: $c_5 H_0^6 = -\frac{2}{3} (-2c_2 + 3c_3 H_0^2 - 3c_4 H_0^4)$;
6. The desired action is $S_\pi = \int d^4x c_2 \mathcal{L}_2 + \dots + c_5 \mathcal{L}_5$.

Notice that if we want the trivial configuration $\pi = 0$ to be stable, so that we can think of our theory as a standard relativistic QFT constructed about a (perturbatively) stable Poincaré invariant vacuum state, then at point 2 above we need $c_2 > 0$. This is compatible with everything else, however.

3 The quantum theory

So far we focused on classical aspects of our solution. However we also have to impose that it make sense quantum-mechanically—that is, that it be within the regime of validity of our effective theory. At first sight, this is not difficult to achieve. The constraints derived in the previous section are linear in $c_n H_0^{2(n-2)}$, and thus can be generically satisfied by choosing

$$c_2 \equiv f^2, \quad c_{n>2} = a_n \frac{f^2}{H_0^{2(n-2)}}. \quad (56)$$

with arbitrary f^2 and with $a_n = O(1)$. Now, f^2 appears as an overall multiplicative constant in the Lagrangian, and therefore its value cannot affect the classical solutions. On the other hand, quantum-mechanically it formally plays the role of $1/\hbar$. The bigger f^2 , the smaller the quantum mechanical effects. Indeed this is made manifest by going to canonical normalization for π : $\pi_c \equiv f \pi$. Then all self-interactions for π_c are suppressed by powers of H_0 and of f . By keeping H_0 fixed and taking $f \gg H_0$, we can make quantum effects arbitrarily weak while keeping a given classical solution unchanged.

Still, the limit $f \gg H_0$ entails not quite an inconsistency—but a peculiarity of our model. The problem lies in the softness of $2 \rightarrow 2$ scattering amplitudes at low energies, as we now explain. It has been pointed in ref. [7], that for an effective theory to be the low-energy limit of a relativistic microscopic theory obeying the standard properties of the S -matrix (analyticity, unitarity, Froissart bound), the $2 \rightarrow 2$ forward scattering amplitude for identical particles cannot go to zero faster than s^2 at low energies—and the coefficient of the s^2 term has to be positive. Now, the galileon grossly violates this, before the conformal completion we adopted. This is because each galilean interaction involves too many derivatives per field (see eqs. (6, 7)). And in fact at tree-level the $2 \rightarrow 2$ amplitude vanishes exactly in the forward limit ($t \rightarrow 0$)—by power counting it should behave like s^3 , but crossing symmetry in the forward limit implies parity under $s \rightarrow -s$.

The conformal completion can in principle fix this, because it supplements the galileon Lagrangian with interaction terms with fewer derivatives per field [4]. For instance with \mathcal{L}_3 there is an associated $(\partial\pi)^4$ term, which has exactly the right structure to yield an s^2 contribution to the scattering amplitude. On the other hand, the conformal version of \mathcal{L}_4 has too many derivatives, and that of \mathcal{L}_5 has too many π 's (as well as too many derivatives), so we can neglect them. So, the part of the Lagrangian that can contribute an s^2 piece to the forward scattering amplitude is (see eqs. (11, 12))

$$\begin{aligned} \mathcal{L} &\supset c_2 \mathcal{L}_2 + c_3 \mathcal{L}_3 \\ &= -\frac{1}{2} c_2 e^{2\pi} (\partial\pi)^2 - \frac{1}{2} c_3 \square\pi (\partial\pi)^2 - \frac{1}{4} c_3 (\partial\pi)^4 \\ &\rightarrow -\frac{1}{2} c_2 (\partial\tilde{\pi})^2 + \frac{1}{4} c_3 (\partial\tilde{\pi})^4 + \dots, \end{aligned} \quad (57)$$

where we got the last line after suitable non-linear field redefinitions, and the dots stand for higher orders in $\tilde{\pi}$ and derivatives thereof.

We thus see that with our conformally invariant Lagrangian, we can have a positive s^2 term in the low-energy forward $2 \rightarrow 2$ scattering amplitude,

$$\mathcal{M}(s, t \rightarrow 0) \sim \frac{c_3}{c_2} s^2 + \dots, \quad (58)$$

for positive c_3 . Notice that $c_3 > 0$ is consistent with our ‘recipe’ of sect. 2.3. The strong coupling scale associated with this term in the amplitude is

$$\Lambda = (c_2^2/c_3)^{1/4} \sim (fH_0)^{1/2}. \quad (59)$$

However, the strong coupling scales associated with higher-order terms can be lower. For instance the cubic interaction $\square\pi(\partial\pi)^2$ used twice in two trilinear vertices connected by a propagator yields an $s^3 + t^3 + u^3$ term (which vanishes in the forward limit but is non-zero otherwise) suppressed by the scale

$$\Lambda' = (c_2^3/c_3^2)^{1/6} \sim (fH_0^2)^{1/3}. \quad (60)$$

This is much lower than Λ for $f \gg H_0$. More systematically, expanding the tree-level Lagrangian in powers of the canonically normalized field π_c , we get interactions of the form

$$\mathcal{L}_{n,m,q} \sim \partial\pi_c \partial\pi_c (\pi_c/f)^n (\partial\pi_c/H_0 f)^{2m} (\partial\partial\pi_c/H_0^2 f)^q, \quad (61)$$

with $1 \leq m+q \leq 3$ and arbitrary n , except for $m=0, q=1$, in which case $n=0$. The associated strong-coupling scales are

$$\Lambda_{n,m,q} = H_0 (f/H_0)^{\frac{n+2m+q}{n+4m+3q}}. \quad (62)$$

For $f \gg H_0$ the lowest of such scales corresponds to $n=m=0$,

$$\Lambda_{\text{strong}} \equiv \Lambda_{0,0,q} = H_0 (f/H_0)^{1/3}, \quad (63)$$

which is exactly the scale Λ' of eq. (60).

Now, the fact that the leading low-energy term in the amplitude is weighed by a scale that is much higher than that weighing higher-order terms is not inconsistent, but suggests that the UV-completion of the model, if it exists, is weakly coupled. In such a case, the different strong-coupling scales $\Lambda_{n,m,q}$ do not correspond to any physical scales in the theory—the theory never becomes strongly coupled—they are just low-energy mirages arising from taking different combinations of a weak coupling and of the true new physics mass-scale of the theory. This is evident for instance in the tree-level exchange of a weakly coupled massive particle at low energies:

$$g^2 \frac{1}{p^2 + M^2} \rightarrow \frac{1}{(M/g)^2} - \frac{p^2}{(M/g^{1/2})^4} + \frac{p^4}{(M/g^{1/3})^6} + \dots, \quad p^2 \ll M^2 \quad (64)$$

(as we will see our case suggests a slightly subtler interpretation than this simple example). Now, because of scale invariance (9) and of eq. (56), the structure of our Lagrangian is

$$\mathcal{L} = f^2 e^{2\pi} \partial\pi \partial\pi F(e^{-\pi} \partial\pi/H_0, e^{-2\pi} \partial\partial\pi/H_0^2) \quad (65)$$

where the function F is dimensionless, and has no small parameter. Then, it is natural to conjecture that the derivative expansion be controlled by the scale

$$\phi(x) \equiv e^{\pi(x)} H_0. \quad (66)$$

That is, if we rewrite our classical Lagrangian as

$$\mathcal{L} = \frac{1}{g^2} \cdot \phi^4 \tilde{F}(\partial\phi/\phi^2, \partial\partial\phi/\phi^3), \quad g^2 \equiv H_0^2/f^2, \quad (67)$$

it is natural to interpret g as a coupling constant—weak, in the limit $f^2 \gg H_0^2$ —and ϕ as a dilaton controlling the mass-scale where the UV-completion enters. If this is correct, then the weakly coupled limit $g \ll 1$ indeed has formally very high strong coupling scale, but this says nothing about where the effective theory is actually breaking down. New physics has to enter at the scale ϕ . For the trivial solution $\pi = 0$ we have $\phi = H_0$, and the effective theory is valid at energies below this. For our deSitter solution (13) we have

$$\phi_{\text{dS}} = \frac{1}{|t|}, \quad (68)$$

and the solution itself is outside the effective theory, in the sense that all terms in \tilde{F} above are of the same order, thus signaling a breakdown of the derivative expansion.

We want to stress that, in fact, we are not forced to this conclusion. If \tilde{F} were an infinite series—the full derivative expansion—then indeed identifying the UV-completion’s mass scale would be unambiguous. For instance one could compute the $2 \rightarrow 2$ tree-level scattering amplitude at all orders in s/ϕ^2 about a $\phi = \text{const}$ solution. Then the radius of convergence of that series—the distance of the first non-analyticities from the origin—would define the scale of new physics. Perhaps more physically, the strong-coupling scale associated with a very high-order term would be very close to ϕ , the extra factor of g being diluted by a tiny fractional power (see eq. (64)). In our case instead \tilde{F} is a finite polynomial, and the identification of the UV-scale is not as straightforward. For $g \ll 1$ quantum effects are small, and the effective theory is unitary up to energies of order of the strong-coupling scale $\Lambda_{\text{strong}} = H_0/g^{1/3}$.

We can however put an upper bound on the UV-completion’s scale by means of dispersion relations, and such a bound turns out to be parameterically smaller than the strong coupling scale—in agreement with the idea that the UV-completion be weakly coupled. If we assume analyticity, unitarity, and that the forward $2 \rightarrow 2$ scattering amplitude $\mathcal{A}(s) \equiv \mathcal{M}(s, t \rightarrow 0)$ is bounded by s^2 at large s , then we have [7]

$$\mathcal{A}''(0) = \frac{2}{\pi} \int_0^\infty \frac{\sigma(s)}{s^2} ds \quad (69)$$

where $\sigma(s)$ is the total cross-section for $\pi\pi$ scattering. With our effective theory we can compute the left-hand side, and the low-energy contribution to the right-hand one. Since the integrand is positive, this equation then yields an upper bound Λ_{UV} on the scale where we should stop trusting our low-energy computations—the cutoff of the effective theory. About the trivial solution $\pi = 0$, $\phi = H_0$, the l.h.s. can be got from eq. (58):

$$\mathcal{A}''(0) \sim \frac{c_3}{c_2^2} = \frac{g^2}{H_0^4}. \quad (70)$$

On the other hand the low-energy cross-section is UV-dominated (as befits a derivatively coupled theory), and in particular at energies above $\Lambda_{0,0,1}^3/\Lambda_{0,1,0}^2 = H_0$ it is dominated by the s^3 piece in the amplitude, thus scaling as

$$\sigma(s) \sim g^4 \frac{s^5}{H_0^{12}}. \quad (71)$$

We thus see that in order not to violate eq. (69), the low-energy contribution has to be cut off below

$$\Lambda_{\text{UV}} = \frac{H_0}{g^{1/4}}. \quad (72)$$

This defines the highest possible energy scale our effective theory can be extrapolated to (under the assumptions given above). The UV-completion has to enter below this scale. Notice that Λ_{UV} is parameterically higher than $\phi = H_0$, which was the cutoff scale ‘naturally’ suggested by the structure of eq. (67). This makes our deSitter-like solution consistent with the standard S -matrix properties of well behaved relativistic theories. Notice also that Λ_{UV} is lower than the strong-coupling scale,

$$\Lambda_{\text{strong}} = \frac{H_0}{g^{1/3}}, \quad (73)$$

as it should be for the consistency of our estimate.

In general there may be infinitely many more dispersion relations that can constrain further the structure of the low-energy effective Lagrangian, and set an upper-bound on its cutoff lower than eq. (72). For instance one could consider $n \rightarrow n$ scattering amplitudes with n larger than two. Such constraints are going to be quite cumbersome to work out, and are beyond the scope of the present paper. On the other hand, in the case of weakly coupled UV-completion further information can be extracted from the forward $2 \rightarrow 2$ amplitude $\mathcal{A}(s)$ [7]. We will not take this route either—in fact it is straightforward to show that in our case nothing new is learned on the cutoff of the theory by taking it. Instead, we will consider signal propagation in the presence of weak background fields. This will turn out to be surprisingly powerful.

4 Retaining causality

Ref. [7] identified a class of derivatively coupled theories for which S -matrix positivity along the lines of last section implies the sub-luminality of excitations in the presence of non-trivial background fields. In our case, as we will see in a moment, quite the opposite is true: amplitude positivity forces c_3 to be positive, which in turn makes certain excitations about background fields *super*-luminal, unless the cutoff of the theory is low enough. We will use this to constrain the cutoff of our effective theory. Notice that the analyses of refs. [7, 4] already exhibited the presence of super-luminal excitations for DGP-like and galileon theories about certain background solutions. We find it useful however to provide a self-contained similar analysis here, mainly for two reasons. The first is that we will exhibit superluminality for essentially *any* static weak-field solution to the field equations—which eliminates the possibility that perhaps superluminality about a particular solution just signals that the field cannot be coupled to sources in such a way as to produce that

solution. The second reason is that we will show that superluminality is gone if the UV cutoff of the theory is of order $\langle\phi\rangle$.

To study the propagation of fluctuations about a background solution, it is more convenient to use $\phi = H_0 e^\pi$ as field variable. The Lagrangian becomes (see eq. (57))

$$\mathcal{L} = \frac{1}{g^2} \left[-\frac{1}{2}(\partial\phi)^2 - \frac{\tilde{c}_3}{2} \frac{\square\phi(\partial\phi)^2}{\phi^3} + \frac{\tilde{c}_3}{4} \frac{(\partial\phi)^4}{\phi^4} + \dots \right], \quad (74)$$

where $g = H_0/f$, $\tilde{c}_3 \equiv c_3 H_0^2/f^2$, and the dots stand for higher-order terms, coming from \mathcal{L}_4 and \mathcal{L}_5 . As clear from last section's discussion, the dimensionless coefficient \tilde{c}_3 has to be positive, and of order unity. Now, consider a background solution $\phi_0(x)$, and small fluctuations φ about it. At quadratic order the fluctuation Lagrangian is

$$\begin{aligned} \mathcal{L}_\varphi &= \frac{1}{g^2} \left[-\frac{1}{2}(\partial\varphi)^2 \left(1 + \tilde{c}_3 \frac{\square\phi_0}{\phi_0^3} - \tilde{c}_3 \frac{(\partial\phi_0)^2}{\phi_0^4} \right) + \tilde{c}_3 \left(\frac{\partial^\mu\partial^\nu\phi_0}{\phi_0^3} - 2\frac{\partial^\mu\phi_0\partial^\nu\phi_0}{\phi_0^4} \right) \partial_\mu\varphi\partial_\nu\varphi + \dots \right] \\ &\equiv -\frac{1}{2g^2} G^{\mu\nu} \partial_\mu\varphi\partial_\nu\varphi + \dots, \end{aligned} \quad (75)$$

where the dots stand for non-derivative terms (i.e. mass terms) for φ , as well as for contributions from the terms already neglected in eq. (74), and we defined the effective inverse metric $G^{\mu\nu}$ as

$$G^{\mu\nu} = \eta^{\mu\nu} \left(1 + \tilde{c}_3 \frac{\square\phi_0}{\phi_0^3} - \tilde{c}_3 \frac{(\partial\phi_0)^2}{\phi_0^4} \right) - 2\tilde{c}_3 \left(\frac{\partial^\mu\partial^\nu\phi_0}{\phi_0^3} - 2\frac{\partial^\mu\phi_0\partial^\nu\phi_0}{\phi_0^4} \right) \quad (76)$$

The reason why we focus on the kinetic action for φ and neglect non-derivative terms, is that the causal structure is defined by the highest-derivative terms: according to Leray's theorem [8], for a Lagrangian of the form (75), φ excitations travel along the light-cone of the effective metric $G_{\mu\nu}$.

We are thus led to analyze the light-cone of the effective metric (76) in the presence of a non-trivial $\phi_0(x)$. To be as general as possible, we consider a generic weak vacuum solution $\phi_0(x)$. By 'weak' we mean that the non-linearities of our action are negligible. Also, for simplicity we restrict to a static configuration, so that the solution obeys

$$\nabla^2\phi_0 \simeq 0. \quad (77)$$

The physical situation we have in mind is that we have stationary, localized sources for ϕ , and that away from the sources the field is in its linear regime. We generically expect there to be a stationary point for ϕ_0 ⁴. We zoom in on that point, and choose it as the origin of our coordinate system. The solution can be Taylor expanded about the origin

$$\phi_0(\vec{x}) \simeq \Phi_0 \left(1 + \frac{1}{2} A_{ij} \frac{x^i x^j}{L^2} \right), \quad (78)$$

⁴This is guaranteed if the source arrangement is minimally inventive. Also such an assumption is not necessary to proceed, but makes the discussion neater. Indeed a space gradient term, corresponding to a term linear in x^j in eq. (78), gives just a negative definite contribution to G_{ij} in eq. (80), making the light cone even wider.

where Φ_0 is the field value at the origin, L is the typical scale over which the solution varies, and A_{ij} is a dimensionless matrix with entries of order unity and which, because of Laplace's equation (77), is traceless. As we will see, this last property is crucial for our argument.

Now, to impose, as we assumed, that the solution be in a linear regime, we have to demand

$$L \gg 1/\Phi_0 . \quad (79)$$

Indeed the size of non-linearities in eq. (74) for the terms explicitly kept is of order ∂^2/ϕ^2 , which for a configuration like (78) is precisely $1/(L\Phi_0)^2$. The neglected terms in eq. (74) are suppressed by even higher powers of $\partial^2/\phi^2 \sim 1/(L\Phi_0)^2$ —that is why we are allowed to neglect them in the first place. Not surprisingly, the size of non-linearities $1/(L\Phi_0)^2$ is also controlling the corrections to $\eta^{\mu\nu}$ in the inverse metric (76). As long as this is small, the latter can be straightforwardly inverted to yield the effective metric

$$G_{\mu\nu} \simeq \eta_{\mu\nu} + 2\tilde{c}_3 \left(\frac{\partial_\mu \partial_\nu \phi_0}{\phi_0^3} - 2 \frac{\partial_\mu \phi_0 \partial_\nu \phi_0}{\phi_0^4} \right) , \quad (80)$$

where we neglected the correction proportional to $\eta_{\mu\nu}$, because it does not affect the light-cone aperture. That is, at this order we can factor it out of the metric, as an overall conformal factor, which is irrelevant for our question.

To check whether the light-cone of $G_{\mu\nu}$ includes directions that are super-luminal with respect to the true Minkowski light-cone defined by $\eta_{\mu\nu}$, we contract $G_{\mu\nu}$ with a generic null direction $n^\mu = (1, \hat{n})$. If n^μ turns out to be time-like for $G_{\mu\nu}$ at least for some direction \hat{n} , we know that φ signals can travel superluminally along \hat{n} . For the moment we consider small distances from the origin, $x \ll L$, so that the denominators in (80) can be taken as constant for our solution. We have

$$G_{\mu\nu} n^\mu n^\nu \simeq \frac{2\tilde{c}_3}{(L\Phi_0)^2} \left[A_{ij} - 2 \frac{(A_{ik}x^k)(A_{jl}x^l)}{L^2} \right] \hat{n}^i \hat{n}^j . \quad (81)$$

The second term in brackets is always a negative definite matrix, being minus the tensor product of $A \cdot \vec{x}$ with itself. It thus gives a superluminal contribution in all directions—that is, it tends to make n^μ time-like w.r.t. $G_{\mu\nu}$ regardless of \hat{n} . However at small distances from the origin, $x \ll L$, the second term is negligible with respect to the first. Now, A_{ij} , being traceless, has at least one negative eigenvalue. The corresponding eigenvector, defines a direction \hat{n} for which the first term also gives a superluminal shift in the propagation speed, of order

$$\Delta c \sim \frac{1}{(L\Phi_0)^2} \quad (82)$$

(recall that both \tilde{c}_3 and A_{ij} are of order one.) We are thus led to the conclusion that a generic linear solution $\phi_0(x)$ allows for superluminal excitations, because the effective metric perturbations propagate in has a light-cone that is broader than that defined by $\eta_{\mu\nu}$. We stress that this result does not depend on the detailed form of the coupling of ϕ to each individual source. For instance on whether it couples through a ‘monopole’ ϕJ , or via dipole $\partial_\mu \phi J^\mu$ or even via higher multipoles. This generalizes the proof of superluminality given in ref. [4] by considering spherical solutions induced by monopole coupling.

The one we just obtained is a purely classical result: at the classical level, to ascertain the existence of superluminal signals, it is enough to analyze the effective metric (Leray’s theorem). Upon canonical quantization the causal structure does not change [9], but since we have an effective theory with a cutoff, to be consistent we have to demand that the superluminal effect we isolated be measurable within the effective theory. The superluminal shift (82) in the propagation speed entails superluminal propagation of actual signals if, for a mode of given momentum k propagating from the origin all the way to L , the gain over a strictly luminal signal is at least of order of one wavelength. That is, we need

$$\Delta c k L \sim \frac{k}{\Phi_0} \frac{1}{L \Phi_0} \gtrsim 1. \quad (83)$$

We are restricting to a region of size L around the origin, because that is where our approximations hold. We will comment below on what happens outside this region. Recall that for the solution to be linear, we want $L \Phi_0$ larger than one. We thus see that the superluminal effect is unobservable if we restrict to fluctuations with momenta smaller than Φ_0 . Therefore, if we postulate that the cutoff of our effective theory is the local value of ϕ , like the structure of our Lagrangian (67) naturally suggested, we cannot consistently conclude that the causal structure of the theory is different than the usual relativistic one.

Nothing new is learned by relaxing the $x \ll L$ restriction. For a generic linear solution, with a generic arrangement of sources, the structure of the solution will deviate significantly from eq. (78) beyond $x \sim L$, and full knowledge of the solution will be needed to conclude anything about signal propagation. However since eq. (78) is a good vacuum solution for $x \gg L$ as well—non-linear effects are down by powers of $L^4/(\Phi_0^2 x^6)$ —one could study how signals propagate on it for distances much longer than L . The problem is that the superluminal effect is proportional to the leading non-linearity, and so the former also rapidly goes to zero at very large x . That is, for a solution like eq. (78) the superluminal effect is concentrated in a region of size L around the origin.

In summary, if we assume that our effective theory arises as the low-energy limit of a microscopic theory with the standard relativistic causal structure, we are *forced* to bound its regime of validity by the local value of ϕ . Under those assumptions our stable NEC-violating solution of sect. 2 is therefore untenable.

5 Back to the amplitude

In hindsight, it is natural to ask whether the S -matrix encodes any signs of the superluminal effect we unveiled. In simpler examples, like those considered in ref. [7], sub-luminality of excitations about background fields is tied to positivity of forward $2 \rightarrow 2$ amplitudes at order s^2 . In our case instead, as we discussed, such positivity is obeyed so long as c_3 is positive, yet we have super-luminality regardless of c_3 ’s sign. In fact, our super-luminal effect relies on the cubic interaction $\square\pi(\partial\pi)^2$ whose contributes to $2 \rightarrow 2$ amplitudes is of order s^3 and vanishes in the forward limit, being proportional to $s^3 + t^3 + u^3$. It thus looks like that in order to isolate the term responsible for super-luminality, we need to consider non-forward amplitudes.

Some form of positivity still holds for non-forward $2 \rightarrow 2$ amplitudes $\mathcal{M}(s, t)$. Indeed following

Martin⁵ [10], unitarity together with standard properties of Legendre polynomials imply that for real s

$$\frac{\partial^n}{\partial t^n} \text{Im } \mathcal{M}(s + i\epsilon, 0) \geq 0 \quad \forall n, \quad (84)$$

which is a generalization of the optical theorem, $\text{Im } \mathcal{M}(s + i\epsilon, 0) \propto \sigma_{\text{tot}} > 0$. If there is a mass-gap, $\mathcal{M}(s + i\epsilon, t)$ is analytic in the t -plane, for $|t|$ below threshold⁶. This means that we can Taylor-expand $\mathcal{M}(s + i\epsilon, t)$ in t about the origin, and use the inequalities (84) to conclude

$$\text{Im } \mathcal{M}(s + i\epsilon, t) > 0 \quad \text{for } t > 0, \quad (85)$$

so long as t (now real and positive) is below threshold. We can then run a standard dispersion argument (see e.g. [7])—this time using analyticity in the s -plane—and show that

$$\frac{\partial^2}{\partial s^2} \mathcal{M}(0, t) = \frac{2}{\pi} \int ds \text{Im } \mathcal{M}(s + i\epsilon, t) \left[\frac{1}{s^3} + \frac{1}{(s+t)^3} \right] > 0 \quad \text{for } t > 0, \quad (86)$$

where the integral is taken along the cut on the positive real axis, from the threshold to infinity.

Now, in our case there is no mass-gap—the particles we are scattering are themselves massless—and we have non-analyticities at arbitrarily low s and t . However, we already argued that the UV-completion of our model must be weakly coupled, with coupling $\lambda \ll 1$. We can then consider the scattering amplitude at tree-level in λ . Such an amplitude has non-analyticities associated with the exchange of the UV-completion’s massive states, but it does not have low-energy cuts associated with multi-massless particle states—these enter at higher order in λ [7]. It thus makes sense to apply the above dispersion relation to the tree-level amplitude, even in the absence of a mass gap. The mass-scale playing the role of the gap is then the UV cutoff of the effective theory, where the UV-completion enters and the first non-analyticities show up.

Now, in our case we assumed that the cutoff of the theory is parameterically higher than H_0 , perhaps as high as eq. (72). At intermediate energies, $H_0 \ll E \ll \Lambda_{\text{UV}}$, the amplitude is dominated by the $\mathcal{O}(s^3)$ terms, and we can neglect the $\mathcal{O}(s^2)$ ones. The former gets a contribution from the Lagrangian term we are trying to isolate, $\square\pi(\partial\pi)^2$, but also directly from the quartic galilean interaction contained in \mathcal{L}_4 . Indeed we can perform a non-linear field redefinition that eliminates the cubic interaction and yields a quartic term of the form

$$\partial\pi\partial\pi\partial\partial\pi\partial\partial\pi, \quad (87)$$

which is precisely the structure of \mathcal{L}_4 . Therefore, unless the \mathcal{L}_4 contribution accidentally vanishes, which it does not, we already see that our attempt to connect super-luminality with violation of amplitude positivity, fails: amplitude positivity can be fixed by a suitable choice of c_4 , which as we argued does not affect signal propagation in the presence of weak background fields. However, if we want our Lagrangian to describe NEC-violating stable solutions, c_4 is not completely arbitrary. It may be that once we impose the conditions of sect. 2.3, positivity turns out to be violated.

⁵For a recent application see also [11].

⁶More precisely, as proven in ref. [10], analyticity in (s, t) holds on the product space $(|t| < M) \times (\text{cut } s \text{ plane})$.

Computing the $\mathcal{O}(s^3)$ tree-level amplitude is straightforward. The result is

$$\mathcal{M}(s, t) = \frac{3}{4} \frac{1}{c_3^2} (c_3^2 - 2c_2c_4) st(s+t). \quad (88)$$

Positivity in the form of eq. (86) is obeyed iff

$$c_3^2 > 2c_2c_4. \quad (89)$$

Such an inequality is in fact *implied* by our conditions of sect. 2.3, as manifest from the following simple algebraic manipulation:

$$0 \leq (c_3 - 2c_2/H_0^2)^2 = c_3^2 - 4(c_3c_2/H_0^2 - c_2^2/H_0^4) < c_3^2 - 2c_2c_4, \quad (90)$$

where in the last step we used item 4. of sect. 2.3.

In conclusion, by assuming that our effective theory can be extrapolated to energies parameterically higher than H_0 we find no contradiction with amplitude positivity, even in the non-forward case. Conversely, absence of superluminal excitations demands the theory to be cut off at H_0 .

6 Discussion

Our effective theory allows for an isotropic NEC-violating solution that is stable against small fluctuations. Moreover, signal propagation on this solution is exactly luminal. This is a nice counterexample to the theorem of ref. [1], which here is violated because one of the hypotheses, subtly, is. The relevant part of the Lagrangian formally involves higher-derivative terms, yet the corresponding field equations are second-order.

Still, if the goal was to find a completely sensible relativistic QFT that can violate the NEC in the semi-classical regime, then we failed. From the standard EFT standpoint, our theory *is* perfectly sensible, and the relevant NEC solution is within its regime of validity. Moreover, our theory obeys standard positivity of $2 \rightarrow 2$ amplitudes, even in the non-forward case—which is a stronger requirement than purely low-energy consistency [7]. Nevertheless, if we start from the vacuum and excite a generic weak field configuration, small excitations about this background will now propagate super-luminally. The presence of super-luminal excitations signals that the causal structure of the theory is not the standard Lorentz-invariant one. As a consequence our effective theory cannot arise as the low energy limit of a Lorentz-invariant UV-complete local QFT, for which micro-causality has to hold as an exact (i.e., non-perturbative) operator statement [7, 9].

Now, we can avoid super-luminality if the cutoff of our theory is in fact much lower than expected. However for such a low cutoff, our NEC-violating solution is outside the effective theory and thus cannot be trusted. That is: either *(i)* we believe that the cutoff is high enough to make room for the NEC-violating solution, but then we have to accept the existence of super-luminal excitations about other solutions, also inside the effective theory; or *(ii)* we believe that the cutoff is low enough to forbid superluminal propagation of signals, but then our NEC-violating solution is gone as well. Essentially, absence of superluminal excitations demands that all non-linearities be small: if the cutoff of the theory is the local value of ϕ , as required by sub-luminality, then all non-linear terms in eq. (67) have to be small for the solution to be within the effective theory, despite

what suggested by purely low-energy EFT considerations and even by dispersion relations. This makes the tension we found particularly robust. It is not an accident of our particular solution, but rather it is a property of any solutions that crucially rely on non-linearities. Since the NEC is obeyed by the free part of the Lagrangian, *all* NEC-violating classical solutions are outside the regime of the effective theory.

We thus see that there still is a strong tie linking the NEC and subluminality. It is less direct than in the absence of higher-derivative terms, where it is the NEC-violating solutions themselves that support superluminal excitations [1], but it appears to be equally tight. It would be interesting however to understand whether this relation is an accident of our Lagrangian, or whether instead it is a more general, and thus deeper, fact. For instance, we adopted a conformal modification of the galileon theory in order to respect positivity at order s^2 in the scattering amplitude. However this requirement by itself does not single out the conformal group $SO(4, 2)$ as the symmetry group of our theory. We chose it because then with such a high degree of symmetry, finding very symmetric solutions and studying their properties is immediate. However there are probably other equally convenient choices—for instance the 5D Poincaré group $ISO(4, 1)$ [4]—and in principle nothing stops us from lowering the number of symmetry generators, as long as galilean invariance is an approximate symmetry in the UV. Note that for superluminality to happen, one just needs a non-vanishing galilean cubic term. In our theory c_3 has to be strictly positive for the NEC-violating solution to be well behaved (sect. 2.3) and for the forward amplitude to be positive at order s^2 (sect. 3). It is entirely possible that disentangling the coefficient of the cubic galilean term from others' could eliminate superluminality altogether while retaining the good features of our theory as well as of the NEC-violating solution.

From a more general field theoretical perspective, we find it interesting that in our theory relativistic dispersion relations involving the $2 \rightarrow 2$ scattering amplitude are not enough to discover superluminality in the presence of sources. In the simple cases of ref. [7], the link was so direct that forward dispersion relations were enough to imply sub-luminality about any background solution. In our case instead, even *non-forward* ones are obeyed by a theory with superluminal excitations. We expect that at some level of complexity (for instance one could consider $n \rightarrow n$ scattering amplitudes, with $n > 2$) dispersion relations should encompass the requirement of subluminality about any background fields. Indeed the sources turning on these backgrounds may be viewed as fictitious ones—used to construct the generating functional $W[J]$ from which the n -point functions $\langle \phi(x_1) \dots \phi(x_n) \rangle$ are then straightforwardly derived. It is clear then, that the propagation properties of ϕ fluctuations about some background field *are* encoded in these n -point functions, and thus in the S -matrix elements. It would be interesting to understand for a generic theory the S -matrix equivalent of the sub-luminality requirement.

Acknowledgements

A.N. and E.T. would like to thank the Institute of Theoretical Physics of the Ecole Polytechnique Fédérale de Lausanne for hospitality during this project. The work of R.R. is partially supported by the Swiss National Science Foundation under contract No. 200021-116372 and 200021-125237.

References

- [1] S. Dubovsky, T. Gregoire, A. Nicolis and R. Rattazzi, “Null energy condition and superluminal propagation,” *JHEP* **0603**, 025 (2006) [arXiv:hep-th/0512260].
- [2] N. Arkani-Hamed, H. C. Cheng, M. A. Luty and S. Mukohyama, “Ghost condensation and a consistent infrared modification of gravity,” *JHEP* **0405**, 074 (2004) [arXiv:hep-th/0312099].
- [3] P. Creminelli, M. A. Luty, A. Nicolis and L. Senatore, “Starting the universe: Stable violation of the null energy condition and non-standard cosmologies,” *JHEP* **0612**, 080 (2006) [arXiv:hep-th/0606090].
- [4] A. Nicolis, R. Rattazzi and E. Trincherini, “The galileon as a local modification of gravity,” *Phys. Rev. D* **79**, 064036 (2009) [arXiv:0811.2197 [hep-th]].
- [5] M. A. Luty, M. Porrati and R. Rattazzi, “Strong interactions and stability in the DGP model,” *JHEP* **0309**, 029 (2003) [arXiv:hep-th/0303116].
- [6] A. Nicolis and R. Rattazzi, “Classical and quantum consistency of the DGP model,” *JHEP* **0406**, 059 (2004) [arXiv:hep-th/0404159].
- [7] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, “Causality, analyticity and an IR obstruction to UV completion,” *JHEP* **0610**, 014 (2006) [arXiv:hep-th/0602178].
- [8] See e.g. R. M. Wald, “General Relativity,” *Chicago, Usa: Univ. Pr. (1984) 491p*
- [9] S. Dubovsky, A. Nicolis, E. Trincherini and G. Villadoro, “Microcausality in Curved Space-Time,” *Phys. Rev. D* **77**, 084016 (2008) [arXiv:0709.1483 [hep-th]].
- [10] A. Martin, “Extension of the axiomatic analyticity domain of scattering amplitudes by unitarity. 1,” *Nuovo Cim. A* **42**, 930 (1965).
- [11] L. Vecchi, “Causal vs. analytic constraints on anomalous quartic gauge couplings,” *JHEP* **0711**, 054 (2007) [arXiv:0704.1900 [hep-ph]].