Symmetries from locality. III. Massless spin-2 gravitons and time translations

Mark P. Hertzberg[*] and Jacob A. Litterer[†]

Institute of Cosmology, Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA

(Received 15 August 2020; accepted 27 September 2020; published 13 October 2020)

We relax the assumption of time translation and Lorentz boost symmetry in theories involving massless spin-2 gravitons, while maintaining a basic notion of locality that there is no instantaneous signaling at a distance. We project out longitudinal modes, leaving only 2 degrees of freedom of the graviton. Our previous work, which assumed time translation symmetry, found that the Lorentz boost symmetry is required to ensure locality at leading order. In this work, without assuming time translations or Lorentz boosts, we show that locality of the exchange action between matter sources demands that massless spin-2, at leading order, organizes into Einstein-Hilbert plus a Gauss-Bonnet term with a prefactor that is constrained to be a particular function of time; while in the matter sector we recover time translation and Lorentz boost symmetry. Finally, we comment on whether the time dependence of the Gauss-Bonnet prefactor may be forbidden by going to higher order in the analysis and we mention that other possibilities are anticipated if graviton mass terms are included.

DOI: 10.1103/PhysRevD.102.085007

I. INTRODUCTION

The modern picture of cosmology is remarkably successful. However, it possesses a range of curious features; this includes the need for dark matter, dark energy, coincidences, fine-tuning problems, and, more recently, possible tension with different measurements of the Hubble parameter [1]. It has sometimes been suggested that one or more of these issues may be addressed by modifying or deforming the underlying rules of gravitation, namely general relativity.

But any theoretically sensible deformation of general relativity is difficult to come by. This is because there is a theorem [2–4]: *the unique local, unitary, Poincaré invariant theory of a species of massless spin-2 particles is, at large distances, general relativity.* (Multiple gravitons with higher order interactions can lead to inconsistencies [5–7].) Hence to escape the conclusion that we necessarily have general relativity in a local and unitary way, one must do one (or more) of the following: (i) consider small distances, (ii) add extra degrees of freedom, (iii) deform Poincaré symmetry. The first option (i) is not interesting to resolve puzzles in cosmology, which involve phenomena on the largest of scales; so we shall not pursue this possibility here. In the second option (ii) one may (a) add a mass to the graviton (which takes the number of degrees of freedom from two helicities to five polarizations). However, this possesses various strong coupling problems and no known UV completion (arguably due to problems with causality), so we shall also ignore this possibility here. Or one may (b) add other types of particles, especially spin-0 particles, which can potentially mediate long range forces if they are extremely light. Such interactions are tightly constrained by solar system tests and are unlikely to solve the above conceptual problems anyhow (one can just add scalars as part of the matter sector to act as dark matter of course); so we shall not pursue this possibility here either.

This leaves us with investigating option (iii), which has been investigated in the literature [8–20], and is the focus of this paper. In our previous papers [21–23] we considered deforming Lorentz boosts, while maintaining all other parts of the Poincaré symmetry, namely rotation symmetry, spatial translation symmetry, and time translation symmetry. In [22] we considered spin-1 particles (electromagnetism) and in [23] we considered spin-2 particles (gravitation). By imposing a very basic notion of locality, that there is no instantaneous signaling, we showed that these particles must couple to conserved sources, associated with a symmetry. In the case of spin-1 we showed it must couple to a conserved current, which requires the familiar $U(1)$ symmetry, and for spin-2 we showed it must couple to a conserved 2-index symmetric current (energy-momentum tensor), which requires Lorentz boost symmetry. Since the latter recovers

[*]mark.hertzberg@tufts.edu
[†]jacob.litterer@tufts.edu

Lorentz boosts and in turn the full Poincaré symmetry involving spin-2, then it is anticipated to recover the full structure of general relativity just from locality of spin-2.

For potential cosmological applications it is important to move beyond the above assumptions. Most importantly, cosmology involves an expanding universe, i.e., a background space-time that breaks time translation invariance. While in the context of general relativity this is associated with spontaneous breakdown of time translations, there remains several mysteries as mentioned earlier. Therefore it motivates exploring setups in which time translation symmetry is explicitly broken. This work is a first step in this direction. Our particular setup is as follows: we will assume, again, a preferred frame with rotation invariance (which allows for a notion of particle spin) and translation invariance in space, but we will not assume Lorentz boosts or time translation invariance. As a first investigation into this problem, we will assume the graviton is massless. This may seem a very mild assumption since we are going to project down to only the two polarizations of the spin-2 graviton and we know that Poincaré symmetry plus unitarity would dictate that it would need to be massless. But since we are not assuming Poincaré symmetry in this work, this argument is no longer valid. In fact in our previous paper [23] where we did not assume Lorentz boosts, we allowed for mass terms for the graviton, but went on to show that locality required them to vanish. If we now give up both Lorentz boost and time translation invariance, we can again have mass terms for the graviton, and for the greatest generality we should only use some principle like locality to dictate what terms are allowed. However, it turns out this issue is rather complicated. So, for simplicity, we will assume the graviton is massless here, and leave the inclusion of mass terms for future work. We note that all current data is indeed consistent with our assumption of a massless graviton [24]. However, we know that when expanding around a Friedmann-Robertson-Walker (FRW) background space-time in standard general relativity, mass terms appear; so our work here will not be able to recover perturbations around FRW. Instead we view the present work as a solid starting point to explore the explicit breakdown of time translations, albeit not the most general. To be clear, we are especially interested in explicit breakdown, rather than spontaneous, which is already a well studied problem.

The structure of this paper is similar to the previous papers [22,23], especially to the gravitation paper [23], and we encourage readers to read these accompanying papers. Since many details are already laid out there, our description here will be relatively brief, with a focus on the construction and primary results. Our paper is organized as follows: We first recap the fact that standard general relativity is local. We then discuss a family of generalizations that break time translations and Lorentz boosts. We then focus on classes of theories that cut down to two polarizations of the graviton. In each case we impose locality, defined here as avoiding instantaneous signaling, and find the same resulting theory. We believe this resulting theory is valid regardless of the particular choice of how to cut down the degrees of freedom. We find that we recover general relativity (GR) to the leading order we are working, plus an additional set of terms that involve a very specific form of time translation symmetry violation. We then comment on what is anticipated to occur when working to higher order.

II. THEORIES OF MASSLESS SPIN-2

A. Recap of Poincaré invariant massless spin-2

Our goal in this work is to study and generalize theories of massless spin-2 particles. We will begin with the standard theory of general relativity, defined around a flat background, and then generalize this in a particular fashion that we will describe in the next subsection. So it is useful to recap the structure of general relativity (the unique local, unitary, Poincaré invariant theory of massless spin-2). To study the leading order interactions, we expand the metric $g_{\mu\nu}$ around a flat background as

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}, \tag{1}$$

where $\eta_{\mu\nu}$ is the metric of Minkowski space-time and $h_{\mu\nu}$ is the spin-2 field whose purpose is to describe interactions of spin-2 particles in a local way. To project out the unphysical degrees of freedom, one needs the gauge redundancy: $h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\alpha_{\nu)}$. It is well known that this leads to the unique Poincaré invariant quadratic kinetic term for gravitons, which is the Einstein-Hilbert term expanded to quadratic order (up to field redefinitions and boundary terms)

$$\mathcal{L}_{\rm kin} = \frac{1}{2}(\eta^{\alpha\beta}\partial_\alpha h^{\mu\nu}\partial_\beta h_{\mu\nu} - \eta^{\mu\nu}\partial_\mu h^{(4)}\partial_\nu h^{(4)}) + \partial_\mu h^{\mu\nu}\partial_\nu h^{(4)} - \partial_\mu h^{\mu\alpha}\partial_\nu h^{\nu}_{\alpha}, \tag{2}$$

where $h^{(4)}$ is the four-dimensional trace $h^{(4)} = \eta^{\mu\nu}h_{\mu\nu}$. Furthermore, in order to maintain the gauge invariance at leading order, the graviton must couple to a symmetric 2-index conserved current, the energy-momentum tensor, as $\mathcal{L}_{\rm int} = -\frac{1}{2}\kappa h_{\mu\nu}T^{\mu\nu}$. Here the coupling κ is $\kappa \equiv \sqrt{32\pi G_N}$, where G_N is Newton's constant. To quadratic order in $h_{\mu\nu}$, the full action is $\mathcal{L}_{\rm GR}{}^{(2)} = \mathcal{L}_{\rm kin} + \mathcal{L}_{\rm int}$.

To leading order, i.e., ignoring the induced influence of gravitation on the matter sector itself, $T^{\mu\nu}$ obeys $\partial_\mu T^{\mu\nu} = 0$. Since the graviton is massless and mediates long ranged forces, it is highly nontrivial that it leads to local interactions. (We note that just because a theory is Lorentz invariant does *not* imply locality; it is trivial to construct Lorentz invariant theories that are nonlocal. For example, consider the Lagrangian for a scalar field $\mathcal{L} = (\partial\phi)^2/2 - (\partial\phi)^4/\Lambda^4 + \cdots$; this leads to superluminality around nontrivial ϕ backgrounds

[25].) In fact it can be shown that if the sources were to disobey the null energy condition $T_{\mu\nu}n^\mu n^\nu \geq 0$, then in fact the graviton could itself be superluminal around certain kinds of background configurations. Instead if we assume the null energy condition, the theory of general relativity, with standard leading order interactions, is in fact local.

As an important manifestation of this, we can consider the tree-level graviton exchange between a pair of matter sources. In harmonic (de Donder) gauge $\partial_\mu h^\mu_\nu = \frac{1}{2}\partial_\nu h^{(4)}$ the equations of motion are known to become very simple $\Box h_{\mu\nu} = -\frac{\kappa}{2}(T_{\mu\nu} - \frac{1}{2}T^{(4)}\eta_{\mu\nu})$, with $\Box = \partial_t^2 - \nabla^2$. One can then readily check if the tree-level exchange action between sources is local. To do so, one first solves for the particular solutions $h_{\mu\nu} = -\frac{\kappa}{2\Box}(T_{\mu\nu} - \frac{1}{2}T^{(4)}\eta_{\mu\nu})$. The corresponding tree-level graviton exchange action between sources is half the interaction term $-\frac{1}{4}\kappa h_{\mu\nu}T^{\mu\nu}$, giving the result

$$\frac{8\mathcal{L}_{\rm ex}}{\kappa^2} = T_{ij}\frac{T_{ij}}{\Box} - \frac{T}{2}\frac{T}{\Box} + \frac{T_{00}}{2}\frac{T_{00}}{\Box} - 2T_{0i}\frac{T_{0i}}{\Box} + T_{00}\frac{T}{\Box}. \quad (3)$$

Here we have broken up the energy-momentum tensor into a spatial tensor T_{ij}, spatial vector $T_{0i} = T_{i0}$, and spatial scalar T_{00} (along with $T \equiv \delta_{ij}T^{ij}$). This decomposition will be useful for comparison to our later analyses where we break Lorentz symmetry explicitly. This exchange action is local, despite the presence of the inverse operators on the right-hand side. That is because the inverse operators are the wave operator, which is associated with retarded wave propagation, as opposed to the inverse Laplacian, which is associated with instantaneous action at a distance. Later when we deform several aspects of the Poincaré symmetry, we shall see that almost all deformations will in fact lead to inverse Laplacians and instantaneous action at a distance. So by imposing locality, we will be able to derive that almost all deformations are forbidden, as we saw in our previous pair of papers [22,23].

B. Spatially invariant massless spin-2

Let us make clear what symmetries we will explicitly break and which remain. We assume a preferred frame that has rotation invariance and spatial translation invariance. However, as in our previous pair of papers [22,23], we do not assume Lorentz boosts. Moreover, we now do not assume time translations. Maintaining rotation will be important in this work. In particular, this ensures we have a notion of particle spin and our goal will be to build an interacting theory of spin-2 particles. Since we are interested in maintaining an explicit notion of locality, we will use the formalism that makes locality manifest; namely we can embed the spin-2 particles into fields. In principle we only need a 3×3 symmetric matrix to describe this, which we call h_{ij}, which transforms as a tensor under rotations. But locality will in fact require us to also introduce a nondynamical scalar ϕ and a nondynamical vector ψ_i. In the Lorentz invariant case, these can be organized into a 4×4 symmetric matrix field $h_{\mu\nu}$ that transforms as a tensor under Lorentz transformations (up to gauge transformations). In the non-Lorentz invariant, but rotationally invariant, case we can just refer to its components as

$$h_{00} \equiv \phi, \qquad h_{0i} = h_{i0} \equiv \psi_i, \qquad h_{ij} \quad (4)$$

(where ϕ and the Newtonian potential ϕ_N are related by $\phi = 2\phi_N/\kappa$). Also, the 3-trace is $h \equiv \delta^{ij}h_{ij}$.

To leading order, we need to form some appropriate generalization of the above dimension 4 operators that describe the quadratic terms of Eq. (2). In this work we are giving up both Lorentz boosts and time translations; this means there are many ways we can generalize the above theory of Poincaré invariant massless spin-2. A concrete generalization is the following: we can write out each of the terms of Eq. (2) in terms of its rotationally covariant building blocks h_{ij}, $\psi_i = h_{0i}$, and $\phi = h_{00}$. Then we can insert a different function in front of every term as follows:

$$\begin{aligned}\mathcal{L} = &-2A\dot{\psi}_i\partial_j h_{ij} + 2B\dot{h}\partial_i\psi_i - C\partial_i\phi\partial_i h + D\partial_i\phi\partial_j h_{ij} \\ &- E\partial_i h\partial_j h_{ij} - F(\partial_i\psi_i)^2 + G\partial_j h_{ij}\partial_k h_{ik} + H\partial_j\psi_i\partial_j\psi_i \\ &- \frac{I}{2}\dot{h}^2 + \frac{J}{2}\partial_i h\partial_i h + \frac{K}{2}\dot{h}_{ij}\dot{h}_{ij} - \frac{L}{2}\partial_k h_{ij}\partial_k h_{ij} + \mathcal{L}_{\rm int},\end{aligned} \quad (5)$$

where the coefficients are allowed to *a priori* be arbitrary functions of time, $A(t), B(t), \ldots, L(t)$. The Poincaré invariant theory of massless spin-2 is recovered by setting all these coefficients to 1. Although the above 12 kinetic terms are the full set of dimension 4 terms that describe the graviton in the Poincaré invariant case, it is not the full set allowed if we give up Lorentz and time translation symmetry as we are doing here. In particular one can also add a pair of terms $\sim(\partial_i\phi)^2$ and $\sim\dot{\phi}\partial_i\psi_i$. However it is relatively simple to show (as we mention in [23]) that they lead to nonlocality, so they shall be ignored here.

There are lower dimension quadratic terms that could be included too. In particular one could include several different dimension 2 mass terms and dimension 3 single-derivative terms. In the time translationally invariant theory that we analyzed in Ref. [23], we addressed these possible terms. In particular we allowed for mass terms, at least in part, and we mentioned that one could then perform field redefinitions to remove the dimension 3 terms (generating dimension 5 terms that are higher order and ignorable in this analysis); we then showed that mass terms led to nonlocality. Similarly, here we could include such lower dimension terms. Indeed this is important to obtain perturbations around FRW in general relativity. However, for simplicity, this will not be our focus here. Instead we will focus on massless spin-2 and assume the leading quadratic terms are dimension 4, as they are in general relativity. We suspect that if one sets the masses to zero then in fact

the dimension 3 terms must vanish in order to maintain locality, though we do not have a proof to present here. We leave the important case of including time-dependent mass terms and time-dependent dimension 3 terms to future work.

The leading order interaction allowed involves a single graviton coupled to matter. Poincaré invariance requires the coupling to the energy-momentum tensor $\sim h_{\mu\nu}T^{\mu\nu}$, as we reviewed in the previous subsection. We generalize this to

$$\mathcal{L}_{\rm int} = -\frac{1}{2}\kappa h_{\mu\nu}\mathcal{T}^{\mu\nu}, \tag{6}$$

where $\mathcal{T}^{\mu\nu}$ is a general source which does not *a priori* have any relationship to an energy-momentum tensor. In fact since we do not *a priori* assume Lorentz boosts or time translation invariance there does not *a priori* even exist a conserved, not to mention symmetric, energy-momentum tensor. We note that we will again specify the coupling as κ, which is useful for power counting. One could promote the coupling to be a function of time $\kappa = \kappa(t)$; however, without loss of generality, we can absorb this time dependence into the sources instead. Since we are only assuming spatial rotation invariance, we again find it most convenient to decompose it into its scalar, vector, and tensor (under rotations) pieces, as

$$\mathcal{T}^{00} \equiv \rho, \qquad \mathcal{T}^{0i} = \mathcal{T}^{i0} \equiv p_i, \qquad \mathcal{T}^{ij} \equiv \tau_{ij}. \tag{7}$$

This class of theories of massless spin-2 is specified by 12 functions of time $A(t), B(t), \ldots, L(t)$ (in addition to the strength of gravity κ). As mentioned above, the Lorentz invariant and time translationally invariant limit (general relativity) is where all these quantities are set to 1. Hence we have a large space of deformations from general relativity; namely 12 new functions of time. However, several of these functions can be eliminated by redefining the fields ϕ, ψ_i, and h_{ij} and exploiting the degeneracy in meaning between the scalar sources ρ and the trace of τ_{ij}; this leaves eight independent functions.

The classical equations of motion that follow from the action (5) are

$$\tilde{\kappa}\rho = -C\nabla^2 h + D\partial_i\partial_j h_{ij} \tag{8}$$

$$\tilde{\kappa}p_i = B\partial_i\dot{h} + H\nabla^2\psi_i - F\partial_i\partial_j\psi_j - \frac{\partial}{\partial t}(A\partial_j h_{ij}) \tag{9}$$

$$\begin{aligned}\tilde{\kappa}\tau_{ij} = {}& 2\delta_{ij}\frac{\partial}{\partial t}(B\partial_k\psi_k) - \delta_{ij}\frac{\partial}{\partial t}(I\dot{h}) + \frac{\partial}{\partial t}(K\dot{h}_{ij}) \\ & - A\partial_{(i}\dot{\psi}_{j)} + D\partial_i\partial_j\phi - E\partial_i\partial_j h - E\delta_{ij}\partial_k\partial_l h_{kl} \\ & - C\delta_{ij}\nabla^2\phi + G\partial_k\partial_{(i}h_{j)k} + J\delta_{ij}\nabla^2 h - L\nabla^2 h_{ij},\end{aligned} \tag{10}$$

where $\tilde{\kappa} \equiv -\kappa/2$. We can see breaking time translation symmetry results in one additional term in Eq. (9) proportional to $\dot{A}$ and three additional terms in Eq. (10) proportional to $\dot{B}$, $\dot{I}$, $\dot{K}$, as compared to GR in which $A = B = \cdots = L = 1$ are all constants.

The sources $\mathcal{T}^{\mu\nu}$ are not obviously required to be conserved here. As we did in [23], we parametrize the violation of source conservation by a scalar σ and vector w_i as

$$\sigma \equiv \partial_i p_i + \dot{\rho}_r \tag{11}$$

$$w_i \equiv \partial_j\tau_{ij} + \dot{p}_{r,i}, \tag{12}$$

where $\dot{\rho}_r \equiv (A/D)\dot{\rho}$ and $\dot{p}_{r,i} \equiv (A/H)\dot{p}_{r,i}$. In the GR limit we have $\partial_\mu\mathcal{T}^{\mu\nu} = 0$, which means $\sigma = w_i = 0$, but these are generally nonzero when Lorentz and time translation symmetry is abandoned. In fact we can use the classical equations of motion to express these violations of conservation in terms of the fields as

$$\begin{aligned}\tilde{\kappa}\sigma = {}& (H-F)\nabla^2\partial_i\psi_i + \left(B - \frac{AC}{D}\right)\nabla^2\dot{h} \\ & + \left(A\frac{\dot{D}}{D} - \dot{A}\right)\partial_i\partial_j h_{ij} - \frac{A}{D}\dot{C}\nabla^2 h\end{aligned} \tag{13}$$

$$\begin{aligned}\tilde{\kappa}w_i = {}& \left(2B - A - \frac{AF}{H}\right)\partial_i\partial_j\dot{\psi}_j + 2\dot{B}\partial_i\partial_j\psi_j \\ & + \left(\frac{AB}{H} - I\right)\partial_i\ddot{h} + \left(\frac{A\dot{B}}{H} - \dot{I}\right)\partial_i\dot{h} \\ & + (J-E)\partial_i\nabla^2 h + (G-L)\nabla^2\partial_j h_{ij} - \frac{A\ddot{A}}{H}\partial_j h_{ij} \\ & + \left(K - \frac{A^2}{H}\right)\partial_j\ddot{h}_{ij} + \left(\dot{K} - 2\frac{A\dot{A}}{H}\right)\partial_j\dot{h}_{ij} \\ & + (G-E)\partial_i\partial_j\partial_k h_{jk} + (D-C)\partial_i\nabla^2\phi.\end{aligned} \tag{14}$$

This representation is useful because it shows explicitly that σ and w_i vanish in the GR limit $A = B = \cdots = L = 1$, since all terms on the right-hand side are zero, but are nonzero otherwise.

III. RESTRICTING TO 2 DEGREES OF FREEDOM

As is well known, in the Poincaré invariant case the massless unitary representations are individual helicities, and for spin-2 we need to keep both left- and right-handed helicities to build a theory with local interactions. Furthermore, this uniquely specifies the free theory of Eq. (2) as it carries the gauge redundancy needed to cut down to these 2 degrees of freedom: $h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\alpha_{\nu)}$. Then one can construct all the interactions, order by order in κ, to build up the full Einstein-Hilbert action.

In contrast, it is sometimes suggested that the existence of the graviton can be *derived* by promoting the global

Poincaré symmetry to the gauge symmetry of diffeomorphisms. But this is incorrect, as any gauge symmetry can be trivially introduced by the Stueckelberg trick. Instead, the only known way to introduce the graviton within the framework of effective field theory is to simply *postulate* it.

Now in the present work, since we are giving up Lorentz boost and time translation invariance, but still studying massless spin-2, there is no longer an unambiguous consequence for the number of degrees of freedom. However, we will once again just postulate the existence of the two polarizations of the graviton in this work. We are motivated by the following: (i) we wish to make only small deformations to general relativity, while additional degrees of freedom could be viewed as large changes; (ii) all current evidence of gravitational waves is consistent with only 2 degrees of freedom [26]; and (iii) in the Lorentz invariant case additional degrees of freedom in massive gravity have been argued to suffer strong coupling and causality problems [25,27–29].

Since the theory of Eq. (5) with prefactor functions $A(t), B(t), \ldots, L(t)$ that are *a priori* allowed to be arbitrary functions of time, there is no longer any gauge redundancy in the theory. This means that it can propagate more than 2 degrees of freedom. Hence to project down to only 2 degrees of freedom for the graviton, one has to impose some constraints on the fields and/or the prefactor functions. We consider two options, which encompass the most general set of possibilities: (A) set $\partial_i\psi_i = 0 = \partial_j h_{ij}$; (B) fix parameters such that $\partial_i\psi_i$ and $\partial_j h_{ij}$ are directly determined by the equations of motion. Both of these examples may seem like a form of "gauge fixing," but we emphasize that since there is no gauge redundancy assumed, they are instead a choice of theory with different physical predictions. In our previous work of Ref. [23] we considered a third option as well. However, after enforcing locality, this third option was readily seen to be a special case of the second, so we will not consider it here.

Nevertheless, by demanding the tree-level exchange action is local, we find at the leading order to which we are working that both cases (A) and (B) will collapse to the same theory; namely GR with some additional terms. As in our previous work, we also do not *a priori* impose conservation laws on the sources $\mathcal{T}^{\mu\nu}$. However, locality will restrict its form tremendously, as we will describe.

A. Theory A: Transverse constraint

In this and the next section we describe two explicit ways to cut down to 2 degrees of freedom in the above class of theories. This is analogous to the first two theories we studied in [23] (while the third theory studied in [23] was found to be only a special case of the second theory once locality was imposed, so we will not repeat that analysis here). Moreover, we believe that our results are general and apply to any such choice.

Perhaps the most basic way is to remove any longitudinal modes directly by imposing that they vanish as

$$\partial_i\psi_i = 0, \qquad \partial_j h_{ij} = 0. \tag{15}$$

Note that this choice is not merely a "gauge choice," since the lack of Lorentz symmetry in the theory means there is no gauge redundancy in our field formalism, but a choice of theory. We will only need to work with the classical theory in this analysis (which suffices to derive our basic conclusions) and so we can imagine the above is implemented by Lagrange multipliers. The corresponding classical equations of motion are

$$\tilde{\kappa}\rho = -C\nabla^2 h \tag{16}$$

$$\tilde{\kappa} p_i = B\partial_i\dot{h} + H\nabla^2\psi_i \tag{17}$$

$$\begin{aligned}\tilde{\kappa}\tau_{ij} = {} & -\delta_{ij}\frac{\partial}{\partial t}(I\dot{h}) + \frac{\partial}{\partial t}(K\dot{h}_{ij}) - L\nabla^2 h_{ij} - A\partial_{(i}\dot{\psi}_{j)} \\ & - C\delta_{ij}\nabla^2\phi + D\partial_i\partial_j\phi - E\partial_i\partial_j h + J\delta_{ij}\nabla^2 h.\end{aligned} \tag{18}$$

Note the functions F and G no longer appear in the theory.

We can explicitly check on the status of source conservation in this theory. The scalar component of the usual conservation of sources $\partial_\mu T^{\mu\nu} = 0$ is now

$$\partial_i p_i + B\frac{\partial}{\partial t}\left(\frac{\rho}{C}\right) = 0 \tag{19}$$

so the quantity $\int d^3\rho/C$ is conserved, though in our choice of variables it is not simply the integral of the source ρ, but rescaled by a time-dependent coefficient in this conservation equation. This implies that the matter sector must involve some Abelian symmetry; though at this stage we cannot identify its details as we would need more information about the matter sector (we will later see that locality requires it to be related to a space-time symmetry). On the other hand the vector component of the usual conservation of sources $\partial_\mu T^{\mu\nu} = 0$ is rather complicated:

$$\begin{aligned}\partial_j\tau_{ij} + \frac{A}{H}\dot{p}_i = {} & \left(I - \frac{AB}{H}\right)\frac{\partial_i\ddot{\rho}}{C\nabla^2} + \frac{E-J}{C}\partial_i\rho \\ & + \frac{D-C}{\tilde{\kappa}}\partial_i\nabla^2\phi + \left[\frac{AB}{H}\left(\frac{\dot{B}}{B} - \frac{\dot{H}}{H}\right) - \dot{I}\right] \\ & \times \frac{\partial}{\partial t}\frac{\partial_i}{\nabla^2}\left(\frac{\rho}{C}\right).\end{aligned} \tag{20}$$

Enforcing locality.—The basic principle which we are imposing is that of locality: no instantaneous signaling at a distance. However the previous equation will in general spoil this because there are several terms that involve inverse Laplacians. This means that even if the sources

themselves (ρ, etc.) are initially local, they will evolve to ensure sources (p_i, τ_{ij}) are nonlocally distributed in space. This would provide a means to instantaneously signal to a distant observer by having the distant observer situated near one of these effectively nonlocal sources and then disturbing the source. Hence to maintain locality we need to enforce that these terms vanish. This requires the following set of conditions be satisfied:

$$AB = IH, \qquad \frac{\dot{B}}{B} = \frac{\dot{I}}{I} + \frac{\dot{H}}{H}, \qquad D = C, \tag{21}$$

where the first condition here arises from requiring the coefficient of $\partial_i \ddot{\rho}/\nabla^2$ vanish, the second condition here arises from requiring the coefficient of $\partial_i(\rho/C)/\nabla^2$ vanish, and the third condition here arises from requiring the coefficient of $\nabla^2\phi$ vanish [this final condition is needed because ϕ itself is highly nonlocal; see ahead to Eq. (26)]. With these three conditions enforced, the conservation equation becomes local and nearly canonical. To make it appear as canonical as possible it is useful to define a shifted tensor source as

$$\tau_{ij} = \tilde{\tau}_{ij} + \frac{E-J}{D}\delta_{ij}\rho. \tag{22}$$

By then inserting into Eq. (20) with the above trio of conditions, we then obtain

$$\partial_j \tilde{\tau}_{ij} + \frac{A}{H}\dot{p}_i = 0. \tag{23}$$

So, in this theory, locality has enforced that the sources obey almost standard conservation laws. (We note that $\tilde{\tau}_{ij}$ couples linearly to the graviton, just as τ_{ij} does.)

Having ensured that the sources themselves are local, we can now move to another test of locality. Since the graviton is massless (or even if it were very light) it will mediate a long ranged force between matter sources. Hence it can very easily lead to nonlocal interactions. One necessary condition to avoid this is that the exchange interaction between a pair of matter sources, with no external gravitons, is local. (Later we will comment on additional ramifications that can emerge from considering external gravitons.) So it is useful to have the inhomogeneous particular solutions for the gravitational fields

$$\frac{h}{\tilde{\kappa}} = \frac{-\rho}{D\nabla^2} \tag{24}$$

$$\frac{\psi_i}{\tilde{\kappa}} = \frac{p_i}{H\nabla^2} + \frac{B}{H}\frac{\partial_i}{\nabla^4}\frac{\partial}{\partial t}\left(\frac{\rho}{D}\right) \tag{25}$$

$$\frac{\phi}{\tilde{\kappa}} = \frac{-\tau}{2D\nabla^2} + \frac{E+L-3J}{2D^2}\frac{\rho}{\nabla^2} + \frac{1}{2D}\frac{\partial}{\partial t}\left\{(3I-K)\left[\frac{\partial}{\partial t}\left(\frac{\rho}{D\nabla^4}\right)\right]\right\} \tag{26}$$

$$\begin{aligned}\frac{h_{ij}}{\tilde{\kappa}} = {}& \frac{\tau_{ij}}{\tilde{\Box}} + \frac{(\partial_i\partial_j - \delta_{ij}\nabla^2)\tau}{2\tilde{\Box}\nabla^2} + \frac{1}{\tilde{\Box}}\left(\frac{A}{\nabla^2}\frac{\partial}{\partial t}\left(\frac{\partial_{(i}p_{j)}}{H}\right)\right. \\ & + \frac{\delta_{ij}}{2D}(E-J+L)\rho + \left(\frac{3J-3E-L}{2D}\right)\frac{\partial_i\partial_j\rho}{\nabla^2} \\ & + \frac{\delta_{ij}}{2\nabla^2}\frac{\partial}{\partial t}\left[(I-K)\frac{\partial}{\partial t}\left(\frac{\rho}{D}\right)\right] \\ & + \frac{\partial_i\partial_j}{\nabla^4}\left\{2A\frac{\partial}{\partial t}\left[\frac{B}{H}\frac{\partial}{\partial t}\left(\frac{\rho}{D}\right)\right]\right. \\ & \left.\left. + \frac{1}{2}\frac{\partial}{\partial t}\left[(K-3I)\frac{\partial}{\partial t}\left(\frac{\rho}{D}\right)\right]\right\}\right),\end{aligned} \tag{27}$$

where we have defined a wave operator,

$$\tilde{\Box} \equiv K\partial_t^2 + \dot{K}\partial_t - L\nabla^2. \tag{28}$$

This unusual wave operator merits some discussion. First, though it involves functions of time through $K = K(t)$ and $L = L(t)$ it is still a linear operator, and therefore has a Green's function, which lets us continue using its inverse to simply mean convolution with that Green's function. Or, more simply, whatever its functional form, this operator has a matrix representation, and the operations of matrix multiplication and inversion are well defined.

The corresponding tree-level graviton-exchange action between a pair of sources is in fact 1/2 of the interaction term $\mathcal{L}_{\rm ex} = \mathcal{L}_{\rm int}/2 = -\kappa h_{\mu\nu}T^{\mu\nu}/4$. Then to determine under what conditions this exchange action is local, we can now write the interaction Lagrangian, Eq. (6), only in terms of the sources by using the solutions (24)–(27) to eliminate the fields from the interaction Lagrangian. We can then integrate by parts (in the spatial integrals) in the action to replace all divergences using the conservation equations (19) and (23). Integrating by parts in time now introduces a great deal of complexity to the problem because each term in the action is cubic in functions of time, and the operator $\Box$ in Eq. (28) also contains functions of time. The general nonlocal exchange action contains a great many unique terms. To study this in full generality was only feasible with computer algebra software, such as *Mathematica*. We will describe its general form and present a few simple terms as concrete examples.

Schematically, the exchange action now contains terms of the form

$$\begin{aligned}\mathcal{L}_{\rm ex} \supseteq {}& a_1[t]\mathcal{S}\mathcal{D}(a_2[t]\mathcal{S}) + a_3[t]\mathcal{S}\mathcal{D}(a_4[t]\dot{\mathcal{S}}) \\ & + a_5[t]\mathcal{S}\mathcal{D}(a_6[t]\ddot{\mathcal{S}}),\end{aligned} \tag{29}$$

where $a_n[t] = a_n[A(t), B(t), \ldots, L(t)]$ are functions of the coefficients in the action and their time derivatives, $\mathcal{D} = \mathcal{D}[\tilde{\Box}^{-1}, \nabla^2]$ is some differential operator involving at least one power of the inverse wave operator and various (positive or negative) powers of Laplacians, and $\mathcal{S}$ stands

for any of the sources ρ, p_i, τ_{ij}. Note in the static limit ($\dot{A} = \dot{B} = \cdots = \dot{L} = 0$), the second term would be a total derivative in the action, while the other two terms would not necessarily vanish. To enforce locality, for every distinct term in $\mathcal{L}_{\rm ex}$ that contains any inverse Laplacians the functions $a_n[t]$ must be set to zero or a constant to eliminate the nonlocal terms, making them identically zero or a total derivative as appropriate. In most cases each set of terms provides multiple possible constraint options to eliminate the nonlocal parts, such as in the first term of the schematic example, Eq. (29), where if this term were nonlocal, either a_1 or a_2 could be set to zero to eliminate it. Nevertheless, it turns out there is a unique set of conditions that eliminates all nonlocal terms from the exchange action.

We emphasize that any terms in $\mathcal{L}_{\rm ex}$ that have inverse Laplacians would give rise to instantaneous signaling of one source to a distant source; which we wish to avoid in this work. This would be in sharp contrast to all known physical theories (general relativity, electromagnetism, etc.) which avoid instantaneous signaling; and indeed avoid inverse Laplacians in the final exchange action.

Most of the sources can only couple to themselves, as represented in Eq. (29) where each term $\sim\mathcal{SS}$. The only possible mixed terms are $\sim\rho\tau$, and turn out to be among the least numerous. As a simple example, grouping nonlocal terms $\sim\rho\tau$, the exchange action contains

$$\mathcal{L}_{\rm ex} \supset \frac{\tau}{2C}\frac{(I-K)\ddot{\rho}}{\tilde{\Box}\nabla^2} + \frac{I\ddot{\rho}}{2C}\frac{\tau}{\tilde{\Box}\nabla^2} - \frac{\rho}{2C}\frac{K\ddot{\tau}}{\tilde{\Box}\nabla^2} \tag{30}$$

from which we find locality requires $I = K$ in order to remove the first term as it is proportional to $1/\nabla^2$. We also find $\dot{K} = 0$ is a necessary condition for locality, rendering $I = K =$ const. Note, however, this does not completely eliminate every term in Eq. (30), just the nonlocal ones; integrating by parts in time generates local terms higher order in $\Box^{-1}$. The full set of necessary and sufficient conditions to eliminate all nonlocal terms from the tree-level exchange action in the theory of this subsection are found to be

$$\begin{aligned} A = B = c_1, \qquad & C = D = c_2, \\ I = K = c_3, \qquad & H = \frac{c_1^2}{c_3} \\ 2E(t) = J(t) + L(t), \qquad & \dot{E} = \dot{J} = \dot{L} = c_4, \end{aligned} \tag{31}$$

where c_1, c_2, c_3, c_4 are constants (independent of time); c_1, c_2, c_3 are dimensionless, while c_4 has units of inverse time. We see that locality imposes that most of the coefficients obey the standard relations in the Poincaré invariant theory of general relativity. In particular, most of the coefficients are required to be constants. Moreover the conservation equations (19) and (23) are now simple because the coefficients in those equations are required to be time independent. In fact building $\mathcal{T}^{\mu\nu}$ out of the sources ρ, p_i and $\tilde{\tau}_{ij}$, we can easily rescale ρ and p_i by a constant to write $\partial_\mu\tilde{\mathcal{T}}^{\mu\nu} = 0$, so the theory is required to have a canonical conserved source.

The new feature that departs from the Poincaré invariant theory is the functions E, J, and L which can be linear functions of time, and must have the same time dependence. Writing $J = J_0 + c_4 t$ etc., enforcing the conditions (31) leaves the exchange action

$$\begin{aligned} \frac{8\mathcal{L}_{\rm ex}}{\kappa^2} = {} & \tau_{ij}\frac{\tau_{ij}}{\Box} - \frac{\tau}{2}\frac{\tau}{\Box} + \frac{\{\rho\, a_1\tau\}}{4c_2}\frac{}{\Box} + \frac{\rho}{8c_2^2}\frac{a_2\rho}{\Box} - \frac{2c_3}{c_1^2}p_i\frac{Lp_i}{\Box} \\ & + \frac{c_3c_4}{c_2}\rho\left[\frac{\dot{\tau}}{\Box^2} + c_4\frac{\nabla^2\tau}{\Box^3}\right] \\ & - \frac{c_3c_4}{c_2^2}\rho\left[c_4\frac{\rho}{\Box^2} + \frac{(L-3J)\dot{\rho}}{2\Box^2}\right] \\ & + \frac{c_3c_4^2}{2c_2^2}\rho\frac{(3J-5L)\nabla^2\rho}{\Box^3} - \frac{2c_3^2c_4}{c_1^2}p_i\frac{\dot{p}_i}{\Box^2}, \end{aligned} \tag{32}$$

where $a_1 = 3L - J$, $a_2 = 18JL - 3J^2 - 11L^2$, $a_3 = 3J - 5L$, $\{\rho a_1\tau\}/\Box \equiv \rho(a_1\tau/\Box) + \tau(a_1\rho/\Box)$, and the wave operator with K constant has become $\Box \equiv K\partial_t^2 - L\nabla^2$. This is now completely local and the first line has a similar form to the GR action Eq. (3), but with coefficients a_1 linear in time and a_2 quadratic in time. We are left with six unspecified parameters, c_1, c_2, c_3, c_4, J_0 and L_0. In the static case ($c_4 = 0$) the two symmetric $\rho\tau$ terms on the first line combine and the higher order $\Box^{-1}$ terms vanish, in agreement with our previous work on the static case. However, with $c_4 \neq 0$, we are left with a more general—explicitly time-dependent—theory.

Even in the static limit, it is not immediately obvious how Eq. (32) reduces to Eq. (3). As we remarked in [23], if we rewrite the exchange action in terms of the conserved sources and rescale $\rho \to (c_2/L)\rho$, in the static case we then recover Eq. (3) exactly. However, this does not generalize to the present case with $L(t)$ a function of time. Expressing the exchange action (32) in terms of the conserved source $\tilde{\tau}_{ij}$ [from Eqs.(22) and (23)] we obtain

$$\begin{aligned} \frac{8\mathcal{L}_{\rm ex}}{\kappa^2} = {} & \tilde{\tau}_{ij}\frac{\tilde{\tau}_{ij}}{\Box} - \frac{\tilde{\tau}}{2}\frac{\tilde{\tau}}{\Box} + \frac{\{\rho\, L\tilde{\tau}\}}{2c_2}\frac{}{\Box} + \frac{\rho}{2c_2^2}\frac{L^2\rho}{\Box} - \frac{2c_3}{c_1^2}p_i\frac{Lp_i}{\Box} \\ & + \frac{c_3c_4}{c_2}\left(\rho\frac{\dot{\tau}}{\Box^2} + c_4\rho\frac{\nabla^2\tau}{\Box^3}\right) + \frac{c_3c_4}{c_2^2}\rho\frac{L\dot{\rho}}{\Box^2} + \frac{c_3^2c_4^2}{c_2^2}\rho\frac{\ddot{\rho}}{\Box^3} \\ & - 2\frac{c_3^2c_4}{c_1^2}p_i\frac{\dot{p}_i}{\Box^2}. \end{aligned} \tag{33}$$

In this form $J(t)$ drops out, so there are really only five constants here: c_1, c_2, c_3, c_4, and L_0 (we may, as in the next section, label $L_0 \equiv c_5$). Unlike in the static case, factors of t (from $L = L_0 + c_4 t$) can only appear next to source terms inside the $\Box^{-1}$. To appear to the left in any of these terms,

outside the $\Box^{-1}$, requires some integration by parts in time, leaving only $\sim\dot{L} = c_4$ outside the $\Box^{-1}$. Thus in the two terms symmetric in ρ and $\tilde{\tau}$, in one term L comes with the ρ and in the other L comes with the $\tilde{\tau}$. This means with $L(t)$ a function of time we cannot rescale ρ to make the first line look more like Eq. (3) as we could in the static case. Our interpretation of the above result will come in Sec. IV.

B. Theory B: Constraint from equations of motion

In the above section we cut down to 2 degrees of freedom of the graviton by imposing constraints directly on the fields in Eq. (15). A secondary approach is to impose some conditions on the prefactor functions so that the equations of motion include constraints that cut down the degrees of freedom directly. By studying the full theory with all 12 unknown functions $A(t), \ldots, L(t)$, and using the general equations for the nonconserved sources, Eq. (13) and (14), we can impose the constraints on functions

$$A = B = c_1, \qquad C = D = c_2, \qquad I = K = \frac{A^2}{H}. \tag{34}$$

It is simple to show that this now fixes $\partial_i\psi_i$ and $\partial_i h_{ij}$ by the equations of motion, leaving the theory with the desired 2 degrees of freedom. In particular the classical equations of motion enforce that they are given by the sources as

$$\partial_i\psi_i = \frac{\tilde{\kappa}\sigma}{(H-F)\nabla^2} \tag{35}$$

$$\begin{aligned}\partial_i h_{ij} = \frac{\tilde{\kappa}}{(G-L)\nabla^2}\bigg[&\left(\frac{2E-G-J}{2G+J-2E-L}\right)\frac{\partial_j q}{\nabla^2} + w_j \\ &- \frac{K}{A}\frac{\partial_j\dot{\sigma}}{\nabla^2} + \left(\frac{\dot{F}}{F} - \frac{\dot{H}}{H}\right)\frac{AF}{H(H-F)}\frac{\partial_j\sigma}{\nabla^2} \\ &+ \frac{J-E}{D}\partial_j\rho + \frac{\dot{K}}{A}p_j\bigg].\end{aligned} \tag{36}$$

In the expression for $\partial_i h_{ij}$ we have introduced the following function of the sources:

$$\begin{aligned}q \equiv \partial_j w_j - \frac{A}{H}\dot{\sigma} + \frac{J-E}{D}\nabla^2\rho - \frac{\dot{K}}{D}\dot{\rho} \\ + \left[\left(\frac{\dot{F}}{F} - \frac{\dot{H}}{H}\right)\frac{AF}{H(H-F)} + \frac{\dot{K}}{A}\right]\sigma\end{aligned} \tag{37}$$

which gives a nice form for the second derivative,

$$\partial_i\partial_j h_{ij} = \left(\frac{\tilde{\kappa}}{2G-2E+J-L}\right)\frac{q}{\nabla^2}. \tag{38}$$

Enforcing locality.—As in the previous section, we can now solve the equations of motion to obtain the inhomogeneous solutions for the fields, given in the Appendix. We again use these to write the exchange action only in terms of the sources, integrating by parts to replace divergences using the definitions of q, σ, and w_i. Once again, because of the cubic dependence on functions of time, integrating by parts to simplify under the fewest independent variables expands the total number of terms enormously. The general form is similar to that described in Eq. (29).

Now that we have general σ and w_i [as opposed to the restricted form in Theory (A)], there are more possible mixed terms than only $\sim\rho\tau$; there are also terms with $\tau\sigma$, $\rho\sigma$, $\rho\partial_i w_i$, $\tau\partial_i w_i$, and $p_i w_i$. It is advantageous to first look at the $w_i w_i$ and $\sigma\sigma$ terms involving inverse Laplacians to see if locality requires some constraint that simplifies all these other terms involving σ and w_i. Looking for nonlocal terms in $w_i w_i$ and $\sigma\sigma$, the exchange action contains

$$\mathcal{L}_{\rm ex} \supset 2\sigma\frac{1}{\tilde{\Box}\nabla^2}\left[\frac{FKL\sigma}{H(c_1^2 - FK)}\right] + 2w_i\frac{1}{\tilde{\Box}\nabla^2}\left[\frac{Gw_i}{G-L}\right] \tag{39}$$

plus terms higher order in ∇^{-2} and derivatives of σ and w_i. If σ and w_i are general functions, then the theory is immediately nonlocal. To eliminate these terms, the only possible constraints on the coefficients $A(t), \ldots, L(t)$ would impose some of these vanish [e.g., $G(t) = 0$], leading to an overly trivial theory, which we decline to do. The most direct way to avoid this problem is to impose that σ and w_i vanish, i.e., $\sigma = w_i = 0$. In previous work [23] we allowed for σ and w_i to themselves be given by derivatives of other local functions in order to maintain locality by having the Laplacians cancel out. But this only leads to ultralocal terms in the action, which were found to be associated with a completely decoupled, and hence irrelevant, sector. We will ignore this here as it does not represent a real modification to GR. Instead by focusing on the important case of only coupled sectors, we are forced by locality to impose $\sigma = w_i = 0$ [we can also have $w_i = (E-J)\delta_{ij}\rho/D$, but we can easily shift from τ_{ij} to $\tilde{\tau}_{ij}$ to eliminate this, as we did earlier in Eq. (22); so we can ignore that without loss of generality].

Thus locality imposes that the sources are conserved and the theory is now similar in structure to Theory (A) of the previous section, though the coefficients F and G are still present. The full set of conditions to make the exchange action of Theory (B) local are

$$\begin{aligned}&\sigma = 0, \qquad w_i = 0, \qquad I = K = c_3, \\ &E = G = J = L = c_5 + c_4 t, \qquad F = H = c_1^2/c_3.\end{aligned} \tag{40}$$

Again we see that the static case is recovered if $c_4 = 0$ in agreement with Ref. [23]. For the time-dependent case with $c_4 \neq 0$, we have the conditions of Eq. (40). So we are left with five parameters of the theory, as in Theory (A). We therefore recover the same local action Eq. (33) with the same number of parameters. [We note that since $F = H$ and

$G = L$ here, it is compatible with Eqs. (35) and (36), since we also have $\sigma = w_i = 0$. This is clearest seen by multiplying both sides of Eq. (35) by $(H - F)$ and Eq. (36) by $(G - L)$].

IV. DISCUSSION

After enforcing locality, we see that both theories became identical, so we recover the same exchange action. That is, imposing all conditions in Eq. (40), the now local tree-level exchange action of Theory (B) is exactly as in Theory (A) after writing the exchange action in terms of the conserved sources, Eq. (33). We believe this is true for any particular construction that follows from the above class of theories outlined in Sec. II (we can also have decoupled sectors, as we elaborated on in [23], but this is of little interest here).

If we set the coefficient $c_4 = 0$ we recover general relativity exactly (at this order). But with $c_4 \neq 0$ we have a theory of spin-2 with 2 degrees of freedom that explicitly breaks time translation invariance, while maintaining locality in tree-level graviton exchange between conserved matter sources, and is therefore a deviation from general relativity.

A. Gauss-Bonnet

Analyzing only the exchange action as we have in the previous two sections lets us identify nonlocal interactions, but obfuscates the structure of the theory. Returning to the full Lagrangian, after enforcing locality by Eq. (40) what we are left with is the Lagrangian of Eq. (5) with all coefficients constant, plus a term $\propto c_4$ that breaks time translation symmetry. If we set $c_4 = 0$, we recover time translations, and we can also absorb the constant coefficients by a redefinition of the fields, allowing us to recover GR with its full Poincaré symmetry. In the case of interest, with $c_4 \neq 0$, the theory breaks time translation symmetry and in turn the Lorentz boost symmetry of the full theory.

For ease of notation and without loss of generality, we set all $c_1 = c_2 = c_3 = c_5 = 1$. This leaves only c_4 as the adjustable parameter of the theory. We can then write the full Lagrangian of the theory (5) as

$$\mathcal{L} = \mathcal{L}^{(2)}_{\rm GR} + c_4 t\left(\frac{1}{2}\partial_i h \partial_i h - \frac{1}{2}\partial_k h_{ij}\partial_k h_{ij} + \partial_j h_{ij}\partial_k h_{ik} - \partial_i h \partial_j h_{ij}\right). \quad (41)$$

Let us compare this to the Gauss-Bonnet action [30] with an inserted prefactor function $f(t)$:

$$\Delta\mathcal{L}_{\rm GB} = f(t)\sqrt{-g}(R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}). \quad (42)$$

At the leading order in $h_{\mu\nu}$, to which we are working, and after performing some integrations by parts, we can identify

$$\ddot{f} = -c_4 t. \quad (43)$$

That is, the additional terms in Eq. (41) are nothing other than the Gauss-Bonnet action with the function $f(t)$ determined by Eq. (43).

If f is a constant, the action Eq. (42) is a surface term in four dimensions, so it is usually not included in GR even though this particular sum of curvature squared terms satisfies the same symmetries as the Einstein-Hilbert action. While we did not begin our analysis by explicitly adding the Gauss-Bonnet action with a function of time $f(t)$, these terms are contained in the general action of Eq. (5) to quadratic order in $h_{\mu\nu}$. After enforcing locality we recover Einstein-Gauss-Bonnet gravity with $f(t)$ related to the graviton speed $v_g^2 = L(t)$ by Eq. (43).

Since $\Delta\mathcal{L}_{\rm GB}$ is itself invariant under the linearized gauge transformations $h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\alpha_{\nu)}$ and both Theories (A) and (B) end up with a conserved source, $\partial_\mu \tilde{\mathcal{T}}^{\mu\nu} = 0$, the full action of Eq. (41) enjoys the linearized gauge redundancy of GR at this leading order. One can then gauge fix in any desired fashion [e.g., as in Theory (A)] and thus obtain Eq. (33) for the tree-level exchange action between matter sources.

So the general conceptual features of our final results are: the imposition of locality (avoiding instantaneous signaling) is so strong it forces our theory into a familiar gauge redundant form, even though we did not assume gauge redundancy as a starting point in Eq. (5). In particular, in addition to the quadratic Einstein-Hilbert term, which is well known to be gauge invariant, we have also found an allowed Gauss-Bonnet term with a prefactor that depends on time. This in fact is the most general action at the order we are working (albeit with a particular function of time) which is gauge invariant. A general conceptual feature of the sources $\mathcal{T}^{\mu\nu}$ will be discussed in the final subsection.

B. Broken symmetries at this order

The general local action we find for spin-2 with 2 degrees of freedom, Eq. (41), has the linearized gauge invariance of GR, but the presence of the explicit function of time $f(t)$ prevents time translation symmetry, and its relation to the graviton speed by Eq. (43) [i.e., $v_g^2 = L(t)$, $\ddot{f} = -\dot{L}$] prevents us from recovering the usual Lorentz boost invariance. We do not have a simple intuition as to why locality of spin-2 prefers to include the Gauss-Bonnet terms in the action with a prefactor $f(t)$ that can be at most cubic in time, when any function of time would maintain gauge invariance at this order.

In previous work [7] some of us found for any nontrivial $f(t)$ the theory of Eq. (41) allows for superluminal gravitons. This is still true here, but in this work we demanded only a general notion of locality, that there is no instantaneous signaling. So the theory is local in this sense, though allows for gravitons with a time-dependent speed

which can in general be superluminal, i.e., it can be faster than the particle speeds of the matter sector, including photons.

Nonzero c_4 characterizes deviations from Poincaré invariance. To be compatible with observations the value of c_4 (which has units of inverse time) must be very small. As an example of a bound, the recent LIGO observations of merging neutron stars showed that the graviton's speed matched the photon's speed to an accuracy of $|v_g - 1| \lesssim 10^{-15}$ [31] (we set the speed of light to 1 here). Since the photons and gravitons traveled from this event for $t_* \sim 10^8$ years to Earth, we can translate this into the direct bound on the coefficient c_4 of

$$|c_4| \lesssim |v_g - 1|/t_* \sim 10^{-23} \text{ years}^{-1}. \tag{44}$$

Since c_4 is dimensionful it is not obvious what its *a priori* expected value is. On the other hand, since this is such a small number in most fundamental units, it suggests that its true value may very well be exactly zero.

C. Derived symmetries and higher order

One of our primary findings is that locality requires the conservation of matter sources $\partial_\mu \tilde{T}^{\mu\nu} = 0$. Since there is only one nontrivially conserved 2-index tensor, then this must be the energy-momentum tensor $\tilde{T}^{\mu\nu} = T^{\mu\nu}$. As we detailed at the end of [23], the existence of a conserved symmetric energy-momentum tensor relies both on Lorentz and time translation invariance. Hence we have found that locality enforces those symmetries are obeyed by the matter sector.

On the other hand, our final action allows for the breaking of time translations in the gravity sector. At the above level of analysis, this is not an inconsistency because we have only worked to leading order; in particular, we only need the *matter sector* to exhibit time translations for $\partial_\mu T^{\mu\nu} = 0$, and this is true so long as we ignore back-reaction from the graviton. Nevertheless this does suggest that if one were to demand locality at higher order, then it is possible that the time translation violation, as measured by the parameter c_4, would in fact have to vanish.

Moreover, one could reasonably anticipate this as follows: so far we have only allowed the external sources to be part of the matter sector and we found that such a matter sector must exhibit time translations in order to maintain locality. In addition, however, one should also consider the external sources to be gravitons. In this case, it seems plausible that a similar analysis would demand that the energy-momentum tensor of the graviton itself needs to be conserved, which would then demand time translations in the graviton sector, giving $c_4 = 0$ and the full Poincaré symmetry of general relativity. Another hint is the following: we have remarked that locality has enforced that the final action of Eq. (41) is gauge invariant to this leading order. It is plausible that at higher order, locality will again demand gauge invariance; this would presumably require f to be a constant to recover the usual nonlinear diffeomorphism invariance of general relativity. But we do not have a proof of these issues.

An important direction for future work is to include mass terms and single derivative terms in the starting action. Although we could show that mass terms needed to vanish in order to maintain locality in the time translation invariant case of [23] (and the single derivative terms could then be converted by field redefinitions into higher dimension operators), they may play an important role in the time-dependent case. This is especially important in that such terms are known to exist when expanding around an FRW background in ordinary general relativity. We leave this important issue for future work.

ACKNOWLEDGMENTS

We would like to thank McCullen Sandora for very helpful input. M. P. H. is supported in part by National Science Foundation Grants No. PHY-1720332 and No. PHY-2013953.

APPENDIX: EQUATIONS OF MOTION IN THEORY B (SEC. III B)

In this Appendix we report on the inhomogeneous particular solutions to the equations of motion (8)–(10) with the conditions of Theory (B) to cut down to 2 degrees of freedom, Eqs.(35) and (36):

$$\phi = \frac{-\tau}{2C\nabla^2} + \frac{\partial}{\partial t}\left[\left(\frac{3I-K}{2C\nabla^4}\right)\frac{\partial}{\partial t}\left(\frac{\rho}{C} - \frac{q}{(2G-2E+J-L)\nabla^2}\right)\right] + \left(\frac{E-3J+L}{2C^2}\right)\frac{\rho}{\nabla^2} + \left(\frac{2G-4E+3J-L}{2G-2E+J-L}\right)\frac{q}{2C\nabla^4} + \frac{2A}{C}\frac{\partial}{\partial t}\left[\frac{\sigma}{(H-F)\nabla^4}\right] \tag{A1}$$

$$\psi_i = \frac{p_i}{H\nabla^2} + \frac{F}{H(H-F)}\frac{\partial_i\sigma}{\nabla^4} + \frac{A}{H}\frac{\partial}{\partial t}\left\{\frac{1}{(G-L)\nabla^4}\left[w_i - \frac{K}{A}\frac{\partial_i\dot{\sigma}}{\nabla^2} + \left(\frac{\dot{F}}{F} - \frac{\dot{H}}{H}\right)\frac{AF}{H(H-F)}\frac{\partial_i\sigma}{\nabla^2} + \frac{\dot{K}}{A}p_i\right] + \frac{A}{C}\left(1 + \frac{J-E}{G-L}\right)\frac{\partial_i\rho}{\nabla^4} - \frac{A}{G-L}\frac{\partial_i q}{\nabla^6}\right\} \tag{A2}$$

$$
\begin{aligned}
h_{ij} = \frac{1}{\tilde{\Box}}\Bigg\{ & \tau_{ij} + \left(\frac{\partial_i\partial_j}{\nabla^2} - \delta_{ij}\right)\frac{\tau}{2} + \frac{\delta_{ij}}{2}\left(\frac{E+L-J}{C}\right)\rho + \frac{\partial_i\partial_j}{2\nabla^2}\left[\frac{L(L-3J)-G(J+L)+E(G+3L)}{C(G-L)}\right]\rho \\
& + \frac{\partial_i\partial_j}{\nabla^4}\frac{\partial}{\partial t}\left[2K\frac{\partial}{\partial t}\left(\left\{1-\frac{J-E}{G-L}\right\}\frac{\rho}{C}\right) - \frac{K\dot{\rho}}{C}\right] + \frac{\delta_{ij}}{2}\left(\frac{2G-2E+3J-L}{2G-2E+J-L}\right)\frac{q}{\nabla^2} \\
& + \frac{\partial_i\partial_j}{2\nabla^4}\left[\frac{G(2G-2E+J)+3L(G-2E+J)-L^2}{(G-L)(2G-2E+J-L)}\right]q + \frac{\partial_i\partial_j}{\nabla^6}\frac{\partial^2}{\partial t^2}\left[\frac{K(4E-3G-2J+L)q}{(G-L)(2G-2E+J-L)}\right] \\
& + \frac{2\partial_i\partial_j}{\nabla^4}\left[A\frac{\partial}{\partial t}\left(\frac{F\sigma}{H(H-F)}\right) + \frac{GK\dot{\sigma}}{A(G-L)} - \left(\frac{\dot{F}}{F}-\frac{\dot{H}}{H}\right)\frac{AFG\sigma}{H(H-F)(G-L)} - A\frac{\partial}{\partial t}\left(\frac{\sigma}{H-F}\right)\right] \\
& + \frac{2\partial_i\partial_j}{\nabla^6}\frac{\partial}{\partial t}\left[K\frac{\partial}{\partial t}\left(\frac{1}{G-L}\left\{\frac{-K\dot{\sigma}}{A} + \left(\frac{\dot{F}}{F}-\frac{\dot{H}}{H}\right)\frac{AF\sigma}{H(H-F)}\right\}\right)\right] + A\frac{\partial_{(i}}{\nabla^2}\frac{\partial}{\partial t}\left(\frac{p_{j)}}{H}\right) \\
& + \frac{A^2\partial_{(i}}{\nabla^4}\frac{\partial^2}{\partial t^2}\left[\frac{1}{G-L}\left(w_{j)} + \frac{\dot{K}}{A}p_{j)}\right)\right] - \frac{G\partial_{(i}}{(G-L)\nabla^2}\left(w_{j)} + \frac{\dot{K}}{A}p_{j)}\right)\Bigg\},
\end{aligned}
\tag{A3}
$$

where as above $\tilde{\Box} \equiv \partial_t(K\partial_t) - L\nabla^2$, and we can take the trace to obtain

$$
h = \frac{-\rho}{C\nabla^2} + \frac{q}{(2G-2E+J-L)\nabla^4}. \tag{A4}
$$

[1] L. Verde, T. Treu, and A. G. Riess, Tensions between the early and the late Universe, Nat. Astron. **3**, 891 (2019).

[2] S. Weinberg, Photons and gravitons in S matrix theory: Derivation of charge conservation and equality of gravitational and inertial mass, Phys. Rev. **135**, B1049 (1964).

[3] S. Weinberg, Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations, Phys. Rev. **138**, B988 (1965).

[4] S. Deser, Self-interaction and gauge invariance, Gen. Relativ. Gravit. **1**, 9 (1970).

[5] R. M. Wald, Spin-2 fields and general covariance, Phys. Rev. D **33**, 3613 (1986).

[6] M. P. Hertzberg, Constraints on gravitation from causality and quantum consistency, Adv. High Energy Phys. **2018**, 2657325 (2018).

[7] M. P. Hertzberg and M. Sandora, General relativity from causality, J. High Energy Phys. 09 (2017) 119.

[8] G. Amelino-Camelia, Relativity in space-times with short distance structure governed by an observer independent (Planckian) length scale, Int. J. Mod. Phys. D **11**, 35 (2002).

[9] J. Magueijo and L. Smolin, Lorentz Invariance with an Invariant Energy Scale, Phys. Rev. Lett. **88**, 190403 (2002).

[10] V. A. Kostelecky, Gravity, Lorentz violation, and the standard model, Phys. Rev. D **69**, 105009 (2004).

[11] J. Collins, A. Perez, D. Sudarsky, L. Urrutia, and H. Vucetich, Lorentz Invariance and Quantum Gravity: An Additional Fine-Tuning Problem? Phys. Rev. Lett. **93**, 191301 (2004).

[12] D. Mattingly, Modern tests of Lorentz invariance, Living Rev. Relativity **8**, 5 (2005).

[13] P. Horava, Quantum gravity at a Lifshitz point, Phys. Rev. D **79**, 084008 (2009).

[14] J. Khoury, G. E. J. Miller, and A. J. Tolley, On the origin of gravitational Lorentz covariance, Classical Quantum Gravity **31**, 135011 (2014).

[15] J. Collins, A. Perez, and D. Sudarsky, Lorentz invariance violation and its role in quantum gravity phenomenology, in *Approaches to Quantum Gravity*, edited by D. Oriti (Cambridge University Press, Cambridge, 2006), p. 528–547.

[16] C. Charmousis, G. Niz, A. Padilla, and P. M. Saffin, Strong coupling in Horava gravity, J. High Energy Phys. 08 (2009) 070.

[17] A. Papazoglou and T. P. Sotiriou, Strong coupling in extended Horava-Lifshitz gravity, Phys. Lett. B **685**, 197 (2010).

[18] T. Griffin, K. T. Grosvenor, C. M. Melby-Thompson, and Z. Yan, Quantization of Horava gravity in 2 + 1 dimensions, J. High Energy Phys. 06 (2017) 004.

[19] J. W. Moffat, Superluminary universe: A possible solution to the initial value problem in cosmology, Int. J. Mod. Phys. D **02**, 351 (1993).

[20] E. Pajer, D. Stefanyszyn, and J. Supeł, The boostless bootstrap: Amplitudes without Lorentz boosts, arXiv:2007.00027.

[21] M. P. Hertzberg and M. Sandora, Special relativity from soft gravitons, Phys. Rev. D **96**, 084048 (2017).

[22] M. P. Hertzberg and J. A. Litterer, Symmetries from locality. I. Electromagnetism and charge conservation, Phys. Rev. D **102**, 025022 (2020).

[23] M. P. Hertzberg, J. A. Litterer, and M. Sandora, Symmetries from locality. II. Gravitation and Lorentz boosts, Phys. Rev. D **102**, 025023 (2020).

[24] M. Tanabashi *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **98**, 030001 (2018).
[25] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, Causality, analyticity and an IR obstruction to UV completion, J. High Energy Phys. 10 (2006) 014.
[26] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116**, 061102 (2016).
[27] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz, Effective field theory for massive gravitons and gravity in theory space, Ann. Phys. (Amsterdam) **305**, 96 (2003).
[28] S. Deser and A. Waldron, Acausality of Massive Gravity, Phys. Rev. Lett. **110**, 111101 (2013).
[29] S. Deser, M. Sandora, A. Waldron, and G. Zahariade, Covariant constraints for generic massive gravity and analysis of its characteristics, Phys. Rev. D **90**, 104043 (2014).
[30] D. Lovelock, The Einstein tensor and its generalizations, J. Math. Phys. (N.Y.) **12**, 498 (1971).
[31] B. P. Abbott *et al.* (LIGO Scientific, Virgo, Fermi-GBM, and INTEGRAL Collaborations), Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A, Astrophys. J. Lett. **848**, L13 (2017).