CALT-TH 2014-150

Covariant constraints for generic massive gravity and analysis of its characteristics

S. Deser[ℭ], M. Sandora[𝔊,𝔇], A. Waldron[𝔐] and G. Zahariade[𝔇]

[ℭ] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125; Physics Department, Brandeis University, Waltham, MA 02454.*
deser@brandeis.edu

[𝔊] *CP3-Origins, Center for Cosmology and Particle Physics Phenomenology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*

[𝔇] *Department of Physics University of California, Davis CA 95616, USA*
mesandora@ucdavis.edu, zahariad@ucdavis.edu

[𝔐] *Department of Mathematics University of California, Davis CA 95616, USA*
wally@math.ucdavis.edu

Abstract

We perform a covariant constraint analysis of massive gravity valid for its entire parameter space, demonstrating that the model generically propagates five degrees of freedom; this is also verified by a new and streamlined Hamiltonian description. The constraint's covariant expression permits computation of the model's caustics. Although new features such as the dynamical Riemann tensor appear in the characteristic matrix, the model still exhibits the pathologies uncovered in earlier work: superluminality and likely acausalities.

1 Introduction

Massive gravity (mGR) models defined in terms of a fiducial metric have been intensely studied in recent years in the hope of providing an observationally viable, finite range, extension of Einstein's general relativity (GR) [1]. This spate of activity occurred despite the fact that no definitive analysis of fiducial massive gravity (fmGR[1]) propagation and causal properties valid for its full parameter range had been undertaken; this is our aim: Our findings bolster earlier ones of both acausality and superluminality. The key technical advance enabling these computations is the first covariant degree of freedom (DoF) analysis valid for the model's full parameter range. We will also present an improved Hamiltonian analysis as a check on these findings.

It was realized long ago that interacting higher spin $s \geq 1$ fields can suffer from a variety of inconsistencies. The first issue is that the field theoretic propagating DoF of the interacting theory may not match those of its free limit. As shown in [2], generic massive gravity theories fail at this first hurdle. However, even models passing this first consistency barrier—in particular fmGR—may still propagate unphysical modes. This phenomenon was first observed in the context of the canonical quantum commutators of charged spin 3/2 fields; they were found to be pathological in EM backgrounds [3]. That pathology was later traced back to the underlying kinetic structure of the theory. The latter was studied by searching for superluminal shock wave solutions to the underlying PDEs [4] and extended to spin 2 in [5]. Shock waves propagate on characteristic surfaces, off of which the evolution of all physical variables is no longer determined. This explains why zero and negative norm states appear in canonical commutators. In background-independent GR, the characteristic surfaces encode the causal structure of the theory and are not fatal *per se*. However, if one takes this viewpoint (thus abandoning fmGR consistency as a spin 2 field theory in its fiducial background), there remains the further requirement that solutions with local[2] closed timelike curves (CTCs) be absent. These are notoriously difficult to avoid in models with field-dependent characteristic matrices [6].

The first fmGR model was given by Zumino in 1970 [7], by setting one of the two dynamical metrics of the, then new, bimetric "f-g" theories of Isham, Salam and Strathdee [8] to a fixed (fiducial) background and requiring the free limit to be the massive, $s = 2$ Fierz–Pauli (FP) theory. However, it was soon realized [2] that mGR models generically included an additional, sixth, ghost-like, zero helicity, field theoretic DoF. Furthermore, even (linear)

[1] These models have also been dubbed "dRGT mGR" or "ghost-free mGR". We prefer the more descriptive fmGR title, because these models are not ghost-free (see below) but rather only avoid the obvious sixth, zero helicity, ghostlike excitation. Moreover their inconsistencies can be traced back to the external-fiducial background.

[2] The less pernicious global CTCs of Gödel type are in principle still permitted.

FP theory was found to predict incorrect results for bending of light in its vanishing mass limit [9]. Subsequently it was argued that this difficulty could be an artefact of the linearized limit–setting the interaction strength to zero before the massless limit could cause the faulty light-bending predictions [10]. Alas, in the absence of a consistent interacting massive model, this suggestion was very difficult to verify (although it was shown that a similar mechanism for the FP model in cosmological backgrounds, interchanging limits of vanishing mass and cosmological constant did cure the light-bending disease [11]). This set the stage for effective field theorists to apply the decoupling limit (large Planck mass $M_{\rm P}$, small graviton mass m and constant $m^2 M_{\rm P}$) technique to study fmGR's dangerous zero helicity sector. Remarkably, they recovered Zumino's original fmGR model plus two further extensions as candidate ghost-free theories [1].

At this point, a frenzy of mGR activity ensued (see the reviews [12]); but some darker clouds had gathered on the horizon: An intricate, $(3+1)$, ADM constraint analysis verified that the fmGR models propagated five field theory DoF but cast little light on its kinetic structure [13], except that it was rather complicated–to be precise, various implicit field redefinitions were needed, yielding a potentially pathological symplectic current. Indeed, already in the decoupling limit superluminalities had been detected [14]. This indicated that the difficulties faced by other (finite tower) higher spin models would likely befall fmGR. Indeed, a second order shock analysis discovered fmGR superluminalities, at least for a one-dimensional subspace of its allowed parameter values [15]; this result was extended to a two-dimensional subspace in [16]. These were later shown to be consequences of superluminal behavior detected via a first order computation of the model's characteristic matrix [17] (see also [18]). Worse still, this first order computation showed how to use superluminality to embed closed timelike loops and thus violate microcausality. Hand in hand with those results, it was also discovered that fmGR possessed no consistent partially massless limit [16, 19]. [Since partially massless theories were originally discovered by demanding lightlike propagation [20], and underlie the cosmological solution to the light-bending problem [11], this constitutes strong evidence against fmGR consistency.]

In this article we extend earlier constraint and propagation analyses to the full fmGR parameter range. This requires not only toil but significant additional grist, because previous covariant constraint analyses failed when applied to the remaining direction in the fmGR parameter space. This was because—seemingly non-removable—terms appeared in the putative scalar constraint that involved the full dynamical Riemann tensor. Given that previous (3+1) constraint analyses for this case were rather implicit [13], absence of the field theoretic ghost in this corner of fmGR parameter space was not confirmed by the first covariant analysis of [21]. [Other groups have investigated the full parameter space, but only for specialized field configurations and agree with our result [22].] We solve that problem by using completely covariant methods. This also allows us to compute the characteristic matrix for the full fmGR parameter space; indeed the formerly troublesome Riemann terms imply a new dependence of the characteristic matrix on dynamical curvatures.

The characteristic matrix is a powerful tool for examining consistency of models. Ultimately, fmGR proponents would need to show non-vanishing of its determinant to save the model from pathology. This criterion could be used both to discover a preferred parameter choice and to determine a preferred fiducial background. [The freedom to choose by hand the fiducial metric in order to fit data implies a massive loss of predictability.] However, in [17] fatal acausalities for very general field configurations and independent of choice of fiducial metric were uncovered, so the range of physical viability of fmGR theories is likely to be highly limited at best. In this article, we content ourselves with exhibiting the characteristic method at work for some simple examples around flat fiducial backgrounds.

Our article is structured as follows. Our covariant constraint analysis is given in Section 2 and the characteristic matrix is computed in Section 3. We analyze the characteristic matrix for propagation pathologies in Section 4. Our conclusions, where we discuss fmGR's last vestige of applicability as an effective field theory as well as related models such as the bimetric theory where the fiducial background is promoted to a dynamical field, are given in Section 5. In Appendix A we present the linear limit of our first order, covariant constraint analysis, while Appendix B gives a rapid sketch of the model's frame-like Hamiltonian, description from which the DoF count can also be checked.

2 Covariant constraint analysis

At their genesis, bimetric [8] and massive gravity [7] were originally formulated in terms of vierbeine e^m and f^m. Both these fields are dynamical for the bimetric theory, while for mGR, f^m is taken to be a fiducial background (*e.g.*, $\bar{g}_{\mu\nu} := f_\mu{}^m \eta_{mn} f_\nu{}^n$). In these terms, the statement of the model and its constraint analysis are rather simple.

Throughout this Section, unless explicitly noted, we will use a differential form notation where wedge products are assumed whenever obvious. The action is a sum of Einstein–Hilbert and mass terms:

$$S_{\rm fmGR} := S_{\rm EH} + S_m \ ,$$

where

$$\begin{aligned} S_{\rm EH} \ &:= \ -\frac{1}{4}\int \epsilon_{mnrs}\, e^m e^n \, \left[d\omega^{rs} + \omega^r{}_t \omega^{ts}\right] \ , \\ S_m \ &:= \ m^2 \int \epsilon_{mnrs}\, e^m \left[\frac{\beta_0}{4} e^n e^r e^s + \frac{\beta_1}{3} e^n e^r f^s + \frac{\beta_2}{2} e^n f^r f^s + \beta_3 f^n f^r f^s\right] \ . \end{aligned}$$

Note that β_0 parameterizes a standard cosmological term (which will be required to obtain the FP linearized limit, even around flat, Minkowski, backgrounds), while a β_4 term made from four fiducial vierbeine only contributes an irrelevant additive constant, so has been omitted. It is known that the model can be linearized around fiducial Einstein backgrounds with cosmological constant $\bar{\Lambda}$ only if the model's parameters obey

$$\frac{\bar{\Lambda}}{3!} = m^2 \left(\beta_0 + \beta_1 + \beta_2 + \beta_3\right) . \tag{1}$$

Its linearized limit (see Appendix A) is then the FP theory with mass

$$m_{\rm FP}^2 := m^2(\beta_1 + 2\beta_2 + 3\beta_3) . \tag{2}$$

The model's dynamical fields are the vierbein and spin-connection one-forms (e^m, ω^{mn}) whose variations give equations of motion:

$$\begin{aligned} \mathcal{T}^m &:= \nabla e^m \approx 0 , \\ \mathcal{G}_m &:= G_m - m^2 t_m \approx 0 . \end{aligned} \tag{3}$$

Equations which hold on-shell are written using the weakly vanishing notation $\approx$ and a calligraphic font will be used for weakly vanishing quantities. The first equation implies vanishing torsion so that the spin-connection weakly equals the Levi-Civita one. The second equation is the standard Einstein equation modified by the mass term. In the above, we have denoted the exterior covariant derivative with respect to ω^{mn} by ∇ so that for any Lorentz vector-valued form σ^m

$$\nabla\sigma^m := d\sigma^m + \omega^m{}_n\sigma^n .$$

Moreover we have defined the Einstein three-form

$$G_m := \frac{1}{2}\epsilon_{mnrs}e^n R^{rs} , \tag{4}$$

where the two-form $R^{mn} := d\omega^{mn} + \omega^m{}_t\omega^{tn}$ is the Riemann curvature associated to the connection. The dual of the display (4) is the Einstein tensor. Finally the *mass stress-tensor* is encoded by the three-form [23]

$$t_m := \epsilon_{mnrs}\left[\beta_0 e^n e^r e^s + \beta_1 e^n e^r f^s + \beta_2 e^n f^r f^s + \beta_3 f^n f^r f^s\right] . \tag{5}$$

To analyze the model's constraints we need a notion of timelike evolution. For that, one assumes invertibility of the dynamical vierbein and in turn of the metric $ds^2 = e^m \otimes e_m$, which is taken to have signature $(-,+,+,+)$. [Our analysis easily extends to arbitrary dimensions

$d \geq 3$, see footnote 8.] Hence, for any choice of timelike[3] evolution parameter t we can decompose a p-form θ (with $p < 4$) as

$$\theta := \boldsymbol{\theta} + \mathring{\theta}\,, \tag{6}$$

where $\mathring{\theta} \wedge dt = 0$. Thus $\boldsymbol{\theta}$ is the purely spatial part of the form θ. The beauty of this notation is that the purely spatial $\boldsymbol{\mathcal{P}} \approx 0$ part of any on-shell relation $\mathcal{P} \approx 0$ polynomial in (∇, e, ω) is a constraint because it necessarily contains no time derivatives. Our analysis proceeds as follows. The forty first order equations of motion for forty fields ultimately describe (at least generically) ten propagating fields, so five physical DoF. To establish this result in a simple covariant formalism, we therefore need to find thirty constraints, *i.e.*, weak relations not involving time derivatives of dynamical fields. Sixteen of these are given directly by the equations of motion themselves and are thus primary constraints. Evolving these gives ten secondary constraints whose evolution in turn yields the final four tertiary constraints.

2.1 Primary constraints

The spatial parts of the equations of motion (3) give sixteen primary constraints

$$\begin{aligned} \boldsymbol{\mathcal{T}}^m &\approx 0\,, \\ \boldsymbol{\mathcal{G}}_m &\approx 0\,. \end{aligned} \tag{7}$$

In terms of dynamical fields, these read

$$\boldsymbol{\nabla} \boldsymbol{e}^m := \boldsymbol{d}\boldsymbol{e}^m + \boldsymbol{\omega}^m{}_n \boldsymbol{e}^n \approx 0\,,$$

$$\tfrac{1}{2}\epsilon_{mnrs}\boldsymbol{e}^n \left(\boldsymbol{d}\boldsymbol{\omega}^{rs} + \boldsymbol{\omega}^n{}_t \boldsymbol{\omega}^{ts}\right) \approx m^2 \epsilon_{mnrs} \left(\beta_0 \boldsymbol{e}^n \boldsymbol{e}^r \boldsymbol{e}^s + \beta_1 \boldsymbol{e}^n \boldsymbol{e}^r \boldsymbol{f}^s + \beta_2 \boldsymbol{e}^n \boldsymbol{f}^r \boldsymbol{f}^s + \beta_3 \boldsymbol{f}^n \boldsymbol{f}^r \boldsymbol{f}^s\right)\,.$$

2.2 Secondary constraints

In principle we could compute secondary constraints by brute force by taking a time derivative of the primary constraints (7). That computation is vastly simplified by considering their exterior covariant derivatives. The purely spatial part of this is of course not a new constraint, but the remainder, modulo the field equations can possibly yield new, secondary, constraints.

[3]Here we mean timelike with respect to the dynamical metric, although none of the constraints found in this Section depend essentially on the choice of foliation of the underlying spacetime manifold.

2.2.1 The symmetry constraint

On-shell, $G_m \wedge e_n$ is equal to the volume form multiplied by the Einstein tensor and thus symmetric under interchange of m and n. Indeed,

$$G_{[m}e_{n]} = \frac{1}{2}\epsilon_{mnrs}e^r \nabla \mathcal{T}^s \approx 0 \ .$$

This leads to six secondary constraints

$$t_{[m}e_{n]} = \frac{1}{m^2}\left(\frac{1}{2}\epsilon_{mnrs}e^r \nabla \mathcal{T}^s - \mathcal{G}_{[m}e_{n]}\right) \approx 0 \ .$$

Using the Schouten identity[4], this gives

$$M^{mn}\mathcal{F} \approx 0 \ , \tag{8}$$

in terms of the two-forms

$$\mathcal{F} := e^m f_m \quad \text{and} \quad M^{mn} := \beta_1 e^m e^n + 2\beta_2 e^{[m} f^{n]} + 3\beta_3 f^m f^n \ .$$

In (8) the operator M^{mn} maps two-forms to antisymmetric Lorentz tensors (multiplied by the volume form) and is therefore generically invertible. Thus, in the following we will assume

$$\mathcal{F} \approx 0 \tag{9}$$

(hence the calligraphics) even if this is technically not implied for all regions of parameter space[5] [24]. We call this the *symmetry constraint*.

2.2.2 The vector constraint

The Einstein tensor's Bianchi identity implies that diffeomorphism invariant metric DoF must be coupled to divergence-free sources. In the fmGR context, this yields a constraint. Here computing the covariant curl of the Einstein three-form, using $\nabla R^{mn} \equiv 0$, gives

$$\nabla G_m = \frac{1}{2}\epsilon_{mnrs}\mathcal{T}^n R^{rs} \approx 0 \ .$$

[4]This (tautological) identity states $v_m\epsilon_{nrs\ldots} = v_n\epsilon_{mrs\ldots} + v_r\epsilon_{nms\ldots} + v_s\epsilon_{nrm\ldots} + \cdots$.

[5]For example, when $\beta_1 = \beta_3 = 0$ the operator above is not invertible. In the cases $\beta_2 = \beta_3 = 0$ and $\beta_1 = \beta_2 = 0$ as well as $\beta_2 = \beta_1\lambda$, $\beta_3 = \frac{\beta_1\lambda^2}{3}$ (so long as $e^m + \lambda f^m$ is a basis of the cotangent space) equation (8) does imply $\mathcal{F} \approx 0$.

This leads to the four constraints

$$\nabla t_m = \frac{1}{m^2}\left(\frac{1}{2}\epsilon_{mnrs}\mathcal{T}^n R^{rs} - \nabla \mathcal{G}_m\right) \approx 0 \ . \tag{10}$$

We can write these constraints explicitly because

$$\begin{aligned}\nabla t_m = &\ \epsilon_{mnrs}\mathcal{T}^n\left(3\beta_0 e^r e^s + 2\beta_1 e^r f^s + \beta_2 f^r f^s\right) + \ \epsilon_{mnrs} M^{nr} K^{st} f_t \\ = &\ \epsilon_{mnrs}\left[\mathcal{T}^n\left(3\beta_0 e^r e^s + 2\beta_1 e^r f^s + \beta_2 f^r f^s\right) + \mathcal{F}\left(\beta_1 e^n + \beta_2 f^n\right) K^{rs}\right] - \frac{1}{2}\epsilon_{nrst} M^{nr} K^{st} f_m \ .\end{aligned}$$

Here we have defined the contorsion

$$K^m{}_n := \omega^m{}_n - \bar{\omega}^m{}_n \ ,$$

where $\bar{\omega}^m{}_n$ is the fiducial Levi-Civita spin connection. The contorsion measures the difference between dynamical and fiducial spin-connections, thus

$$\nabla f^m = K^m{}_n f^n \ .$$

Hence, using invertibility of the fiducial vierbein, we finally[6] have the *vector constraint*

$$\mathcal{V} := \epsilon_{mnrs} M^{mn} K^{rs} \approx 0 \ . \tag{11}$$

2.3 Tertiary constraints

We must now compute the time evolution of the ten secondary constraints, comprised of the six symmetry (8), and four vector (11) constraints. This will lead, respectively, to three and one additional tertiary constraints.

2.3.1 Evolving the symmetry constraint

Since the symmetry constraint is the weak vanishing of the two-form $\mathcal{F}$, we can simply take the covariant curl of (8) to generate tertiary constraints:

$$\nabla \mathcal{F} = \mathcal{T}^m f_m + K_{mn} e^m f^n \approx 0 \, ,$$

which yield the three-form, *curled symmetry constraint*:

$$\mathcal{K} := K_{mn} e^m f^n \approx 0 \ . \tag{12}$$

[6] For the reason mentioned above, the equation $\boldsymbol{\nabla}\boldsymbol{\mathcal{T}}^m = \boldsymbol{R}^m{}_n \boldsymbol{e}^n \approx 0$ yields no further secondary constraints since it is the spatial derivative of $\boldsymbol{\mathcal{T}}^m \approx 0$.

This three-form might seem to constitute four new constraints, but exactly as above, its purely spatial part is just the spatial derivative of the symmetry constraint and hence not new. Therefore, in the notation of (6), we have three new constraints

$$\mathring{K}^{mn}\boldsymbol{e}_m\boldsymbol{f}_n + \boldsymbol{K}^{mn}\mathring{e}_m\boldsymbol{f}_n + \boldsymbol{K}^{mn}\boldsymbol{e}_m\mathring{f}_n \approx 0 \ .$$

2.3.2 Evolving the vector constraint

Since the vector constraint is the weak vanishing of the three-form $\mathcal{V}$, we take the curl of (11) to generate the final, scalar, tertiary constraint:

$$\nabla\mathcal{V} = \epsilon_{mnrs}\left[2\mathcal{T}^m\left(\beta_1 e^n + \beta_2 f^n\right)K^{rs} - \left(2\beta_2 e^m + 6\beta_3 f^m\right)K^{nr}K^s{}_t f^t + M^{mn}\nabla K^{rs}\right] \approx 0 \ .$$

At first glance, the above equation is not obviously a constraint because the last term, involving the curl of the contorsion, could contain a time derivative of the dynamical spin connection. To see that this is not the case, we begin with an identity:

$$\nabla K^{mn} \equiv R^{mn} - \bar{R}^{mn} + K^m{}_t K^{tn} \ , \tag{13}$$

where the two-form $\bar{R}^{mn}$ is the fiducial Riemann curvature. This shows that dangerous time derivatives can only arise via the dynamical Riemann curvature R^{mn}. However, the equations of motion (3) tell us that the Einstein tensor $G_{\mu\nu}$ weakly equals terms involving no time derivatives (namely the mass stress tensor). Moreover, standard identities for the Riemann tensor all hold weakly, in particular its divergence is related to the curl of the Einstein tensor by

$$\nabla^\mu R_{\mu\nu\rho\sigma} \approx 2\nabla_{[\rho}\Big(G_{\sigma]\nu} - \frac{1}{2}g_{\sigma]\nu}G_\mu{}^\mu\Big) \, . \tag{14}$$

Hence, on-shell, the quantity $\partial_t R_{0\nu\rho\sigma}$ generically has at most one time derivative on dynamical fields[7], so the only dangerous Riemann components $R_{0\nu\rho\sigma}$ have none and hence, in turn, nor does the curvature R^{mn}. [This simple covariant argument is also readily verified by writing out the equations of motion (3) in an explicit $3+1$ split for any choice of time coordinate.] Therefore

$$\mathcal{S} := \epsilon_{mnrs}\Big[M^{mn}\nabla K^{rs} - \left(2\beta_2 e^m + 6\beta_3 f^m\right)K^{nr}K^s{}_t f^t\Big] \approx 0 \tag{15}$$

is a constraint equation. This is the long-sought *scalar constraint*. In the special case $\beta_3 = 0$, the curl of the contorsion in the above display is traced with the *dynamical* vierbein and thus can be converted to a trace of the Riemann tensor, *i.e.* the Einstein tensor, which can be

[7] In more detail, $\nabla^\mu R_{\mu\nu\rho\sigma} = \nabla^0 R_{0\nu\rho\sigma} + \cdots = g^{00}\dot{R}_{0\nu\rho\sigma} + \cdots$, where the "$\cdots$" terms involve at most one time derivative of the dynamical fields (e^m, ω^{mn}).

handled directly by the equations of motion. In contrast, when $\beta_3 \neq 0$, the dangerous Weyl part of Riemann is only traced with fiducial vierbeine so one must rely on equation (14) to prove that $\mathcal{S}$ is a constraint. This explains why previous works [21, 15, 16, 24] failed to find a covariant expression valid for the entire $(\beta_0, \beta_1, \beta_2, \beta_3)$ parameter space.

It will be useful to have a more explicit expression for the scalar constraint. For that, we employ equations (4,13), and the Schouten identity (see footnote 4) to rewrite it as

$$\begin{aligned}\mathcal{S} \;=\; &\epsilon_{mnrs}\left(\beta_1 e^m e^t - 2\beta_2 e^{(m} f^{t)} - 3\beta_3 f^m f^t\right) K^{nr} K^s{}_t \\ &+ 2\;(\;\beta_1 e^m + 2\beta_2 f^m)\, G_m + 3\epsilon_{mnrs}\beta_3 f^m f^n R^{rs} \\ &- 2\,(2\beta_2 e^m + 3\beta_3 f^m)\,\bar{G}_m - \epsilon_{mnrs}\beta_1 e^m e^n \bar{R}^{rs} \approx 0 \;,\end{aligned} \tag{16}$$

where $\bar{G}_m := \frac{1}{2}\epsilon_{mnrs} f^n \bar{R}^{rs}$ is the background Einstein tensor. Here one can exchange the dynamical Einstein tensor for the mass stress tensor $G_m \approx m^2 t_m$, whose explicit expression (algebraic in dynamical fields) is given in (5). Moreover, remember that for the stubborn case $\beta_3 \neq 0$, the term involving the Riemann tensor (weakly) does not depend on time derivatives of dynamical variables. The above expression coincides with known results for the covariant scalar constraint for $\beta_3 = 0$ [15, 16]. Also, specializing to the case $\beta_0 = \beta_1 = \beta_3 = 0$ and choosing the partially massless tuning of β_2 to the background cosmological constant [20], only the terms involving the square of the contorsion remain. These are precisely the obstruction to a partially massless limit of massive gravity [16, 19].

At this point, so long as we can establish that the thirty constraints found so far are independent, the model describes no more than five physical DoF[8]. For the subspace of parameter space given by models which linearize to FP, this is essentially guaranteed (see Appendix A). The possibility that fewer DoF propagate, especially in special limits, such as the massless Einstein or a putative partially massless limit remains. The former of course, holds, but the latter possibility was ruled out in [15, 16]. It could also be that, for parameter branches where the symmetry constraint is not guaranteed, fewer DoF propagate.

3 The Characteristic Matrix

For a system of coupled, first order PDEs, we must ask whether, given initial data, higher derivatives of fields are determined. This question is addressed by the system's characteristic

[8] In $d \geq 3$ dimensions the model propagates $\frac{1}{2}(d+1)(d-2)$ physical DoF, which can be seen as follows: There are $d^2 + \frac{1}{2}d^2(d-1)$ dynamical vielbeine and spin connections. These are subject to $d + \frac{1}{2}d(d-1)(d-2)$ primary, $d + \frac{1}{2}d(d-1)$ secondary, and $1 + \frac{1}{2}(d-1)(d-2)$ tertiary constraints. This leaves $(d+1)(d-2)$ first order propagating fields which yields the quoted DoF count.

matrix which can be computed by studying shocks. In particular, the characteristic surface is defined by vanishing of the corresponding determinant. Shocks propagate along this surface–where all higher derivatives can no longer be determined. In particular, spacelike characteristic surfaces signal superluminal shock propagation. These foretell doom for the model viewed as a theory of a spin 2 field in a fiducial background. This method also allows us to study whether the model can access the escape route taken by (background independent) GR, whose characteristics determine its physical causal structure. Thus we now study fmGR shocks along surfaces with timelike normal vectors. These determine the model's characteristic matrix from its equations of motion and constraints, as given in the previous Section. More precisely, we study the characteristics of the following set of *seventy-six* first order PDEs

$$\begin{aligned}
\mathcal{T}^m &:= \nabla e^m \approx 0\ , \\
\mathcal{G}_m &:= G_m - m^2 t_m \approx 0\,, \\
\mathcal{R}^{mn} &:= R^{mn} - d\omega^{mn} - \omega^m{}_t \omega^{tn} \approx 0\ .
\end{aligned} \tag{17}$$

The first forty of these are familiar from the initial set (3), while the remaining thirty-six (trivial) equations have been introduced in order to also treat the Riemann curvature as an *independent variable* and thus handle efficiently the curl of the contorsion in the scalar constraint (15). Hence there are seventy-six dynamical fields $(e^m, \omega^{mn}, R^{mn})$.

We begin our study by assuming the existence of a spacelike, with respect to the dynamical[9] metric $g_{\mu\nu}$, characteristic surface Σ; this can be thought of as the world-sheet of a shock-wavefront propagating at superluminal speeds. More precisely it is characterized as a surface where the first derivatives of the dynamical fields suffer discontinuities in the direction of the normal ξ^μ to Σ; the discontinuity of any quantity q across this surface will be denoted by

$$[\,q\,] := q\Big|_{\Sigma+} -\ q\Big|_{\Sigma-}\ .$$

In particular, for the dynamical fields

$$\begin{aligned}
[\partial_\mu e_\nu{}^m] &:= \xi_\mu \mathfrak{e}_\nu{}^m\,, \\
[\partial_\mu \omega_\nu{}^{mn}] &:= \xi_\mu \mathfrak{w}_\nu{}^{mn}\,, \\
[\partial_\mu R_{\rho\sigma}{}^{mn}] &:= \xi_\mu \mathfrak{R}_{\rho\sigma}{}^{mn}\ .
\end{aligned}$$

Since Σ is spacelike, the normal vector obeys

$$\xi^\mu \xi_\mu = -1\,;$$

[9]One might also choose fiducially spacelike surfaces. This does noes not alter the superluminality conclusions below. Our choice enables us to also study dynamical acausalities.

there is no loss of generality in normalizing ξ. Throughout indices will be manipulated using the dynamical metric and vierbein. The forms $(\mathfrak{e}^m, \mathfrak{w}^{mn}, \mathfrak{R}^{mn})$ are the tensors that characterize the shock-wave profile.

The discontinuities in the field equations (17), along with those of their constraints (see the preceding Section), determine whether spacelike characteristic surfaces and concomitant superluminalities are permitted: Computing the discontinuity across Σ of the equations of motion and gradients of constraints gives a linear homogeneous system of equations in the shock-wave profiles of the form

$$\chi \begin{pmatrix} \mathfrak{e}^m \\ \mathfrak{w}^{mn} \\ \mathfrak{R}^{mn} \end{pmatrix} \approx 0 \ . \tag{18}$$

Here χ is called the *characteristic matrix*; if it is invertible, space-like characteristics are excluded. Note that a field-dependent characteristic matrix usually foretells non-invertibility and thus superluminality.

3.1 The strategy

The characteristic matrix analysis is streamlined by introducing a natural orthonormal basis $(\xi := \xi_\mu dx^\mu, \varepsilon^i)$ for the cotangent spaces along the characteristic hypersurface. Any tensorial quantity can be expressed in terms of this basis and its dual, for example a one-form becomes

$$\theta = -\theta_o \, \xi + \theta_i \varepsilon^i := \boldsymbol{\theta} + \mathring{\theta} \, .$$

With these definitions, $\xi^o = 1$ and $\xi^i = 0$, while $\xi_o = -1$. Moreover $g_{oo} = -1$, $g_{ij} = \delta_{ij}$ and $g_{oi} = 0$. This split into timelike and spatial parts defined by the shock wave-front allows us to adopt a notation similar to that of the previous section for differential forms: any p-form θ (with $p < 4$) can be decomposed as

$$\theta := \boldsymbol{\theta} + \mathring{\theta} \, ,$$

where

$$\mathring{\theta} \wedge \xi = 0 \, .$$

Thus $\boldsymbol{\theta}$ is the purely spatial part of the form θ.

In the above basis (modulo judicious field redefinitions) the characteristic equation (18) will take the block form

$$\left(\begin{array}{c|c|c} \mathbf{1} & 0 & 0 \\ \hline \star & \mathbf{1} & \star \\ \hline \star & \star & \star \end{array}\right) \left(\begin{array}{c} \mathfrak{O} \\ \hline \mathfrak{T} \\ \hline \mathfrak{E} \end{array}\right) \approx 0 \ ,$$

where $(\mathfrak{O}, \mathfrak{T}, \mathfrak{E})$ are linear combinations of the shock-wave profiles $(\mathfrak{e}^m, \mathfrak{w}^{mn}, \mathfrak{R}^{mn})$. The form above implies that the *null part* $\mathfrak{O} = 0$ and allows us to solve for the *trivial part* $\mathfrak{T}$ as functions of the *essential part* $\mathfrak{E}$. This gives the *reduced characteristic equation*

$$\widehat{\chi}\, \mathfrak{E} \approx 0 \, . \tag{19}$$

3.2 The null and trivial parts

To obtain the null part of the characteristic equation, our first step is to compute the shock in the equations of motion (17):

$$\begin{aligned} [\mathcal{T}^m] &= \xi\, \mathfrak{e}^m \approx 0 \, , \\ [\mathcal{G}_m] &= \epsilon_{mnrs} e^n \xi\, \mathfrak{w}^{rs} \approx 0 \, , \\ [\mathcal{R}^{mn}] &= \xi\, \mathfrak{w}^{mn} \approx 0 \, . \end{aligned}$$

In general, vanishing of the wedge product of the one-form ξ and a k-form X implies $X_{i_1 i_2 \ldots i_k} = 0$; thus the first and third relation above imply

$$\mathfrak{e}_i{}^m \approx 0 \approx \mathfrak{w}_i{}^{mn} \, ,$$

while the second gives no new information. There are further relations contributing to the null part of the characteristic equation; to obtain those, we first notice that the curl of the trivial equation in (17) gives $\nabla \mathcal{R}^{mn} = \nabla R^{mn} \approx 0$, whose shock yields

$$[\nabla \mathcal{R}^{mn}] = \xi\; \mathfrak{R}^{mn} \approx 0 \, ,$$

so that

$$\mathfrak{R}_{ij}{}^{mn} \approx 0 \, .$$

Further relations on the shock-wave profiles can be obtained by studying the discontinuities in the gradients of the constraints. Our analysis is further simplified by using variables that maximize the null part of the characteristic equation. We can indeed use the variable $f_{\mu\nu} = e_\nu{}^m f_{\mu m}$ instead of the dynamical vierbein, so long as the fiducial vierbein is taken to be invertible. Calling its shock-wave profile $\mathfrak{f}_{\mu\nu}$, we have

$$[\partial_\mu f_{\nu\rho}] := \xi_\mu \mathfrak{f}_{\nu\rho} \quad \text{and} \quad \mathfrak{f}_{\nu\rho} = \mathfrak{e}_\mu{}^m f_{\nu m} \, ,$$

because all fiducial quantities are assumed to be smooth across Σ. Since the symmetry constraint (8) then says

$$\mathcal{F}_{\mu\nu} = f_{[\mu\nu]} \approx 0 \, ,$$

taking the shock of its gradient we have

$$-\xi^\mu\left[\partial_\mu \mathcal{F}_{\mu\nu}\right] = \left[\partial^o \mathcal{F}_{\mu\nu}\right] = \mathfrak{f}_{[\mu\nu]} \approx 0\,.$$

In turn, since $\mathfrak{e}_i{}^m \approx 0$, it follows that of the shock-wave profiles $\mathfrak{f}_{\mu\nu}$, only $\mathfrak{f}_{oo} \not\approx 0$. This has several very useful consequences, in particular

$$\mathfrak{f}_{\mu\nu} = \xi_\mu \xi_\nu \mathfrak{f}_{oo} \quad \text{and} \quad \mathfrak{e}_{\mu\nu} = \xi_\mu l_{o\nu} \mathfrak{f}_{oo}\,,$$

so that

$$\left[\partial^o g_{\mu\nu}\right] = 2\mathfrak{e}_{(\mu\nu)} = 2 l_{o(\mu}\xi_{\nu)} \mathfrak{f}_{oo}\,,$$

where $l^\mu{}_m$ is the inverse fiducial vierbein. Hence the shock in the Christoffel symbols is

$$\left[\Gamma^\rho_{\mu\nu}\right] = \xi_\mu \xi_\nu l_o{}^\rho \mathfrak{f}_{oo}\,.$$

This allows us to compute the remaining $\mathfrak{R}_{\mu\nu}{}^{mn}$ shock-wave profiles in terms of $\mathfrak{f}_{oo}$. For that we study the shock in the relation (14). Because the shock in the gradient of the vierbein is proportional to $\mathfrak{f}_{oo}$, the same applies for the shock of the gradient of the mass stress tensor, so we define

$$\left[\partial^o t_{mn}\right] := \tau_{mn} \mathfrak{f}_{oo}\,,$$

where $t_m := \frac{1}{3!}\epsilon_{nrst} t_m{}^n e^r \wedge e^s \wedge e^t$. The tensor τ_{mn} is easily computed and we find

$$\tau_m{}^n = \frac{1}{2}\epsilon_{mrst}\epsilon^{nlpq}\, M_{pq}{}^{st}\xi^r \xi_l\ ,$$

where $M_{pq}{}^{mn} = M_{\mu\nu}{}^{mn} e^\mu{}_p e^\nu{}_q$ are the components of the two-form M^{mn} in the dynamical vierbein basis. Turning to the shock in the relation (14), we use the above to obtain

$$\mathfrak{R}_{o\nu\rho\sigma} \approx \left[l_o{}^\kappa \left(R_{\nu\kappa\rho\sigma} + \xi_\nu R_{o\kappa\rho\sigma}\right) + 2m^2 \xi_{[\rho}\left(\tau_{\sigma]\nu} - \frac{1}{2} g_{\sigma]\nu}\tau^\kappa{}_\kappa\right)\right]\mathfrak{f}_{oo}\,. \tag{20}$$

As a consistency check, one can verify that $\mathfrak{R}_{oo\rho\sigma} = 0$ requires

$$\tau_{o\nu} \approx 0\,,$$

which holds because mass stress tensor is weakly conserved, $\nabla^\mu t_{\mu\nu} \approx 0$; the shock of this relation gives precisely the above. Decomposing the relation (20) gives

$$\mathfrak{R}_{oioj} \approx \left[l_o{}^\kappa R_{i\kappa oj} - m^2\left(\tau_{ij} - \frac{1}{2} g_{ij}\tau^k{}_k\right)\right]\mathfrak{f}_{oo} \quad \text{and} \quad \mathfrak{R}_{oijk} \approx l_o{}^\kappa R_{i\kappa jk}\mathfrak{f}_{oo}\,.$$

This completes the determination of the null and trivial parts of the shock-wave profiles. At this juncture, the only independent shock-wave profiles are $(\mathfrak{f}_{oo}, \mathfrak{w}_{omn})$; these constitute the essential part of the shock-wave profiles and are subject to the reduced characteristic equation (19).

3.3 The reduced characteristic matrix

To compute the reduced characteristic matrix, we begin by searching for relations on the shock-wave profiles $\mathfrak{w}_{omn}$. These come from the shocks of the gradients of the curled symmetry $[\partial^o \mathcal{K}] \approx 0$, and vector $[\partial^o \mathcal{V}] \approx 0$, constraints. These equations can be written in condensed differential form notation upon noticing that

$$\mathfrak{e}^m = \mathring{\mathfrak{e}}^m = \xi\, l_o{}^m \mathfrak{f}_{oo} \quad \text{and} \quad \mathfrak{w}^{mn} = \mathring{\mathfrak{w}}^{mn} = -\xi\, \mathfrak{w}_o{}^{mn} \ .$$

Indeed they are given by

$$\begin{aligned} l_o{}^m \boldsymbol{K}_{mn} \boldsymbol{f}^n \mathfrak{f}_{oo} + \boldsymbol{e}^m \boldsymbol{f}^n \mathfrak{w}_{omn} \approx 0 \ , \\ 2\epsilon_{mnrs} l_o{}^m (\beta_1 \boldsymbol{e}^n + \beta_2 \boldsymbol{f}^n) \boldsymbol{K}^{rs} \mathfrak{f}_{oo} - \epsilon_{mnrs} \boldsymbol{M}^{mn} \mathfrak{w}_o{}^{rs} \approx 0 \ . \end{aligned} \tag{21}$$

It remains only to shock the scalar constraint (16). The key ingredient for this computation is the Riemann tensor shock (20); we find

$$\begin{aligned} &\epsilon_{mnrs} \left(\beta_1 \boldsymbol{e}^m \boldsymbol{e}^t - 2\beta_2 \boldsymbol{e}^{(m} \boldsymbol{f}^{t)} - 3\beta_3 \boldsymbol{f}^m \boldsymbol{f}^t\right) \left(\boldsymbol{K}^{nr} \mathfrak{w}_o{}^s{}_t - \boldsymbol{K}^s{}_t \mathfrak{w}_o{}^{nr}\right) \\ &+2\epsilon_{mnrs} \left(\beta_1 l_o{}^{[m} \boldsymbol{e}^{t]} - \beta_2 l_o{}^{(m} \boldsymbol{f}^{t)}\right) \boldsymbol{K}^{nr} \boldsymbol{K}^s{}_t \mathfrak{f}_{oo} \\ &+2m^2 \epsilon_{mnrs} l_o{}^m \Big(4\beta_0\beta_1 \boldsymbol{e}^n \boldsymbol{e}^r \boldsymbol{e}^s + 3(\beta_1^2 + 2\beta_0\beta_2) \boldsymbol{e}^n \boldsymbol{e}^r \boldsymbol{f}^s \\ &\qquad\qquad +6\beta_1\beta_2 \boldsymbol{e}^n \boldsymbol{f}^r \boldsymbol{f}^s + (\beta_1\beta_3 + 2\beta_2^2) \boldsymbol{f}^n \boldsymbol{f}^r \boldsymbol{f}^s\Big) \mathfrak{f}_{oo} \\ &-3\epsilon_{mnrs} \beta_3 \boldsymbol{f}^m \boldsymbol{f}^n \left(\boldsymbol{\rho}^{rs} + 2m^2 \xi^{[r} \left[\boldsymbol{\tau}^{s]} - \frac{1}{2} \tau^t{}_t \boldsymbol{e}^{s]}\right]\right) \mathfrak{f}_{oo} \\ &-4\beta_2 l_o{}^m \bar{\boldsymbol{G}}_m \mathfrak{f}_{oo} - 2\epsilon_{mnrs} \beta_1 l_o{}^m \boldsymbol{e}^n \bar{\boldsymbol{R}}^{rs} \mathfrak{f}_{oo} \approx 0 \ , \end{aligned} \tag{22}$$

where the one forms ρ^{mn} and τ_m are defined by $\rho^{mn} = \rho_\nu{}^{mn} dx^\nu := l_o{}^\mu R_{\mu\nu}{}^{mn} dx^\nu$ and $\tau_m = \tau_{\mu m} dx^\mu := \frac{1}{2} \epsilon_{mrst} \epsilon_\mu{}^{\nu\alpha\beta} \xi^r \xi_\nu M_{\alpha\beta}{}^{st} dx^\mu$.

Assembling the system of equations (21,22) into matrix form determines the 7×7 reduced characteristic matrix $\hat{\chi}$ as in (19) where the essential part is $\mathfrak{E} = (\mathfrak{w}^{mn}, \mathfrak{f}_{oo})$. This matrix encodes all the necessary information about the well-posedness of the initial value problem for the system of PDEs (17) and hence also (3).

4 Flat fiducial propagation analysis

The characteristic matrix is a powerful tool for analyzing the fmGR parameter space to sort out inconsistent theories. This is because characteristic surfaces signal a loss of hyperbolicity as well as superluminal shock propagation over a dynamical mean field solution. If our aim were to establish complete fmGR consistency, we would have to (i) calculate the determinant of the reduced characteristic matrix $\hat{\chi}$ determined by equations (21) and (22) and (ii) prove that it cannot vanish weakly for any configuration of fields. Of course, one might hope that this singled out a special choice of parameters. This computation is rather involved, and in any case counterexamples for subsets of the parameter space are already known [24]. A discussion of how to generally construct zeros of the characteristic determinant, superluminalities and even how to embed closed timelike curves is given in [17]. Therefore, to illustrate the method, we shall restrict ourselves to analyzing some extremely simple physical configurations that already further restrict the allowed parameters.

If we take both background and fiducial metrics flat and aligned, $g_{\mu\nu} = \bar{g}_{\mu\nu} = \eta_{\mu\nu}$, a vanishing characteristic matrix would signal superluminal shocks in the FP theory[10]. Since the (mean field) contorsions and curvatures vanish, the system (21) immediately yields

$$\mathfrak{w}_{omn} \approx 0\,.$$

The scalar shock is also extremely simple:

$$\Big[m^2(\beta_1 + 2\beta_2 + 3\beta_3)\Big]^2\, \mathfrak{f}_{oo} \approx 0\,.$$

We immediately recognize the left hand side to be $m_{\rm FP}^4$, so non-vanishing FP mass rules out superluminality[11] here.

However, not every field configuration is healthy since *a priori* in a theory of a dynamical metric propagating in a fiducial background, the two lightcones are not guaranteed to be compatible. A simple case is a flat fiducial background and flat dynamical mean field that are not Lorentz-related, for example

$$d\bar{s}^2 = -dt^2 + dx^2 + dy^2 + dz^2 \quad \text{and} \quad ds^2 = -dz^2 + dx^2 + dy^2 + dt^2\,.$$

This configuration solves the equations of motion (3) iff

$$\beta_0 + \beta_1 + \beta_2 + \beta_3 = 0 = \beta_1 + 2\beta_2 + 3\beta_3\,.$$

The first condition coincides with the usual one required for the background to solve the equations of motion (1), while the second implies vanishing FP mass (2). One might therefore already rule out the parameter choice $\beta_1 + 2\beta_2 + 3\beta_3 = 0$ because the interacting DoF count does not equal that of the free (massless spin 2) limit. We shall instead rule out this theory on grounds of superluminality. Consider a putative characteristic constant-z surface[12] (so the

[10] Essentially, the shock in this case is a small perturbation of a continuous Minkowski mean field.
[11] When $m_{\rm FP} = 0$, the superluminal modes are pure gauge in the linear FP system.
[12] We label $f^0 = dt$, $f^i = dx^i$ and $e^0 = dz$, $e^1 = dx$, $e^2 = dy$ and $e^3 = dt$.

normal vector $\xi^\mu\partial_\mu = \frac{\partial}{\partial z}$). Then the curled symmetry shock (the first of the equations (21)) here implies $\mathfrak{w}_{o12} = 0$, $\mathfrak{w}_{o23} = \mathfrak{w}_{o02}$ and $\mathfrak{w}_{o13} = \mathfrak{w}_{o01}$. In turn, the vector constraint shock implies

$$(\beta_1 + 2\beta_2 + 3\beta_3)\mathfrak{w}_{o0i} \approx 0\,.$$

Since the equations of motion already imply the vanishing coefficient of the shock wavefronts $\mathfrak{w}_{o0i}$ in the above, these are not determined so the characteristic matrix has a vanishing determinant, and the model is indeed superluminal.

We can also probe whether spacetimes in which the fiducial and dynamical metrics have different speeds of light can lead to superluminalities. For this we take a dynamical metric ansatz,

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2\,,$$

and a Minkowski fiducial metric. Again this configuration solves the equations of motion (3) iff

$$\beta_0 + \beta_1 + \beta_2 + \beta_3 = 0 = (\beta_1 + 2\beta_2 + 3\beta_3)(c-1)\;.$$

Since $c = 1$ reduces to the previous FP situation, we consider $c \neq 1$ (so that $m_{\rm FP}$ must vanish) and then study a constant-t putative characteristic surface with $\xi^\mu\partial_\mu = \frac{\partial}{\partial t}$. Since both metrics are flat, we note that all contorsions vanish, and ${f_o}^i = 0$. Then the shocked curl of the symmetry constraint becomes a homogeneous, trivially invertible system, forcing $\mathfrak{w}_{oij} = 0$. The vector constraint's shock now reduces to

$$(\beta_1 + 2\beta_2 c + 3\beta_3 c^2)\,\mathfrak{w}_{ooi} \approx 0\,.$$

Clearly, for generic βs, there are values of c such that the coefficient above is zero, hence we already detect superluminalities with non-zero shock-wave profile $\mathfrak{w}_{ooi}$. Alternatively, keeping c generic, there then exists some combination of the βs such that the coefficient of $\mathfrak{f}_{oo}$ in the shock of the scalar constraint (22) vanishes. In other words, we can find models with superluminality for any value of c.

Analysis of more complicated solutions with non-flat fiducial backgrounds will harness the full power of the reduced characteristic matrix calculated in the previous section, but these simple examples already demonstrate the mechanism responsible for superluminal propagation. Introducing more general fiducial field dependence will generically only make matters worse.

5 Conclusion

We have performed a definitive analysis of full generic fmGR's propagation properties. By employing a first order Palatini formalism, we were able at last to obtain the explicit covariant constraints' form, valid for the theory's full parameter space. This result then enabled us to compute the characteristic matrix for all parameters. A new feature here, in the hitherto unprobed third parameter direction, is the appearance of the full dynamical Riemann curvature in the field-dependent characteristic matrix (which, in previous studies that were limited to subsets of parameter space, only involved metrics/vierbeine and contorsions).

The characteristic matrix is a powerful tool for analyzing any theory. In particular, if it is field-dependent there are many potential difficulties. It is intimately related to the kinetic structure and hence canonical commutators of the quantum version of the theory. Hence, a degenerate characteristic matrix implies zero and negative norm states [3]. It also determines characteristic surfaces, where predictability is lost and along which shocks propagate. Thus spacelike characteristic surfaces are very dangerous for any model. It is even possible to use them to detect micro-acausality (local CTCs). All these pathologies[13] are known features of fmGR [17, 15, 16, 14].

Our main causality result is the characteristic matrix itself, which encodes all this information. There is one last fmGR glimmer of hope, namely that for some distinguished choice of fiducial background and parameters, the characteristic matrix could be non-degenerate. This seems highly unlikely, since already counterexamples are known for broad classes of backgrounds and field configurations (see [17]), the very simplest examples of which were exhibited in section 4. These showed not only how easy it is to construct pathological solutions but also why models depending on fiducial backgrounds lead to an enormous loss of predictability: Even supposing that the model has a limited viability as an effective theory, one would have to first choose a background *by hand* and then check that it supports well-defined propagation for the spacetime region being considered. Without a principle for choosing the background, observational predictability is clearly imperiled.

Another issue, to which we gave little focus, is that for some regions in parameter space, the model has different branches [24] because the symmetry constraint is not the unique solution to (8) (in a second order metric formulation there is a similar issue related to the existence of square roots of the endomorphism $g_{\mu\nu}\bar{g}^{\nu\rho}$ used to define the mass term). Models with branches can suffer both jumps in DoF counts and loss of unique evolution.

The above list of fmGR pathologies suggests that a possible panacea could be the original bimetric model where the fiducial metric is dynamical [8]. A characteristic analysis for the bimetric theory is currently unavailable, but since the causal structures of two dynamical metrics are guaranteed to conflict with one another, there seems little hope for consistency here either. Also, various studies have indicated that the bimetric theory possesses no partially massless limit [27] (even though its linearization does [28]), which is strong evidence

[13]While we have discussed matter couplings, it should be noted that they present additional problems [25]. Nor have we considered fmGR's strong coupling pitfalls [26].

against models of this type.

Acknowledgements

The work of S.D. was supported in part by NSF PHY-1266107 and DOE DE-SC0011632 grants and that of G.Z. in part by DOE Grant DE-FG03-91ER40674. We thank C. Deffayet, K. Hinterbichler, K. Izumi, N. Kaloper, J. Mourad and Y.C. Ong for discussions.

A The Fierz–Pauli limit

To check that the thirty constraints found in Section 2 are independent, we review why this at least holds true in their linear limit. We first expand the dynamical fields around fiducial ones,

$$e^m := f^m + h^m\,, \quad \omega^{mn} := \overline{\omega}^{mn} + K^{mn}\,,$$

where the background is Einstein:

$$\bar{G}_m = \frac{1}{3!}\,\bar{\Lambda}\,\epsilon_{mnrs} f^n f^r f^s\,.$$

Here $\bar{G}_m := \frac{1}{2}\epsilon_{mnrs} f^n \bar{R}^{rs}$ and $\bar{\omega}^{mn}$ encodes the fiducial Levi-Civita connection $\bar{\nabla}$. As already discussed, for this background to be a solution, we must require that

$$\frac{\bar{\Lambda}}{3!} = m^2\left(\beta_0 + \beta_1 + \beta_2 + \beta_3\right)\,.$$

The linearized equations of motion are then

$$\begin{aligned}
\frac{1}{2}\epsilon_{mnrs} f^n \bar{\nabla} K^{rs} \quad &\approx \quad m^2\,\left[3\beta_0 + 2\beta_1 + \beta_2\right]\,\epsilon_{mnrs}\, f^n f^r h^s\,,\\
\bar{\nabla} h^m + f_n K^{nm} \quad &\approx \quad 0\,.
\end{aligned} \tag{23}$$

Just as for their non-linear counterparts (3), the spatial parts of the above field equations yield sixteen primary constraints. Next we find six secondary constraints from symmetry of the linearized Einstein tensor, and a further four secondary constraints by computing the curl of the first equation in (23) (the linearized vector constraint)

$$f^m h_m \approx 0 \approx \frac{3}{4}\,m^2_{\rm FP}\,\epsilon_{mnrs} f^m f^n K^{rs}\,.$$

Here we have defined, as earlier, the FP mass,

$$m_{\rm FP}^2 := m^2(\beta_1 + 2\beta_2 + 3\beta_3)\,.$$

The linearized vector constraint becomes an identity precisely at $m_{\rm FP} = 0$, where the model describes massless gravitons.

The remaining four constraints are tertiary and found from the curls of the secondary constraints:

$$f^m f^n K_{mn} \approx 0 \approx \frac{1}{2}\, m_{\rm FP}^2(\bar\Lambda - \frac{3}{2} m_{\rm FP}^2)\epsilon_{mnrs} f^m f^n f^r h^s\,.$$

Notice that at the value $m_{FP}^2 = \frac{2}{3}\bar\Lambda$, the last—linearized scalar—constraint is elevated to a gauge invariance. Indeed this is precisely the partially massless tuning found in [20], an invariance is known not to survive in the full nonlinear theory [16, 19]. Finally, as promised, observe that all constraints are independent. In particular, they imply that, for $\{m_{\rm FP}^2 \neq 0, \frac{2}{3}\bar\Lambda\}$, the dynamics are described by a symmetric tensor $h_{\mu\nu}$ that is (fiducially) trace- and divergence-free.

B Hamiltonian Analysis

We now give an account of the Hamiltonian analysis of fmGR in Palatini formalism[14]. This computation was first performed for pure gravity in [30] (see also [31]). Writing the fmGR action as an integral $S = \int \mathscr{L}$ over a sum of volume forms $\mathscr{L} := \mathscr{L}_{\rm EH} + \mathscr{L}_m$ where

$$\begin{aligned} \mathscr{L}_{\rm EH} &:= -\frac{1}{4}\,\epsilon_{mnrs}\, e^m e^n\, \left[d\omega^{rs} + \omega^r{}_t \omega^{ts}\right]\,, \\ \mathscr{L}_m &:= m^2 \epsilon_{mnrs}\, e^m \left[\frac{\beta_0}{4} e^n e^r e^s + \frac{\beta_1}{3} e^n e^r f^s + \frac{\beta_2}{2} e^n f^r f^s + \beta_3 f^n f^r f^s\right]\,, \end{aligned} \tag{24}$$

our first task is to decompose these into a $3+1$ split. For that, we employ the notations of Section 2 and define

$$\mathscr{L} =: dt \wedge \boldsymbol{L}\,,$$

with $\boldsymbol{L} := \boldsymbol{L}_{\rm EH} + \boldsymbol{L}_m$. In addition we call

$$\mathring{e}^m =: dt\, N^m\,, \quad \mathring{f}^m =: dt\, \bar N^m \quad \text{and} \quad \mathring{\omega}^{mn} =: dt\, w^{mn}\,.$$

Then (up to surface terms)

$$\boldsymbol{L} = -\frac{1}{4}\epsilon_{mnrs}\, \boldsymbol{e}^m \boldsymbol{e}^n\, \dot{\boldsymbol{\omega}}^{rs} - N^m \boldsymbol{\mathcal{G}}_m - \frac{1}{2} w^{mn} \epsilon_{mnrs} \boldsymbol{\mathcal{T}}^r \boldsymbol{e}^s - \boldsymbol{H}(\boldsymbol{e}^m)\,,$$

[14]See [29] for a three-dimensional Palatini-based fmGR Hamiltonian analysis.

where $\boldsymbol{\mathcal{G}}_m$ and $\boldsymbol{\mathcal{T}}^m$ are defined in equation (7) and

$$\boldsymbol{H}(\boldsymbol{e}^m) = m^2\epsilon_{mnrs}\,\boldsymbol{e}^m\left[\frac{\beta_1}{3}\boldsymbol{e}^n\boldsymbol{e}^r + \beta_2\boldsymbol{e}^n\boldsymbol{f}^r + 3\beta_3\boldsymbol{f}^n\boldsymbol{f}^r\right]\bar{N}^s\,.$$

Upon integrating the time derivative by parts, the action has twelve canonical pairs, $\boldsymbol{e}^m$ and their conjugate momenta $\frac{1}{2}\epsilon_{mnrs}\,\boldsymbol{e}^n\,\boldsymbol{\omega}^{rs}$. These are ostensibly subject to ten constraints imposed by the Lagrange multipliers N^m and w^{mn}. For GR, it was shown that this model is equivalent, upon integrating out the six ω^{mn}, to the standard ADM form involving six canonical pairs built from spatial metrics and their momenta, but still subject to four diffeomorphism constraints [30], thus yielding two physical DoF. For our purposes, however, instead of returning to a metric-based ADM formulation, it is advantageous to decompose the model such that its dependence on the spatial dreibeine $\boldsymbol{e}^a$ (splitting flat indices $m = (0,a)$) manifests three-dimensional coordinate and Lorentz invariance. To that end, recall that the Einstein–Hilbert action (24) takes its familiar $S_{EH} = \int\sqrt{-g}\,R$ form upon algebraically integrating out the spin connection *a la* Palatini, by solving the torsion constraint

$$0 = de^m + \omega^m{}_n e^n\,. \tag{25}$$

Instead of integrating out the entire ω^{mn}, which would also return us to the metric ADM formulation, we solve the above condition only for the spatial spin connections[15]. This is achieved by further decomposing the dynamical fields according to

$$\boldsymbol{e}^m := (\boldsymbol{M},\boldsymbol{e}^a)\,,\quad \mathring{e}^m := (N,N^a)\,dt\,,\quad \boldsymbol{\omega}^{0a} := \boldsymbol{P}^a\,,\quad \mathring{\omega}^{0a} := u^a\,dt\,.$$

The variables N and N^a correspond to ADM lapse and shift variables. We will call $\boldsymbol{M} := M_i dx^i$ the *shaft* while $\boldsymbol{P}^a := P_i{}^a dx^i$ will become nine canonical momenta. Thus we must now solve nine of the twenty-four torsion constraints:

$$0 = \boldsymbol{d}\boldsymbol{e}^a + \boldsymbol{\omega}^a{}_b\boldsymbol{e}^b + \boldsymbol{P}^a\boldsymbol{M}\,.$$

To that end, we henceforth assume invertibility of the dreibein $\boldsymbol{e}^a := e_i{}^a dx^i$ and use it to manipulate three-dimensional indices. The torsion solutions are then

$$\boldsymbol{\omega}^{ab}(\boldsymbol{e},\boldsymbol{M},\boldsymbol{P}) = \boldsymbol{\omega}^{ab}(\boldsymbol{e}) + M^{[a}\boldsymbol{P}^{b]} - P^{[ab]}\boldsymbol{M} - P^{[a}{}_c M^{b]}\boldsymbol{e}^c =: \boldsymbol{\omega}^{ab}(\boldsymbol{e}) + \boldsymbol{K}^{ab}\,.$$

From now on, we denote the three-dimensional Levi-Civita connection based on the Levi-Civita spin connection $\boldsymbol{\omega}(\boldsymbol{e})$ by $\boldsymbol{\nabla}$, and $\boldsymbol{\nabla}_{\boldsymbol{K}}$ is its torsionful counterpart based on the spin

[15] This parallels the canonical analysis of "Palatini" Maxwell theory: there $\boldsymbol{B}$ is solved for in terms of $\boldsymbol{dA}$, but $\boldsymbol{E}$ is kept independent.

connection in the above display (do not confuse $\boldsymbol{K}$ with the analogous quantity introduced earlier). The respective curvature two-forms will be denoted by $\boldsymbol{R}^{ab}$ and $\boldsymbol{R}^{ab}_{\boldsymbol{K}}$. Substituting this solution into the lagrangian $\boldsymbol{L}$ yields

$$\boldsymbol{L} = \boldsymbol{P}_a \dot{\tilde{\boldsymbol{e}}}^a + \frac{1}{2}\epsilon_{abc}\boldsymbol{M}\boldsymbol{e}^a\dot{\boldsymbol{\omega}}^{bc}(\boldsymbol{e},\boldsymbol{M},\boldsymbol{P}) - N\boldsymbol{\mathcal{G}}_0 - N^a\boldsymbol{\mathcal{G}}_a - \frac{1}{2}w^{ab}\boldsymbol{\mathcal{J}}_{ab} - \boldsymbol{H}(\boldsymbol{e},\boldsymbol{M})\,, \tag{26}$$

where

$$\begin{aligned}
\boldsymbol{\mathcal{G}}_0 \;&= -\frac{1}{2}\epsilon_{abc}\Big(\boldsymbol{e}^a[\boldsymbol{R}^{bc}_{\boldsymbol{K}} + \boldsymbol{P}^b\boldsymbol{P}^c] - 2m^2[\beta_0\boldsymbol{e}^a\boldsymbol{e}^b\boldsymbol{e}^c + \beta_1\boldsymbol{e}^a\boldsymbol{e}^b\boldsymbol{f}^c + \beta_2\boldsymbol{e}^a\boldsymbol{f}^b\boldsymbol{f}^c + \beta_3\boldsymbol{f}^a\boldsymbol{f}^b\boldsymbol{f}^c]\Big)\,,\\
\boldsymbol{\mathcal{G}}_a \;&= -\frac{1}{2}\epsilon_{abc}\Big(2\boldsymbol{e}^b\boldsymbol{\nabla}_{\boldsymbol{K}}\boldsymbol{P}^c - \boldsymbol{M}\big[\boldsymbol{R}^{bc}_{\boldsymbol{K}} + \boldsymbol{P}^b\boldsymbol{P}^c - 2m^2(3\beta_0\boldsymbol{e}^b\boldsymbol{e}^c + 2\beta_1\boldsymbol{e}^b\boldsymbol{f}^c + \beta_2\boldsymbol{f}^b\boldsymbol{f}^c)\big]\\
&\qquad\qquad + 2m^2\bar{\boldsymbol{M}}\big[\beta_1\boldsymbol{e}^b\boldsymbol{e}^c + 2\beta_2\boldsymbol{e}^b\boldsymbol{f}^c + 3\beta_3\boldsymbol{f}^b\boldsymbol{f}^c\big]\Big)\,,\\
\boldsymbol{\mathcal{J}}_{ab} &= \tilde{\boldsymbol{e}}_a\boldsymbol{P}_b - \tilde{\boldsymbol{e}}_b\boldsymbol{P}_a - \epsilon_{abc}\big[\boldsymbol{e}^c\,\boldsymbol{d}\boldsymbol{M} - (\boldsymbol{\nabla}_{\boldsymbol{K}}\boldsymbol{e}^c)\,\boldsymbol{M}\big]\,,\\
\boldsymbol{H} \;&= -m^2\epsilon_{abc}\left[\boldsymbol{M}\boldsymbol{M}^{ab}\bar{N}^c - \boldsymbol{e}^a\Big(\frac{\beta_1}{3}\boldsymbol{e}^b\boldsymbol{e}^c + \beta_2\boldsymbol{e}^b\boldsymbol{f}^c + 3\beta_3\boldsymbol{f}^b\boldsymbol{f}^c\Big)\bar{N} + \boldsymbol{e}^a\left(\beta_2\boldsymbol{e}^b + 6\boldsymbol{f}^b\right)\bar{\boldsymbol{M}}\bar{N}^c\right].
\end{aligned}$$

Here we have introduced the two-form $\tilde{\boldsymbol{e}}_a := -\frac{1}{2}\epsilon_{abc}\boldsymbol{e}^b\boldsymbol{e}^c$ (where $\epsilon_{abc} := \epsilon^0{}_{abc}$), which may equivalently be viewed as the dual of the (densitized) inverse dreibein; this relation may be inverted for $\boldsymbol{e}^a(\tilde{\boldsymbol{e}})$. For the fiducial vierbein, we have defined $\mathring{f}^m := \bar{N}^m = (\bar{N}, \bar{N}^a)$ while $\bar{\boldsymbol{M}}$ is the fiducial shaft. Also note that the triplet of auxiliary fields u^a completely decouples because nine of the torsion constraints have been solved.

Equation (26) is the key to our Hamiltonian analysis: The quartet of auxiliary fields $N^m = (N, N^a)$ play the *rôle* of the shift and lapse Lagrange multipliers in standard ADM, and we shall henceforth so refer to them. The first term in (26) is the Darboux form for nine canonical pairs $(\tilde{\boldsymbol{e}}_a, \boldsymbol{P}^a)$; however, this is complicated by the presence of the second term that potentially involves time derivatives of the dreibeine, shaft and canonical momenta. In Einstein gravity, this difficulty is easily circumvented by using a local diffeomorphism to gauge away the shaft [30]. In an fmGR setting, that route is closed to us; instead therefore, we integrate out the shift Lagrange multipliers N^a. This imposes three relations $\boldsymbol{\mathcal{G}}^a = 0$, which we can *generically* solve for the three components of the shaft $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{e},\boldsymbol{P})$. Configurations where these relations do not determine the shaft are, of course, intimately related to the model's propagation difficulties. Hence, at this point $\boldsymbol{\omega}^{ab} = \boldsymbol{\omega}^{ab}(\boldsymbol{e},\boldsymbol{P})$ and the Lagrangian becomes

$$\boldsymbol{L} = \boldsymbol{P}^a\dot{\tilde{\boldsymbol{e}}}_a + \frac{1}{2}\boldsymbol{M}(\boldsymbol{e},\boldsymbol{P})\epsilon_{abc}\boldsymbol{e}^a\dot{\boldsymbol{\omega}}^{bc}(\boldsymbol{e},\boldsymbol{P}) - N\boldsymbol{\mathcal{G}}_0(\boldsymbol{e},\boldsymbol{P}) - \frac{1}{2}w^{ab}\boldsymbol{\mathcal{J}}_{ab}(\boldsymbol{e},\boldsymbol{P}) - \boldsymbol{H}(\boldsymbol{e},\boldsymbol{P})\,.$$

Given that the symplectic terms now depend only on $(\boldsymbol{e}^a, \boldsymbol{P}^a)$ the model has maximally nine canonical pairs. The three Lagrange multipliers w^{ab} impose three constraints, which ought remove three of these pairs (these constraints are algebraically solvable[16] for the antisymmetric part of the canonical momenta $\boldsymbol{P}_{[ij]}$). The model will then reduce to six canonical pairs, subject to the single "Hamiltonian"[17] constraint $\boldsymbol{\mathcal{G}}_0 = 0$ imposed by the lapse Lagrange multiplier. This computation (modulo checking that the secondary constraint structure is correct) thus shows in 3+1 form that the model generically describes five DoF.

References

[1] C. de Rham, G. Gabadadze, Phys. Rev. D**82** (2010) 044020, [1007.0443v2 [hep-th]]; C. de Rham, G. Gabadadze, A. J. Tolley, Phys. Rev. Lett. **106** (2011) 231101, [1011.1232v2 [hep-th]]; Phys. Lett. B**711** (2012) 190, [1107.3820v1 [hep-th]]; C. de Rham, G. Gabadadze, A. J. Tolley, JHEP **1111** (2011) 093, [1108.4521v2 [hep-th]].

[2] D. Boulware, S. Deser, Phys. Rev. D **6** (1972) 3368; Phys. Lett. B **40**(2) (1972) 227.

[3] K. Johnson and E. C. G. Sudarshan, Annals Phys. **13** (1961) 126.

[4] G. Velo and D. Zwanziger, Phys. Rev. **186** (1969) 1337; Phys. Rev. **188** (1969) 2218;

[5] M. Kobayashi and A. Shamaly, Phys. Rev. D **17** (1978) 2179 ; Prog. Theor. Phys. **61** (1979) 656; S. Deser, V. Pascalutsa and A. Waldron, Phys. Rev. D **62** (2000) 105031, [hep-th/0003011]; I. L. Buchbinder, D. M. Gitman and V. D. Pershin, Phys. Lett. B **492** (2000) 161, [hep-th/0006144]; S. Deser and A. Waldron, Nucl. Phys. B **631** (2002) 369, [hep-th/0112182].

[6] H. Chen, J. M. Nester, H. -J. Yo, *Acausal PGT Modes and the Nonlinear Constraint Effect*, A. Phys. Pol. B, **29** (1998) 961; H. -J. Yo, J. M. Nester, Int. J. Mod. Phys. D **8** (1999) 459, [gr-qc/9902032v1]; Int. J. Mod. Phys. D **11** (2002) 747, [gr-qc/0112030v1].

[7] B. Zumino, *Effective Lagrangians and Broken Symmetries*, Brandeis Univ. Lectures on Elementary Particles and Quantum Field Theory, MIT Press Cambridge (Mass. , S. Deser, M. Grisaru and H. Pendleton eds.), Vol.**2** (1970) 437.

[16] Strictly, this is manifestly the case before setting $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{e}, \boldsymbol{P})$, but we could have first integrated out w^{ab} and then the shift, so the only real solvability concern is in that constraint.

[17] Here we use the modifier Hamiltonian to refer to the timelike diffeomorphism constraint of GR in an ADM approach; the massive model, of course, still has a non vanishing Hamilitonian $\boldsymbol{H}(\boldsymbol{e}, \boldsymbol{P})$ on the constraint surface.

[8] A. Salam and J. Strathdee, Phys. Rev. **184**, 1750 and 1760 (1969); C. J. Isham, A. Salam and J. Strathdee, Phys. Lett. B **31** (1970) 300.

[9] H. van Dam and M. J. G. Veltman, Nucl. Phys. B **22**, 397 (1970); V. I. Zakharov, JETP Lett. **12**, 312 (1970); Pisma Zh. Eksp. Teor. Fiz. **12**, 447 (1970); P. van Nieuwenhuizen, Phys. Rev. D **7**, 2300 (1973).

[10] A. I. Vainshtein, Phys. Lett. B **39** (1972) 393.

[11] I. I. Kogan, S. Mouslopoulos and A. Papazoglou, Phys. Lett. B **503**, 173 (2001) [hep-th/0011138]; M. Porrati, Phys. Lett. B **498**, 92 (2001) [hep-th/0011152]; S. Deser and A. Waldron, Phys. Lett. B **501**, 134 (2001) [hep-th/0012014].

[12] K. Hinterbichler, Rev. Mod. Phys. **84**, 671 (2012) [arXiv:1105.3735 [hep-th]]; C. de Rham, *Massive Gravity,* arXiv:1401.4173 [hep-th].

[13] S. F. Hassan, R. A. Rosen, Phys. Rev. Lett. **108** (2012) 041101, [1106.3344v3 [hep-th]]; JHEP**04** (2012) 123, [1111.2070v1 [hep-th]]; S. .F. Hassan, R. A. Rosen, A. Schmidt-May, JHEP**02** (2012) 026, [1109.3230v2 [hep-th]]; S. Hassan, A. Schmidt-May, M. von Strauss, Phys. Lett. B **715** (2012) 335, [1203.5283v2 [hep-th]]. See however, L. Alberte, A. H. Chamseddine and V. Mukhanov, JHEP **1104**, 004 (2011) [arXiv:1011.0183 [hep-th]]; A. H. Chamseddine and V. Mukhanov, JHEP **1108**, 091 (2011) [arXiv:1106.5868 [hep-th]].

[14] C. Burrage, C. de Rham, L. Heisenberg and A. J. Tolley, JCAP **1207** (2012) 004, [1111.5549 [hep-th]]; P. de Fromont, C. de Rham, L. Heisenberg and A. Matas, *Superluminality in the Bi- and Multi-Galileon*, [1303.0274 [hep-th]]; A. Gruzinov, *All Fierz-Paulian Massive Gravity Theories Have Ghosts or Superluminal Modes*, [1106.3972 [hep-th]].

[15] S. Deser, A. Waldron, Phys. Rev. Lett. **110** (2013) 111101, [1212.5835v3 [hep-th]].

[16] S. Deser, M. Sandora and A. Waldron, Phys. Rev. D 87 (2013) 101501, [1301.5621 [hep-th]].

[17] S. Deser, K. Izumi, Y. C. Ong and A. Waldron, Phys. Lett. B **726**, 544 (2013) [arXiv:1306.5457 [hep-th]]; *Superluminal Propagation and Acausality of Nonlinear Massive Gravity,* Proceedings, Conference in Honor of the 90th Birthday of Freeman Dyson: Singapore, 2013 World Scientific (2014) arXiv:1312.1115 [hep-th]. See also S. Deser, K. Izumi, Y. C. Ong and A. Waldron, *Problems of Massive and Bimetric Gravities*, in preparation.

[18] K. Izumi and Y. C. Ong, Class. Quant. Grav. **30**, 184008 (2013) [arXiv:1304.0211 [hep-th]].

[19] C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, Phys. Rev. D **88**, no. 2, 024003 (2013) [arXiv:1302.0025 [hep-th]]. C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, *Evidence for and Obstructions to Non-Linear Partially Massless Gravity,* [1302.0025 [hep-th]].

[20] S. Deser and R. I. Nepomechie, Annals Phys. 154 (1984) 396 ; Phys. Lett. B132 (1983) 321; S. Deser and A. Waldron, Phys. Rev. Lett. **87**, (2001) 031601 [hep-th/0102166]; Nucl. Phys. B **607**, (2001) 577 [hep-th/0103198]; Phys. Lett. B **508** (2001) 347 [hep-th/0103255] and **513**, (2001) 137 [hep-th/0105181]; Phys. Lett. B **513**, (2001) 137 [hep-th/0105181]; Phys. Lett. B **603**, (2004) 30 [hep-th/0408155]; Phys. Rev. D 74 (2006) 084036 [hep-th/0609113]; C. de Rham and S. Renaux-Petel, "Massive Gravity on de Sitter and Unique Candidate for Partially Massless Gravity," arXiv:1206.3482 [hep-th]; S. Deser, E. Joung and A. Waldron, Phys. Rev. D **86**, 104004 (2012) [arXiv:1208.1307 [hep-th]].

[21] C. Deffayet, J. Mourad and G. Zahariade, JCAP **1301** (2013) 032, [1207.6338 [hep-th]].

[22] J. Mourad, K. Noui and D. A. Steer, arXiv:1407.2920 [hep-th]; J. Mourad and D. A. Steer, JCAP **1406**, 058 (2014) [arXiv:1405.1862 [hep-th]]; JCAP **1312**, 004 (2013) [arXiv:1310.6560 [hep-th]].

[23] K. Hinterbichler and R. A. Rosen, JHEP **1207**, 047 (2012) [arXiv:1203.5783 [hep-th]].

[24] C. Deffayet, J. Mourad and G. Zahariade, JHEP **1303**, 086 (2013) [arXiv:1208.4493 [gr-qc]].

[25] S. Deser and A. Waldron, Phys. Rev. D **89**, 027503 (2014) [arXiv:1310.2675 [gr-qc]].

[26] C. Burrage, N. Kaloper and A. Padilla, Phys. Rev. Lett. **111**, no. 2, 021802 (2013) [arXiv:1211.6001 [hep-th]].

[27] S. Deser, M. Sandora and A. Waldron, Phys. Rev. **D88**, 081501 (2013) 1306.0647 [hep-th]; E. Joung, W. Li and M. Taronna, *No unitary theory of PM spin two and gravity,* arXiv:1406.2335 [hep-th]; Y. M. Zinoviev, *Massive spin-2 in the Fradkin-Vasiliev formalism. I. Partially massless case,* arXiv:1405.4065.

[28] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Higher Derivative Gravity and Conformal Gravity From Bimetric and Partially Massless Bimetric Theory,* arXiv:1303.6940 [hep-th].

[29] M. Bañados, C. Deffayet and M. Pino, Phys. Rev. D **88**, no. 12, 124016 (2013) [arXiv:1310.3249 [hep-th]].

[30] S. Deser and C. J. Isham, Phys. Rev. D **14**, 2505 (1976).

[31] J. M. Charap and J. E. Nelson, J. Phys. A **16**, 1661 (1983); Class. Quant. Grav. **3**, 1061 (1986); M. Charap Class. Quantum Grav. **4** 247 (1987); J. M. Charap, M. Henneaux and J. E. Nelson, Class. Quant. Grav. **5**, 1405 (1988).