


## Dynamical formulation of ghost-free massive gravity

Claudia de Rham,<sup>\*</sup> Jan Kozuszek<sup>†</sup>, Andrew J. Tolley,<sup>‡</sup> and Toby Wiseman<sup>§</sup>

*Theoretical Physics Group, Blackett Laboratory, Imperial College, London SW7 2AZ, United Kingdom*

 (Received 27 March 2023; accepted 3 October 2023; published 25 October 2023)

We present a formulation of ghost-free massive gravity with flat reference metric that exhibits the full nonlinear constraint algebraically, in a way that can be directly implemented for numerical simulations. Motivated by the presence of higher order operators in the low-energy effective description of massive gravity, we show how the inclusion of higher-order gradient (dissipative) terms leads to a well-posed formulation of its dynamics. The formulation is presented for a generic combination of the minimal and quadratic mass terms (the phenomenologically interesting case) on any background. For concreteness, we then focus on the numerical evolution of the minimal model for spherically symmetric gravitational collapse of scalar field matter. This minimal model does not carry the relevant interactions to switch on an active Vainshtein mechanism, at least in spherical symmetry, thus we do not expect to recover usual general relativity behavior even for small graviton mass. Nonetheless we may ask what the outcome of matter collapse is for this gravitational theory. Starting with small initial data far away from the center, we follow the matter through a nonlinear regime as it falls towards the origin. For sufficiently weak data the matter disperses. However for larger data we generally find that the classical evolution breaks down resulting in the theory becoming infinitely strongly coupled without the presence of an apparent horizon shielding this behavior from an asymptotic observer.

DOI: [10.1103/PhysRevD.108.084052](https://doi.org/10.1103/PhysRevD.108.084052)

### I. INTRODUCTION

Current and upcoming cosmological observations, event horizon mapping, and gravitational wave detections offer a unique opportunity to test the laws of gravity in unprecedented situations. While Einstein's theory of general relativity (GR) has proven to be in outstanding agreement with all observations to date, existing cosmological challenges and the need for an ultimate high-energy completion of GR have motivated the search for alternative frameworks. Even if GR provides the ultimate description of gravity on low-energy scales, the measure of success requires comparison with alternatives against which GR can be meaningfully tested. This is particularly important when observations and detections require the use of templates or priors through which assumptions about the underlying framework have to be made. With this aim in mind, and driven by the potential of tackling the

cosmological constant problem and the physics underlying the nature of the dark sector, a plethora of alternatives to GR have been formulated in the past two decades. While most of these models propose a modification of gravity through the introduction of additional modes (typically scalar fields), nonminimally coupled either to gravity or matter, a genuine modification of the graviton at low-energy (the IR) has proven more challenging. Large extradimensional models of gravity provided a first class of explicit realizations, where the structure of the graviton was genuinely modified in the IR or at large (cosmological) distances. In particular the Dvali-Gabadadze-Porrati model of gravity introduced in 2000 proposes a model where the graviton appears as a broad resonance of light massive modes from a four-dimensional perspective [1–6] dubbed “soft massive gravity.” The center and sharpness of this resonance was then further controlled by adjusting the size, scale, topology and the number of extra dimensions [7–12]. Attempts to define a theory with zero width—“a hard massive gravity”—have a long history and proposals motivated by extra dimensions were given in [13–16].

The first explicit attempts to construct a four-dimensional formulation of (hard) massive gravity were proposed in the 1970s, but the presence of a ghost at a low-energy scale, highlighted in [17–19], appeared to plague every explicit realization. Formulating massive gravity with the use of Stückelberg fields, as first introduced by Delbourgo and Salam in 1975, [20] proved particularly

<sup>\*</sup>c.de-rham@imperial.ac.uk

<sup>†</sup>j.kozuszek21@imperial.ac.uk

<sup>‡</sup>a.tolley@imperial.ac.uk

<sup>§</sup>t.wiseman@imperial.ac.uk

*Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.*

insightful in understanding the origin of this ghost [21] and ultimately led to a framework where it could be eradicated all together, leading to the development of “ghost-free massive gravity” (sometimes referred to as dRGT massive gravity) [22,23]. The absence of ghosts has not only been proven using the Stückelberg fields, but generalized to a multitude of different formalisms [24–39], confirming the existence of secondary constraints [40–42] (see also [43] for a review). The form of the constraint was derived on arbitrary backgrounds [44–49], including on spherically symmetric ones as will be relevant for the explicit numerical study presented here [50–55].

In what follows, we shall use the “vielbein-inspired” or symmetric vielbein formulation of massive gravity [28] which utilizes a 10 component vierbein to describe the geometry. This formalism exhibits the full nonlinear scalar constraint as presented in [35,45,46] in a way which can be directly applied to numerical evolution. In [35] the scalar constraint was explicitly identified for the minimal and quadratic models of massive gravity as a first order derivative scalar equation, derived from the Einstein equations with mass terms. It was shown to be more subtle for the cubic mass term, which cannot be expressed in a covariant way in the vielbein language. However this cubic mass term yields a specific coupling between the helicity-0 and -2 graviton modes which typically leads to other types of instabilities, and thus does not give rise to realistic phenomenology, and we will not consider it further here [43,56,57]. For perturbations about a general background it was shown in [45,46] that this scalar constraint explicitly removes the unwanted Boulware-Deser ghost, leaving only the five expected dynamical degrees of freedom. In this work, one of our aims is to formulate this constraint locally and use it to explicitly eliminate the unwanted variables, rather than working with a first order differential equation to be solved on every timeslice. We provide this algebraic phrasing of the scalar constraint by performing a  $(3 + 1)$  decomposition and then identifying appropriate momenta. In these variables the constraint will simply become algebraic in the time-time component of the vierbein, and furthermore for the simple scalar field matter we employ, will be either a quadratic or cubic equation in that vierbein component, depending on which mass terms one takes.

For concreteness, the numerical results derived in this work will be for the minimal model, for which a Vainshtein mechanism [58] is not expected to occur (unless one relies on the helicity-one interactions [59], which are absent in the spherically symmetric case we shall consider). Nonetheless it is a model of gravity with a dynamical spacetime, and thus a natural question is what its behavior is for collapse of matter. Does it resemble GR in the sense that it forms black holes for sufficiently nonlinear collapse? Or is its behavior unlike GR, with naked singularity formation? In what follows we shall provide answers to these basic questions.

Applications to the quadratic model, which for cosmological asymptotic conditions, is expected to have a working Vainshtein mechanism, and hence yield behavior similar to GR for low graviton masses, will be explored in further studies.

Current tests of GR, direct and indirect detections of gravitational waves and astrophysical/cosmological observations already provide interesting bounds on the graviton mass, [60], however the strongest constraints remain very model dependent. Model-independent bounds typically rely on the propagation of gravitational waves or modification of the dispersion relation, leading to a bound of the graviton mass which remains many orders of magnitude away from the phenomenologically interesting region (tackling the cosmological constant problem or the origin of dark energy requires a graviton mass of order of the Hubble parameter today,  $m \sim H_0 \sim 10^{-32}$  eV, while model-independent constraints on the graviton mass bound it to be  $\lesssim 10^{-22}$  eV). To better improve these bounds, an outstanding open question is what is the precise behavior of black holes in massive gravity, and in particular, what is the effect of the graviton mass on the production of gravitational waves and the resulting waveform?

Since the curvature invariant related to any realistic astrophysical black hole is dozens of orders of magnitude above the graviton mass,<sup>1</sup> we would expect the graviton mass to be utterly irrelevant to the dynamics of black holes and to the production of gravitational waves during inspiral and black hole mergers. However this argument relies on the existence of a smooth decoupling of scales. Such a decoupling would only occur if an efficient Vainshtein mechanism is in place to screen out the effect of the additional graviton polarizations. In practice, the presence of such a screening mechanism has been challenging to prove formally other than in specific configurations [56,61–91]. Another major challenge to making further progress is the fact that the constraint that prevents the presence of a ghost in massive gravity also prevents the existence of highly symmetric exact solutions. This feature has inhibited the existence of exact homogeneous and isotropic (cosmological) solutions on all scales [92], where solutions can appear arbitrarily close to Friedmann-Robertson-Lemaître-Walker on scales of the order of the observable Universe but departure from homogeneity or isotropy must emerge on large distance scales beyond the current cosmological horizon. A similar feature plagues the search for black hole solutions, where Birkhoff’s theorem is broken and the constraint prevents the existence of perfectly static and spherically symmetric solutions [93,94], (aside from solutions that exhibit a physical singularity at the horizon) see also [51–53,57,95–103] for other black hole solutions in massive (bi)gravity. As pointed out in [104],

<sup>1</sup>For  $m \sim H_0$ , only a black hole of the size of the Universe would carry a curvature invariant of order of the graviton mass.

a spherically symmetric nonsingular black hole solution can nonetheless accommodate an asymptotic Yukawa-like behavior if a small time dependence (scaling as the graviton mass) is included. For a graviton mass of the order of the Hubble parameter today, this would correspond to a time dependence that only manifests itself on time scales of the order of the age of the Universe and in practice the solutions are locally indistinguishable from standard Schwarzschild solutions. These solutions were obtained perturbatively about the black hole horizon. Exploring more precisely some of the features of these solutions was pioneered in [105] but deriving an explicit exact solutions has remained challenging.

It is worth emphasizing that the physical relevance of the Kerr black hole in GR stems firstly from the fact that black holes form in generic matter collapse, as encapsulated in the singularity theorems, and second, that due to the Uniqueness theorems, it is the only vacuum black hole solution. In massive gravity no such theorems are currently known, and thus even if candidate black hole end states for collapse are found, their physical relevance hinges on whether it is indeed these solutions that form dynamically from gravitating matter. In this work, we therefore initiate some of the first steps towards understanding what the end state of dynamical collapse is in the dRGT theory of gravity. We do so by solving numerically for spherically symmetric solutions in nonlinear massive gravity by considering the spherical gravitational collapse of a “lump” of matter (described by a massless scalar field) in minimal massive gravity.

Before going to the core of the formulation and the numerical framework, it is worth clarifying that just like GR should be seen as the leading order term in an infinite effective field theory (EFT) expansion [106], the same applies to massive gravity. At best, massive gravity is only ever expected to represent a low-energy description of gravity, and should always be seen as the leading order contribution in an infinite EFT expansion where the inclusion of higher order operators is unavoidable [107–109]

$$S_{\text{mGR}} = \int d^4x \sqrt{-g} \left[ \frac{1}{16\pi G_N} \left( R[g] + \frac{m^2}{2} \sum_n \alpha_n \mathcal{L}_n[\mathcal{K}] \right) + \Lambda_{\text{cutoff}}^4 \mathcal{L}^{\text{EFT}} \left( \frac{\nabla_\mu}{\Lambda_{\text{cutoff}}}, \frac{\Lambda_3^3}{\Lambda_{\text{cutoff}}^3} \mathcal{K}_\nu^\mu, \frac{R^\mu{}_{\alpha\nu\beta}}{\Lambda_{\text{cutoff}}^2} \right) \right], \quad (1)$$

where  $\Lambda_{\text{cutoff}}$  is the cutoff of the EFT and  $\Lambda_3 = (M_{\text{Pl}} m^2)^{1/3} \sim (m^4/8\pi G_N)^{1/6}$ . The Lagrangian terms  $\mathcal{L}_n$  are the “total derivative polynomials” (dRGT mass terms) introduced in [23] and  $\mathcal{K} \sim \delta - \sqrt{g^{-1}f}$  are the building blocks of the ghost-free mass term (reminiscent of the extrinsic curvature in models of massive gravity arising from extra dimensions [110]). An example of a higher order operator is a term  $\mathcal{K}^2$  which is not of the dRGT mass term form. From the EFT point of view it is natural to include

such a contribution, but it will come suppressed by the cutoff,

$$\Delta\mathcal{L} \sim \frac{\Lambda_3^6}{\Lambda_{\text{cutoff}}^2} \mathcal{K}^2. \quad (2)$$

Such a term appears to induce a ghost, but one whose mass is at the scale  $\Lambda_{\text{cutoff}}$  which renders it harmless.<sup>2</sup>

For the theory of massive gravity to make sense, it should be ghost-free up to the cutoff  $\Lambda_{\text{cutoff}}$  which should be parametrically larger than the graviton mass  $m$ . In principle the cutoff can be separate from the strong coupling scale  $\Lambda_3$  at which naive perturbative unitarity of the truncation to the leading  $\mathcal{L}_n$  operators breaks down [111,112]. For the theory to enjoy a standard (Wilsonian-like) and weakly coupled high-energy completion one would need the cutoff to be parametrically smaller than the scale  $\Lambda_3$  [107], however it is typically expected that theories that have a Vainshtein mechanism be embedded in alternative completions [112–117]. In particular, owing to the existence of a nonrenormalization theorem [118,119] that protects the ghost-free structure of the Lagrangian, situations where the higher order operators enter at a scale  $\Lambda_{\text{cutoff}}$  parametrically larger than the strong coupling scale  $\Lambda_3$  can be considered, with  $\Lambda_{\text{cutoff}}$  potentially close to the Planck scale. Ultimately, the scale at which new physics enters, and how it manifests itself through the higher order operators encapsulated in  $\mathcal{L}^{\text{EFT}}$ , depends on the precise details of the UV completion (Wilsonian or not, weakly coupled or not), but for the low-energy theory of massive gravity to make sense, low-energy observables should be independent of these details.

At the level of the classical continuum partial differential equations (PDEs) describing the truncation to the leading low energy theory, short distance modes will explore this region that is out of control of the EFT. Unlike in the case of GR, as we discuss later, we believe it is unlikely that this truncation to the leading dRGT terms makes sense as a theory with a well-posed initial-value formulation in its own right. In order to perform numerical calculations of this continuum leading low energy theory, we wish to accommodate such short distance excursions while remaining as agnostic as possible to the precise operators entering in  $\mathcal{L}^{\text{EFT}}$ . We cannot expect to simply ignore the issue, as if the leading low energy theory is ill-posed without high order terms, then one would not expect to have a good continuum limit when refining a numerical discretization for it. If one is unable to refine a numerical approximation to a continuum limit it is then unclear what the status of that numerical calculation is. Thus here we take the conservative view that we should start with a well-posed theory *before* considering a numerical discretization, so that a

<sup>2</sup>The parametric scaling is reflecting the fact that in the decoupling limit  $M_{\text{Pl}} \rightarrow \infty$ ,  $m \rightarrow 0$ ,  $\mathcal{K}_{\mu\nu} \sim \partial_\mu \partial_\nu \pi / \Lambda_3^3$  with  $\pi$  the helicity zero mode of the graviton, hence  $\Delta\mathcal{L} \sim (\square\pi)^2 / \Lambda_c^2$ .



good continuum limit does exist.<sup>3</sup> In order to achieve this for dRGT we introduce specific higher derivative terms that are convenient for numerical simulation, and natural in our  $(3 + 1)$  decomposition. We are able to prove the resulting continuum PDE system is well posed. Since we include these terms within the  $(3 + 1)$  split, they are not of the form expected from a Lorentz invariant UV completion, but nonetheless mimic the dissipative effect of the operators expected to enter in  $\mathcal{L}^{\text{EFT}}$  at the level of the dynamical equations, while ensuring that the resulting low-energy physics is insensitive to the scale at which they enter (so long as the scale is sufficiently large). Provided gradients remain below the scale of these new operators, they will be irrelevant for the dynamics, and this may be explicitly checked. This strategy complements that formulated in [120–128] for other theories of modified gravity or scalar/vector(-tensor) EFTs involving screening. An important outcome of this work is the existence of a manifestly well-posed initial-value formulation of the dynamics of massive gravity.

The rest of this work is organized as follows: We start by providing a brief review of ghost-free massive gravity in Sec. II, formulating the constraint algebraically in a symmetric “vielbein-inspired” language. Then in Sec. III, as a first step towards formulating the dynamics of massive gravity in a way that is amenable to numerical simulations, we perform a  $3 + 1$  space and time decomposition of the dynamical variables and their associated dynamical equations and constraints. Second, we are able to identify specific momenta such that the second class constraints in the system (the vector and scalar constraints) can be solved algebraically for components of the vierbein and its momenta. Specifically the scalar constraint becomes an algebraic relation for the time-time component of our symmetric vierbein. The remaining second order degrees of freedom precisely account for the five degrees of freedom in the theory. This formulation does not require any spacetime symmetry and is nonperturbative (i.e. not dependent on an expansion about a background). From this point on we focus on the simplest theory, that with minimal mass term. In Sec. IV we briefly outline an alternative harmonic formulation of the theory where the vector constraint is automatically satisfied if it is imposed on the initial Cauchy surface. We then introduce higher derivative dissipative terms motivated by the higher-order EFT operators present in (1) in Sec. V. These terms are convenient in our  $(3 + 1)$  formulation, and we argue the resulting PDEs then have a well-posed initial value formulation. In Sec. VI we demonstrate that our formulation may be implemented numerically in a straightforward

<sup>3</sup>An alternative perspective is to regard the numerical discretization itself as a cutoff, which cannot be completely removed. While perhaps a valid viewpoint, it is then difficult to assess the accuracy of the numerical calculations, and hence this is not the approach taken here.

manner by studying spherical collapse. We diagnose under which conditions the system evolves smoothly and when one hits pathological regions, where the EFT will become strongly coupled and lose predictability. We also diagnose under which conditions our results remain insensitive to the scale at which higher order EFT (or dissipative effects) kick in. A summary of our results and outlooks for further work are discussed in Sec. VII. Further details on the numerical implementation, diffusion effects and convergence are provided in Appendix B.

While massive gravity can be formulated in any number of dimensions, throughout this work we focus for concreteness on four spacetime dimensions and use mainly  $+$  signature. The relation between the reduced Planck scale and Newton’s constant is given by  $8\pi G_N = M_{\text{Pl}}^{-2}$  and we work with one or the other depend on context. For most of this work we will choose units where  $8\pi G_N = M_{\text{Pl}}^{-2} = 1$  but reintroduce dimensions whenever needed.

## II. BRIEF REVIEW OF dRGT MASSIVE GRAVITY

As for any massive field, the notion of mass requires a reference, which we shall denote as the reference metric  $f_{\mu\nu}$ . In principle massive gravity can be formulated for any reference metric  $f_{\mu\nu}$  [129], however the notion of mass is only unambiguous when dealing with representations of the Lorentz group or (anti-)de Sitter group, and so it makes sense to restrict to maximally symmetric spacetimes. For the remainder we make the standard choice that the reference metric is Minkowski. In that case, massive gravity in unitary gauge respects a global version of Lorentz invariance and admits Minkowski spacetime as a vacuum solution.

The usual dynamical metric is denoted as  $g_{\mu\nu}$ , and it is this metric that matter is chosen to couple to, preserving the weak equivalence principle. The global structure of the dynamical metric can be very different than the reference metric, but for comparison with GR in this work we shall require it to be asymptotically flat. As indicated in the introduction, restricting to four spacetime dimensions the ghost-free massive gravity action (1) can be formulated in terms of “total derivative polynomials”  $\mathcal{L}_n[\mathcal{K}]$  given in terms of the Levi-Civita symbols  $\varepsilon$  by [23]

$$\mathcal{L}_n[\mathcal{K}] = \varepsilon^{a_1 \dots a_4} \varepsilon_{b_1 \dots b_4} \mathcal{K}^{b_1}_{a_1} \dots \mathcal{K}^{b_n}_{a_n} \delta^{b_{n+1}}_{a_{n+1}} \dots \delta^{b_4}_{a_4}. \quad (3)$$

In particular we have

$$\mathcal{L}_0[\mathcal{K}] = 4!, \quad (4)$$

$$\mathcal{L}_1[\mathcal{K}] = 3![\mathcal{K}], \quad (5)$$

$$\mathcal{L}_2[\mathcal{K}] = 2!([\mathcal{K}]^2 - [\mathcal{K}^2]), \quad (6)$$

$$\mathcal{L}_3[\mathcal{K}] = ([\mathcal{K}]^3 - 3[\mathcal{K}][\mathcal{K}^2] + 2[\mathcal{K}^3]), \quad (7)$$

$$\mathcal{L}_4[\mathcal{K}] = ([\mathcal{K}]^4 - 6[\mathcal{K}]^2[\mathcal{K}^2] + 3[\mathcal{K}^2]^2 + 8[\mathcal{K}][\mathcal{K}^3] - 6[\mathcal{K}^4]), \quad (8)$$

where square brackets represent the trace of tensors (taken with respect to the dynamical metric). The building block  $\mathcal{K}^\mu{}_\nu$  defined as

$$\mathcal{K}^\mu{}_\nu = \delta^\mu{}_\nu - E^\mu{}_\nu = g^{\mu\alpha}(g_{\alpha\nu} - E_{\alpha\nu}), \quad (9)$$

is constructed out of the symmetric vielbein  $E_{\mu\nu} = E_{(\mu\nu)}$  [28], which is defined from the metric and reference metric as

$$g_{\mu\nu} = (f^{-1})^{\alpha\beta} E_{\alpha\mu} E_{\beta\nu}, \quad (10)$$

where  $(f^{-1})^{\mu\nu}$  is the inverse to the Minkowski reference metric. We may equivalently write the relation as

$$f^\mu{}_\nu = (g^{-1})^{\mu\alpha} f_{\alpha\nu} = E^\mu{}_\alpha E^\alpha{}_\nu, \quad (11)$$

so that symbolically, we may write  $E^\mu{}_\nu = \sqrt{f^\mu{}_\nu}$ . We also define  $(E^{-1})^{\mu\nu}$  as the inverse to  $E_{\mu\nu}$  in the sense that  $(E^{-1})^{\mu\alpha} E_{\alpha\nu} = \delta^\mu{}_\nu$ , so that

$$(E^{-1})^\mu{}_\nu = \sqrt{g_{\mu\alpha}(f^{-1})^{\alpha\nu}} = \sqrt{(f^{-1})^\mu{}_\nu}. \quad (12)$$

Omitting for now the EFT contributions that enter at the cutoff scale, and including coupling to matter, the formulation of massive gravity we shall be interested in is given by

$$S_{\text{mGR}} = \frac{1}{16\pi G_N} \int d^4x \sqrt{-g} \left( R[g] + \frac{m^2}{2} \sum_n \alpha_n \mathcal{L}_n[\mathcal{K}] \right) + S^{(\text{matter})}[g, \psi_i], \quad (13)$$

where  $R$  is the standard scalar curvature of the dynamical metric  $g_{\mu\nu}$  and matter fields  $\psi_i$  only couple to the physical metric  $g$ . When perturbing about flat spacetime, each Lagrangian  $\mathcal{L}_n[\mathcal{K}]$  is order  $n$  in fluctuations, which allows us to establish the order at which each interaction enters. We see that  $\mathcal{L}_0$  simply gives a cosmological constant. The terms  $\mathcal{L}_1$  and  $\mathcal{L}_2$  both contribute to the mass of the graviton. The remaining terms  $\mathcal{L}_3$  and  $\mathcal{L}_4$  do not contribute to the mass, but generate interaction terms for the massive gravitons. Since  $\mathcal{L}_1$  contains a tadpole it must be combined appropriately with  $\mathcal{L}_0$  to remove it and ensure that  $g_{\mu\nu} = f_{\mu\nu}$  is a vacuum solution, the condition being  $\alpha_0 = -\alpha_1/4$ . Expanding to quadratic order, the condition that the mass of the graviton is the parameter  $m$  above is  $\alpha_1 = 2(1 - \alpha_2)/3$ . Thus we may parametrize the theory with the mass  $m$  and the dimensionless couplings  $(\alpha_2, \alpha_3, \alpha_4)$ , fixing  $\alpha_0$  and  $\alpha_1$  using these two conditions. However the linear combination of terms  $\sqrt{-g}(\mathcal{L}_0 - 4\mathcal{L}_1 + 6\mathcal{L}_2 - 4\mathcal{L}_3 + \mathcal{L}_4) = 4!$  is

simply a constant, so it is not dynamical. This implies the theory is unchanged shifting the dimensionless parameters as

$$(\alpha_2, \alpha_3, \alpha_4) \rightarrow (\alpha_2 + 6c, \alpha_3 - 4c, \alpha_4 + c) \quad (14)$$

and thus the invariant dimensionless couplings are the combinations  $\tilde{\alpha} = (2\alpha_2 + 3\alpha_3)$  and  $\tilde{\beta} = (\alpha_3 + 4\alpha_4)$ .

The minimal model corresponds to the special case which can be written using only the terms  $\mathcal{L}_0$  and  $\mathcal{L}_1$ —this is parametrized just by the mass  $m$ , with  $\alpha_0$  and  $\alpha_1$  determined by the mass and tadpole conditions. Here we focus on the theory with both minimal and quadratic mass terms, which can be written using only the terms  $\mathcal{L}_0$ ,  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , so this massive gravity takes the form (still omitting the higher order operators for now),

$$S = \frac{1}{16\pi G_N} \int d^4x \sqrt{-g} \left( R[g] - m_1^2(2[E] - 6) - \frac{m_2^2}{2}([E]^2 - [E^2] - 6) \right) + S^{(\text{matter})}[g, \psi_i], \quad (15)$$

with the two following mass terms

$$m_1^2 = m^2(1 + 2\alpha_2) \quad \text{and} \quad m_2^2 = -2m^2\alpha_2, \quad (16)$$

so that the theory is parametrized by the graviton mass  $m$  and the coupling  $\alpha_2$ , and the minimal model corresponds to  $\alpha_2 = 0$ , while the quadratic model corresponds to  $\alpha_2 = -1/2$ . In terms of the invariant couplings,  $\tilde{\alpha} = 2\alpha_2$  and  $\tilde{\beta} = 0$ .

We note that the theory with minimal and quadratic mass terms has vanishing invariant coupling  $\tilde{\beta} = (\alpha_3 + 4\alpha_4) = 0$ . The most general ghost free massive gravity may be written using the cubic interaction term  $\mathcal{L}_3$  too, and has  $\tilde{\beta} \neq 0$ . However this cubic interaction term generically generates instabilities that render it incompatible with phenomenology and thus we do not consider it here [43,56,57].

The resulting Einstein equation is

$$\mathcal{E}_{\mu\nu} \equiv G_{\mu\nu} + m_1^2 M_{\mu\nu}^{(1)} + m_2^2 M_{\mu\nu}^{(2)} - 8\pi G_N T_{\mu\nu} = 0, \quad (17)$$

where the mass terms contributions are given by

$$M_{\mu\nu}^{(1)} = -E_{\mu\nu} + [E]g_{\mu\nu} - 3g_{\mu\nu}, \quad (18)$$

$$M_{\mu\nu}^{(2)} = \frac{1}{2} E_{\mu\alpha} E^\alpha{}_\nu - \frac{1}{2} [E] E_{\mu\nu} - \frac{1}{4} ([E]^2 - [E^2]) g_{\mu\nu} - \frac{3}{2} g_{\mu\nu}. \quad (19)$$

Using the contracted Bianchi identity and matter-stress energy conservation we may take the divergence of the Einstein equation to derive

$$V_\mu = \nabla^\nu (m_1^2 M_{\mu\nu}^{(1)} + m_2^2 M_{\mu\nu}^{(2)}) = 0, \quad (20)$$

which we refer to as the ‘‘vector equation.’’<sup>4</sup>

Diffeomorphism invariance can be made explicit through the use of Stückelberg fields. However for now it will prove more convenient to choose ‘‘unitary gauge’’ for the massive theory, meaning that we commit to the particular coordinate choices of Cartesian coordinates  $x^\mu = (t, x, y, z)$ , or later in Sec. VI when we consider spherical symmetry, spherical coordinates  $x^\mu = (t, r, \theta, \phi)$ . Then formulated in these unitary gauges the reference metric is given by  $f_{\mu\nu} = (f^{-1})^{\mu\nu} = \eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$  in the Cartesian case or  $f_{\mu\nu} = \text{diag}(-1, +1, +r^2, +r^2 \sin^2(\theta))$  in the spherical case. Thus until Sec. VI we will take  $f_{\mu\nu} = \eta_{\mu\nu}$ , and from now we also adopt units such that  $8\pi G_N = 1$ .

### A. Linear perturbations

Considering linear fluctuations about flat space, by expanding the dynamical metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (21)$$

we recover the standard Fierz-Pauli mass term in Minkowski,

$$m_1^2 M_{\mu\nu}^{(1)} + m_2^2 M_{\mu\nu}^{(2)} = \frac{1}{2} m^2 (h_{\mu\nu} - h \eta_{\mu\nu}), \quad (22)$$

where indices here, and in what follows in this perturbative discussion, are raised and lowered with the Minkowski metric  $\eta$ , and the graviton mass (squared) is given by  $m^2 = m_1^2 + m_2^2$ . Assuming the matter fields are all covariantly coupled to the dynamical metric, the resulting matter stress-energy tensor  $T_{\mu\nu}$  is then conserved,  $\partial^\mu T_{\mu\nu} = 0$ . At the linear level, the Bianchi identity (20) hence imposes the condition

$$\partial_\mu h^\mu{}_\nu = \partial_\nu h. \quad (23)$$

Note that unlike in GR, we have already set unitary gauge at this stage and there is therefore no additional gauge choice available. In particular the condition (23) appears as a constraint and not as a gauge choice. At the linearized level, this constraint forces the Ricci scalar to vanish, irrespectively of the trace of the stress-energy tensor. This indicates that the linearized theory is unable to properly capture coupling with external matter sources. This is at the origin of the infamous van-Dam-Veltman-Zakharov discontinuity [130,131], whose resolution lies in the contribution from the nonlinear interactions [58] we shall discuss

in Sec. II B. For now, carrying on with the linear analysis, the vanishing of the linearized Ricci scalar implies the following constraints

$$h = -\frac{2}{3m^2} T, \quad \partial^\rho h_{\rho\mu} = -\frac{2}{3m^2} \partial_\mu T. \quad (24)$$

As for the dynamical equations, they are given by the linearized Einstein field equations (17), after substituting the linearized condition (23), leading to the massive wave equation

$$-\frac{1}{2} (\nabla^2 - m^2) h_{\mu\nu} + \partial_{(\mu} \chi_{\nu)} = T_{\mu\nu} - \frac{1}{3} \eta_{\mu\nu} T, \quad (25)$$

with  $\nabla^2 = \partial^\rho \partial_\rho$  and  $\chi_\mu = \partial^\rho h_{\rho\mu} - \frac{1}{2} \partial_\mu h = \frac{1}{2} \partial_\mu h$ .

The five linear constraints (24) on  $\partial^\rho h_{\rho\mu}$  and  $h$  reduce the number of dynamical degrees of freedom from 10 down to 5. The special structure of the Fierz-Pauli term ensures the existence of an algebraic constraint on  $h$  rather than it obeying a wavelike equation, which would yield an additional ghostly degree of freedom. The existence of an analogous algebraic constraint in the full nonlinear massive gravity was indicated in [23–25] and proven more generically in various languages in [26–38], among others, together with the existence of a secondary constraint [40–42], (see also [43] for a review). The form of the constraint was derived on arbitrary backgrounds, including on spherically symmetric ones as will be relevant to the study presented here [51–55]. While the formulation of this primary constraint and the secondary one that follows has been well-established by now, how to implement it efficiently and in a well-posed way for the study of numerical evolution has proven more challenging. Part of the purpose of this paper is to present the constraint in the full nonlinear massive gravity in a formalism that can be directly implemented in a numerical evolution. Before carrying on with the full nonlinear analysis in Sec. II C, we first briefly discuss the physics behind the Vainshtein mechanism which signals a breakdown of linear perturbations and show how nonlinear interactions play an essential role when considering the small mass limit of massive gravity.

### B. Decoupled modes and Vainshtein

To better capture the essence of the Vainshtein mechanism, it is convenient to first identify the five propagating degrees of freedom more explicitly by writing the metric fluctuation as

$$h_{\mu\nu} = a_{\mu\nu} + \frac{1}{m^2} \partial_{(\mu} A_{\nu)} + \frac{1}{2} \pi \eta_{\mu\nu} + \frac{1}{m^2} \partial_\mu \partial_\nu \pi. \quad (26)$$

The helicity-2, helicity-1 and helicity-0 degrees of freedom are then encapsulated in  $a_{\mu\nu}$ ,  $A_\mu$  and  $\pi$ , respectively. This formulation enjoys two local invariances,

<sup>4</sup>Had we not set unitary gauge from the outset and kept the Stückelberg fields arbitrary, the following vector equation would simply arise as the Stückelberg equation of motion.

$$\begin{aligned}
a_{\mu\nu} &\rightarrow a_{\mu\nu} + \partial_{(\mu} v_{\nu)} & a_{\mu\nu} &\rightarrow a_{\mu\nu} + \frac{\lambda}{2} \eta_{\mu\nu} \\
A_\mu &\rightarrow A_\mu - m^2 v_\mu, & A_\mu &\rightarrow A_\mu + \partial_\mu \lambda \\
\pi &\rightarrow \pi & \pi &\rightarrow \pi - \lambda
\end{aligned} \quad (27)$$

for  $v_\mu$  and  $\lambda$  a Lorentz covector and scalar, respectively.

Now that these gauge invariances are made explicit, we can see that the gauge invariant rank-2 tensor  $a_{\mu\nu}$  carries  $10 - 4 \times 2 = 2$  degrees of freedom (same as a massless graviton),  $A_\mu$  carries  $4 - 1 \times 2 = 2$  degrees of freedom (same as any other gauge invariant vector field), and the last degree of freedom is carried by  $\pi$ , yielding a total of five degrees of freedom.

Endowed with these two sets of gauge invariances, we can freely pick the equivalent of the respective harmonic gauges (i.e. a slightly modified de Donder gauge for  $a_{\mu\nu}$  and modified Lorenz gauge for  $A_\mu$ ). Through appropriate gauge transformation, we can always set<sup>5</sup>

$$\partial^\rho a_{\rho\mu} - \frac{1}{2} \partial_\mu a = -\frac{1}{2} A_\mu, \quad \partial \cdot A = -m^2(a + 3\pi). \quad (28)$$

In this gauge, the degrees of freedom entirely decouple and lead to the following set of three (dynamical) wave equations,

$$\begin{aligned}
-\frac{1}{2}(\nabla^2 - m^2)a_{\mu\nu} &= T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T, \\
-\frac{1}{2}(\nabla^2 - m^2)A_\mu &= 0, \quad -\frac{1}{2}(\nabla^2 - m^2)\pi = \frac{1}{3}T.
\end{aligned} \quad (29)$$

Expressed in this form, it is clear that the helicity-0 mode  $\pi$  remains coupled to matter (at the linear level), even when we take the massless limit,  $m \rightarrow 0$  holding  $T_{\mu\nu}$  fixed. This explains why Eq. (25) does not lead to the same linearized equations as in GR,  $-\frac{1}{2}\nabla^2 h_{\mu\nu} + \partial_{(\mu}\chi_{\nu)} = T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T$ , due to the couplings to the trace of the stress tensor being different. The coupling of the helicity-0 mode to matter is responsible for an additional contribution of  $T/6$  even in the small mass limit. However it is also clear from the constraint (24) on  $h$  that the linearized theory breaks down in the massless limit. This is the essence of the Vainshtein mechanism pointed out in [58]. Accounting for nonlinear contributions under special conditions, GR was recovered nonperturbatively in the massless limit of the Dvali-Gabadadze-Porrati model [1] in [61] and the same mechanism was proven to occur in dRGT massive gravity [22,65].

More precisely when taking the  $m^2 \rightarrow 0$  limit, while keeping the stress energy source fixed, we may approximately solve the Einstein equations with a GR solution if

<sup>5</sup>Any gauge transformation with parameters  $v_\mu$  and  $\lambda$  satisfying  $(\nabla^2 - m^2)v_\mu = 0$  and  $(\nabla^2 - m^2)\lambda = 0$  preserve that gauge choice, so there are some residual gauge freedom one could further set.

we can find a coordinate system so that the vector equation  $V_\mu = 0$  is satisfied. We may regard this vector equation as a ‘‘gauge condition’’ for the GR solution, and at least locally we have the correct number of coordinate degrees of freedom to solve it. From the linear analysis we also know that generally the linear response differs to that of GR. This implies that while the solution of (17),  $\mathcal{E}_{\mu\nu} = 0$ , resembles that of GR in the massless limit it is in a gauge where  $|E_\nu^\mu - \delta_\nu^\mu| \gg 1$  so it is a nonlinear deformation of Minkowski at the level of the metric (even though curvatures may be small). In this regime, the theory becomes classically ‘‘strongly coupled’’ in the sense that standard perturbation theory breaks down, without indicating a failure of predictability [112]. In particular, subleading terms in the effective theory (1) involving higher derivatives of  $\mathcal{K}_\nu^\mu$  potentially become important (particularly if no Vainshtein resummation occurs), even though higher curvature terms may remain small. Far outside the nonlinear Vainshtein radius, the metric tends to Minkowski and the linear theory then applies, and hence leads to features that differ from GR. The matching region around the Vainshtein radius is then both a nonlinear deformation of Minkowski, as well as having non-GR behavior, and is subtle to track down precisely. Beyond its behavior in the decoupling limit of the theory, it has remained challenging to follow this transition precisely other than for static and spherical symmetric situations as well as in other theories of massive gravity [62].

The Vainshtein region where nonlinearities are important can be estimated by determining when the graviton mass  $m$  is negligible compared to curvature invariants,  $m^2 \lesssim R$ . For a compact matter source of mass  $M$ , the corresponding Vainshtein radius  $R_V \sim (R_G/m^2)^{1/3}$  where  $R_G \sim GM$  is the Schwarzschild radius associated with that  $M$ . We note that the same nonlinear terms that give rise to the Vainshtein mechanism and a smooth massless limit towards GR also lead to a breaking of perturbative unitarity at the scale  $\Lambda_3 \sim (M_{\text{Pl}} m^2)^{1/3}$ , with  $M_{\text{Pl}}$  the Planck mass. While this scale is naively very low, it is redressed within the Vainshtein radius, so that radiative corrections to the theory are irrelevant on scales where gravity may be probed [112].

Since all the nonlinear pure helicity-0 interactions vanish for the minimal model, it is not expected to exhibit a standard Vainshtein mechanism (at the very least not without existing the helicity-1 modes nontrivially [59]). There is therefore little known about the nonlinear behavior of this minimal theory in response to matter, even in the small mass limit.

### C. Vector equation and scalar constraint

We now return to the full nonlinear theory and review the scalar and vector constraints. We emphasize that the existence of these constraints has been discussed extensively in the literature and here we simply review these using the symmetric vierbein formulation. In the next



section we will see that these constraints can be formulated algebraically in a  $(3 + 1)$  decomposition nonperturbatively by appropriately identifying momentum variables.

In order to reveal the spin-1 and spin-0 constraints it is natural to consider (diffeomorphism) variations of the metric taking the form

$$\delta g_{\mu\nu} = \nabla_{(\mu} v_{\nu)}. \quad (30)$$

Now the action varies to give

$$\delta_v S = -\frac{1}{2} \int d^4x \sqrt{-g} v_\alpha V^\alpha, \quad (31)$$

where  $V_\mu$  is the vector defined in (20). The equation of motion from varying  $v_\mu$  is then the same as the vector constraint  $V_\mu = 0$ . In terms of the vector  $\xi^\mu$  defined as

$$\xi_\mu = E_{\mu\alpha} \eta^{\alpha\beta} V_\beta, \quad (32)$$

the constraint then takes the remarkably simple form,

$$\xi^\mu = V^{\mu\alpha\beta\sigma} \partial_{[\alpha} E_{\beta]\sigma} = 0, \quad (33)$$

with

$$V^{\mu\alpha\beta\sigma} = -2m_1^2 g^{\mu\alpha} (E^{-1})^{\beta\sigma} + m_2^2 (E^{\mu\alpha} (E^{-1})^{\beta\sigma} + g^{\mu\alpha} g^{\beta\sigma} - E^\rho{}_\rho g^{\mu\alpha} (E^{-1})^{\beta\sigma}), \quad (34)$$

so we may view this as a linear constraint on the components of  $\partial_{[\alpha} E_{\beta]\sigma}$ .

Now we consider varying the action with respect to a scalar mode  $\pi$  taking the form of a diffeomorphism combined with a conformal transformation and shift involving the reference metric,

$$\delta g_{\mu\nu} = \frac{\pi}{2} (m_1^2 g_{\mu\nu} + m_2^2 E_{\mu\nu}) + \nabla_{(\mu} \tilde{v}_{\nu)}, \quad \tilde{v}_\mu = E_{\mu\alpha} \eta^{\alpha\beta} \partial_\beta \pi. \quad (35)$$

This leads to the following variation of the action:

$$\delta_\pi S = \frac{1}{2} \int d^4x \sqrt{-g} \pi \Pi = 0, \quad (36)$$

where

$$\Pi = \frac{1}{2} (m_1^2 g^{\mu\nu} + m_2^2 E^{\mu\nu}) \mathcal{E}_{\mu\nu} + \nabla \cdot \xi. \quad (37)$$

Some comments on this variation are in order. Linearizing about flat space, so  $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$  and  $E_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{2} h_{\mu\nu}$ , we simply recover the spin-0 part of (26), namely  $\delta g_{\mu\nu} = \frac{\pi}{2} m^2 \eta_{\mu\nu} + \partial_\mu \partial_\nu \pi$ , but have now identified the fully non-linear equivalent excitation about an arbitrary background.

We note that in the ‘‘vielbein-like’’ language, this perturbation takes the simple form

$$\delta E_{\mu\nu} = \frac{\pi}{4} (m_1^2 E_{\mu\nu} + m_2^2 \eta_{\mu\nu}) + \frac{1}{2} \tilde{\nabla}_\mu \tilde{\nabla}_\nu \pi, \quad (38)$$

where the covariant derivative  $\tilde{\nabla}$  is taken with the connection  $\tilde{\Gamma}^\alpha{}_{\mu\nu}$  expressed in terms of the standard dynamical metric connection by the relation,

$$\Gamma^\alpha{}_{\mu\nu} = E^\alpha{}_\beta (E^{-1})^\gamma{}_\mu \tilde{\Gamma}^\beta{}_{|\nu\rangle\gamma} + (E^{-1})^{\alpha\beta} \partial_{(\mu} E_{|\nu)\beta}. \quad (39)$$

In terms of  $E_{\mu\nu}$ , the scalar equation can be explicitly written as

$$\Pi = \frac{1}{2} (m_1^2 \Pi^{(1)} + m_2^2 \Pi^{(2)}) = 0, \quad (40)$$

where we have defined,

$$\begin{aligned} \Pi^{(1)} &= A_{(1)}^{\alpha\beta\gamma\mu\nu\rho} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\rho} + m_1^2 (3[E] - 12) \\ &\quad + m_2^2 \left( \frac{1}{2} [E]^2 - \frac{1}{2} [E^2] - 6 \right) - T, \end{aligned} \quad (41)$$

$$\begin{aligned} \Pi^{(2)} &= A_{(2)}^{\alpha\beta\gamma\mu\nu\rho} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\rho} + m_1^2 (-[E^2] + [E]^2 - 3[E]) \\ &\quad + m_2^2 \left( \frac{1}{2} [E^3] - \frac{3}{4} [E][E^2] + \frac{1}{4} [E]^3 - \frac{3}{2} [E] \right) - E^{\mu\nu} T_{\mu\nu}. \end{aligned} \quad (42)$$

Let us assume the matter is such that the stress tensor does not involve derivatives of the metric, as for example is the case for (minimally coupled) scalar or vector fields, Yang-Mills theories or perfect fluids. Then the scalar equation  $\Pi = 0$  never involves terms with more than one derivative acting on the metric (or equivalently on  $E_{\mu\nu}$ ), and thus is a constraint equation. Furthermore, the one derivative terms are determined by the tensors,

$$\begin{aligned} A_{(1)}^{\alpha\beta\gamma\mu\nu\rho} &= \eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta} - 2(E^{-1})^{\rho[\alpha} g^{\beta][\mu} (E^{-1})^{\nu]\gamma} \\ &\quad + 4(E^{-1})^{\gamma[\alpha} g^{\beta][\mu} (E^{-1})^{\nu]\rho}, \end{aligned} \quad (43)$$

$$\begin{aligned} A_{(2)}^{\alpha\beta\gamma\mu\nu\rho} &= [E] \left( \frac{1}{2} \eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta} - (E^{-1})^{\rho[\alpha} g^{\beta][\mu} (E^{-1})^{\nu]\gamma} \right. \\ &\quad \left. + 2(E^{-1})^{\gamma[\alpha} g^{\beta][\mu} (E^{-1})^{\nu]\rho} \right) - 2\eta^{\gamma\rho} g^{\alpha[\mu} E^{\nu]\beta} \\ &\quad + (E^{-1})^{\gamma\rho} g^{\nu[\alpha} g^{\beta]\mu} - 4(E^{-1})^{\rho[\mu} g^{\nu][\alpha} g^{\beta]\gamma} \\ &\quad \left. + 2(E^{-1})^{\rho[\alpha} E^{\beta][\mu} (E^{-1})^{\nu]\gamma} - 4(E^{-1})^{\gamma[\alpha} E^{\beta][\mu} (E^{-1})^{\nu]\rho} \right). \end{aligned} \quad (44)$$

An important point that will be relevant shortly is that due to the derivatives of  $E_{\mu\nu}$  entering only via the combination



$\partial_{[\alpha}E_{\beta]\gamma}$ , the scalar constraint contains no time derivatives of  $E_{tt}$  at all. However it does contain spatial derivatives of  $E_{tt}$ .

### III. 3 + 1 DYNAMICAL FORMULATION

We now employ a 3 + 1 decomposition and since the map  $E_{\mu\nu} \rightarrow g_{\mu\nu}$  is explicit, it will prove convenient to work with  $E_{\mu\nu}$  as our dynamical variable. Given a novel choice for momentum variables, this 3 + 1 decomposition will allow us to solve the vector and scalar constraints explicitly. Our starting point is the action, which written in terms of  $E_{\mu\nu}$  takes the rather elegant form,

$$S = \int d^4x |\det E| \left( -\frac{1}{2} A_{(1)}^{\alpha\beta\gamma\mu\nu\sigma} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\sigma} - m^2 \mathcal{L}_{\text{mass}} + \mathcal{L}_{\text{matter}} \right). \quad (45)$$

Note the derivative term is identical to the one entering in  $\Pi^{(1)}$ . This is because, in the absence of matter, the terms containing derivatives of the  $E$  matrix in  $\Pi^{(1)}$  are simply equal to  $-R + 2\nabla \cdot \xi^{(1)}$  with this last term being a total divergence. Hence we see the Einstein-Hilbert term in the action is just given by the derivative terms in  $\sim |\det E| \Pi^{(1)}$  without matter.

Consider now the canonical conjugate momenta to  $E_{\mu\nu}$ . First note that the action contains no momentum conjugate to  $E_{tt}$  since  $\dot{E}_{tt}$  does not appear in the Lagrangian. Then the canonical momenta conjugate to  $E_{ti}$  and  $E_{ij}$  are given by

$$\pi^i = |E| A_{(1)}^{it\mu\nu\sigma} \partial_{[\mu} E_{\nu]\sigma}, \quad \pi^{ij} = -|E| A_{(1)}^{ij\mu\nu\sigma} \partial_{[\mu} E_{\nu]\sigma}, \quad (46)$$

where we use the notation  $|E| = |\det E|$ . In what follows we will work with the simpler momentum variables,

$$P_i = \partial_{[t} E_{i]t}, \quad P_{ij} = \partial_{[t} E_{i]j}, \quad (47)$$

where the approximation  $\simeq$  is understood to mean up to terms that only involve spatial derivatives acting on  $E_{\mu\nu}$ . Hence we may regard these four equations as linear constraints for the four momentum variables  $\tilde{P}$  and  $P_i$ , and at least near flat space we may invert this linear system to solve for these momenta. These four momenta then depend on all the metric components  $E_{\mu\nu}$ , including  $E_{tt}$ , through the components of  $V^{\mu\alpha\beta\sigma}$ . They also depend linearly on spatial derivatives of metric components through  $\partial_{[i} E_{j]\sigma}$ , but crucially they do not depend on derivatives of  $E_{tt}$ .

which are linearly related to  $\pi^i$  and  $\pi^{ij}$  with coefficients depending only on  $E_{\mu\nu}$ . An important point we return to later is that when the action is written in these variables, there are then no derivatives of  $E_{tt}$  at all—the only derivatives that enter above are spatial ones, and these can only occur in the combination  $P_i = \partial_{[t} E_{i]t}$ .

Now using the (spatial part of the) reference metric we may decompose the spatial components  $E_{ij}$  and our momenta  $P_{ij}$  into their traceless parts,  $\tilde{E}_{ij}$  and  $\tilde{P}_{ij}$ , and trace parts  $\tilde{E}$  and  $\tilde{P}$ , as

$$E_{ij} = \tilde{E}_{ij} + \tilde{E} \delta_{ij}, \quad P_{ij} = \tilde{P}_{ij} + \tilde{P} \delta_{ij}, \\ \delta^{ij} \tilde{E}_{ij} = \delta^{ij} \tilde{P}_{ij} = 0. \quad (48)$$

We now regard the upper triangular components of the symmetric spatial traceless  $\tilde{E}_{ij}$  (so  $j \geq i$ ) as the dynamical variables of our massive gravity theory, in the sense that they have second order time evolution equations. As we will shortly discuss, the remaining components  $\tilde{E}$  and  $E_{tt}$  have first order evolution equations from the vector equation, and the last component  $E_{tt}$  is algebraically determined (at least for conventional matter) in terms of the other variables by the scalar constraint. Thus we may write coordinates on the phase space as

$$(E_{ti}, \tilde{E}, \tilde{E}_{ij}, P_i, \tilde{P}, \tilde{P}_{ij}), \quad (49)$$

and then  $E_{tt}$  is a function of these phase space variables and their first derivatives which we can regard as an auxiliary variable. We now explicitly show how this works.

#### A. Vector equation

Focusing on the vector equation,  $\xi_\mu = 0$ , given in Eq. (33) then performing the 3 + 1 decomposition in phase space variables (49) and expanding about flat space, so  $E_{\mu\nu} \simeq \eta_{\mu\nu}$ , we can write

$$V^{i\alpha\beta\sigma} \partial_{[\alpha} E_{\beta]\sigma} = 2V^{t[ti]t} P_i + 2V^{t[ti]j} \tilde{P} + 2V^{t[ti]j} \tilde{P}_{ij} + V^{tij\sigma} \partial_{[i} E_{j]\sigma} \simeq 6m^2 \tilde{P} \\ V^{i\alpha\beta\sigma} \partial_{[\alpha} E_{\beta]\sigma} = 2V^{i[tj]t} P_j + 2V^{i[tj]j} \tilde{P} + 2V^{i[tj]k} \tilde{P}_{jk} + V^{ijk\sigma} \partial_{[j} E_{k]\sigma} \simeq -2m^2 P_i, \quad (50)$$

#### B. Scalar equation

Now we turn to the scalar constraint  $\Pi = 0$  given in (40), again assuming our matter is of a conventional type so that while the stress tensor depends on the metric, it does not explicitly involve metric connection terms. As already observed above, this equation only depends on first derivatives of  $E_{\mu\nu}$ . These enter through the combination  $\partial_{[\mu} E_{\nu]\sigma}$ , so as noted above, there are no  $\dot{E}_{tt}$  terms. Furthermore spatial gradients of  $E_{tt}$  come in the structure  $\partial_{[t} E_{i]t}$ , and hence are replaced with the momenta  $P_i$ . Therefore, we

may write the scalar constraint in terms of our phase space variables (49), and their derivatives, together with  $E_{tt}$  so that it contains no derivatives of  $E_{tt}$  at all.

The dependence on the momenta is quadratic and will be given more explicitly below. First note that  $(E^{-1})^{\mu\nu}$  can be written as

$$(E^{-1})^{\mu\nu} = \frac{1}{|E|} Q^{\mu\nu}, \quad (51)$$

where each component  $Q^{\mu\nu}$  is a polynomial in those of  $E_{\mu\nu}$ , and linear in each one. Hence given the form of  $A_{(1)}^{\alpha\beta\gamma\mu\nu\sigma}$  above we might have naively expected a quartic expansion of its derivative terms of the form,

$$|E|^4 A_{(1)}^{\alpha\beta\gamma\mu\nu\sigma} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\sigma} = C_4 E_{tt}^4 + C_3 E_{tt}^3 + C_2 E_{tt}^2 + C_1 E_{tt} + C_0, \quad (52)$$

with the coefficients  $C_A$  depending on the components of  $E_{\mu\nu}$  other than  $E_{tt}$ , together with the spatial gradients  $\partial_{[i} E_{j]k}$  and also all the momenta,  $P_i$ ,  $\tilde{P}$  and  $\tilde{P}_{ij}$ , but no derivatives of these. Likewise taking  $|E|^5 A_{(2)}^{\alpha\beta\gamma\mu\nu\sigma}$  we might have expected a quintic expansion in the component  $E_{tt}$ . However, as we explain in detail in Appendix A, due to the index antisymmetries of these two tensors,  $A_{(1,2)}^{\alpha\beta\gamma\mu\nu\sigma}$ , in fact we find simpler quadratic and cubic expansions going as

$$\begin{aligned} |E|^2 A_{(1)}^{\alpha\beta\gamma\mu\nu\sigma} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\sigma} &= C'_2 E_{tt}^2 + C'_1 E_{tt} + C'_0, \\ |E|^3 A_{(2)}^{\alpha\beta\gamma\mu\nu\sigma} \partial_{[\alpha} E_{\beta]\gamma} \partial_{[\mu} E_{\nu]\sigma} &= C''_3 E_{tt}^3 + C''_2 E_{tt}^2 + C''_1 E_{tt} + C''_0. \end{aligned} \quad (53)$$

Since the mass terms have similar structures, then for certain types of matter this constraint may determine  $E_{tt}$  as the root of a polynomial. As an example, consider matter that is a canonical scalar field  $\phi$  with potential  $V(\phi)$ , so

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} g_{\mu\nu} ((\partial\phi)^2 + V(\phi)). \quad (54)$$

Now restricting ourselves to the case of a minimal mass term (so  $m_2 = 0$ ), we can always scale the scalar constraint by the determinant of  $E$ , and consider the constraint  $|E|^2 \Pi = 0$ . The quadratic gradient term takes the form above. For the remaining terms, the explicit  $E_{tt}$  dependence of the stress tensor term that enters takes an identical form,

$$|E|^2 (m_1^2 (3|E| - 12) - g^{\mu\nu} T_{\mu\nu}) = D'_2 E_{tt}^2 + D'_1 E_{tt} + D'_0, \quad (55)$$

so in this minimal case with a canonical scalar field the scalar constraint is in fact a simple algebraic quadratic

polynomial in  $E_{tt}$ . We will give its explicit form in the case of spherical symmetry in our later numerical example, but emphasize that this reduction to a quadratic condition does not require any symmetry. Including also the nonminimal mass term, so  $m_1, m_2 \neq 0$ , then considering  $|E|^3 \Pi = 0$  yields an algebraic cubic equation in  $E_{tt}$  for such scalar field matter.

Since the scalar constraint is algebraic in  $E_{tt}$  we may wonder whether we can solve it for real  $E_{tt}$ . Near flat spacetime,  $E_{\mu\nu} \simeq \eta_{\mu\nu}$ , the scalar constraint reduces to the form in linear theory as written earlier in (24), so in our variables,

$$m^2 (-E_{tt} + 3\tilde{E}) = -\frac{2}{3} T, \quad (56)$$

and thus (given a nonzero mass) near flat spacetime we may always solve this for  $E_{tt}$ . However when the geometry deforms away from flat spacetime nonlinearly in  $E_{\mu\nu}$  it is then an interesting question whether these algebraic relations can be solved for  $E_{tt}$  (such that it is real). Since the EFT breaks down when the algebraic relation can no longer be solved for real  $E_{tt}$ , this simply indicates a sensitivity on UV physics at that point. We will return to this issue in our explicit numerical example later.

### C. Physical degrees of freedom

Starting with our phase space coordinates,  $(\tilde{E}, E_{ti}, \tilde{E}_{ij}, \tilde{P}, P_i, \tilde{P}_{ij})$  and auxiliary  $E_{tt}$ , the evolution of  $\tilde{E}$ ,  $E_{ti}$  and  $E_{ij}$  is determined by

$$\begin{aligned} \dot{\tilde{E}} &= 2\tilde{P} + \frac{1}{3} \partial^j E_{ti}, & \dot{E}_{ti} &= 2P_i + \partial_i E_{tt}, \\ \dot{\tilde{E}}_{ij} &= 2\tilde{P}_{ij} + \partial_i E_{tj} - \frac{1}{3} \delta_{ij} \partial^k E_{tk}. \end{aligned} \quad (57)$$

However, we have now seen that  $E_{tt}$  is determined algebraically by the phase space variables through the scalar constraint. This statement is fully nonlinear and valid about any background. Further we have seen that the vector equation  $\xi_\mu = 0$  gives linear constraints on  $\tilde{P}$  and  $P_i$ , the coefficients in these linear equations again depending algebraically on the auxiliary variable  $E_{tt}$ . Thus the evolution of  $\tilde{E}$ ,  $E_{ti}$  is first order in time, and determined by this vector equation. Hence we may reduce to the physical phase space of the theory, the component which enjoys a second order dynamics,

$$(\tilde{E}_{ij}, \tilde{P}_{ij}), \quad (58)$$

together with a first order dynamics for  $(\tilde{E}, E_{ti})$ , given by the first two equations above in (57), and the non-dynamical auxiliary  $E_{tt}$  by solving the scalar and vector constraints simultaneously for  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ .

Writing  $P_A = (P_i, \tilde{P})$ , and taking  $X$  to be the set of variables,  $X = \{\tilde{E}, E_{ii}, \tilde{E}_{ij}\}$ , then schematically this system takes the form,

$$\begin{aligned} M^{AB}(E_{ii}, X)P_A P_B + M^A(E_{ii}, X, \partial_i X, \tilde{P}_{ij})P_A + M(E_{ii}, X, \partial_i X, \tilde{P}_{ij}) &= 0, \\ Q^{AB}(E_{ii}, X)P_B + Q^A(E_{ii}, X, \partial_i X, \tilde{P}_{ij}) &= 0, \end{aligned} \quad (59)$$

the former equation being the scalar constraint, the latter the vector one, and  $\partial_i X$  denotes the set of spatial derivatives of  $X$ . We emphasize that this system is local to a spatial location  $x^i$  since  $P_A$  and  $E_{ii}$  only enter algebraically, not through their derivatives. Thus given the data  $X$  on a time slice, hence we also have  $\partial_i X$ , then at any point  $x^i$  on the slice we can consider this relatively simple system and solve for  $P_A$  and  $E_{ii}$ . Explicitly one may solve the momenta as  $P_B = -(Q^{-1})_{BA}Q^A$ . From the earlier Eq. (50) we see we may always solve this linear system near flat spacetime. Then having solved for the  $P_A$  we substitute them into the scalar equation to obtain a somewhat more complicated algebraic relation for  $E_{ii}$  of the form,

$$\Pi(E_{ii}, X, \partial_i X, \tilde{P}_{ab}) = 0. \quad (60)$$

Thus given data on a time slice,  $\tilde{E}$ ,  $E_{ii}$ ,  $\tilde{E}_{ij}$  and  $\tilde{P}_{ab}$ , we can then compute  $\tilde{P}$ ,  $P_i$  and  $E_{ii}$  from this. Having solved for these we have the evolution of the first order variables  $E_{ii}$  and  $\tilde{E}$ , and what remains to close the dynamics is to obtain equations for the time derivatives  $\dot{\tilde{P}}_{ij}$  which come from components of the Einstein equations. Then the second order dynamics ( $\tilde{E}_{ij}, \tilde{P}_{ij}$ ) gives the physical degrees of freedom of the theory, and given that  $\tilde{E}_{ij}$  is the traceless part of the spatial components  $E_{ij}$ , correctly accounts for the five expected degrees of freedom.

#### D. Evolution equations

To obtain the second order evolution equations we consider the components of the Einstein equation with one index raised, which we decompose as

$$H_\mu = \mathcal{E}^t_\mu, \quad \mathcal{E}^i_j = \tilde{\mathcal{E}}^i_j + \tilde{\mathcal{E}}\delta^i_j, \quad (61)$$

so that  $\tilde{\mathcal{E}}^i_i = 0$ . We note that since  $g_{\mu\alpha}\mathcal{E}^\alpha_\nu = g_{\nu\alpha}\mathcal{E}^\alpha_\mu$  then  $\mathcal{E}^i_t$  is determined by a linear combination of the above components.

Now we consider the equation  $\tilde{\mathcal{E}}^i_j$  when it is written using our phase space variables ( $E_{ii}, \tilde{E}, \tilde{E}_{ij}, P_i, \tilde{P}, \tilde{P}_{ij}$ ) and  $E_{ii}$ . The terms with time derivatives in the equations  $\tilde{\mathcal{E}}^i_j$  then take the form,

$$\tilde{\mathcal{E}}^i_j = \mathcal{F}^i_j \dot{\tilde{P}} + \mathcal{F}^{ai}_j \dot{P}_a + \mathcal{F}^{abi}_j \dot{\tilde{P}}_{ab} + \mathcal{F}^{ij}_j \dot{E}_{ii} + \dots, \quad (62)$$

where the coefficients depend on these phase space variables and  $E_{ii}$ , and the ellipses ... are terms with no

time derivatives (when written in these variables). Near flat spacetime one simply finds,  $\tilde{\mathcal{E}}^i_j \simeq \delta^{ia}\dot{\tilde{P}}_{aj} + \dots$ . We note that one might have naively expected terms going as  $(\dot{E}_{ii})^2$  as this would be a two derivative term, but we see from the form of the action in (45) that it contains no time derivatives of  $E_{ii}$ , and thus in the equations of motion one can find at most one time derivative of  $E_{ii}$ .

Now since we impose vanishing constraints,  $\xi_\mu$  and  $\Pi$ , which contain no time derivatives in our variables, we may also differentiate these with respect to time. Rewriting the time derivatives of  $\tilde{E}$ ,  $E_{ii}$  and  $\tilde{E}_{ij}$  that are generated in terms of our momenta  $\tilde{P}$ ,  $P_i$  and  $\tilde{P}_{ij}$ , they then have a similar form to the above, so

$$\begin{aligned} 0 = \dot{\xi}_\mu &= \mathcal{G}_\mu \dot{\tilde{P}} + \mathcal{G}^a_\mu \dot{P}_a + \mathcal{G}^{ab}_\mu \dot{\tilde{P}}_{ab} + \mathcal{G}'_\mu \dot{E}_{ii} + \dots, \\ 0 = \dot{\Pi} &= \mathcal{H}\dot{\tilde{P}} + \mathcal{H}^a \dot{P}_a + \mathcal{H}^{ab} \dot{\tilde{P}}_{ab} + \mathcal{H}' \dot{E}_{ii} + \dots \end{aligned} \quad (63)$$

At every point on the Cauchy surface, Eqs. (62) and (63) form a linear system which we may solve for  $\dot{\tilde{P}}$ ,  $\dot{P}_a$ ,  $\dot{\tilde{P}}_{ab}$  and  $\dot{E}_{ii}$ . Now having solved for  $\dot{\tilde{P}}_{ab}$  we may use the definition of this momentum, in Eq. (57), to find  $\dot{\tilde{E}}_{ab}$ , closing our dynamical system. We note that while the solution of the above system gives  $\dot{E}_{ii}$ ,  $\dot{P}_a$  and  $\dot{\tilde{P}}$ , since we are solving for these algebraically at every point, this information is redundant.

#### E. Hamiltonian and momentum initial constraints

While we have the expected five components of  $\tilde{E}_{ij}$  which enjoy second order dynamics, we also have a first order dynamics in  $\tilde{E}$  and  $E_{ii}$  and one may wonder how to make sense of the data on the initial surface associated to these variables. However, as we now argue, just as in GR we have to satisfy the analog of the Hamiltonian and momentum constraints on initial data, and we can take these to determine these four components  $\tilde{E}$  and  $E_{ii}$  in the initial data. Thus whilst these variables obey first order equations, since their initial data is constrained, there is no physical dynamics associated to them.

Recall in usual GR the Hamiltonian and momentum constraints are the Einstein equations  $(\mathcal{E}_{(\text{GR})})^t_\mu$ . Due to the contracted Bianchi identity and matter stress-energy conservation, we have

$$-\partial_t(\mathcal{E}_{(\text{GR})}^t)_{\mu} = \partial_i(\mathcal{E}_{(\text{GR})}^i)_{\mu} + \Gamma^{\beta}_{\beta\alpha}(\mathcal{E}_{(\text{GR})}^{\alpha})_{\mu} - \Gamma^{\alpha}_{\mu\beta}(\mathcal{E}_{(\text{GR})}^{\beta})_{\alpha}. \quad (64)$$

Provided the remaining Einstein equations  $(\mathcal{E}_{(\text{GR})}^i)_j$  hold, we see that  $\partial_t(\mathcal{E}_{(\text{GR})}^t)_{\mu} = c_{\mu}^{i\nu}\partial_i(\mathcal{E}_{(\text{GR})}^t)_{\nu} + c_{\mu}^{\nu}(\mathcal{E}_{(\text{GR})}^t)_{\nu}$  for some coefficients  $c_{\mu}^{i\nu}$ ,  $c_{\mu}^{\nu}$ . This, however, does not guarantee the constraints are satisfied, only that if they are initially true they will remain true. Thus we impose them as constraints on the initial data.

Returning to our massive gravity theory, having imposed the  $\tilde{\mathcal{E}}^i_j$ ,  $\xi_{\mu}$  and  $\Pi$  equations we may consider the analogous relations for the remaining Einstein equations  $H_{\mu}$  and  $\tilde{\mathcal{E}}$  in our decomposition (61). First, given  $\tilde{\mathcal{E}}^i_j = 0$ , we may write

$$g^{ij}g_{jk}\mathcal{E}^k_t = g^{k\mu}g_{\mu\nu}H_{\mu} + \tilde{\mathcal{E}}g^{ij}g_{jt}, \quad (65)$$

and invert this as

$$\mathcal{E}^i_t = \omega^{i\mu}H_{\mu} + \omega^i\tilde{\mathcal{E}}, \quad (66)$$

which implicitly defines the coefficients  $\omega^i$  and  $\omega^{i\mu}$ . At least near flat space this inversion is possible, and  $\omega^{ij} \simeq -\delta^{ij}$ , with  $\omega^{it} = \omega^i \simeq 0$ . Using this expression we may write the vector and scalar constraint equations we are imposing as

$$0 = \nabla_{\nu}\mathcal{E}^{\nu}_t = \dot{H}_t + \omega^{i\nu}\partial_i H_{\nu} + \omega^i\partial_i\tilde{\mathcal{E}} + a_t^{\nu}H_{\nu} + a_t\tilde{\mathcal{E}}, \quad (67)$$

$$0 = \nabla_{\nu}\mathcal{E}^{\nu}_j = \dot{H}_j + \partial_j\tilde{\mathcal{E}} + a_j^{\nu}H_{\nu} + a_j\tilde{\mathcal{E}}, \quad (68)$$

$$0 = \Pi = b^{\mu}H_{\mu} + b\tilde{\mathcal{E}}, \quad (69)$$

recalling that the vector equation comes from the divergence of the Einstein equations, and if it is satisfied, so  $\xi_{\mu} = 0$ , then from Eq. (37)  $\Pi$  takes the above form when the equations  $\tilde{\mathcal{E}}^i_j = 0$  are satisfied. Here the various coefficients  $a_{\mu}^{\nu}$ ,  $a_{\mu}$ ,  $b^{\mu}$  and  $b$  are functions of the metric variables and the connection, explicitly given as

$$\begin{aligned} a_t^{\nu} &= \Gamma^{\rho}_{t\rho}\delta_t^{\nu} - \Gamma^{\nu}_{tt} + \omega^{i\nu}\Gamma^k_{ik} + \partial_i\omega^{i\nu} & a_j^{\nu} &= \Gamma^{\rho}_{t\rho}\delta_j^{\nu} - \Gamma^{\nu}_{jt} - \Gamma^t_{ji}\omega^{i\nu} & b^{\mu} &= \frac{1}{2}m_1^2\delta_t^{\mu} + \frac{1}{2}m_2^2(E^{\mu}_t + E^t_j\omega^{j\mu}) \\ a_t &= \Gamma^k_{ik}\omega^i - \Gamma^k_{ik} + \partial_i\omega^i & a_j &= \Gamma^t_{jt} - \Gamma^t_{ji}\omega^i & b &= \frac{3}{2}m_1^2 + \frac{1}{6}m_2^2(\tilde{\mathcal{E}} + 3E^t_j\omega^j) \end{aligned}$$

Noting that we may eliminate  $\tilde{\mathcal{E}}$  in terms of  $H_{\mu}$  using the scalar constraint equation, which at least near flat space where  $b \simeq \frac{3}{2}m_1^2 + \frac{1}{6}m_2^2$  will be possible, leaves the vector equation of the form

$$0 = \dot{H}_{\mu} + c_{\mu}^{i\nu}\partial_i H_{\nu} + c_{\mu}^{\nu}H_{\nu}. \quad (70)$$

We see that this is precisely the same structure as in conventional GR in Eq. (64) above—to ensure that the full Einstein equation system  $\mathcal{E}^{\mu}_{\nu}$  is solved when evolving just the traceless spatial part  $\tilde{\mathcal{E}}^i_j$  and imposing the constraints  $\xi_{\mu}$  and  $\Pi$ , it is sufficient and necessary to impose the Hamiltonian  $H_t$  and momentum  $H_i$  equations on initial data. Once imposed on the initial surface, they remain true for all time without leading to any further constraints or needing to be imposed at each time.

These Hamiltonian and momentum constraints depend on the variables  $E_{ti}$  and  $\tilde{\mathcal{E}}$ , and thus we can think of using the freedom in their initial values to solve these initial constraints. Then, as claimed above, locally the free initial data is just that for the five second order degrees of freedom  $(\tilde{\mathcal{E}}_{ij}, \tilde{P}_{ij})$ .

## F. Summary

Let us now summarize our evolution scheme. Starting with data on a Cauchy slice comprising the physical phase space variables  $(\tilde{\mathcal{E}}_{ij}, \tilde{P}_{ij})$  and first order variables  $(\tilde{\mathcal{E}}, E_i)$ :

- (i) At each point on the Cauchy slice the constraint system (59) can be solved for  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ .
- (ii) These then yield the time derivatives of the first order variables  $(\dot{\tilde{\mathcal{E}}}, \dot{E}_i)$  using Eq. (57).
- (iii) Now at each point the combination  $\tilde{\mathcal{E}}^i_j$ ,  $\dot{\xi}_{\mu}$  and  $\dot{\Pi}$  forms a linear system which may be solved for  $\dot{\tilde{P}}_{ab}$  (and also  $\dot{\tilde{P}}, \dot{P}_a$  and  $\dot{E}_{tt}$ , but these are not required). Then from (57) this yields the time derivatives of our dynamical variables  $\tilde{E}_{ij}$ .

If the data on the initial slice is chosen so that the Hamiltonian and momentum constraints  $H = \mathcal{E}^t_t$  and  $H_i = \mathcal{E}^t_i$  vanish, in addition to the vector and scalar constraints, then  $H_{\mu}$  will remain zero under the evolution.

## IV. HARMONIC FORMULATION FOR THE MINIMAL THEORY

Before turning to the issue of well posedness, and demonstrating the above 3 + 1 decomposition scheme can be used to perform numerical gravitational collapse evolutions, we pause to briefly mention a different, elegant formulation of the dRGT theory where the vector constraint, if satisfied on the initial Cauchy surface, is automatically satisfied for all times, and leads to no secondary constraint. We do this here only for the case of the minimal mass term. It can be done also for the quadratic mass term, but we do not detail this here as this alternate formulation is



not our focus, and the resulting form is considerably more complicated than for the minimal mass term case.

For the minimal mass term, so setting  $m_2 = 0$ , there is a natural dynamical formulation that appears similar to the harmonic formulation for GR. For GR we may reformulate the (trace reversed) Einstein equation as (see for example [132] for a discussion of this),

$$R_{\mu\nu}^H \equiv R_{\mu\nu} - \nabla_{(\mu} v_{\nu)} - \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) = 0, \quad (71)$$

which we can term the ‘‘harmonic’’ Einstein equation, where we have introduced terms involving a vector field constructed as,

$$v^\mu = g^{\alpha\beta} (\Gamma^\mu_{\alpha\beta} - \bar{\Gamma}^\mu_{\alpha\beta}), \quad (72)$$

where  $\bar{\Gamma}^\mu_{\alpha\beta}$  is the connection of a smooth reference metric  $\bar{g}_{\mu\nu}$ . We note that since we have a difference of connections  $v^\mu$  transforms correctly as a vector field globally. Now the principal part of the equation is

$$(R_{\mu\nu}^H)_{\text{PP}} = -\frac{1}{2} g^{\alpha\beta} \partial_\alpha \partial_\beta g_{\mu\nu}, \quad (73)$$

so that all components of the metric propagate according to the wave operator of the geometry itself. Thus initial data for the problem is  $g_{\mu\nu}$  and  $\dot{g}_{\mu\nu}$ . However clearly generic initial data does not evolve to satisfy our original Einstein equation. The key point is that the contracted Bianchi identity and matter stress-energy conservation implies

$$\nabla^2 v^\mu - R^\mu_{\nu} v^\nu = 0, \quad (74)$$

so that provided  $v^\mu = 0$  and  $\dot{v}^\mu = 0$  on an initial Cauchy surface, then  $v^\mu$  will remain zero. If we choose  $g_{\mu\nu}$  and  $\bar{g}_{\mu\nu}$  and their first time derivatives to agree on the initial surface then  $v^\mu = 0$  there. We may further choose our initial data  $g_{\mu\nu}$  and  $\dot{g}_{\mu\nu}$  so that the four conditions  $\dot{v}^\mu = 0$  hold. These are precisely the usual Hamiltonian and momentum constraints. Then evolving for some choice of reference metric will yield a solution in the generalized harmonic gauge  $v^\mu = 0$ .

In massive gravity we naturally have a reference metric, and therefore one may wonder whether there is also such a harmonic formulation. For the mass terms considered here this is indeed the case. We modify our Einstein equation (17) (with only the minimal mass term) using the quantity  $\xi^\mu$  in Eq. (33), whose vanishing gives the vector constraint, as follows:

$$\mathcal{E}_{\mu\nu}^H \equiv G_{\mu\nu} - \frac{2}{m_1^2} \left( \nabla_{(\mu} \xi_{\nu)} - \frac{1}{2} g_{\mu\nu} \nabla \cdot \xi \right) + m_1^2 M_{\mu\nu}^{(1)} - T_{\mu\nu} = 0. \quad (75)$$

The principal part of this is

$$(\mathcal{E}_{\mu\nu}^H)_{\text{PP}} = (-\delta_{(\mu}^{\beta} \delta_{\nu)}^{\gamma} (E^{-1})^{\alpha\sigma} + \delta_{\mu}^{\alpha} \delta_{\nu}^{\beta} (E^{-1})^{\gamma\sigma} + E_{\rho(\mu} \delta_{\nu)}^{\beta} \eta^{\rho\sigma} g^{\alpha\sigma} - E_{\rho(\mu} \delta_{\nu)}^{\gamma} \eta^{\sigma\rho} g^{\alpha\beta}) \partial_\alpha \partial_\beta E_{\gamma\sigma} \quad (76)$$

and we see that this vanishes for the trace,  $g^{\mu\nu} \mathcal{E}_{\mu\nu}^H$ . Indeed the trace of this harmonic Einstein equation is precisely the scalar constraint  $\Pi = 0$  (for  $m_2 = 0$ ) and hence it contains no two derivative terms. For perturbations of flat space we see

$$(\mathcal{E}_{\mu\nu}^H)_{\text{PP}} \simeq -\partial^2 E_{\mu\nu} + \partial_\mu \partial_\nu E, \quad (77)$$

so the components of the traceless part of  $E_{\mu\nu}$  obey wave equations. Thus near flat space all but one linear combination of the  $E_{\mu\nu}$  propagate by hyperbolic wave operators, with the remaining part being determined by the scalar constraint equation which is only first order in derivatives. Rather than solving the vector equation  $\xi_\mu = 0$  as a linear constraint on momenta as in our previous 3 + 1 decomposition, instead in this formulation the Bianchi identity implies

$$\nabla^2 \xi^\mu - R^\mu_{\nu} \xi^\nu = 0, \quad (78)$$

so that if  $\xi_\mu = 0$  and  $\dot{\xi}_\mu = 0$  initially then the vector constraint  $\xi_\mu = 0$  remains true under time evolution. The initial condition that  $\xi_\mu = 0$  is just the condition that the vector constraint is imposed on the initial data. Then the condition that  $\dot{\xi}_\mu = 0$ , together with the scalar constraint holding, is the condition that the Hamiltonian and momentum constraints hold initially. Thus in this harmonic formulation, we have a wavelike evolution for all but one linear combination of  $E_{\mu\nu}$ , and that is determined on every time slice by the scalar constraint. The vector constraint is only imposed initially, being automatically satisfied at all times, and no new constraint arises.

## V. A WELL-POSED SHORT DISTANCE COMPLETION

Famously GR admits a well-posed hyperbolic formulation, as do some modified gravity theories, such as the Horndeski [133,134] and Einstein-Aether [135] theories. When dealing with massive gravity, it is understood that even classically, it ought to be treated as a low-energy effective field theory with operators entering at some cutoff scale  $\Lambda_{\text{cutoff}}$  as indicated in (1) and hence only providing a meaningful description of the long wavelength dynamics.<sup>6</sup> Since the issue of well posedness is embedded in the short

<sup>6</sup>Low energy gravity in string theory is another example where the leading low energy supergravity receives classical corrections; the  $\alpha'$  higher derivative terms. However in this case the truncation to the leading supergravity (or at least its bosonic part) will be well posed but not at higher order.

distance behavior of the theory, whether the classical truncation of dRGT massive gravity without the higher order EFT operators is well posed or not, is not a relevant physical question. However, when numerically simulating a theory, it is important to have PDEs that are well posed in the continuum limit. Without a well-posed formulation, it is unclear what a numerical discretization of the system represents—while a finite discretization will give a unique time evolution from initial data, one would not expect to be able to refine the discretization and obtain numerical solutions in a continuum limit. Hence the interpretation of any numerical solution is not well defined (see earlier footnote 3). Even though we are only interested in the long wavelength dynamics of the system which should remain insensitive to the higher order EFT operators included in (1), these operators are of paramount importance to the well posedness of the theory, and hence to its numerical simulation, if the truncation to the leading low energy theory is not well posed itself.

Thus we should ask whether our 3 + 1 formulation is expected to give a well-posed initial value problem. We believe this is generally unlikely. Certainly linearized perturbations of flat spacetime obey a well-posed hyperbolic system. However the scalar constraint changes away from flat space—for linear perturbations about flat space  $E_{tt}$  is simply determined by the stress tensor and  $\tilde{E}$ , as in (56), but nonlinearly it involves terms quadratic in first derivatives of the second order dynamical variables  $\tilde{E}^i_j$  [as seen explicitly in (40)]. Thus solving for  $E_{tt}$  and then substituting its form into the evolution equations for  $\tilde{E}^i_j$  will change the derivative structure, and likely will lead to ill posedness for dynamics away from flat spacetime.

While the truncated theory (very likely) lacks a well-posed initial value formulation, the higher order EFT operators naturally provide a short distance completion to achieve well posedness, see for instance [121–128]. The precise effects of higher order EFT operators on the long-wavelength modes is irrelevant. However on short-distance modes these operators affect the behavior in a way that can naturally lead to well posedness of the theory as a whole. The aim of this work is not to prove that *every* completion leads to well posedness (an unlikely outcome particularly when focusing on a specific formulation and gauge choice), but rather to show that the low-energy EFT we consider *can in principle* be embedded within a well-posed formulation and that low-energy observables are immune to the details of this high-energy-inspired formulation.

Instead of going back to the covariant formulation of higher order operators at the level of the action, a more pragmatic approach we will follow here is to include dissipative contributions directly at the level of our 3 + 1 formulation. These are understood to mimic the effect of higher order (covariant) EFT operators on short distance modes entering at the cutoff scale  $\Lambda_{\text{cutoff}}$ . In doing so, we will need to ascertain that adjusting the precise value of that

scale bears little effects on low-energy physics. This will then ensure that the completed theory has a well posed and diffusive short distance behavior, but the physical scales of interest are insensitive to the details of the regulation for a sufficiently small cutoff length scale. More details on how diffusive or higher order gradient terms arise from the UV completion of related types of theories are found in [126]. We now discuss the concrete inclusion of these terms and their effects on the posedness of the system.

Our dynamical system comprises the fields  $\tilde{E}_{ij}$ ,  $E_j$ ,  $\tilde{E}$  and momenta  $\tilde{P}_{ij}$  once we have algebraically eliminated  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ , and we may write this system as

$$\partial_t \tilde{P}_{ij} = \mathcal{S}_{ij}, \quad \partial_t \tilde{E}_{ij} = \mathcal{U}_{ij}, \quad \partial_t E_i = \mathcal{V}_i, \quad \partial_t \tilde{E} = \mathcal{W}. \quad (79)$$

The latter three relations simply follow directly from the definition of our momenta in (47). The first derives from solving (62) and (63) for  $\partial_t \tilde{P}_{ij}$ .

Focussing on the highest spatial derivative terms in these evolution equations, the equation for the time evolution of the momenta  $\tilde{P}_{ij}$  contains second spatial derivatives of these fields,

$$\mathcal{S}_{ij} = \mathcal{J}_{ij}{}^{klmn} \partial_k \partial_l \tilde{E}_{mn} + \mathcal{J}_{ij}{}^{klm} \partial_k \partial_l \tilde{E}_m + \mathcal{J}_{ij}{}^{kl} \partial_k \partial_l \tilde{E} + \dots \quad (80)$$

where the ellipses include terms with only first spatial derivatives acting on  $\tilde{E}_{ij}$ ,  $E_j$ ,  $\tilde{E}$  and  $\tilde{P}_{ij}$ . The coefficient functions, the  $\mathcal{J}$ s above, depend on the fields and also their first spatial derivatives (as they generally depend on the  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ , which when eliminated introduce first derivatives of the other fields). The remaining evolution equations for the fields  $\tilde{E}_{ij}$ ,  $E_j$ ,  $\tilde{E}$  only contain first order spatial derivatives.

While the Einstein equations are second order in spatial derivatives, the structure above is quite nontrivial in the sense that one might imagine  $\mathcal{S}_{ij}$  should contain spatial derivatives of higher order for two reasons. First, as discussed in Sec. III D it derives not just from the Einstein equations  $\tilde{E}^i_j$ , as in (62), but also from time derivatives of the scalar and vector constraints, as in (63). Second we might imagine  $\mathcal{S}_{ij}$  should contain spatial derivatives of higher order than two once  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$  are eliminated, since we know that these depend quadratically on first derivatives of the fields  $\tilde{E}_{ij}$ ,  $E_j$ ,  $\tilde{E}$ .

To address these points, we recall the fact noted earlier, that the action (45) when written in our momentum variables  $P_i$  and  $P_{ij}$  is only algebraic in  $E_{tt}$ , having no terms with derivatives (time or space) acting on it. Further it is clearly algebraic in the  $P_i$  and  $\tilde{P}$ . Hence being first order in derivatives, when this action is varied to obtain the

Einstein equations, and in particular the components  $\tilde{\mathcal{E}}^i_j$ , and these are written in our momentum variables, these contain at most first derivative terms in  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ . In addition, since the scalar and vector constraints contain no derivatives in  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ , and only first derivatives in the other fields, then solving (62) and (63) for  $\partial_t \tilde{P}_{ij}$  still gives an expression that contains at most second spatial derivatives in  $\tilde{E}_{ij}$ ,  $E_j$ ,  $\tilde{E}$ , and first spatial derivatives in the  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$ . Now finally given the structure (59), so that  $E_{tt}$ ,  $P_i$  and  $\tilde{P}$  depend on first derivatives (albeit quadratically) in the other dynamical fields, when they are eliminated to yield  $\mathcal{S}_{ij}$  they will generate at most second derivative terms.

As motivated by the previous discussion and by the inclusion of higher order operators in our EFT (1), we now simply include additional spatial diffusion terms given by the flat reference metric into each evolution equation as

$$\begin{aligned}\partial_t \tilde{P}_{ij} &= \mathcal{S}_{ij} + \ell^2 \delta^{mn} \partial_m \partial_n \tilde{P}_{ij}, \\ \partial_t \tilde{E}_{ij} &= \mathcal{U}_{ij} + \ell^2 \delta^{mn} \partial_m \partial_n \tilde{E}_{ij},\end{aligned}\quad (81)$$

$$\begin{aligned}\partial_t E_i &= \mathcal{V}_i + \ell^2 \delta^{mn} \partial_m \partial_n E_i, \\ \partial_t \tilde{E} &= \mathcal{W} + \ell^2 \delta^{mn} \partial_m \partial_n \tilde{E},\end{aligned}\quad (82)$$

where the scale  $\ell$  is the length scale associated to the short distance completion. Then for timescales  $T$  and length scales  $L$  such that

$$T \ll 1/\ell^2, \quad L \gg \ell, \quad (83)$$

these diffusion terms will be irrelevant. Now to understand the character of the system we should linearize about a general background, and consider the highest derivative terms for a perturbation about this. We then see the highest derivative terms, which are those of second order, take the form

$$\partial_t \begin{pmatrix} \delta \tilde{P}_{ij} \\ \delta \tilde{E}_{ij} \\ \delta E_i \\ \delta \tilde{E} \end{pmatrix} = \begin{pmatrix} \ell^2 \delta_i^m \delta_j^n \delta^{kl} & \mathcal{J}_{ij}{}^{klmn} & \mathcal{J}_{ij}{}^{klm} & \mathcal{J}_{ij}{}^{kl} \\ 0 & \ell^2 \delta_i^m \delta_j^n \delta^{kl} & 0 & 0 \\ 0 & 0 & \ell^2 \delta_i^m \delta^{kl} & 0 \\ 0 & 0 & 0 & \ell^2 \delta^{kl} \end{pmatrix} \partial_k \partial_l \begin{pmatrix} \delta \tilde{P}_{mn} \\ \delta \tilde{E}_{mn} \\ \delta E_m \\ \delta \tilde{E} \end{pmatrix} + \dots, \quad (84)$$

where the ellipses represent terms that are lower order in derivative terms. In the previous expression, the coefficients  $\mathcal{J}$  are understood to be evaluated on the background. Then on short scales the two derivative terms dominate, and we may think of the coefficient functions in the matrix controlling this term as approximately constant. To elicit the local behavior, we write the perturbation in Fourier space as

$$\delta \tilde{P}_{ij} = a_{ij} e^{-\omega t} e^{ik_m x^m}, \quad \delta \tilde{E}_{ij} = b_{ij} e^{-\omega t} e^{ik_m x^m}, \quad \delta E_i = c_i e^{-\omega t} e^{ik_m x^m}, \quad \delta \tilde{E} = c e^{-\omega t} e^{ik_m x^m} \quad (85)$$

so that on small scales, locally we have

$$\omega \begin{pmatrix} a_{ij} \\ b_{ij} \\ c_i \\ c \end{pmatrix} \simeq \begin{pmatrix} \ell^2 \delta_i^m \delta_j^n k^2 & \mathcal{J}_{ij}{}^{klmn} k_k k_l & \mathcal{J}_{ij}{}^{klm} k_k k_l & \mathcal{J}_{ij}{}^{kl} k_k k_l \\ 0 & \ell^2 \delta_i^m \delta_j^n k^2 & 0 & 0 \\ 0 & 0 & \ell^2 \delta_i^m k^2 & 0 \\ 0 & 0 & 0 & \ell^2 k^2 \end{pmatrix} \begin{pmatrix} a_{mn} \\ b_{mn} \\ c_m \\ c \end{pmatrix} + \dots, \quad (86)$$

with  $k^2 = \delta^{ij} k_i k_j$ . Clearly  $\omega$  is given by the eigenvalues of the matrix on the right-hand side. However its upper triangular form implies that its eigenvalues are simply given by its diagonal entries. Hence we have

$$\omega = \ell^2 k^2, \quad (87)$$

for all the eigenvectors of this system, and thus all the field and momentum perturbations diffuse on small scales, governed by the diffusion length scale  $\ell$ . Thus this diffusive short distance completion has a well-posed initial value formulation, for *any* positive diffusion constant

$\ell > 0$ . Given a setting with physical length and timescales of interest, we may then choose a sufficiently small  $\ell$  so that we expect the dynamics on these physical scales are insensitive to the regulator, while the irrelevant very short scales are controlled by it. In using such a regulator in practice one must confirm that it is indeed the case that the physics is insensitive to the regulator—this is essentially the statement that the leading massive gravity is a good EFT. For example, one could check that the physics on the scales of interest are insensitive to doubling the chosen value of  $\ell$ . We will do exactly this in our later numerical example.

We note that this regularization will work more generally than our case of interest here. Take a dynamical system for a set of fields  $\underline{\phi}$  so that  $\partial_t \underline{\phi} = \underline{F}(\underline{\phi}, \partial_i \underline{\phi}, \partial_i \partial_j \underline{\phi})$  for some functional  $F$  that depends on only up to second derivatives, but may be nonlinear in its arguments, e.g. could contain a term such as  $\sim (\partial_i \underline{\phi})(\partial_j \underline{\phi})(\partial_j \partial_j \underline{\phi})$ . Linearizing about a background  $\underline{\phi} = \underline{\phi}_0 + \delta \underline{\phi}$  will give

$$\begin{aligned} \partial_t \delta \underline{\phi} &= \underline{M}^{ij}(\underline{\phi}_0) \cdot \partial_i \partial_j (\delta \underline{\phi}) + \underline{M}^i(\underline{\phi}_0) \cdot \partial_i (\delta \underline{\phi}) \\ &+ \underline{M}(\underline{\phi}_0) \cdot (\delta \underline{\phi}) + \underline{C}(\underline{\phi}_0). \end{aligned} \quad (88)$$

Now our regularization simply adds diffusion terms as  $\partial_t \underline{\phi} = \underline{F}(\underline{\phi}, \partial_i \underline{\phi}, \partial_i \partial_j \underline{\phi}) + D \partial_i \partial_i \underline{\phi}$  so that the highest derivative terms which determine the principle symbol of the PDE system are  $\partial_t \delta \underline{\phi} \sim \underline{M}^{ij}(\underline{\phi}_0) \cdot \partial_i \partial_j (\delta \underline{\phi}) + D \partial_i \partial_i (\delta \underline{\phi})$ . Writing,  $\delta \underline{\phi}(t, \vec{x}) = e^{\omega t} e^{i k_m x^m} \underline{f}(t, \vec{x})$  with  $\underline{f}$  being slowly time and space varying functions yields

$$\omega \underline{f} = -\underline{M}^{ij}(\underline{\phi}_0) \cdot k_i k_j \underline{f} - D k^2 \underline{f} + \dots \quad (89)$$

so that ... are terms that are subdominant in the large  $k^2$  limit. Now if the matrix  $\underline{M}^{ij}$  is upper triangular for all  $(ij)$ , then the diffusion terms will determine the diagonal components and hence the eigenvalues of this system, so that  $\omega = -Dk^2$  for all the eigenvectors, and the system becomes well posed and of diffusive character for *any*  $D > 0$ . Thus  $D$  can be taken very small, so as not to interfere with the physical length and timescales of interest, but will ensure well posedness on the shortest scales. In practice, when taking some small  $D$  it must be checked that indeed the dynamics on the physical scales are insensitive to the regulator, for instance by adjusting the value of  $D$  and confirming that the results are insensitive to the precise value of  $D$  being used.

We may obtain this structure from relativistic dynamical systems with suitable kinetic terms and equations of motion that are at most second order in derivatives. In a  $3 + 1$  decomposition half the fields  $\underline{\phi}$  will be the fields of the theory,  $\Phi_A$ , and their time derivatives will be determined by their momenta,  $\Pi_A$ , and for suitably conventional kinetic terms these relations will not involve any second spatial derivatives of the fields or momenta. The remaining half of  $\underline{\phi}$  will be these momenta, and their time derivatives will involve second derivatives of the fields, but not of the momenta. Thus we will have

$$\partial_t \begin{pmatrix} \Pi_A \\ \Phi_A \end{pmatrix} = \begin{pmatrix} D \partial_k^2 \delta_{AB} & M_{AB}^{ij} \partial_i \partial_j \\ 0 & D \partial_k^2 \delta_{AB} \end{pmatrix} \begin{pmatrix} \Pi_B \\ \Phi_B \end{pmatrix} + \dots \quad (90)$$

and hence obtain an  $\underline{M}^{ij}$  that takes an upper triangular form. Our massive gravity theory is of this form, except with

additional momentum fields.<sup>7</sup> It is instructive to consider a simple example, a single relativistic scalar field with equation  $\ddot{\phi} - \dot{\phi} - \partial_k^2 \phi = 0$ . This is well posed, and thus obviously does not require regulation, but we may still consider regulating it. Defining a momentum  $\pi = \dot{\phi}$  this can be written in first order form,

$$\partial_t \begin{pmatrix} \pi \\ \phi \end{pmatrix} = \begin{pmatrix} D \partial_k^2 & \partial_k^2 \\ 1 & 1 + D \partial_k^2 \end{pmatrix} \begin{pmatrix} \pi \\ \phi \end{pmatrix}, \quad (91)$$

where we have included the regulation, and clearly see the upper triangular structure for the two derivative terms on the right-hand side. Now eliminating  $\pi$  and rewriting this in second order form we obtain

$$\ddot{\phi} - \dot{\phi} - \partial_k^2 \phi = -D \partial_k^2 \phi + 2D \partial_k^2 \dot{\phi} - D^2 (\partial_k^2)^2 \phi. \quad (92)$$

From the perspective of the second order system we see the regulator has introduced a higher order spatial derivative. Recall we are interested in  $D > 0$  and taking the limit  $D \rightarrow 0$ . For any  $D > 0$ , no matter how small, for sufficiently short scales this four-derivative term dominates and therefore controls the character of the equation. However, in the limit  $D \rightarrow 0$  all the terms on the right-hand side become irrelevant on large scales compared to the two-derivative term of the original equation on the left-hand side. Thus for a usual second order system, our regularization method is equivalent to adding a particular set of higher derivative terms with small coefficients.

The formulation (81) and (82), motivated by the existence of a meaningful completion allows the theory to enjoy a well-posed continuum that can then be discretized and numerically solved. We are taking here a pragmatic (unashamedly artificial) approach which should not be regarded as *the* actual physical completion, such as, for example one arising from integrating out additional massive degrees of freedom.<sup>8</sup> Rather it is a pragmatic proxy for what one would expect to arise. We emphasize that classical dRGT must be completed by *something*, but the precise details of what this completion is, is irrelevant to the description of long wavelength phenomena. Despite being

<sup>7</sup>We note that not all dynamical systems can be regulated in this manner. An example is the first order system  $\dot{\phi} = \partial_x \phi + \lambda (\partial_x \phi) \partial_x^2 \phi$ . For general backgrounds this is not well posed. However being a single equation, so the matrix  $\underline{M}^{ij} \partial_i \partial_j$  is  $1 \times 1$ , and simply given by  $\lambda (\partial_x \phi) \partial_x^2$ , it does not take the required upper triangular form. In order to regulate this one would have to add suitable higher derivative terms, rather than two derivative diffusion, or use a ‘‘fixing’’ scheme as introduced in [120]. We are very grateful to Luis Lehner for valuable discussion on this point.

<sup>8</sup>A specific covariant example of how integrating out additional massive degrees of freedom leads to EFT operators that change the nature of the dispersion relation and ultimately lead to a trivial eigenvalue for the system was presented in [136,137].



more artificial (and not formulated covariantly), our formulation is very attractive from a numerical perspective, being simple, and also is very natural when using a  $(3 + 1)$ -phase space formulation—the decomposition in time naturally defining the frame for diffusion. We also emphasize that with a short distance completion, the theory is not guaranteed to be free from instabilities and pathologies. This is an independent question from that of well posedness. Instabilities or pathologies may still arise in the long wavelength dRGT dynamics, but they will not be associated with arbitrarily short scales, and instead will be associated to the dynamical length scales in the problem, set by the length scale of the graviton mass, as well as the scales included in the initial data. Such instabilities or pathologies, should they exist, would be independent of the irrelevant diffusion terms of the completion, and then interpreted as physical phenomena of the long wavelength description, signaling its breakdown. If such phenomena arise, then a physical short distance completion would be required to continue dynamical evolution, rather than the artificial one we have introduced. We now turn to explicit simulation of the dRGT theory to illustrate the above formulation with its diffusive short distance contributions.

## VI. SPHERICAL COLLAPSE IN THE MINIMAL THEORY

While the minimal theory is not thought to exhibit the non-linearity required to switch on an active Vainshtein mechanism, it is nonetheless interesting to explore what happens under gravitational collapse, even though we do not expect it to exhibit four-dimensional GR-like behavior in the small mass limit. Even when the mass is not small it is still a theory of gravity, by which we mean a theory of a dynamic spacetime, and hence it is interesting to understand its behavior. For example do (exotic) black holes form when matter collapses? Can naked singularities form? In what follows we therefore explore some aspects of its dynamics under the assumption of spherical symmetry and, as expected, we shall indeed see very different dynamics to that of usual GR. The point of the following work is not to disfavor the minimal theory against known gravitational dynamics but rather to show a proof of concept of how the dynamical evolution can be followed through numerically in that simple (minimal) example. The application to the

nonminimal model and to other physically relevant situations will be explored elsewhere.

To proceed, we commit to a specific matter content and consider gravitational collapse of a massless scalar field,  $\Phi$ , in spherical symmetry. The scalar equation of motion is  $\nabla^2\Phi = 0$  and gives the matter stress tensor,

$$T_{\mu\nu} = \partial_\mu\Phi\partial_\nu\Phi - \frac{1}{2}g_{\mu\nu}(\partial\Phi)^2. \quad (93)$$

We use the coordinate invariance of the theory to choose spatial polar coordinates for the metric and Minkowski reference metric,  $x^\mu = (t, r, \theta, \phi)$ , writing

$$E_{\mu\nu} = \begin{pmatrix} c & rh & 0 & 0 \\ b + 2r^2a & 0 & 0 & 0 \\ & r^2(b - r^2a) & 0 & 0 \\ & & r^2(b - r^2a) \sin^2\theta & 0 \end{pmatrix},$$

$$f_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & r^2 & 0 \\ & & & r^2 \sin^2\theta \end{pmatrix}, \quad (94)$$

so that regularity of the metric at the origin implies that  $a, b, c, h$  should be smooth functions of  $r^2$  there. Earlier, our Minkowski reference metric was expressed in Minkowski coordinates. The only change to the previous discussion from the use of spherical spatial coordinates is at the level of the scalar and vector constraints, and in the definition of the momenta in Secs. II C and III, where the spatial partial derivatives,  $\partial_i$ , are now replaced by spatial covariant derivatives with respect to these spherical coordinates in Minkowski. We have

$$\bar{\nabla}_{[t}E_{i]\mu} = \partial_{[t}E_{i]\mu} - \tilde{\Gamma}_{\mu[t}^\nu E_{i]\nu}, \quad (95)$$

where  $\bar{\nabla}$  is the covariant derivative with respect to the reference metric  $f_{\mu\nu}$ , so the Minkowski metric in the spherical spatial coordinates. Since  $\tilde{\Gamma}_{\alpha\beta}^\nu$  vanishes if any of its indices equal time, then the connection terms vanish in the momenta  $P_i = \bar{\nabla}_{[t}E_{i]r}$ . However they do contribute to  $P_{ij} = \bar{\nabla}_{[t}E_{i]j}$  and its trace part  $\tilde{P} = \frac{1}{3}f^{ij}\bar{\nabla}_{[t}E_{i]j}$ . We then find first order dynamical variables and momenta,

$$E_{tr} = rh, \quad P_r = \frac{1}{2}(\partial_t E_{tr} - \partial_r E_{tt}) = \frac{1}{2}(r\dot{h} - \partial_r c),$$

$$\tilde{E} = b, \quad \tilde{P} = \frac{1}{6}(3\partial_t \tilde{E} - \partial_r E_{tr} + f^{ij}\tilde{\Gamma}_{ij}^\nu E_{t\nu}) = \frac{1}{2}\left(\dot{b} - \frac{1}{3}\partial_r h - h\right), \quad (96)$$

with the second order geometric variable being

$$\tilde{E}_{rr} = 2r^2 a, \quad \tilde{P}_{rr} = \frac{1}{2}(\partial_t E_{rr} - \partial_r E_{tr} + \tilde{\Gamma}_{rr}^v E_{tt}) - \tilde{P}f_{rr} = r^2 \dot{a} - \frac{1}{3}r\partial_r h, \quad (97)$$

together with the matter scalar  $\Phi$  and its momentum  $p_\Phi = \dot{\Phi}$ . The remaining component,

$$E_{tt} = c, \quad (98)$$

will be algebraically determined by the scalar constraint.

For convenience we define the following quantities from the momenta,

$$p_h = \frac{2}{r}P_r, \quad p_b = 2\tilde{P}, \quad p_a = \frac{1}{r^2}\tilde{P}_{rr}, \quad (99)$$

so that  $\dot{h} = p_h + \dots$ ,  $\dot{b} = p_b + \dots$  and  $\dot{a} = p_a + \dots$ , where terms with no time derivatives are included in the ellipses  $\dots$ . This will allow us to formulate the evolution equations for the metric functions  $a$ ,  $b$  and  $h$ . We see that the metric functions  $h$  and  $b$  are associated to  $P_i$  and  $\tilde{P}$  and thus have a first order dynamics, since these momenta are linearly determined by the vector constraint and we may algebraically eliminate them. Thus it is the function  $a$  with its momentum  $p_a$  that embodies the second order dynamics in the metric, and  $c$  is determined algebraically.

In this formulation, the vector equation has nontrivial components,

$$\begin{aligned} V_t &= \frac{3p_b(ar^2 + b) + 2hr(5ar + r^2\partial_r a - \partial_r b) - 6ar^4 p_a}{(b - ar^2)(r^2(2ac - h^2) + bc)}, \\ V_r &= \frac{10acr - 2hr^3 p_a - ar^3 p_h + 2hrp_b + brp_h + 2cr^2\partial_r a - 2c\partial_r b}{(b - ar^2)(r^2(2ac - h^2) + bc)}, \end{aligned} \quad (100)$$

which we note are linear in  $p_h$  and  $p_b$ , and the scalar constraint takes the quadratic form,  $a_2 c^2 + a_1 c + a_0 = 0$ , with the precise formulae for the coefficients  $a_i$  given in Appendix A.

These vector and scalar constraints then form an algebraic system for determining  $c$ ,  $p_h$  and  $p_b$  in terms of the other variables. While the scalar constraint takes a form above that is quadratic in  $c$ , it is worth emphasizing that since it depends on  $p_b$  and  $p_h$ , once we solve for these from the vector constraint (which also contains  $c$ ), then the resulting algebraic equation for  $c$  takes a complicated form. Thus in what follows we solve this system for  $c$ ,  $p_b$  and  $p_h$  numerically at each point, rather than analytically eliminating these variables.

The Einstein equation  $\tilde{\mathcal{E}}_r^r$  together with the scalar field equation then determines the second order dynamics giving  $\tilde{P}_{rr}$  and  $p_\Phi$ .

### A. Strong coupling

Since the scalar constraint is a quadratic in  $c$ , we may formally write its solution as

$$c = -\frac{1}{2} \left( \frac{a_1}{a_2} \pm \sqrt{\Delta} \right), \quad \Delta = \frac{a_1^2}{a_2^2} - 4 \frac{a_0}{a_2}. \quad (101)$$

We will call  $\Delta$  the “discriminant” even though it differs in normalization from the usual definition. We call the “positive branch” the solution with the “+” sign and the

“negative branch” that with the “−” sign. Linearizing about flat spacetime,

$$\begin{aligned} c &= -1 - \epsilon\delta c, & b &= 1 + \epsilon\delta b, & h &= \epsilon\delta h, \\ a &= \epsilon\delta a, & \Phi &= \sqrt{\epsilon}\delta\Phi, \end{aligned} \quad (102)$$

then we find

$$\begin{aligned} a_2 &\simeq -3m^2 + \epsilon((\partial_r \delta\Phi)^2 - 21m^2\delta b - 6m^2 r^2 \delta a), \\ a_1 &\simeq -3m^2 - 6\epsilon m^2(2\delta b + r^2 \delta a), \\ a_0 &\simeq -\epsilon(\partial_t \delta\Phi)^2, \\ \Delta &\simeq 1 - \frac{2\epsilon}{3m^2}(9m^2\delta b + 2(\partial_t \delta\Phi)^2 - (\partial_r \delta\Phi)^2). \end{aligned} \quad (103)$$

We recall that we are considering just the minimal mass  $m_2 = 0$  and we will rescale units such that  $m = m_1 = 1$ . Thus  $\Delta \simeq 1$  for small perturbations about flat spacetime, and furthermore,  $c$  is given by the *positive* solution in (101) above.

When nonlinear effects associated to strong coupling conspire so that the vierbein components strongly deviate from their flat spacetime values then two interesting pathologies may occur with the scalar constraint. First  $\Delta$  may appear to become negative, indicating that no (real) solution to the scalar constraint can exist for  $c$ . More precisely as  $\Delta \rightarrow 0^+$  the solution will become infinitely strongly coupled and the EFT breaks down before this

point. This is easy to see if we naively perturb around the solution with  $\Delta = 0$  then we would obtain say  $\delta c \sim \mp \frac{1}{2} \sqrt{\delta \Delta}$  which is inconsistent in perturbation theory if  $\delta \Delta$  has a first order perturbation. Requiring that  $\delta \Delta$  starts at second order imposes a restriction on the variables that is indicative of a degree of freedom being lost, i.e. its kinetic term vanishing. This is the tell-tale sign of infinite strong coupling.

A second feature that may occur is that the quadratic form linearizes with  $a_2 \rightarrow 0$  (with  $a_1$  remaining finite). Whilst we would not normally regard this as a pathology for a quadratic equation, depending on the sign of  $a_1$  this then picks a particular branch. The “correct” branch for  $a_1 > 0$  is the positive branch, and for  $a_1 < 0$  it is the negative branch, and then the usual linear solution is reproduced  $c \simeq -a_0/a_1$  in this limit  $a_2 \rightarrow 0$ . However being on the opposite “wrong” branch implies  $c \simeq -a_1/a_2$  and diverges as  $a_2 \rightarrow 0$ . We note that for flat spacetime we have  $a_1 < 0$  and are on the positive branch. If a situation arose where  $a_2 \rightarrow 0$  with  $a_1$  staying the same sign, then this would correspond to being on the “wrong” branch, and the solution for  $c$  would diverge which in itself indicates infinite strong coupling and the breakdown of the EFT.

We may identify these two pathologies by either  $\Delta$  tending to zero (strictly small values) or alternatively  $\Delta$  diverging positively as  $a_2 \rightarrow 0$ . In this latter case  $c$  would diverge positively. Later we will see that for certain choices of initial data indications of both these behaviors in the nonlinear collapse dynamics.

## B. Initial data

We begin with initial data that is an approximately ingoing pulse of the scalar field, starting initially away from the origin. We must then solve the Hamiltonian and momentum constraints as well as the scalar and vector constraints. We choose the width of the scalar pulse to be approximately the length scale associated to the graviton mass,  $\sim 1/m_1$ . In doing so we depart very much from the phenomenologically interesting regime of massive gravity since this would presuppose a spherical symmetric source of the size of the Hubble radius. For our purposes this is merely a proof of principle that the dynamical formulation we have developed is well defined. It is beyond the scope of this paper to consider the type of hierarchies and boundary conditions needed for phenomenological applications.

Thus from now on we choose units so that  $m_1 = 1$ . For the results we present here we take an initial approximately Gaussian profile for the scalar, localized at a radius of  $r \simeq 2$ , with a momentum profile that in flat spacetime would give a purely ingoing pulse,

$$\Phi(t=0, r) = \frac{A}{A_0} r^4 \exp\left(-\frac{(r^2-2)^2}{10}\right), \quad (104)$$

$$p_\Phi(t=0, r) = \partial_r \Phi(t=0, r) + \frac{1}{r} \Phi(t=0, r). \quad (105)$$

Here  $A_0$  is chosen so that  $A = \max(\Phi(t=0, r))$  is a constant giving the maximum amplitude of the scalar profile. The metric function  $a$  is the one that has second order dynamics, and we initially choose it to have its flat space value (which is zero), and vanishing momentum, so that

$$a(t=0, r) = p_a(t=0, r) = 0. \quad (106)$$

Recalling that the momenta  $p_h$  and  $p_b$  are eliminated using the vector constraint, then it remains to give the metric functions  $c$ ,  $h$  and  $b$  to determine the initial data. Now  $c$  is determined from the scalar constraint, but we must also solve the Hamiltonian and momentum constraints, giving the two conditions for  $h$  and  $b$ . These involve second spatial derivatives of  $b$ , and first derivatives of  $h$ .

We may solve the nonlinear system of scalar, Hamiltonian and momentum constraints by using an iterative relaxation method or Newton’s method—we have implemented both. Since our scalar pulse is quite far from the origin, and we start it with relatively small amplitude, the solution is close to the solution to the linearized system which is easily determined by taking the metric functions close to their flat spacetime values,

$$\begin{aligned} c(t=0, r) &= -1 - \epsilon \delta c, \\ b(t=0, r) &= 1 + \epsilon \delta b, \\ h(t=0, r) &= \epsilon \delta h, \\ \Phi(t=0, r) &= \sqrt{\epsilon} \delta \Phi, \end{aligned} \quad (107)$$

where one then finds  $\delta b$  is determined from the ordinary differential equation,

$$2\delta b'' + \frac{4\delta b'}{r} - 3m^2 \delta b = -\frac{\delta \Phi^2}{2r^2} - \frac{\delta \Phi \delta \Phi'}{r} - \delta \Phi'^2, \quad (108)$$

which may be solved by quadrature with the boundary condition that  $\delta b \rightarrow 0$  as  $r \rightarrow \infty$ , and is regular at the origin. The remaining  $\delta h$  and  $\delta c$  are given algebraically in terms of  $\delta \Phi$  and the solution to  $\delta b$  as

$$\begin{aligned} -3m^2 r^2 \delta c &= 9m^2 r^2 \delta b + 2r \delta \Phi \delta \Phi' + \delta \Phi^2, \\ m^2 r \delta h &= \delta \Phi' \left( \delta \Phi' + \frac{\delta \Phi}{r} \right). \end{aligned} \quad (109)$$

We may use this linearized approximate solution as an initial guess to solve the scalar, Hamiltonian and momentum constraints by an iterative relaxation or Newton’s method.

For convenience we compactify the radial coordinate as  $\tilde{r} = r/(1-r^2)$ , so that the real axis is compactified to the interval  $[0, 1]$  for  $\tilde{r}$ . We then employ sixth order spatial

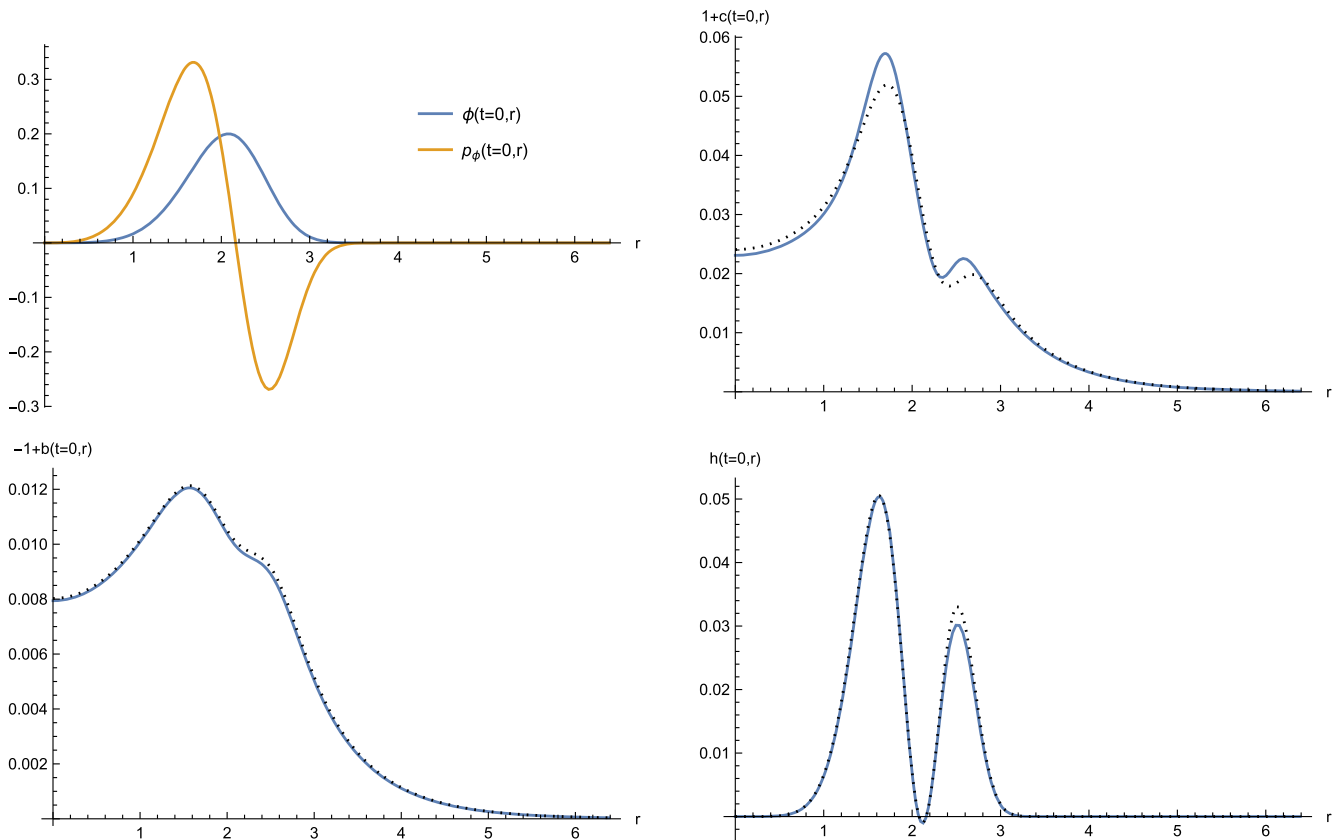


FIG. 1. The top left plot shows the initial radial profile for the scalar  $\Phi$  and its momentum  $p_\Phi$  for amplitude  $A = 0.20$  (in our units where the mass  $m = 1$ ). Solving the scalar, Hamiltonian and momentum constraints with this matter, we then obtain the initial metric functions  $c$ ,  $b$  and  $h$ , shown as solid curves in the top right, bottom left and bottom right frames, respectively. Recall we are choosing the remaining metric function  $a$  and its momentum to vanish. In order to solve this coupled nonlinear system we use relaxation or Newton's method, with an initial guess given by the linear approximation—these linear solutions are shown as dotted curves for each of these metric functions, and since the scalar pulse starts quite far from the axis with relatively small amplitude, it is close to the full nonlinear solution.

differencing. For the iterative relaxation scheme we use a method analogous to Gauss-Seidel for the Poisson equation, solving each of the three equations in turn at each lattice point, then moving to the next until the whole grid has been covered, and then we repeat until convergence is reached. This method, while crude, works well and is straightforward to implement, giving the same results as the Newton solver.

An example of this initial data is given in Fig. 1 for one of the largest amplitudes,  $A = 0.2$  used later in the discussion. In the figure we have plotted both the full nonlinear solution as well as the linearized approximation, which can be seen to be close.

### C. Evolution

We then evolve this initial data by imposing the scalar constraint, vector constraint and evolution equation  $\tilde{\mathcal{E}}^r$ , together with the matter scalar equation. As for finding the initial data, we use the compactified radial coordinate  $\tilde{r} = r/(1 - r^2)$ , and sixth order finite differencing for

spatial derivatives on the interval  $[0, 1]$  in  $\tilde{r}$ . For time derivatives we use an implicit Crank-Nicolson differencing scheme. We solve this implicit system using iterative relaxation. As mentioned above, while we may in principle solve the vector and scalar constraints to eliminate  $c$ ,  $p_h$  and  $p_b$  from the remaining equations, in practice since we implement an implicit scheme which must be solved at each time step anyway, we have found it convenient to simply solve the constraints as part of this implicit system.

Recall that we are using units for which the mass  $m_1 = m = 1$ . We take initial data to be a pulse with approximately unit width. Thus all scales in the problem are comparable, and the parameter we now vary is the scalar pulse amplitude  $A$ . For small amplitude we expect the theory to be well described by linear dynamics, where the pulse will travel in to the origin, pass through it and then disperse to infinity. As we increase the amplitude  $A$  we expect to see nonlinear behavior, and it is this that is our focus. The natural question is whether in this massive theory of gravity we see a horizon form, or some different nonlinear phenomena.



In order to ensure that our continuum  $3 + 1$  system is well posed under time evolution, we complete it at short distance by adding the diffusion terms as discussed in the previous section. Focussing on the long wavelength physics and typical dynamical timescales we will see we are insensitive to this term, provided the diffusion constant,  $D = \ell^2$ , is sufficiently small. Without the diffusion term, so setting  $D = 0$ , we find some short distance instability on the scale of the lattice spacing when employing resolutions greater than  $N > 100$  points, and becoming more severe for evolutions that deviate further from flat space. The instability typically arises near the origin, and it is unclear at this stage whether this is a result of the ill posedness of the continuum equations (without diffusion) or whether it is an artifact of the numerical discretization—recall that even well-posed continuum equations may have lattice scale pathologies depending in detail on the numerical scheme chosen to discretize them.<sup>9</sup> Taking  $D = 0.001$  stabilizes the system for all resolutions considered here, up to the highest we have implemented  $N = 1600$ , and this value of  $D$  is the one used to make the figures presented here unless otherwise stated. As discussed later, by varying the value of this diffusion constant, we can confirm that this value is sufficiently small that it has essentially no impact on the solutions we find, and the diffusion term is irrelevant on the length scales and timescales of interest, i.e. those of order  $\sim \mathcal{O}(1)$  in size.

We are able to simulate for a range of lattice spacings. The data we show here is for  $N = 400$  lattice points, which gives very good accuracy with our sixth order spatial finite difference. A small time step is required for stability of the Crank-Nicholson scheme, and in the data shown we typically take  $\Delta t = 0.00002$ . Refining  $N$  indeed shows that our code converges to a good continuum limit, and we give more details of this convergence in Appendix B. However, since we modify the short distance physics using the diffusion terms, and we regard these short scales as being beyond the validity of our EFT, in practice we find that our resolution of  $N = 400$  is sufficient to reproduce the long distance physics of interest for  $D = 0.001$ . Taking higher resolutions shows our discretization properly approaches the continuum given by the well-posed low energy truncation together with short distance diffusive completion, but refining past  $N = 400$  accesses the scales dominated by this diffusion and does not reveal the low energy physics we are interested in more accurately.

Finally the Hamiltonian and momentum constraints, once satisfied for the initial data, are preserved during the evolution if one takes only the low energy truncation. However, with the diffusion terms added then already at the

level of the continuum PDEs the constraints will no longer be exactly preserved under evolution. In Appendix B we study the violation of these constraints under evolution as we change the resolution,  $N$ , and also the diffusion constant  $D$ . As expected we find that for a given small diffusion constant  $D$ , refining  $N$  makes this constraint violation smaller until some value of  $N$  past which the violation is caused by the diffusion terms at the level of the continuum PDEs and is not due to numerical discretization error. For smaller  $D$ , a larger  $N$  is reached before diffusion dominates the violation, and the smaller the constraint violation becomes. Again for our typical choice of  $D = 0.001$  we find that for  $N = 400$  the constraint violation is small, and is dominated by the diffusion terms rather than numerical discretization error.

#### D. Low amplitude initial data

First let us consider the collapse of a small amplitude scalar pulse, taking small  $A = 0.01$ , so that the subsequent evolution is a weak perturbation of Minkowski spacetime, bearing in mind that this is far from a phenomenologically realistic situation. The initial ingoing pulse of the scalar traverses towards the center, increasing in amplitude as it becomes increasingly focussed. It then reaches the origin, with still a relatively small peak height of  $|\phi| \sim 0.05$ , passes through, and subsequently disperses. In Fig. 2, we show the scalar field as a function of time and the radial coordinate. The behavior is qualitatively similar to that for GR, although obviously in detail differs due to the mass. Also in the figure the vierbein function  $c(t, r)$  is shown, and we see that this metric component deviates from its flat space value of zero with a maximum amplitude of  $\sim 0.02$ . The same is true for the other components. As a result the discriminant function, also shown, remains very close to its Minkowski vacuum value  $\Delta \simeq 1$ .

To examine the nature of the spacetime, we further show that the Ricci scalar deviates from its trivial zero value at order  $\mathcal{O}(10^{-5})$ , and the trace of the stress tensor  $T^\mu{}_\mu$  has amplitude  $\sim 0.05$  as the pulse traverses the origin. Recall that  $m = 1$  in our units, and all scales in the initial data are  $\sim \mathcal{O}(1)$  too. We see that the Ricci scalar is much smaller than this stress tensor trace, as we expect from linear theory where it vanishes for massive gravity. (Recall that in GR we would simply have  $R = -T$  in our units.)

For comparison purposes, we then show in Fig. 3 the collapse of an ingoing pulse with a greater initial amplitude of  $A = 0.04$ . Here we see a stronger nonlinear response to the pulse, although a similar qualitative behavior as the pulse reaches the origin with a height of  $\sim 0.2$  and again then disperses. Now the metric function  $c$  is clearly perturbed from its zero Minkowski value by an  $\mathcal{O}(1)$  amount as the matter traverses the origin. The same is true for the other vierbein components. While the Ricci scalar remains small compared to the characteristic scales of the problem, we see a much stronger response than for

<sup>9</sup>It is interesting to note that the most well-behaved PDE, the diffusion equation itself, is unstable numerically when using explicit time differencing, and even using an implicit Crank-Nicholson scheme it requires the time step to be sufficiently small to avoid lattice scale instabilities.

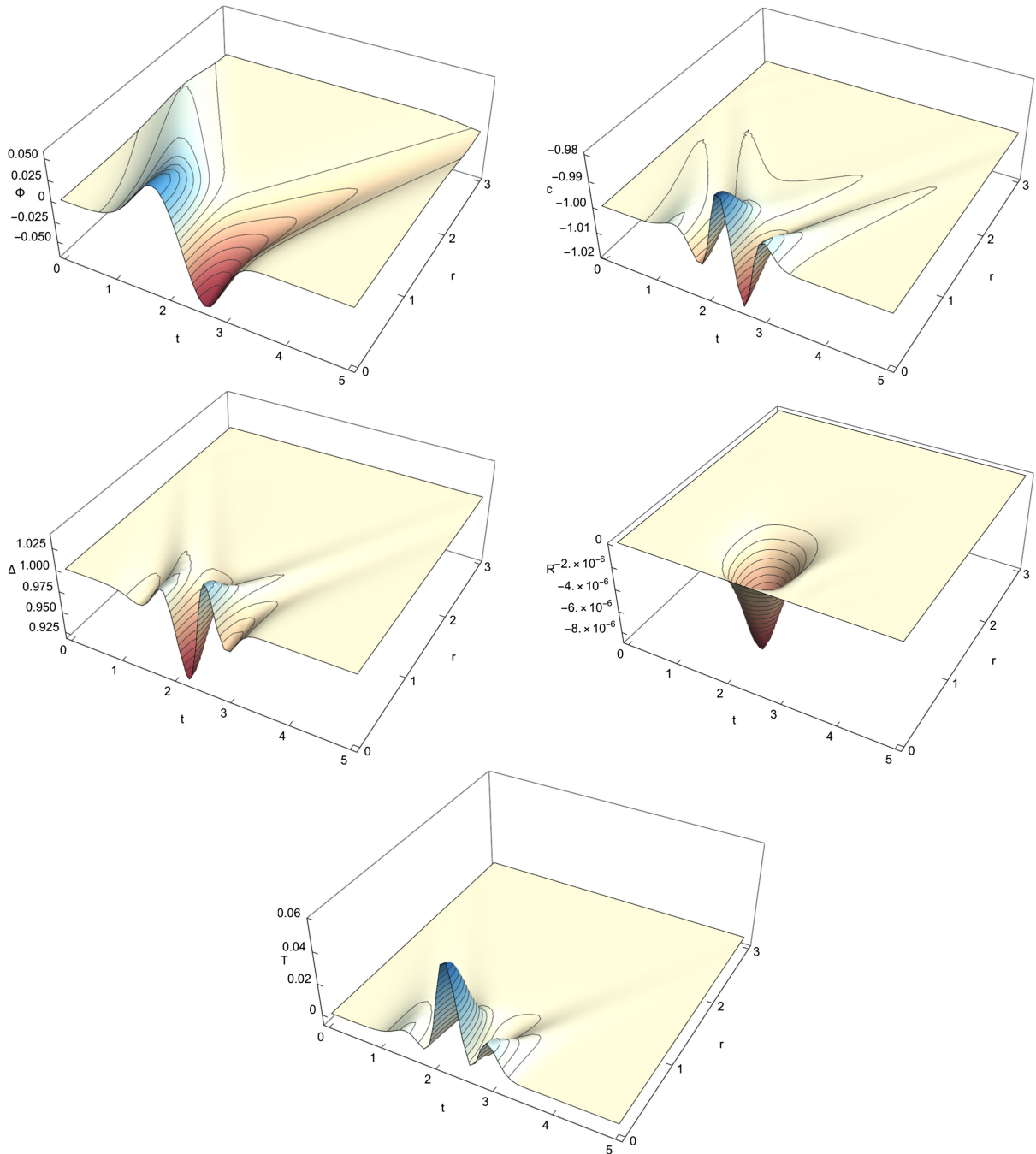


FIG. 2. Figure showing the scalar field for weak initial data with  $A = 0.01$  in the top left-hand frame. During the evolution all the metric components remain close to their values for Minkowski spacetime, with the ingoing scalar pulse passing through the origin and then dispersing. The vierbein component  $c$  is shown in the top right-hand frame, and the other metric components have a response of similar amplitude. The middle left figure shows the discriminant function,  $\Delta$ , for the quadratic scalar constraint. The middle right-hand figure shows the Ricci scalar  $R$  and the bottom plot shows trace of the stress tensor  $T$ . For linear perturbations the Ricci scalar vanishes in massive gravity (unlike GR where  $R = -T$ ).

the pulse with  $A = 0.01$ , reflecting the fact that the Ricci scalar reacts nonlinearly (recall it vanishes in linear theory). We see that the discriminant  $\Delta$  now reaches much smaller values, dipping to  $\sim 0.1$  near the origin as the pulse transits through. As already discussed, if  $\Delta$  becomes very small one may still be able to classically evolve the leading terms of

the EFT, but the full theory will become strongly coupled and is classically no longer under control. We emphasize that precisely the same phenomena happens for GR—in principle we may evolve classical GR right up to a curvature singularity, but we should not trust the full theory once curvatures reach the Planck scale due to strong

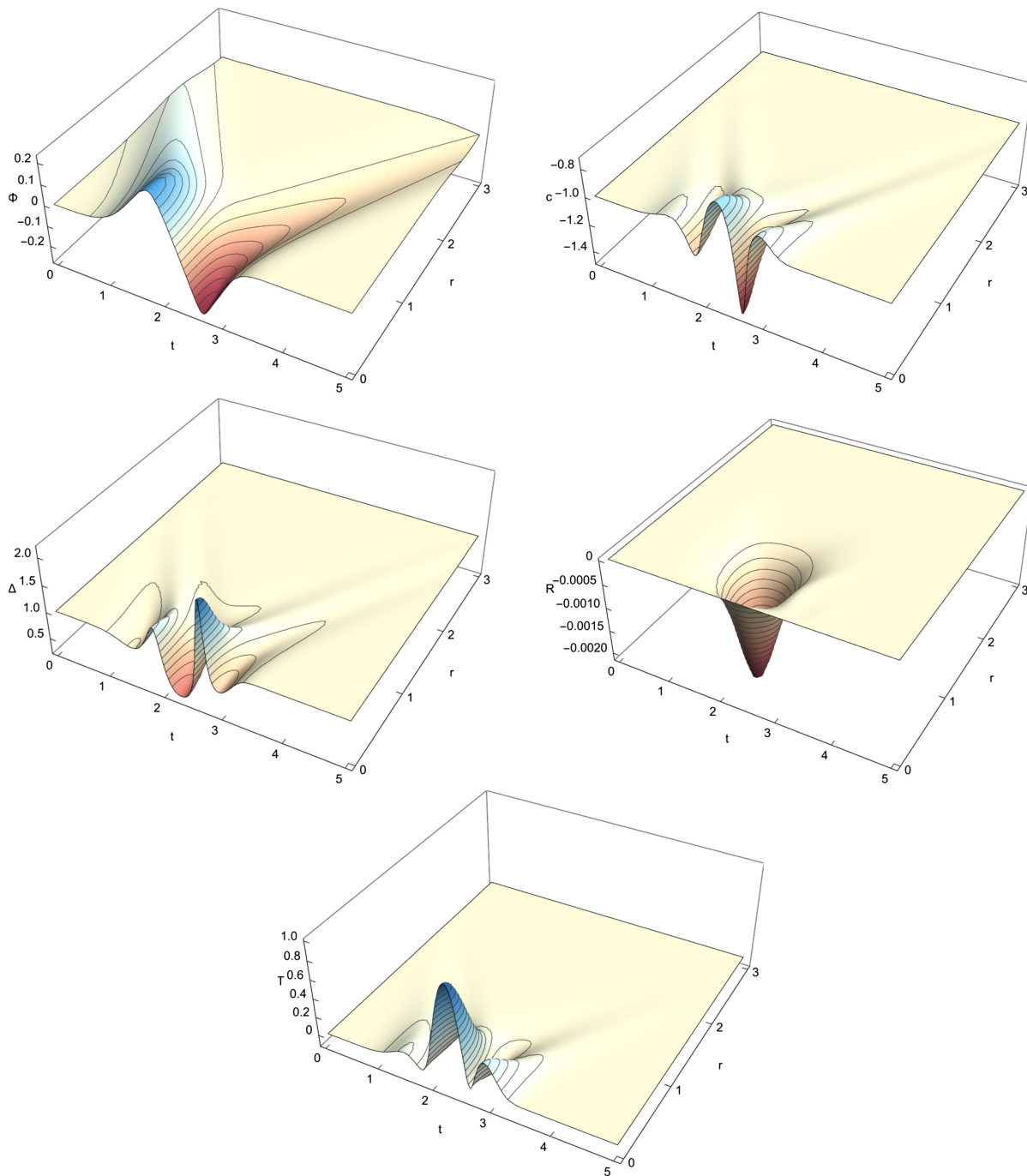


FIG. 3. The same quantities are shown as in the previous figure, now for an ingoing scalar pulse with greater initial amplitude  $A = 0.04$ . We see a greater response in the metric function  $c$ , and correspondingly in  $\Delta$ , and likewise for the trace of the stress tensor. We see a much larger response in the Ricci scalar, of order  $10^3$  larger than for  $A = 0.01$ , reflecting that its dependence on amplitude is nonlinear (recall it vanishes in linear theory).

coupling. For  $\Delta \sim \mathcal{O}(1) \gg \ell^2$ , one can in principle still expect to trust the theory, but as we carry on with stronger initial data, as we shall do shortly, we will reach a point where the EFT truncation can no longer be used to describe the system dynamics.

### E. Large amplitude initial data and a breakdown of evolution leading to strong coupling

We now consider larger amplitudes of initial data. In Figs. 4 and 5 we plot  $\Phi$ ,  $c$ ,  $\Delta$  and also the quadratic coefficient  $a_1$  in the scalar constraint for the amplitudes

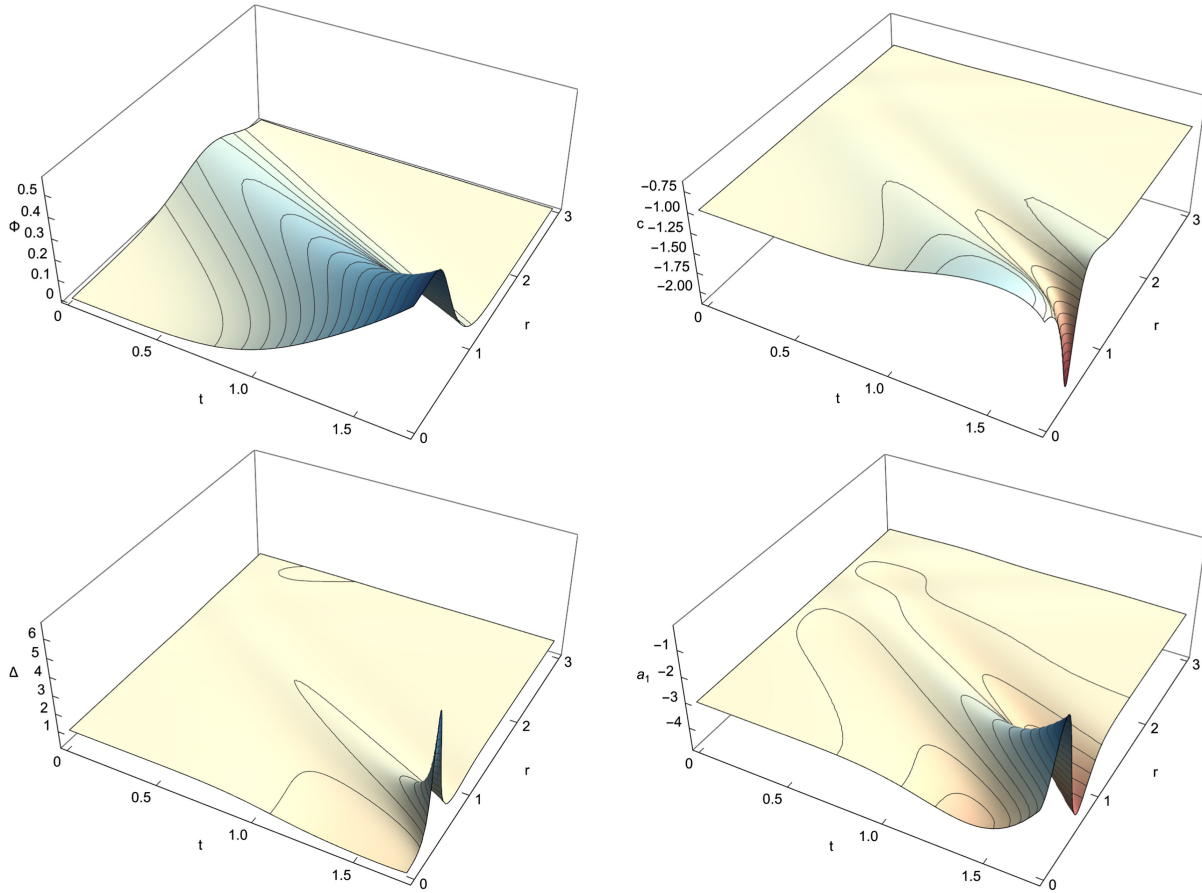


FIG. 4. For an initial amplitude  $A = 0.12$  the evolution breaks after a time  $t \simeq 1.7$ . The top plots show the matter field  $\Phi$  (left) and  $c$  (right), and the bottom ones show  $\Delta$  (left) and the coefficient  $a_1$  from the scalar constraint (right). These are plotted over the whole range where the evolution exists. We see that near the end of the evolution  $\Delta$  appears to be vanishing near the origin, whereas it appears to diverge where the matter pulse is located. The function  $a_1$  is negative everywhere, and hence the divergence in  $\Delta$  corresponds to being on the “wrong” branch of the quadratic solution as the equation linearizes. Indeed we see  $c$  appear to diverge negatively where  $\Delta$  does so. Thus the time evolution ends as the scalar constraint becomes pathological, and cannot be solved for a finite real  $c$ .

$A = 0.12$  and  $A = 0.20$ , respectively. Both numerical evolutions break down *before* the matter pulse reaches the origin, in the sense that the implicit system governing the timesteps cannot be solved. The figures show the evolutions up until the timeslice when the simulation breaks. Looking at the discriminant  $\Delta$  and the quadratic coefficient  $a_1$  we can see why the evolutions break down. In both cases as the matter pulse traverses towards the origin and grows due to the focussing, we see  $\Delta$  become small near the axis, but also appear to diverge positively away from the axis approximately at the location of the peak of the scalar field amplitude. We can see that although the coefficient  $a_1$  is decreasing in the region where  $\Delta$  becomes large, it remains positive, and hence the diverging of  $\Delta$  signifies being on the “wrong” branch—as a result we see that the vierbein function  $c$  appears to diverge in the region where  $\Delta$  does. Since this phenomenon occurs with a change of branch, perturbations become infinitely strongly coupled on that solution and it can hence no longer be trusted before  $c$  even diverges.

It is difficult to definitively diagnose whether the simulation is breaking due to  $\Delta$  diverging where the matter pulse is located, or whether it is due to  $\Delta$  going towards zero near the origin. We believe it is the former. Either way, while in principle we can simulate the truncated EFT up to the point of breakdown, where we can no longer solve the Cauchy problem, the theory will become strongly coupled before that point, and there the details of the microscopic physics will become important and need to be folded in to understand the true behavior.

A natural question is whether this breakdown in evolution of the truncated EFT is also associated to the development of a curvature singularity. As usual, the Einstein equation determines the Ricci tensor of the metric, and from this we may compute the Ricci scalar  $R$  and the invariant  $\text{Ric}^2 = R_{\mu\nu}R^{\mu\nu}$ . However, again as for GR, the Kretschman invariant  $\text{Riem}^2 = R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}$  is not determined by the Einstein equation, but may be computed directly from the metric. We plot these quantities in Fig. 6



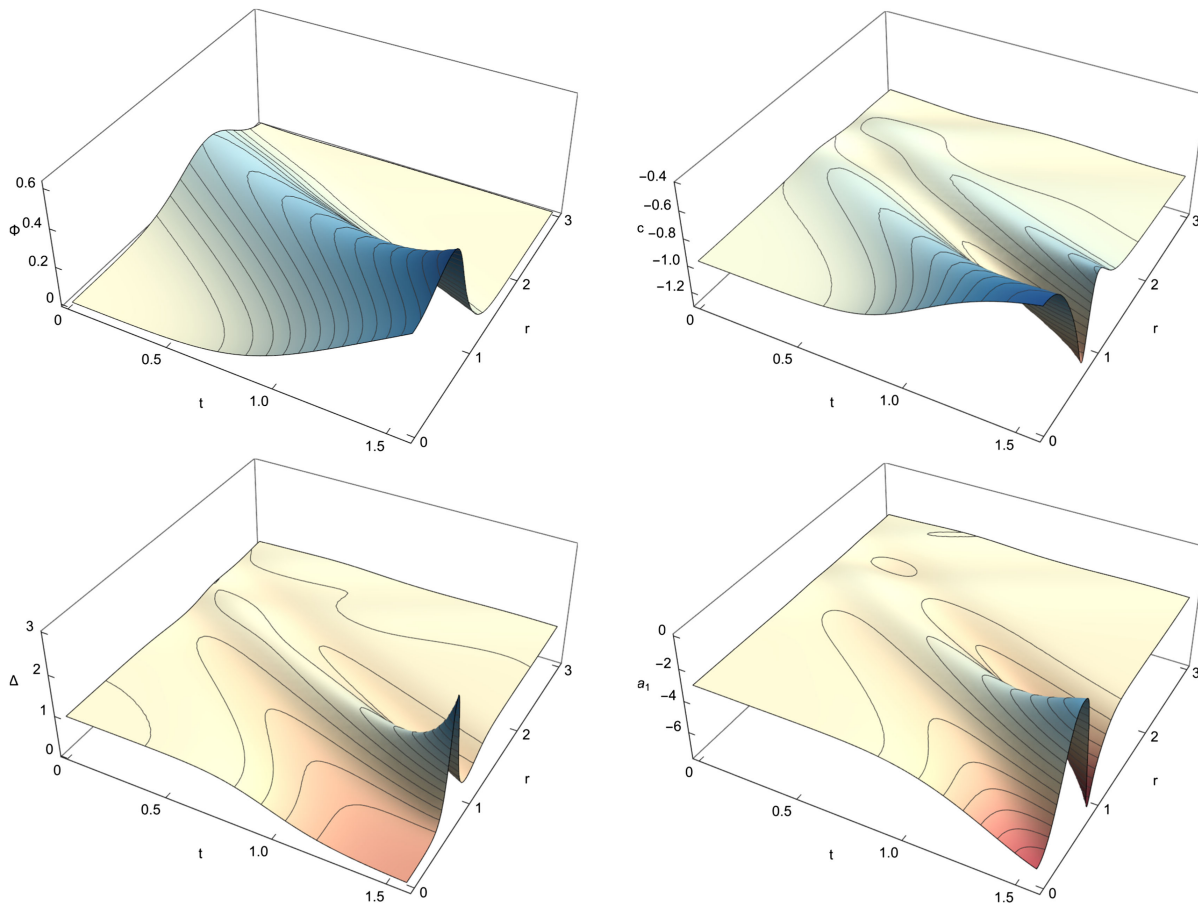


FIG. 5. This figure shows the same plots as in the previous figure for a larger amplitude  $A = 0.2$ . Here we see the time evolution breaks even earlier at  $t \simeq 1.5$ . We see the same behavior in the various functions plotted, and in particular again we see a divergence in  $\Delta$  which we believe ultimately signals the scalar constraint becomes pathological and cannot be satisfied.

for the evolution  $A = 0.12$  and in Fig. 7 for  $A = 0.20$ . In both cases we see the curvature invariants increase near the axis, and the region where  $\Delta$  becomes large. In the case  $A = 0.12$  they do not appear to be strongly diverging. The case  $A = 0.20$  is less clear, as while Ricci and  $\text{Ric}^2$  look well behaved,  $\text{Riem}^2$  appears as if it is quickly increasing as  $\Delta$  looks to diverge, although its value is still relatively small when the evolution breaks down. Thus the  $A = 0.12$  example shown suggests that the pathology may be due to a breakdown in evolution that is not associated to diverging curvature. Such a result would be interesting as it would indicate how strong coupling in the reference metric sector (or its Stückelberg fields) may not necessarily propagate in the standard gravitational sector. The  $A = 0.20$  is less conclusive. Other amplitudes have similar behaviors to the two cases shown, and thus we refrain from making a strong statement about whether there is a curvature blow up.

The theory is invariant under a combined coordinate transformation of the metric and reference metric. Clearly if curvature invariants diverge, then there is no way to remove a singularity by changing coordinates. If instead the

evolution breakdown is not due to curvature diverging, one might naively wonder whether these pathologies are then simply a result of a bad unitary gauge coordinate choice. Also in the Figs. 6 and 7 we show the ratio of determinants of the metric and reference metric,  $\det(g_{\mu\nu})/\det(f_{\mu\nu})$  for the two cases  $A = 0.12$  and  $A = 0.20$ . We see that this ratio appears to be diverging in the region where  $\Delta$ , and consequently  $c$  becomes large. Crucially this ratio of determinants is a coordinate invariant, and hence its divergence signifies this behavior cannot be removed by a coordinate transformation. Hence even if it is the case that curvature remains bounded, the singular behavior we are seeing associated to  $\Delta$  and  $c$  diverging is definitely not a coordinate artifact.

In fact we can argue that neither pathology of the scalar constraint, so the behavior associated to  $\Delta \rightarrow \infty$  or that where  $\Delta$  becomes negative, can be removed by a diffeomorphism. Consider the case of  $\Delta$  diverging, resulting in  $c$  diverging as well (which we believe is driving the breakdown in the evolutions presented here). To render the metric components finite would then require a singular gauge transformation which would then yield a divergent

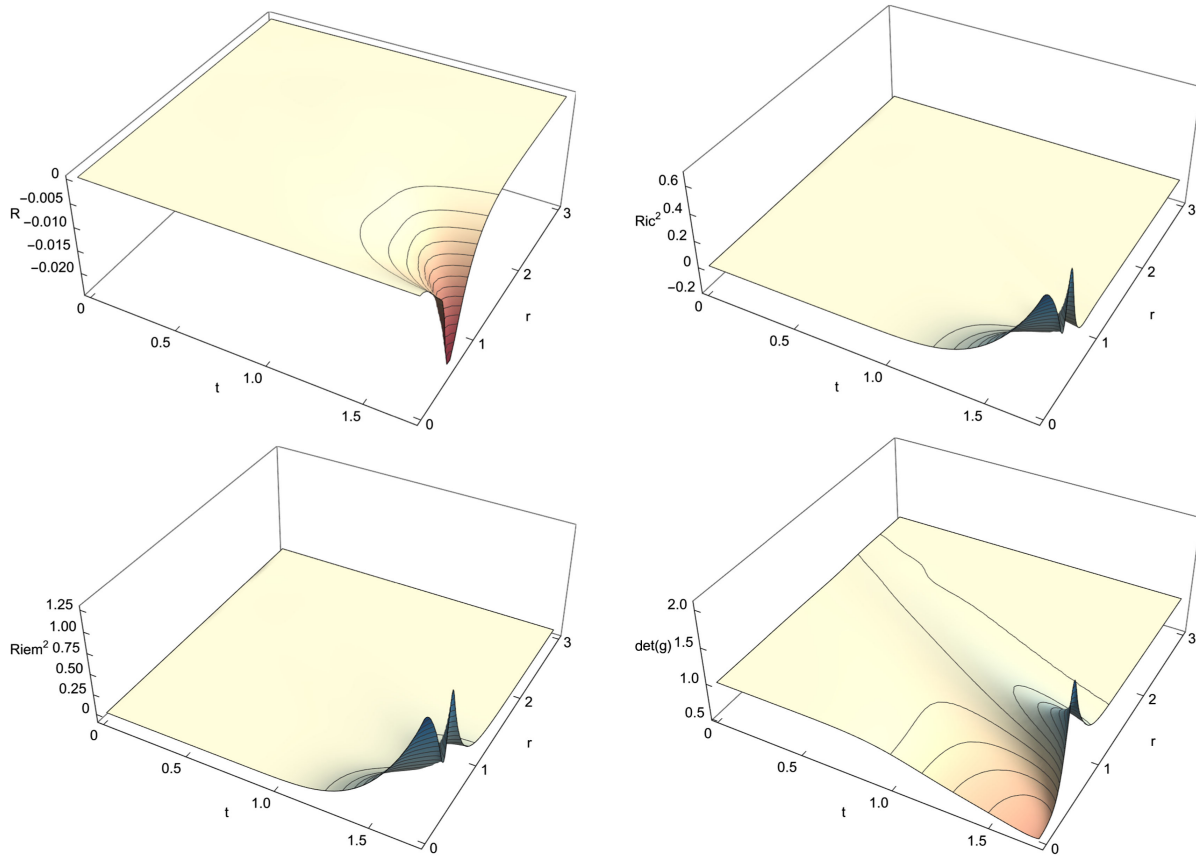


FIG. 6. This figure shows the following curvature invariants for the  $A = 0.12$  evolution: Ricci scalar (top left),  $\text{Ric}^2 = R_{\mu\nu}R^{\mu\nu}$  (top right) and  $\text{Riem}^2 = R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}$  (bottom left). While these increase in the region where the strong coupling develops, they remain quite small. It is therefore unclear whether the pathology in evolution of the truncated EFT is associated to a curvature singularity developing. In the bottom right frame we plot the ratio of determinants  $\det(g_{\mu\nu})/\det(f_{\mu\nu})$ , which is a coordinate invariant, and this appears to become singular, diverging positively in the region where  $\Delta$  and  $c$  look to be blowing up.

reference metric. One can move the divergence between the metric and the reference metric, but there will be no choice of coordinates where both are smooth. Now consider the other case that  $\Delta$  shrinks to zero. Extending past this point presumably requires negative  $\Delta$  and hence complex  $c$  (although it may be bounded in magnitude), and consequently a complex metric. Again a complex coordinate transformation might restore the metric to a real form, but would then render the reference metric to be complex. Thus both pathologies would represent genuine physical breakdown of the EFT, and are not simply coordinate artifacts, even if they do not give rise to curvature blow up.

As the initial scalar amplitude is increased from  $A = 0.2$  the evolution breaks earlier, and as it is decreased the evolution runs longer, with the apparent breakdown in the scalar constraint occurring when the pulse is nearer to the origin. An amplitude of approximately  $A \simeq 0.06$  appears to divide the weak field dispersive behavior from the large amplitude evolutions which break down. This is shown in Fig. 8 where we plot the value of  $\Delta$  at the origin as a function of time for increasing initial amplitude. For the

curves with  $A = 0.06$  and  $0.07$  then  $\Delta$  is only plotted for times up until the evolution breaks down. For these smaller amplitudes it is unclear whether the evolution breaks due to the scalar constraint becoming pathological precisely at the origin, or close to it. In the previous figures, say for  $A = 0.12$ , we clearly see a candidate pathology away from the axis of symmetry. It would be interesting to explore whether there is a critical behavior associated to this transition regime in between dispersion and breakdown of evolution with its associated strong coupling (i.e. examine whether the two behaviors are continuously connected as in a “second order transition”).

As the evolution breaks down it is interesting to see how other physical quantities are affected. An interesting one is the light cone structure governing the scalar field matter, as measured in the rest frame of the reference metric. Since the coordinates we use to evolve are naturally defined by the reference metric, we may determine the null in and outgoing one forms of the inverse metric at a point, and then use the reference metric to convert these to vectors. We plot these in Fig. 9 for both the evolutions with  $A = 0.12$  and

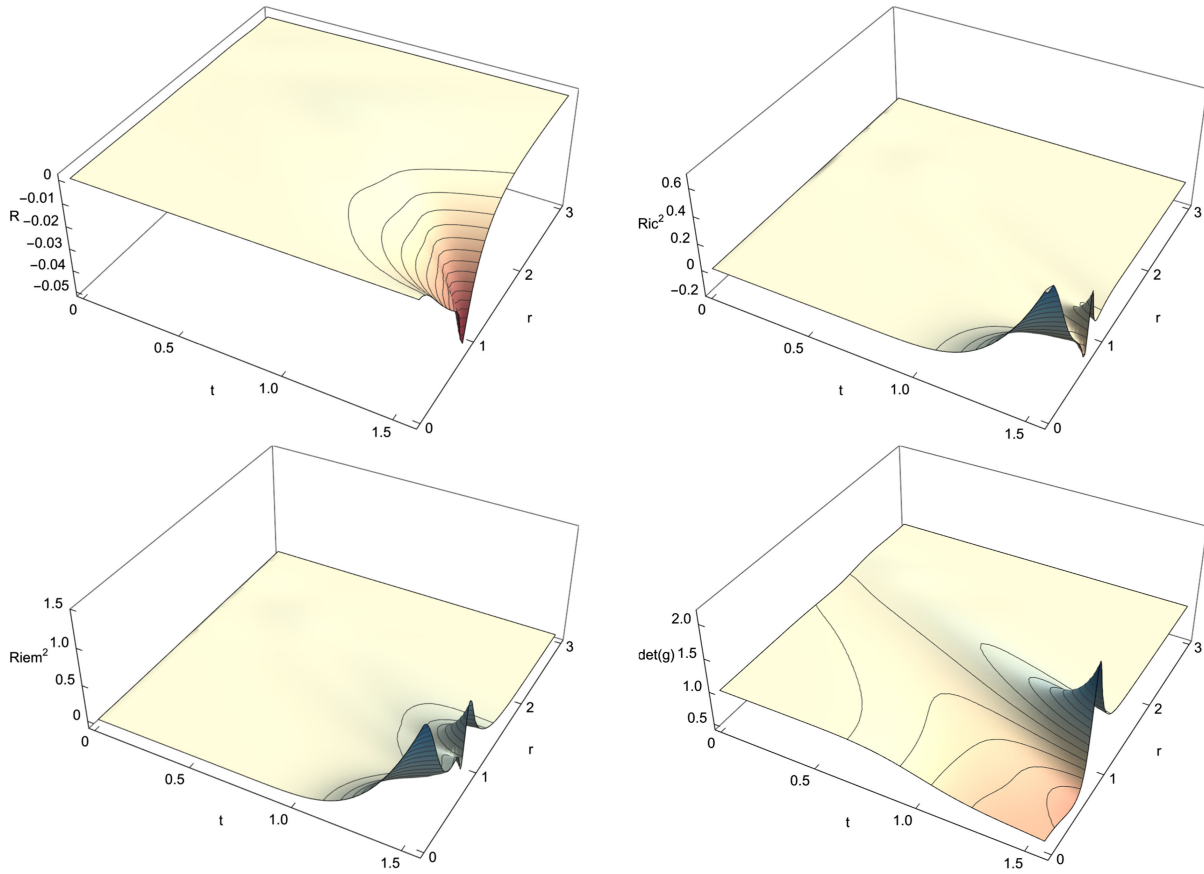


FIG. 7. This figure shows the same curvature and coordinate invariants as in the previous one, now for  $A = 0.20$ . We see that while  $\text{Riem}^2$  remains relatively small, it does appear to grow quickly in the region where  $\Delta$  and  $c$  diverge. Thus in this case it also remains ambiguous whether curvatures remain bounded as the evolution breaks down.

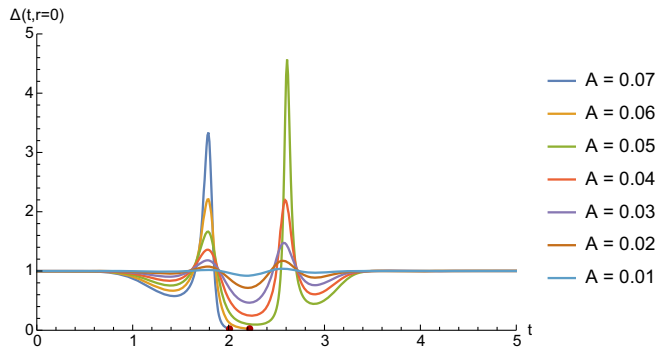


FIG. 8. Plot showing the value of the scalar constraint discriminant,  $\Delta$ , at the origin, as a function of time for evolutions with different initial amplitudes  $A$ . For evolutions with initial  $A \leq 0.05$  the scalar pulse passes through the origin at  $t \sim 2$  and subsequently disperses, and  $\Delta$  returns to its Minkowski value of  $\Delta_{\text{Mink}} = 1$ . However for larger values,  $A = 0.06$  and  $0.07$  (or indeed greater values) the numerical evolution breaks down before  $t \sim 2$  and is associated to strong coupling, and the curves are plotted only up to the point the simulation fails, indicated by the red markers.

0.20. For  $A = 0.12$ , we clearly saw (in earlier Fig. 4) the vierbein component  $c$  appearing to diverge at the point of breakdown where  $\Delta \rightarrow \infty$ , which is away from the origin, and in this region we see the scalar field light cones tip so that the ingoing propagation speed appears to be diverging. We also plot these ingoing speeds, again measured relative to the rest frame of the reference metric, in the figure. We see precisely the same phenomena for  $A = 0.2$ , although there we also see a closing of the light cones near the origin, presumably associated to  $\Delta$  becoming small. One might wonder whether it is this divergence in the matter light cone that is driving the breakdown in the evolution. We believe this is a consequence of the breakdown, due to diverging vierbein components, rather than the factor directly causing it. Here the massive graviton is dynamical in spherical symmetry (unlike the usual graviton of GR, as reflected in Birkhoff's theorem) with its somewhat complicated constrained dynamics. Deducing whether this graviton has a hyperbolic light cone structure is a complicated task which we leave for future investigations, but certainly we have seen that the scalar constraint it must obey becomes

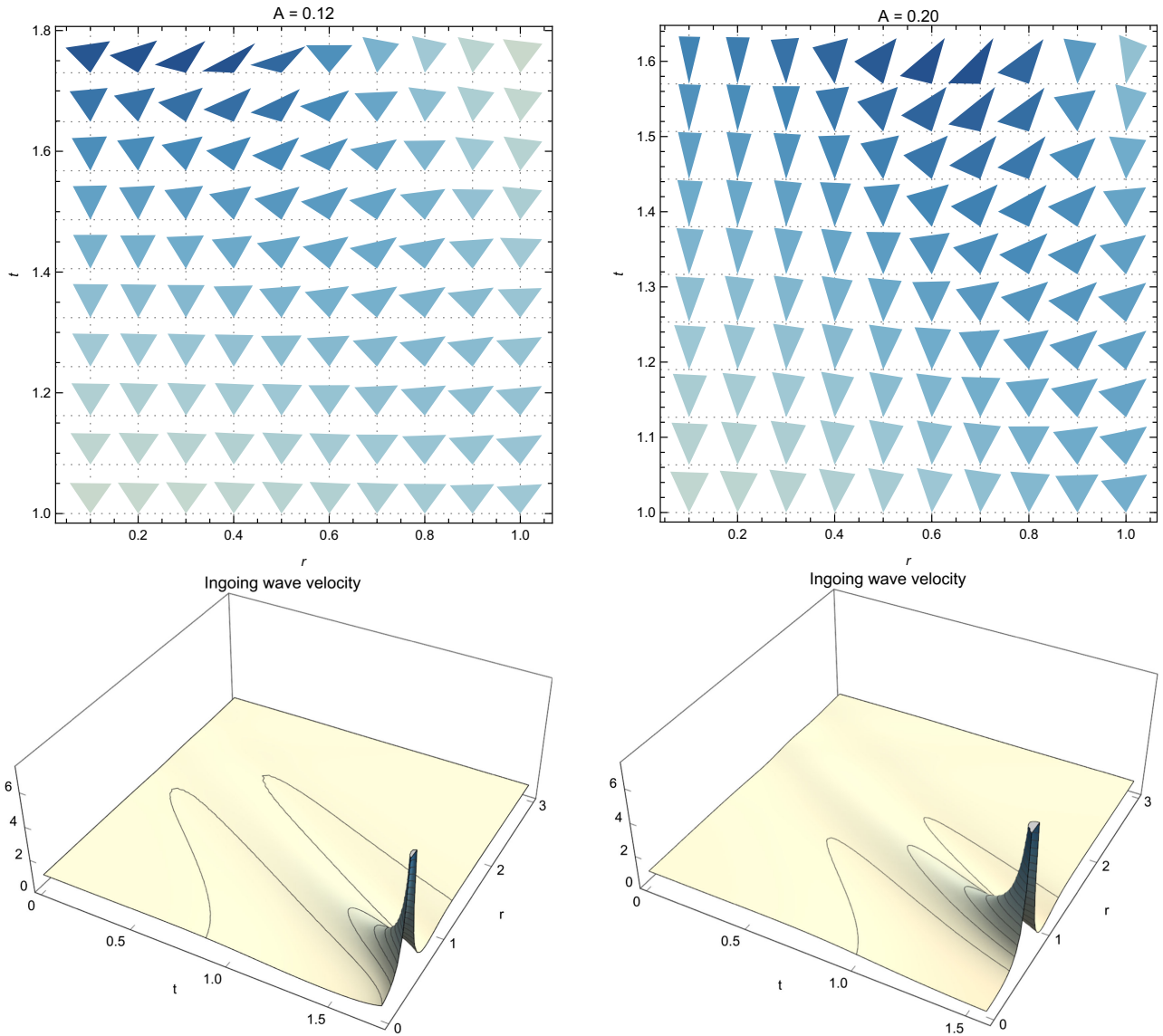


FIG. 9. Plot showing the light cones of the scalar field in the  $t$  and  $r$  coordinates, which are the natural rest frame defined by the reference metric, and also the coordinates we use for the evolution. The upper left plot shows the light cones for the evolution with  $A = 0.12$  until the point the evolution breaks down, and the upper right plot shows the same for  $A = 0.20$ . We see that the ingoing waves speed appears to diverge in the region where  $\Delta \rightarrow \infty$  for these evolutions. For  $A = 0.20$  we also see the light cones narrow at the origin where  $\Delta$  becomes small. In the lower left and right plots we display this ingoing wave speed relative to the reference metric rest frame for  $A = 0.12$  and  $0.20$ , respectively, clearly seeing the divergence at the point where the evolution ends.

pathological at the breakdown, as seen through the behavior of  $\Delta$ . Hence we believe the breakdown is driven by the graviton dynamics rather than that of the matter sector.

### F. Independence on the diffusion terms

As stated earlier, the previous figures were made for evolutions with diffusion constant  $D = \ell^2 = 0.001$ . A key concept is that this diffusion constant must be sufficiently small to ensure good behavior on small scales, whilst being irrelevant for the low energy physics we are interested in. If we try to remove the diffusion altogether, setting  $D = 0$ , simulations where we see reasonable deviations from flat

spacetime break down due to lattice scale instabilities. As with any numerical discretization it is difficult to say whether this is a result of the low energy truncation being ill posed, or whether it is our numerical scheme introducing lattice pathologies. However taking  $D = \ell^2 = 0.001$  controls lattice scales up to the highest resolutions we probed ( $N = 1600$ ), and is sufficiently small that it is irrelevant for the low energy behavior we are studying.

To demonstrate this we may run the same initial data with different diffusion constants to see a good limiting behavior as  $D$  becomes small. In Fig. 10 we plot the function  $\Delta$  at the axis as a function of the time coordinate



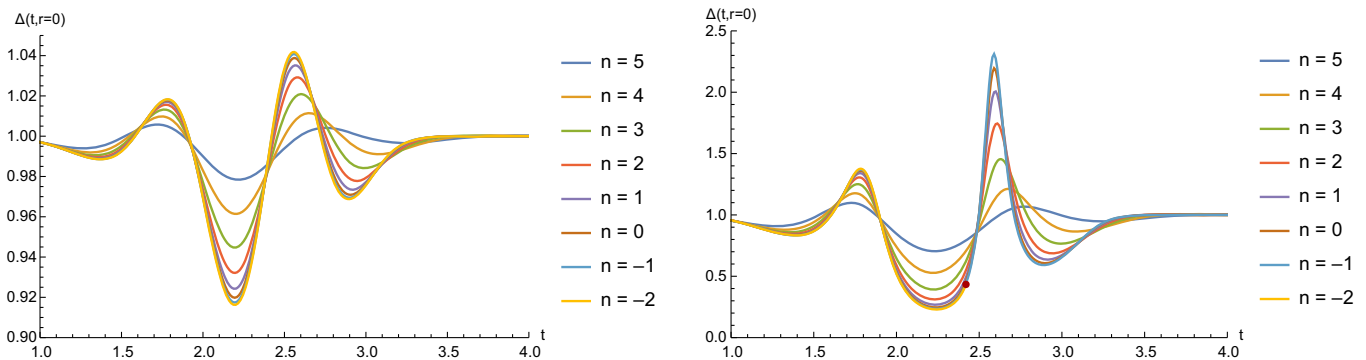


FIG. 10. Figure showing the effect of varying the diffusion coefficient  $D = \ell^2$  on the time dependence of  $\Delta$  evaluated at the origin, for an initial amplitude  $A = 0.01$  (left), or  $A = 0.04$  (right), shown at times  $1 \leq t \leq 4$ . We take values  $D = 2^n \times 10^{-3}$  and show curves for  $-2 \leq n \leq 5$ . We see that as  $D$  approaches zero, so  $n$  gets smaller, the curves appear to converge on a (putative) continuum value. However if  $D$  is taken too small, such as for  $n = -2$  in the right-hand plot (marked by the red dot), then the simulation breaks due to lattice scale instabilities. We see that setting  $n = 0$ , so  $D = 0.001$ , as we have for most of the plots in this work, is sufficiently small to allow long time simulations, and yet keep corrections from the diffusion terms being irrelevant.

for both  $A = 0.01$  and  $A = 0.04$  for a sequence of diffusion constants, each half the previous one (all for a fixed resolution of  $N = 400$  points). We plot the sequence  $D = 2^n \times 10^{-3}$  for integer  $-2 \leq n \leq 5$  so that  $D$  spans the range from  $\sim 0.0002$  up to  $\sim 0.03$ . We see that by eye the resulting curves are very similar for  $n \leq 0$ , so  $D \leq 0.001$ . Indeed the same is true for all the functions plotted above in Figs. 2 through to 7—the effect of the diffusion term for  $D = 0.001$  is nearly everywhere less than percent level and doubling the value of  $D$  results in new figures whose difference from those with the original value cannot be distinguished by eye. The exception to this is regions where the theory becomes strongly coupled. As we can see in Fig. 10, for  $A = 0.04$  near the second peak of  $\Delta$ , the effect of the diffusion is slightly larger, at a few percent. This is consistent with the expectation that where the theory starts to develop strong coupling we should become increasingly sensitive to these higher order operators.

### G. Lack of horizon formation

We have seen for sufficient amplitude initial data we apparently see a breakdown in evolution associated to strong coupling. The dynamics in such a region of strong coupling will then sensitively depend on the short distance completion of the theory. A crucial question is then whether this region is visible to an asymptotic observer. An interesting quantity to compute is the value of  $g_{tt}$ , the time-time component of the metric at the origin. Since the curve  $r = 0$  at the origin is a timelike geodesic, then  $g_{tt} = -c^2$  is a physical quantity that measures the relative redshift/blueshift of this geodesic as seen by an asymptotic observer. In Fig. 11 we plot  $|g_{tt}|$  at the origin against time for increasing amplitudes of initial data, up to  $A = 0.06$  where the evolution just breaks down, apparently at or very near to the origin.

We see a very different pattern to that in GR, where collapsing matter leads to a decrease in  $|g_{tt}|$ , corresponding to a redshift of physics at the origin as seen by an asymptotic observer, and for sufficient amplitudes  $|g_{tt}|$  vanishes as a horizon forms. Instead in this minimal massive theory we see the collapsing matter shell leads to a significant blue shift just before the matter reaches the origin at  $t \sim 2$ , when there is then a redshifting, before another period of even larger blue shift while the matter disperses. Looking at this redshifted period, the minimum value of  $|g_{tt}|$  attained for each evolution does not seem to be

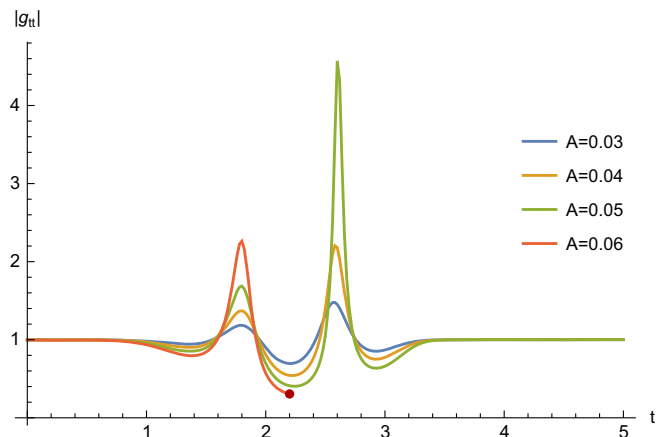


FIG. 11. The behavior of the  $g_{tt}$  component of the physical metric at origin, where it is equal to  $-c^2$ . When  $r = 0$  the off-diagonal terms in the metric vanish, so we can take this to be a diagnostic of the formation of an apparent horizon, which would happen for  $g_{tt} = 0$ . As we see  $|g_{tt}|$  remains positive for the case  $A = 0.06$  where the evolution breakdowns, as indicated by the red marker. It does not appear to tend to zero, but instead we see periods of time where there is a strong blueshifting with  $|g_{tt}|$  becoming quite large for the stronger amplitudes  $A = 0.05$  and  $0.06$ .

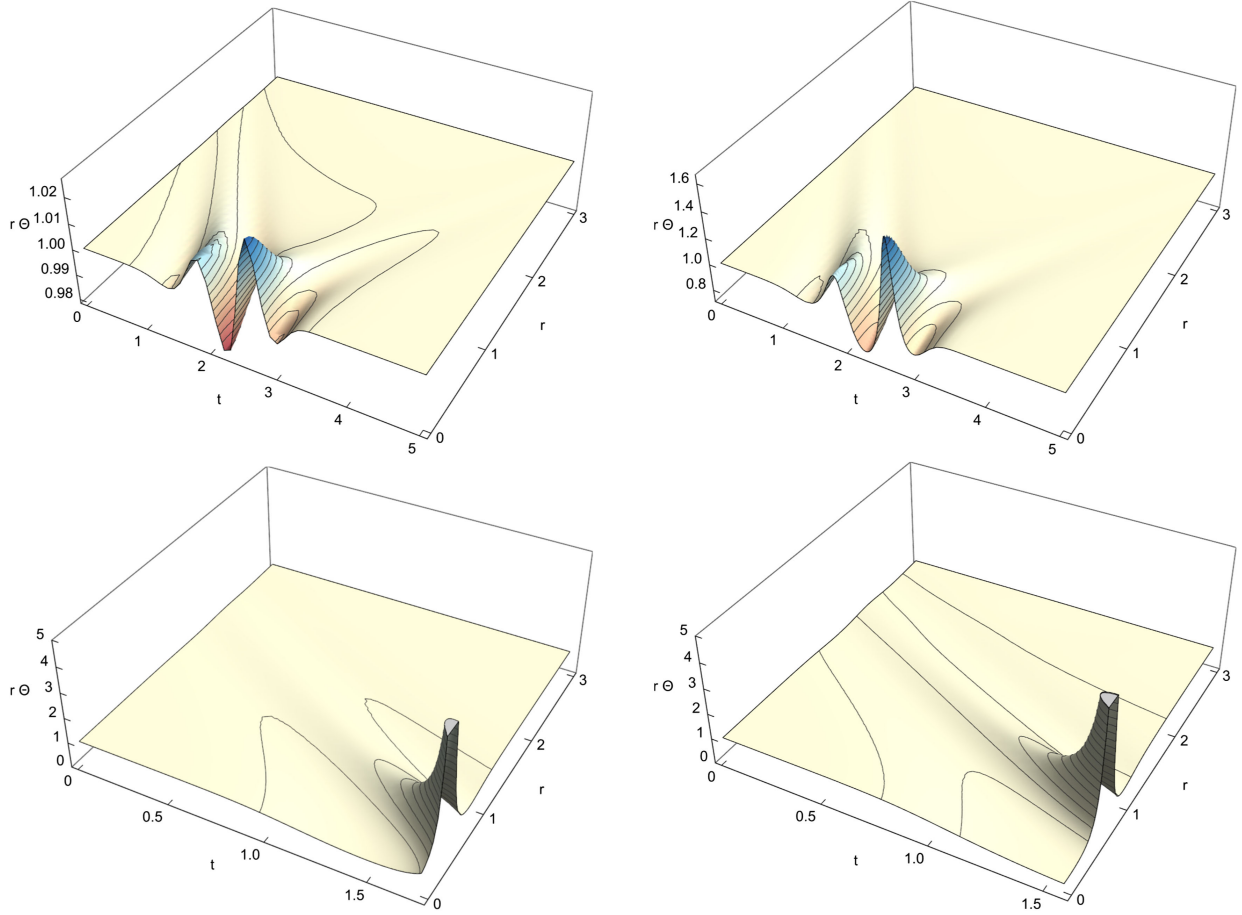


FIG. 12. Plots of the outgoing null expansion  $\Theta$  plotted as  $r\Theta$  to make the behavior more evident. In the top left frame and right frames we plot this for  $A = 0.01$  and  $A = 0.04$ , respectively, and in the bottom left and right frames are plotted the cases  $A = 0.12$  and  $A = 0.2$ . Vanishing  $\Theta$  indicates a trapped surface. Rather than  $\Theta$  becoming small, instead we see  $\Theta$  becoming very large in the region where strong coupling occurs, indicating no horizon is forming. The peak of  $\Theta$ , which is cutoff in the plot, is  $\sim 100$  when evolution stops in the case  $A = 0.12$  and even greater in the case  $A = 0.2$ . This suggests the region of strong coupling will be visible to asymptotic observers.

tending to zero as the amplitude is increased to the limit value of  $A = 0.06$  when the evolution breaks. For  $A = 0.06$  the minimum value is  $|g_{tt}| \simeq 0.3$ , occurring when the evolution breaks at or near the origin, so still somewhat greater than zero. Thus there is no suggestion of a horizon developing.

For larger amplitudes the evolution apparently breaks away from the axis and we should consider whether a trapped surface forms at any radius, not just consider  $g_{tt}$  at the origin. In our spherically symmetric context a trapped surface occurs when the expansion of outward radially directed null rays vanishes. This expansion is computed as

$$\Theta = \frac{1}{2} k^\mu \partial_\mu \log(g_{\theta\theta} g_{\phi\phi}), \quad (110)$$

where  $k^\mu$  is the outward directed null vector with non-vanishing components  $k^t = 1$  and  $k^r$  solving the outward null condition. In Fig. 12 we show this expansion  $\Theta$  as a function of time and radial coordinate for the various

amplitudes displayed as examples here, so  $A = 0.01$ ,  $0.04$ ,  $0.12$  and  $0.2$ . We plot  $r\Theta$  rather than  $\Theta$  as near the origin  $\Theta \sim 1/r$  which obscures the behavior. We see that rather than the expansion becoming smaller and vanishing to give a trapped surface, instead it becomes very large in the region of strong coupling in the cases  $A = 0.12$  and  $A = 0.2$  where the evolution breaks down, presumably reflecting the diverging  $c$  component, and suggesting a blueshifting associated to the strong coupling, rather than redshifting.

An important conceptual point is that for conventional GR and matter the existence of a trapped surface proves the region lies within an event horizon—however for massive gravity it is unclear such a statement would hold. Thus finding a trapped surface or otherwise in this theory cannot prove or disprove the existence of an event horizon—for that presumably one would need to construct the full spacetime and consider the past of  $\mathcal{I}^+$ . However the lack of a trapped surface is suggestive that no event horizon has formed.

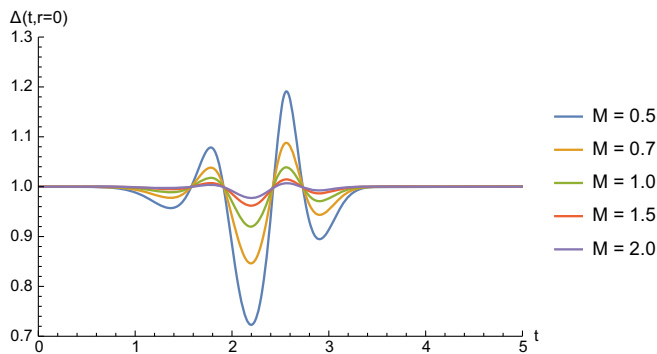


FIG. 13. The effect of varying the graviton mass  $m$  whilst keeping the amplitude  $A = 0.01$  and initial scalar profile constant. We again concentrate on the behavior of  $\Delta$  at the origin, and note that it takes the value  $\Delta = 1$  for flat spacetime for all values of mass. We see that decreasing the mass is analogous to increasing the amplitude, indicating nonlinear effects become increasingly important.

### H. Varying mass

While exploring the full phenomenology of collapse for the minimal theory is beyond the scope of this work, we briefly consider how the behavior changes as we vary the graviton mass from  $m = 1$ , whilst keeping the profile of the initial ingoing scalar pulse the same. In Fig. 13 we show the behavior of  $\Delta$  at the origin as a function of time for the same initial data  $A = 0.01$  and masses ranging from  $m^2 = 4$  to  $m^2 = 0.25$ . For smaller values of  $m^2$  we see the same initial data elicits a stronger response. This presumably reflects the behavior of the linear theory, which breaks down as we take  $m \rightarrow 0$ , again indicating nonlinear effects become more important in the small mass limit for the same matter behavior.

## VII. CONCLUSION

For dRGT massive gravity with its minimal and quadratic mass terms we have provided a  $3 + 1$  dynamical formulation of the theory using a symmetric vierbein. With an appropriate choice of momentum variables the vector and scalar constraints, which are second class and thus must be applied on every time slice, can be solved algebraically for the time-time vierbein component together with certain momenta, yielding a system of evolution equations that can straightforwardly be implemented numerically. With no symmetry assumption this then leaves five second order geometric dynamical variables, which correctly counts the number of massive gravitational degrees of freedom. Our construction cannot be applied if the remaining cubic mass term is included, but we emphasize that phenomenological constraints already rule this cubic term out [43,56,57].

As a low-energy EFT, the truncated theory is not expected to be well-posed on its own. While remaining agnostic as to the details of its high energy embedding, one

can however include spatial diffusion terms to the low-energy EFT which are natural in such a  $3 + 1$  formulation. In the presence of these terms, one can successfully prove that the formulation gives a well-posed initial value problem of diffusive character. Furthermore, the diffusion constant controlling these terms can be taken to be as small as one wishes, and hence can be made sufficiently small to not affect length and timescales of physical interest. We demonstrate this framework by numerically evolving dynamical collapse of massless scalar field matter in spherical symmetry, restricting for simplicity to the theory with minimal mass term. This minimal model is not expected to exhibit a smooth massless limit to GR due to the absence of a Vainshtein mechanism present in more general massive gravity models. Moreover in our numerics we consider sources of comparable size to the graviton Compton wavelength which is far from the regime of phenomenological interest. Nonetheless being an EFT of gravity in its own right it is interesting to ask how the minimal model responds to collapsing matter. We explicitly check our results are insensitive to diffusion terms added to render the continuum PDEs well posed.

We focus on the situation when all scales in the problem—i.e. the mass and the characteristic length scales in the initial data—are approximately equal, so there is no parametric scale separation. Our simulations reveal the expected dispersive behavior for low amplitude initial data. However as the amplitude is increased, and the response becomes nonlinear, we see pathologies occur in the scalar constraint equation, which taking a quadratic form is not guaranteed to have solutions. As a consequence it appears that evolution breaks down before the matter reaches the origin. We have argued that strong coupling will develop in the region before the evolution fails, so that the EFT is breaking down, and evolution past this point would become sensitive to the precise form of the short distance completion. Our results remain ambiguous over whether the breakdown is associated to large curvatures. While we see curvatures are relatively small up until the breakdown, for some cases such as the case  $A = 0.20$  shown above, the curvature does appear to be increasing in the regions where strong coupling appears to develop. Interestingly we see no formation of trapped surfaces, which is suggestive that the region where the strong coupling develops is not shielded from asymptotic observers by a horizon.

In future work we will examine collapse in the theory with quadratic mass term, which is expected to exhibit a Vainshtein mechanism that will yield GR behavior for small graviton masses. Another interesting direction is to refine understanding of the evolution breakdown in the minimal theory. In particular it would be interesting to determine whether it is the divergence of  $\Delta$  (being on the “wrong” branch as the quadratic constraint linearises) or  $\Delta$  being driven negative (associated to the solution for  $c$  becoming complex) that finally breaks the evolution—both

seem to occur in different locations at very similar times. Another question to resolve is whether the curvatures seen by the matter degrees of freedom remain bounded or not. It would also be interesting to find coordinates which excise the region where the theory breaks down, and allow evolution outside the future light cone of the pathology, rather than have to halt evolution on our timeslice as soon as evolution breaks down at one point. Finally, ultimately we would like to know what happens for more phenomenologically relevant physical situations where there is typically a large hierarchy between the size of the source and the Compton wavelength of the graviton. For example, in realistic applications of massive gravity, the graviton Compton wavelength should be taken to be comparable to the Hubble scale, many orders of magnitude larger than the astrophysical scales of collapsing matter distributions.

### ACKNOWLEDGMENTS

We thank Enrico Barausse, Pau Figueras, Luis Lehner, and Mark Trodden for very helpful discussions. We also thank the referee of this paper for useful comments. This work is supported by STFC Consolidated Grant No. ST/T000791/1. C.d.R. is also supported by a Simons Investigator Award No. 690508 and J. K. is funded by an STFC studentship.

### APPENDIX A: SCALAR CONSTRAINT DETAILS

Here we present some further details. First we give an argument for the simplification that leads to the quadratic and cubic forms in Eq. (53). We then give the explicit expressions for the coefficients of the quadratic scalar constraint in the minimal model for our spherically symmetric ansatz.

#### 1. Polynomial expansion in $E_{tt}$ of scalar constraint

We may understand the simplification that we saw in Eq. (53) relative to the naive expectation in (52) by writing the 3 + 1 decomposition of the symmetric vierbein as

$$E_{\mu\nu} = \begin{pmatrix} \phi + n_i n^i & n_i \\ & h_{ij} \end{pmatrix}, \quad (\text{A1})$$

where now indices are raised/lowered with respect to  $h_{ij}$  and its inverse  $h^{ij} = (h_{ij})^{-1}$ . Then the inverse can be written as

$$(E^{-1})^{\mu\nu} = \frac{1}{\phi} n^\mu n^\nu + h^{\mu\nu}, \quad (\text{A2})$$

where  $h^{tt} = h^{ii} = 0$  and  $n^\mu = (1, -n^i)$ . We note that  $|E| = \phi \det h_{ij}$  and the inverse metric is

$$g^{\mu\nu} = \frac{J}{\phi^2} n^\mu n^\nu + \frac{2}{\phi} n^{(\mu} v^{\nu)} + H^{\mu\nu}, \quad v_\mu = \eta_{\mu\nu} n^\nu, \\ J = n^\mu v_\mu, \quad H^{\mu\nu} = h^{\mu\alpha} \eta_{\alpha\beta} h^{\beta\nu}. \quad (\text{A3})$$

Consider  $A_{(1)}^{\alpha\beta\gamma\mu\nu\rho}$  as given in Eq. (43). It comprises a term  $\sim \eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta}$ , which one would naively expect to go as  $\eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta} \sim \frac{1}{\phi^4}$  in an expansion in inverse powers of  $\phi$ . However due to the antisymmetry in the  $[\alpha\beta]$  and  $[\mu\nu]$  index pairs, and the structure of the leading terms in  $g^{\mu\nu}$  above, we see such a term vanishes. Likewise the subleading  $1/\phi^3$  behavior also vanishes, again due to this index antisymmetry. Thus in fact  $\eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta} \sim \frac{1}{\phi^2}$ . Noting that  $|E| \sim \phi$ , then  $|E|^2 \eta^{\gamma\rho} g^{\alpha[\mu} g^{\nu]\beta}$  is a quadratic polynomial in  $\phi$ , and hence is also a quadratic polynomial in  $E_{tt}$  (since  $E_{tt}$  is linear in  $\phi$ ). A similar argument applies to the other two terms in  $A_{(1)}^{\alpha\beta\gamma\mu\nu\rho}$ , showing they also give quadratic contributions in  $\phi$  when multiplied by  $|E|^2$  due to the index antisymmetries. Likewise analogous arguments show that  $|E|^3 A_{(2)}^{\alpha\beta\gamma\mu\nu\rho}$  is a cubic in  $E_{tt}$ .

#### 2. Scalar constraint coefficients

The coefficients  $a_2$ ,  $a_1$  and  $a_0$  of the scalar constraint in the spherically symmetric minimal model of Sec. VI are given by

$$a_2 = -48a^4 m^2 r^8 + 6a^3(8b - 3)m^2 r^6 + a^2 r^2(36b^2 m^2 r^2 - 9bm^2 r^2 + r^2 \partial_r \Phi^2 - 50) \\ - 2ar(r(12b^3 m^2 - 9b^2 m^2 + b \partial_r \Phi^2 + 10r \partial_r a) - 10 \partial_r b) - 12b^4 m^2 + 9b^3 m^2 + b^2 \partial_r \Phi^2 - 2(\partial_r b - r^2 \partial_r a)^2, \quad (\text{A4})$$

$$a_1 = -(b - ar^2)(3m^2(-4a^3 r^6 + b^2 r^2(3a - 8h^2) + bh^2 r^2(5 - 8ar^2) + ah^2 r^4(16ar^2 + 7) + b^3) + 2hr \partial_r \Phi p_\Phi (b - ar^2) \\ + 4rp_h(r(5a + r \partial_r a) - \partial_r b)) + 4hr p_b(\partial_r b - r(5a + r \partial_r a)) + 4hr^3 p_a(r(5a + r \partial_r a) - \partial_r b), \quad (\text{A5})$$



$$\begin{aligned}
a_0 = & -4a^4 r^8 p_\Phi^2 + 2a^3 r^6 (2b p_\Phi^2 + 3h^2 m^2 r^2 + 2hr \partial_r \Phi p_\Phi) + a^2 r^4 (3(-8r^2 p_a p_b + 8r^4 p_a^2 + b^2 p_b^2) \\
& - 2hr(40rp_a + 3b \partial_r \Phi p_\Phi) + h^2(-9bm^2 r^2 + r^2 p_\Phi^2 + r^2(-\partial_r \Phi^2) + 50) + 20hp_b - 12h^4 m^2 r^4) \\
& + b^2(-6r^4 p_a^2 + 6p_b^2 + h^2 r^2(-12h^2 m^2 r^2 + p_\Phi^2 - \partial_r \Phi^2)) + 2ar^2(-b(6r^2 p_a p_b + b^2 p_\Phi^2 - 6p_b^2) \\
& + 2h(r^2 p_a(5b - r^2 p_h - 4r^3 \partial_r a + 4r \partial_r b) + p_b(10b + r^2 p_h + r^3 \partial_r a - r \partial_r b)) + 3(4b - 1)h^4 m^2 r^4 \\
& + h^2 r(-br p_\Phi^2 + r(b \partial_r \Phi^2 + 10r \partial_r a) - 10 \partial_r b)) + 2bhr(2r^2 p_a(rp_h + r^2 \partial_r a - \partial_r b) \\
& - 2p_b(rp_h - 2r^2 \partial_r a + 2 \partial_r b) + 3h^3 m^2 r^3) + 2h^2 r^2(2r^2 p_a p_b + r^4(-p_a^2) - p_b^2 + (\partial_r b - r^2 \partial_r a)^2) \\
& - b^4 p_\Phi^2 + b^3 hr(3hm^2 r + 2 \partial_r \Phi p_\Phi).
\end{aligned} \tag{A6}$$

## APPENDIX B: NUMERICAL DETAILS AND CONVERGENCE

As discussed in the main text, the dynamical evolution is performed with a compactified radial coordinate,  $\tilde{r} = r/(1 - r^2)$  on the interval  $[0, 1]$ . Spatial differences are approximated using sixth order finite differencing, and time derivatives using a Crank-Nicholson scheme so that the system is implicit. The exception to this is that the diffusion terms are differenced using the same spatial scheme, but forward Euler differencing in time. We are not concerned in accurately simulating the diffusion terms which should be irrelevant on the scales we are interested in, and taking second order Crank-Nicholson differencing in time would necessitate an analog of the Courant condition for the time step associated to these terms.

### 1. Iterative solution of the implicit time step

We solve the implicit system as follows: suppose we have the data  $\vec{X} = (\vec{E}_j^i, E_i, \dots)$  on some time slice  $t$ . The equations of motion then allow us to determine the time derivative  $\vec{X}_t$  at  $t$ . With that we make an initial guess of the data at  $t + \Delta t$  as  $\vec{Y} = \vec{X} + \Delta t \vec{X}_t$ . The method then enters an iterative stage with the following steps:

- (i) Estimate the data at  $t + \Delta t/2$  by  $\vec{Z} = (\vec{X} + \vec{Y})/2$ .
- (ii) Compute the equations of motions to find  $\vec{Z}_t$ .
- (iii) Update  $\vec{Y}_{\text{new}} = \vec{Z} + \Delta t \vec{Z}_t/2$ .

This continues until the maximal difference between  $\vec{Y}$  and  $\vec{Y}_{\text{new}}$  becomes smaller than a set error tolerance, which for the results in this paper was taken to be  $10^{-8}$ .

### 2. Convergence

In this section, we examine the convergence of our numerical results as we vary the spatial resolution  $N$  and the diffusion coefficient  $D$ . While there is also numerical error associated to our discretization in time, in practice the time step size we use is sufficiently small that discretization error is dominated by the spatial resolution. To compare the results across the entirety of the spatial slices, we use the  $L^2$  norm in spherical symmetry, given by

$$\|f\|_{L^2} := \left( \int_0^\infty r^2 |f(r)|^2 dr \right)^{1/2}. \tag{B1}$$

As a first check, Fig. 14 examines the behavior of the time-time component of the vierbein  $E_{\mu\nu}$ , so the function  $c$ , as  $N$  is varied between 100 and 1600 by doubling it, keeping the

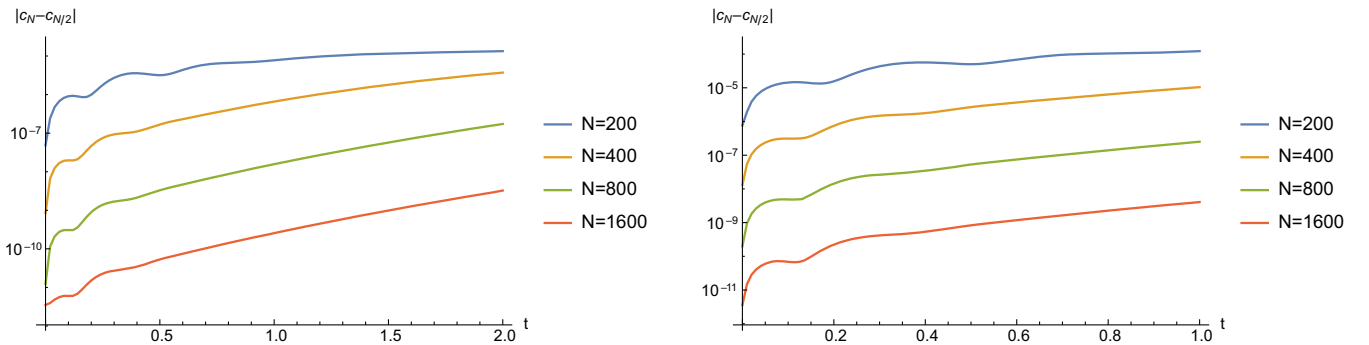


FIG. 14. Figure showing the effect of varying the spatial resolution  $N$  while keeping all other parameters constant, in particular  $D = 0.001$ . The plots display the norm of the difference between the values for the  $c$  vierbein component obtained for a resolution  $N$  and half that resolution, so  $|c_N - c_{N/2}|$ , as a function of time. The left-hand side has  $A = 0.01$ , and the right  $A = 0.04$ . In both cases these differences decrease with resolution, with the decrease being consistent with our sixth order finite differencing—the difference decreases by a factor of  $\sim 2^6$  for a doubling of the resolution  $N$ .

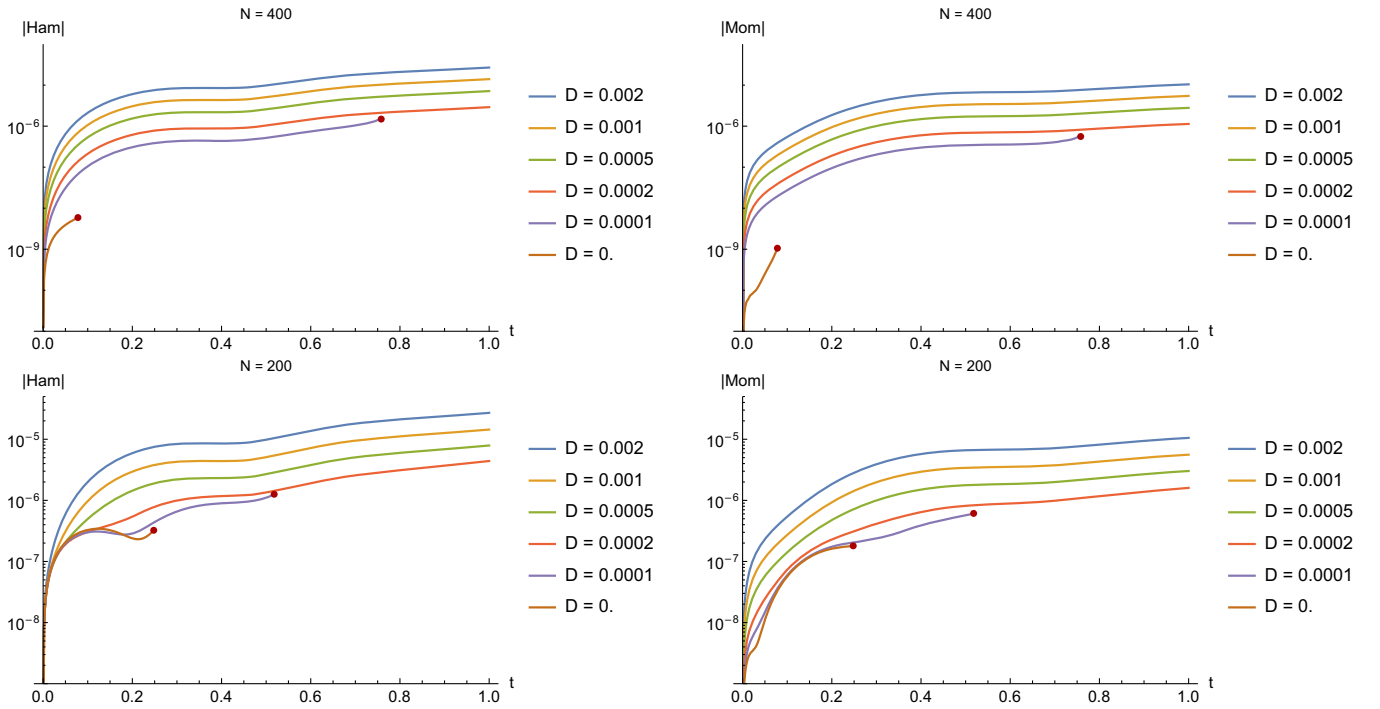


FIG. 15. The behavior of the  $L^2$  norm of the Hamiltonian (left) and momentum (right) constraints against time for  $A = 0.01$  as  $D$  is varied between 0 and 0.002. The top figures have a resolution  $N = 400$ , whilst for the lower ones  $N = 200$ . We see the expected convergence of these constraints to zero as  $D$  is initially decreased, but the simulations become increasingly susceptible to lattice scale instabilities in that limit, and so for  $D$  that is too small the simulations break down, as indicated by the red markers. For  $N = 200$ , eventually there is little benefit in further decreasing  $D$ , signaling that the numerical errors associated with the discretization dominate over the effects of diffusion.

diffusion constant fixed at  $D = 0.001$ . We use the notation that  $c_N$  is the solution for  $c$  at resolution  $N$ , and then plot the norm of the difference of  $c$  between a resolution  $N$  and  $N/2$  as a function of coordinate time, so  $|c_N - c_{N/2}|$  against time  $t$ . In the figure we show this for two initial amplitudes,  $A = 0.01$  and  $A = 0.04$ . We see that this difference decreases as the resolution increases, and is consistent with the sixth order spatial finite differencing we employ—the differences decrease by a factor  $\sim 2^6$  upon doubling the resolution. Thus the solutions are consistent with convergence to a continuum solution. However, we emphasize that this continuum solution is that of massive gravity together with our diffusive term higher order terms, and not of the “bare” system of equations (17).

We now turn to the (non)conservation of the Hamiltonian and momentum constraints during evolution. The reason for this constraint violation is twofold: it includes the errors introduced by the discretization of the PDEs, as well as the effects of the diffusion term not present in the Einstein equations (17). Above in Sec. VIF we have argued that the precise value of the diffusion coefficient  $D$  does not affect the long-range dynamics of our system. In Fig. 15 we further examine the effect of varying  $D$ , for a small amplitude  $A = 0.01$  and resolutions  $N = 200$  and 400. In both cases we see that decreasing the diffusion coefficient

initially leads to an improvement in the constraint violation. However, for the smallest values considered ( $D = 0.0001$  and 0) the simulation breaks before  $t = 1$  due to lattice scale instabilities—some amount of diffusion is indispensable if we want to evolve our system for an appreciable amount of time. Also for  $N = 200$  we see that eventually the curves for decreasing values of  $D$  coincide (before the onset of instability), indicating that the discretization error dominates over the effects of diffusion in causing the constraint violation.

Finally in Fig. 16 we study the effects of varying  $N$  between 50 and 400 for the unstable case  $D = 0$  and the case used throughout this work,  $D = 0.001$ . Again we have set  $A = 0.01$  and evolve until  $t = 1$ .

In the  $D = 0$  case, we find that increasing the resolution always leads to an improvement in the values of the constraints, but also pushes forward the onset of lattice scale instabilities which break the simulation, so that for  $N = 400$  we can only evolve up to  $t \simeq 0.1$ . In contrast, if we set  $D = 0.001$  then all the resolutions are stable in the plotted time period. However, we see that eventually improving the resolution ceases to improve the results, and all curves obtained for  $N \geq 200$  are essentially identical. This is due to the effects of diffusion becoming more important than the discretization error. In other words,

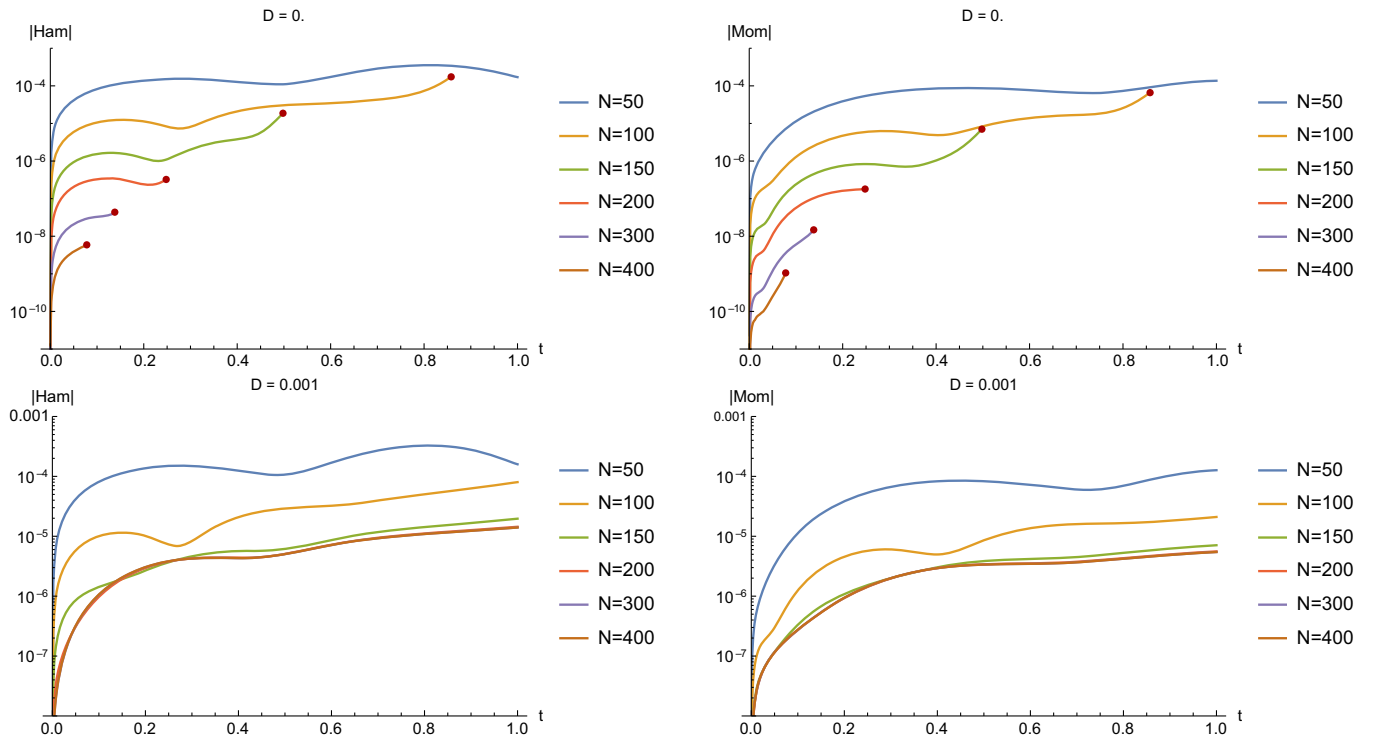


FIG. 16. Plots of the norms of the Hamiltonian and momentum constraints in the unstable case  $D = 0$  (top) and  $D = 0.001$  (bottom), for  $A = 0.01$  and varying spatial resolution  $N$ . With no diffusion, increasing it always improves the conservation of the constraints, at least before the onset of lattice scale instabilities that break the simulations (the times where this occurs are indicated by red markers). Once the diffusive effects are included, for a sufficiently large resolution they dominate over the errors introduced by discretization. In particular for  $D = 0.001$  this happens below  $N = 400$ , which was the case for the majority of the plots in this paper.

we are converging to a solution of the *modified* system of equations, which includes diffusive effects and hence no longer exactly respects the Hamiltonian and momentum constraints. Since there are no appreciable changes to our

results above  $N = 200$ , the choice  $N = 400$ , which was made for most of the plots in this paper, is justified, as further refining the resolution would not have revealed any new physical effects.

- 
- [1] G. Dvali, G. Gabadadze, and M. Porrati, 4-D gravity on a brane in 5-D Minkowski space, *Phys. Lett. B* **485**, 208 (2000).
  - [2] G. Dvali, G. Gabadadze, and M. Porrati, Metastable gravitons and infinite volume extra dimensions, *Phys. Lett. B* **484**, 112 (2000).
  - [3] G. Dvali and G. Gabadadze, Gravity on a brane in infinite volume extra space, *Phys. Rev. D* **63**, 065007 (2001).
  - [4] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and G. Gabadadze, Nonlocal modification of gravity and the cosmological constant problem, [arXiv:hep-th/0209227](https://arxiv.org/abs/hep-th/0209227).
  - [5] G. Gabadadze, ICTP lectures on large extra dimensions, *ICTP Lect. Notes Ser.* **14**, 77 (2003).
  - [6] G. Gabadadze, Cargese lectures on brane induced gravity, *Nucl. Phys. B, Proc. Suppl.* **171**, 88 (2007).
  - [7] G. Gabadadze and M. Shifman, Softly massive gravity, *Phys. Rev. D* **69**, 124032 (2004).
  - [8] C. de Rham, G. Dvali, S. Hofmann, J. Khoury, O. Pujolàs, M. Redi, and A. J. Tolley, Cascading gravity: Extending the Dvali-Gabadadze-Porrati model to higher dimension, *Phys. Rev. Lett.* **100**, 251603 (2008).
  - [9] C. de Rham, S. Hofmann, J. Khoury, and A. J. Tolley, Cascading gravity and degravitation, *J. Cosmol. Astropart. Phys.* **02** (2008) 011.
  - [10] T. Kobayashi, T. Shiromizu, and C. de Rham, Curvature corrections to the low energy effective theory in 6D regularized braneworlds, *Phys. Rev. D* **77**, 124012 (2008).
  - [11] C. de Rham, J. Khoury, and A. J. Tolley, Flat 3-brane with tension in cascading gravity, *Phys. Rev. Lett.* **103**, 161601 (2009).
  - [12] C. de Rham, J. Khoury, and A. J. Tolley, Cascading gravity is ghost free, *Phys. Rev. D* **81**, 124027 (2010).
  - [13] G. Gabadadze, General relativity with an auxiliary dimension, *Phys. Lett. B* **681**, 89 (2009).

- [14] C. de Rham, Massive gravity from Dirichlet boundary conditions, *Phys. Lett. B* **688**, 137 (2010).
- [15] C. de Rham and G. Gabadadze, Self-tuned massive spin-2, *Phys. Lett. B* **693**, 334 (2010).
- [16] L. Berezhiani and M. Mirbabayi, Generalized framework for auxiliary extra dimensions, *Phys. Lett. B* **701**, 654 (2011).
- [17] D. Boulware and S. Deser, Can gravitation have a finite range?, *Phys. Rev. D* **6**, 3368 (1972).
- [18] C. Aragone and S. Deser, Consistency problems of spin-2 gravity coupling, *Nuovo Cimento Soc. Ital. Fis.* **57B**, 33 (1980).
- [19] P. Creminelli, A. Nicolis, M. Papucci, and E. Trincherini, Ghosts in massive gravity, *J. High Energy Phys.* **09** (2005) 003.
- [20] R. Delbourgo and A. Salam, The Stuckelberg formalism for spin two, *Lett. Nuovo Cimento* **12**, 297 (1975).
- [21] C. Deffayet and J.-W. Rombouts, Ghosts, strong coupling and accidental symmetries in massive gravity, *Phys. Rev. D* **72**, 044003 (2005).
- [22] C. de Rham and G. Gabadadze, Generalization of the Fierz-Pauli action, *Phys. Rev. D* **82**, 044020 (2010).
- [23] C. de Rham, G. Gabadadze, and A. J. Tolley, Resummation of massive gravity, *Phys. Rev. Lett.* **106**, 231101 (2011).
- [24] C. de Rham, G. Gabadadze, and A. J. Tolley, Helicity decomposition of ghost-free massive gravity, *J. High Energy Phys.* **11** (2011) 093.
- [25] C. de Rham, G. Gabadadze, and A. J. Tolley, Ghost free massive gravity in the Stückelberg language, *Phys. Lett. B* **711**, 190 (2012).
- [26] S. Hassan and R. A. Rosen, Resolving the ghost problem in non-linear massive gravity, *Phys. Rev. Lett.* **108**, 041101 (2012).
- [27] S. Hassan, A. Schmidt-May, and M. von Strauss, Proof of consistency of nonlinear massive gravity in the Stückelberg formulation, *Phys. Lett. B* **715**, 335 (2012).
- [28] K. Hinterbichler and R. A. Rosen, Interacting spin-2 fields, *J. High Energy Phys.* **07** (2012) 047.
- [29] J. Kluson, Hamiltonian analysis of 1 + 1 dimensional massive gravity, *Phys. Rev. D* **85**, 044010 (2012).
- [30] J. Kluson, Note about Hamiltonian structure of non-linear massive gravity, *J. High Energy Phys.* **01** (2012) 013.
- [31] J. Kluson, Comments about Hamiltonian formulation of non-linear massive gravity with Stückelberg fields, *J. High Energy Phys.* **06** (2012) 170.
- [32] A. Golovnev, On the Hamiltonian analysis of non-linear massive gravity, *Phys. Lett. B* **707**, 404 (2012).
- [33] D. Comelli, M. Crisostomi, F. Nesti, and L. Pilo, Degrees of freedom in massive gravity, *Phys. Rev. D* **86**, 101502 (2012).
- [34] J. Kluson, Remark about Hamiltonian formulation of non-linear massive gravity in Stückelberg formalism, *Phys. Rev. D* **86**, 124005 (2012).
- [35] C. Deffayet, J. Mourad, and G. Zahariade, Covariant constraints in ghost free massive gravity, *J. Cosmol. Astropart. Phys.* **01** (2013) 032.
- [36] C. Deffayet, J. Mourad, and G. Zahariade, A note on ‘symmetric’ vielbeins in bimetric, massive, perturbative and non perturbative gravities, *J. High Energy Phys.* **03** (2013) 086.
- [37] J. Kluson, Note about Hamiltonian formalism for general non-linear massive gravity action in Stückelberg formalism, *Int. J. Mod. Phys. A* **28**, 1350160 (2013).
- [38] D. Comelli, F. Nesti, and L. Pilo, Massive gravity: A general analysis, *J. High Energy Phys.* **07** (2013) 161.
- [39] C. Deffayet, J. Mourad, and G. Zahariade, A new way to count degrees of freedom in dRGT massive gravity, in *Proceedings of the 13th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories* (2015), pp. 1249–1251.
- [40] S. Hassan and R. A. Rosen, Confirmation of the secondary constraint and absence of ghost in massive gravity and bimetric gravity, *J. High Energy Phys.* **04** (2012) 123.
- [41] J. Kluson, Non-linear massive gravity with additional primary constraint and absence of ghosts, *Phys. Rev. D* **86**, 044024 (2012).
- [42] C. de Rham and A. J. Tolley, Vielbein to the rescue? Breaking the symmetric vielbein condition in massive gravity and multigravity, *Phys. Rev. D* **92**, 024024 (2015).
- [43] C. de Rham, Massive gravity, *Living Rev. Relativity* **17**, 7 (2014).
- [44] M. Mirbabayi, A proof of ghost freedom in de Rham-Gabadadze-Tolley massive gravity, *Phys. Rev. D* **86**, 084006 (2012).
- [45] L. Bernard, C. Deffayet, and M. von Strauss, Consistent massive graviton on arbitrary backgrounds, *Phys. Rev. D* **91**, 104013 (2015).
- [46] L. Bernard, C. Deffayet, A. Schmidt-May, and M. von Strauss, Linear spin-2 fields in most general backgrounds, *Phys. Rev. D* **93**, 084020 (2016).
- [47] L. Bernard, C. Deffayet, and M. von Strauss, Massive graviton on arbitrary background: Derivation, syzygies, applications, *J. Cosmol. Astropart. Phys.* **06** (2015) 038.
- [48] C. Mazuet and M. S. Volkov, Massive gravitons in arbitrary spacetimes, *Phys. Rev. D* **96**, 124023 (2017).
- [49] C. Mazuet and M. S. Volkov, Massive spin-2 field in arbitrary spacetimes—The detailed derivation, *J. Cosmol. Astropart. Phys.* **07** (2018) 012.
- [50] D. Comelli, M. Crisostomi, F. Nesti, and L. Pilo, Spherically symmetric solutions in ghost-free massive gravity, *Phys. Rev. D* **85**, 024044 (2012).
- [51] M. S. Volkov, Self-accelerating cosmologies and hairy black holes in ghost-free bigravity and massive gravity, *Classical Quantum Gravity* **30**, 184009 (2013).
- [52] M. S. Volkov, Hairy black holes in the ghost-free bigravity theory, *Phys. Rev. D* **85**, 124043 (2012).
- [53] M. S. Volkov, Hairy black holes in theories with massive gravitons, *Lect. Notes Phys.* **892**, 161 (2015).
- [54] M. S. Volkov, Energy in ghost-free massive gravity theory, *Phys. Rev. D* **90**, 124090 (2014).
- [55] M. S. Volkov, Energy in ghost-free massive gravity theory, in *Proceedings of the 14th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories* (2017), Vol. 2, pp. 1505–1508.
- [56] G. Chkareuli and D. Pirtskhalava, Vainshtein mechanism in  $\Lambda_3$ —theories, *Phys. Lett. B* **713**, 99 (2012).



- [57] L. Berezhiani, G. Chkareuli, C. de Rham, G. Gabadadze, and A. Tolley, Mixed Galileons and spherically symmetric solutions, *Classical Quantum Gravity* **30**, 184003 (2013).
- [58] A. I. Vainshtein, To the problem of nonvanishing gravitation mass, *Phys. Lett.* **39B**, 393 (1972).
- [59] S. Renaux-Petel, On the Vainshtein mechanism in the minimal model of massive gravity, *J. Cosmol. Astropart. Phys.* **03** (2014) 043.
- [60] C. de Rham, J. T. Deskins, A. J. Tolley, and S.-Y. Zhou, Graviton mass bounds, *Rev. Mod. Phys.* **89**, 025004 (2017).
- [61] C. Deffayet, G. Dvali, G. Gabadadze, and A. I. Vainshtein, Nonperturbative continuity in graviton mass versus perturbative discontinuity, *Phys. Rev. D* **65**, 044026 (2002).
- [62] C. Deffayet, Spherically symmetric solutions of massive gravity, *Classical Quantum Gravity* **25**, 154007 (2008).
- [63] E. Babichev, C. Deffayet, and R. Ziour, The Vainshtein mechanism in the decoupling limit of massive gravity, *J. High Energy Phys.* **05** (2009) 098.
- [64] E. Babichev, C. Deffayet, and R. Ziour, The recovery of general relativity in massive gravity via the Vainshtein mechanism, *Phys. Rev. D* **82**, 104008 (2010).
- [65] C. de Rham, G. Gabadadze, L. Heisenberg, and D. Pirtskhalava, Cosmic acceleration and the helicity-0 graviton, *Phys. Rev. D* **83**, 103516 (2011).
- [66] C. de Rham and L. Heisenberg, Cosmology of the Galileon from massive gravity, *Phys. Rev. D* **84**, 043503 (2011).
- [67] K. Koyama, G. Niz, and G. Tasinato, Strong interactions and exact solutions in non-linear massive gravity, *Phys. Rev. D* **84**, 064033 (2011).
- [68] N. Kaloper, A. Padilla, and N. Tanahashi, Galileon hairs of dyson spheres, Vainshtein's coiffure and hirsute bubbles, *J. High Energy Phys.* **10** (2011) 148.
- [69] A. V. Belikov and W. Hu, Equivalence principle violation in weakly Vainshtein-screened systems, *Phys. Rev. D* **87**, 084042 (2013).
- [70] T. Hiramatsu, W. Hu, K. Koyama, and F. Schmidt, Equivalence principle violation in Vainshtein screened two-body systems, *Phys. Rev. D* **87**, 063525 (2013).
- [71] F. Sbisà, G. Niz, K. Koyama, and G. Tasinato, Characterising Vainshtein solutions in massive gravity, *Phys. Rev. D* **86**, 024033 (2012).
- [72] R. Kimura, T. Kobayashi, and K. Yamamoto, Vainshtein screening in a cosmological background in the most general second-order scalar-tensor theory, *Phys. Rev. D* **85**, 024023 (2012).
- [73] R. Gannouji and M. Sami, Vainshtein mechanism in Gauss-Bonnet gravity and Galileon ether, *Phys. Rev. D* **85**, 024019 (2012).
- [74] E. Babichev, C. Deffayet, and G. Esposito-Farese, Constraints on shift-symmetric scalar-tensor theories with a Vainshtein mechanism from bounds on the time variation of  $G$ , *Phys. Rev. Lett.* **107**, 251102 (2011).
- [75] C. de Rham, A. J. Tolley, and D. H. Wesley, Vainshtein mechanism in binary pulsars, *Phys. Rev. D* **87**, 044025 (2013).
- [76] C. de Rham, A. Matas, and A. J. Tolley, Galileon radiation from binary systems, *Phys. Rev. D* **87**, 064024 (2013).
- [77] A. Padilla and P. M. Saffin, Classical duals, legendre transforms and the Vainshtein mechanism, *J. High Energy Phys.* **07** (2012) 122.
- [78] A. De Felice, R. Kase, and S. Tsujikawa, Vainshtein mechanism in second-order scalar-tensor theories, *Phys. Rev. D* **85**, 044059 (2012).
- [79] Y.-Z. Chu and M. Trodden, Retarded Green's function of a Vainshtein system and Galileon waves, *Phys. Rev. D* **87**, 024011 (2013).
- [80] M. Andrews, Y.-Z. Chu, and M. Trodden, Galileon forces in the solar system, *Phys. Rev. D* **88**, 084028 (2013).
- [81] L. Berezhiani, G. Chkareuli, and G. Gabadadze, Restricted Galileons, *Phys. Rev. D* **88**, 124020 (2013).
- [82] K. Koyama, G. Niz, and G. Tasinato, Effective theory for the Vainshtein mechanism from the Horndeski action, *Phys. Rev. D* **88**, 021502 (2013).
- [83] E. Babichev and C. Deffayet, An introduction to the Vainshtein mechanism, *Classical Quantum Gravity* **30**, 184001 (2013).
- [84] B. Li, G.-B. Zhao, and K. Koyama, Exploring Vainshtein mechanism on adaptively refined meshes, *J. Cosmol. Astropart. Phys.* **05** (2013) 023.
- [85] F. Dar, C. De Rham, J. T. Deskins, J. T. Giblin, and A. J. Tolley, Scalar gravitational radiation from binaries: Vainshtein mechanism in time-dependent systems, *Classical Quantum Gravity* **36**, 025008 (2019).
- [86] P. Brax, L. Heisenberg, and A. Kuntz, Unveiling the Galileon in a three-body system: Scalar and gravitational wave production, *J. Cosmol. Astropart. Phys.* **05** (2020) 012.
- [87] G. Lara, M. Bezares, M. Crisostomi, and E. Barausse, Robustness of kinetic screening against matter coupling, *Phys. Rev. D* **107**, 044019 (2023).
- [88] M. Bezares, R. Aguilera-Miret, L. ter Haar, M. Crisostomi, C. Palenzuela, and E. Barausse, No evidence of kinetic screening in simulations of merging binary neutron stars beyond general relativity, *Phys. Rev. Lett.* **128**, 091103 (2022).
- [89] M. Shibata and D. Traykova, Properties of scalar wave emission in a scalar-tensor theory with kinetic screening, *Phys. Rev. D* **107**, 044068 (2023).
- [90] C. J. E. Ter Haar, Phenomenological aspects of gravity beyond general relativity, Ph.D. thesis, SISSA, Trieste, 2022.
- [91] A. Dima, Testing the gravitational phenomenology of compact objects: Superradiance, scalarization and screening mechanisms, Ph.D. thesis, SISSA, Trieste, 2021.
- [92] G. D'Amico, C. de Rham, S. Dubovsky, G. Gabadadze, D. Pirtskhalava, and A. J. Tolley, Massive cosmologies, *Phys. Rev. D* **84**, 124046 (2011).
- [93] C. Deffayet and T. Jacobson, On horizon structure of bimetric spacetimes, *Classical Quantum Gravity* **29**, 065009 (2012).
- [94] L. Berezhiani, G. Chkareuli, C. de Rham, G. Gabadadze, and A. Tolley, On black holes in massive gravity, *Phys. Rev. D* **85**, 044024 (2012).
- [95] A. Gruzinov and M. Mirbabayi, Stars and black holes in massive gravity, *Phys. Rev. D* **84**, 124019 (2011).
- [96] K. Koyama, G. Niz, and G. Tasinato, Analytic solutions in non-linear massive gravity, *Phys. Rev. Lett.* **107**, 131101 (2011).
- [97] Y.-F. Cai, D. A. Easson, C. Gao, and E. N. Saridakis, Charged black holes in nonlinear massive gravity, *Phys. Rev. D* **87**, 064001 (2013).

- [98] G. Tasinato, K. Koyama, and G. Niz, Exact solutions in massive gravity, *Classical Quantum Gravity* **30**, 184002 (2013).
- [99] R. Brito, V. Cardoso, and P. Pani, Massive spin-2 fields on black hole spacetimes: Instability of the Schwarzschild and Kerr solutions and bounds on the graviton mass, *Phys. Rev. D* **88**, 023514 (2013).
- [100] I. Arraut, On the black holes in alternative theories of gravity: The case of nonlinear massive gravity, *Int. J. Mod. Phys. D* **24**, 1550022 (2015).
- [101] H. Kodama and I. Arraut, Stability of the Schwarzschild–de Sitter black hole in the dRGT massive gravity theory, *Prog. Theor. Exp. Phys.* **2014**, 023E02 (2014).
- [102] R. Gervalle and M. S. Volkov, Asymptotically flat hairy black holes in massive bigravity, *Phys. Rev. D* **102**, 124040 (2020).
- [103] R. Berens, L. Krauth, and R. A. Rosen, Gravitational collapse in massive gravity in de Sitter spacetime, *Phys. Rev. D* **105**, 064057 (2022).
- [104] R. A. Rosen, Non-singular black holes in massive gravity: Time-dependent solutions, *J. High Energy Phys.* **10** (2017) 206.
- [105] R. A. Rosen, Black hole mechanics for massive gravitons, *Phys. Rev. D* **98**, 104008 (2018).
- [106] S. Weinberg, On the development of effective field theory, *Eur. Phys. J. H* **46**, 6 (2021).
- [107] C. de Rham, S. Melville, and A. J. Tolley, Improved positivity bounds and massive gravity, *J. High Energy Phys.* **04** (2018) 083.
- [108] C. de Rham, S. Melville, A. J. Tolley, and S.-Y. Zhou, Positivity bounds for massive spin-1 and spin-2 fields, *J. High Energy Phys.* **03** (2019) 182.
- [109] C. de Rham, L. Heisenberg, and A. J. Tolley, Spin-2 fields and the weak gravity conjecture, *Phys. Rev. D* **100**, 104033 (2019).
- [110] C. de Rham, A. Matas, and A. J. Tolley, Deconstructing dimensions and massive gravity, *Classical Quantum Gravity* **31**, 025004 (2014).
- [111] U. Aydemir, M. M. Anber, and J. F. Donoghue, Self-healing of unitarity in effective field theories and the onset of new physics, *Phys. Rev. D* **86**, 014025 (2012).
- [112] C. de Rham and R. H. Ribeiro, Riding on irrelevant operators, *J. Cosmol. Astropart. Phys.* **11** (2014) 016.
- [113] G. Dvali, G. F. Giudice, C. Gomez, and A. Kehagias, UV-completion by classicalization, *J. High Energy Phys.* **08** (2011) 108.
- [114] C. de Rham, M. Fasiello, and A. J. Tolley, Galileon duality, *Phys. Lett. B* **733**, 46 (2014).
- [115] C. de Rham, L. Keltner, and A. J. Tolley, Generalized Galileon duality, *Phys. Rev. D* **90**, 024050 (2014).
- [116] K. Kampf and J. Novotny, Unification of Galileon dualities, *J. High Energy Phys.* **10** (2014) 006.
- [117] L. Keltner and A. J. Tolley, UV properties of Galileons: Spectral densities, [arXiv:1502.05706](https://arxiv.org/abs/1502.05706).
- [118] C. de Rham, G. Gabadadze, L. Heisenberg, and D. Pirtskhalava, Non-renormalization and naturalness in a class of scalar-tensor theories, *Phys. Rev. D* **87**, 085017 (2013).
- [119] C. de Rham, L. Heisenberg, and R. H. Ribeiro, Quantum corrections in massive gravity, *Phys. Rev. D* **88**, 084058 (2013).
- [120] J. Cayuso, N. Ortiz, and L. Lehner, Fixing extensions to general relativity in the nonlinear regime, *Phys. Rev. D* **96**, 084043 (2017).
- [121] G. Allwright and L. Lehner, Towards the nonlinear regime in extensions to GR: Assessing possible options, *Classical Quantum Gravity* **36**, 084001 (2019).
- [122] L. Bernard, L. Lehner, and R. Luna, Challenges to global solutions in Horndeski’s theory, *Phys. Rev. D* **100**, 024011 (2019).
- [123] R. Cayuso and L. Lehner, Nonlinear, noniterative treatment of EFT-motivated gravity, *Phys. Rev. D* **102**, 084008 (2020).
- [124] G. Lara, M. Bezares, and E. Barausse, UV completions, fixing the equations, and nonlinearities in k-essence, *Phys. Rev. D* **105**, 064058 (2022).
- [125] P. Figueras and T. França, Black hole binaries in cubic Horndeski theories, *Phys. Rev. D* **105**, 124004 (2022).
- [126] M. Gerhardinger, J. T. Giblin, Jr., A. J. Tolley, and M. Trodden, Well-posed UV completion for simulating scalar Galileons, *Phys. Rev. D* **106**, 043522 (2022).
- [127] N. Franchini, M. Bezares, E. Barausse, and L. Lehner, Fixing the dynamical evolution in scalar-Gauss-Bonnet gravity, *Phys. Rev. D* **106**, 064061 (2022).
- [128] E. Barausse, M. Bezares, M. Crisostomi, and G. Lara, The well-posedness of the Cauchy problem for self-interacting vector fields, *J. Cosmol. Astropart. Phys.* **11** (2022) 050.
- [129] S. Hassan, R. A. Rosen, and A. Schmidt-May, Ghost-free massive gravity with a general reference metric, *J. High Energy Phys.* **02** (2012) 026.
- [130] H. van Dam and M. J. G. Veltman, Massive and massless Yang-Mills and gravitational fields, *Nucl. Phys.* **B22**, 397 (1970).
- [131] V. Zakharov, Linearized gravitation theory and the graviton mass, *JETP Lett.* **12**, 312 (1970).
- [132] T. Wiseman, Numerical construction of static and stationary black holes, in *Black Holes in Higher Dimensions*, edited by G. T. Horowitz (2012), pp. 233–270, [arXiv:1107.5513](https://arxiv.org/abs/1107.5513).
- [133] G. Papallo and H. S. Reall, On the local well-posedness of Lovelock and Horndeski theories, *Phys. Rev. D* **96**, 044019 (2017).
- [134] A. D. Kovács and H. S. Reall, Well-posed formulation of Lovelock and Horndeski theories, *Phys. Rev. D* **101**, 124003 (2020).
- [135] O. Sarbach, E. Barausse, and J. A. Preciado-López, Well-posed Cauchy formulation for Einstein-ether theory, *Classical Quantum Gravity* **36**, 165007 (2019).
- [136] C. de Rham and S. Melville, Gravitational rainbows: LIGO and dark energy at its cutoff, *Phys. Rev. Lett.* **121**, 221101 (2018).
- [137] C. de Rham, The gravitational rainbow beyond Einstein gravity, *Int. J. Mod. Phys. D* **28**, 1942003 (2019).