Towards the nonlinear regime in extensions to GR: assessing possible options

Gwyneth Allwright$^1$ and Luis Lehner$^1$

$^1$Perimeter Institute for Theoretical Physics, Waterloo, ON, N2L2Y5, Canada

Testing General Relativity and exploring possible departures has received further input with the possibility to do so through gravitational waves emitted in strongly gravitating/highly dynamical scenarios and also through the availability of exquisitely sensitive cosmological observations. However, most extensions suffer from severe pathologies at the mathematical level which have stymied a thorough exploration of putative theories. With the aid of a model problem which captures typical pathologies, we explore suggested methods to control them. We find that the approach that modifies the equations to control higher-order gradients is both robust and efficient.

I. INTRODUCTION

The availability of increasingly sensitive cosmological and astrophysical data is accompanied by further scrutiny on General Relativity (GR) as the theory of gravity governing the universe (e.g. [1–6]). GR continues to successfully meet tests so far enabled by such data; indeed, examples include GR’s successful role in ΛCDM (e.g. [7]), binary pulsar timing (e.g. [2]), and gravitational waves from compact binary systems (e.g. [8]). Nevertheless GR should eventually show cracks [54] and the search for possible deviations together with their physical implications, are greatly enhanced through theoretical analysis of putative extensions to GR. Of particular interest is the highly dynamical, strongly gravitating regime which, arguably, presents the likeliest scenario to detect such deviations.

The analysis in such regimes is however hindered by several facts. First, it is yet unclear which specific theory (theories) one should focus on. Many theories have been proposed motivated by quantum gravity ideas, explorations of specific violations of fundamental principles of GR, alternatives to dark matter/dark energy, etc. While many of them are certainly appealing from academic reasons, no subset of theories has yet arisen as a preferred one.

Second, and at a mathematical level, many such theories do not lend themselves to defining well-posed problems –especially in the regime of interest. That is, given suitable initial conditions, a unique solution can be determined which depends continuously on such conditions [9]. Indeed, with the exception of just a few theories (within the subset of scalar-tensor or scalar-vector-tensor theories), which have been explored nonlinearly in compact binary mergers [10–13], lack of well-posedness presents a severe obstruction to the study of such theories. Of particular relevance is the understanding of compact objects, their stability and behavior in compact binary mergers, as well as cosmologies near the big-bang/or bounce, both of which are nonlinear regimes that naturally involve relativistic speeds and strong gravitational fields.

Of course, mathematical difficulties have been identified before in some theories and to different degrees. For instance, efforts have been directed towards avoiding so-called Ostrogradski’s instabilities (e.g. [14–17]). Such instability, when present, invariably leads to ill-posed problems. Unfortunately, even in theories free of such instability, well-posedness is far from guaranteed, as several aspects still need to be determined (see, e.g. [18]).

One traditional way to assess well-posedness –which relies on energy estimates– involves understanding the structure of the (principal part of the) evolution equations and assess whether they are: (i) symmetric/strongly hyperbolic, (ii) merely weakly hyperbolic or even if (iii) the system can transition from hyperbolic to elliptic within the domain of interest. Property (i), together with appropriate initial and boundary conditions, is essentially sufficient to guarantee well-posedness, irrespective of the lower-order terms of the equations. Property (ii), however, requires a careful analysis of lower-order terms and, with the exception of rare circumstances and in rather simple problems, ill-posedness typically follows in this case. As examples, one could mention that in Horndenski theories, weak hyperbolicity in all but a narrow corner has been recently demonstrated [19] and similarly concerning aspects have been discussed for gravitational Dynamical Chern Simons theory [20]. Also, numerical simulations in theories invoking a Vainshtein-like mechanism have explicitly shown pitfalls presented by some of the mathematical roadblocks [21] aluded. Last, property (iii) indicates there exist severe obstructions to determining the solution beyond where/when the hyperbolic/elliptic transition takes place.

A second hurdle, even with a strongly/symmetric hyperbolic system, is the fact that nonlinearities –in the principal part– typically imply that characteristics can cross. Once they do, uniqueness of the solution is lost and hence well-posedness. Such difficulties are already encountered at the level of hydrodynamical equations but, in such cases, further requirements –the so called Rankine-Hugoniot conditions– are invoked to single out a unique solution and restore well-posedness. As far as we know, the analog of such conditions in gravitational theories has yet to be developed. Finally, there exist theories for which partial differential equations theory is still to be developed. An example of such a case is given by the “simple looking” equation \( \Box \phi = \lambda (\Box \phi)^p + \ldots \) with \( p > 1 \), for which there is no mathematical guidance on how to even formally define an initial value problem.

The above discussion paints, at first sight, a rather
bleak picture with regards to the degree to which most extensions to GR can be thoroughly analyzed. But it might not be the case in practice. On one hand, in perturbative regimes –i.e. cosmology applications– the study of the linear perturbations with respect to a preferred solution allows from introducing an ultraviolet cutoff that circumvents many of the aforementioned problems. Thus, specific predictions can be drawn to test for possible deviations from GR and, with them hopefully draw lessons for the full theory. At the nonlinear level however, the situation is more uncertain. In such case, nonlinearities could be responsible for runaway energy cascades to the UV, rendering effective cutoffs delicate to impose without severely affecting the physics. Moreover, as such regimes require numerical simulations—which at the level of truncation error typically do source all frequencies allowed by the computational grid—sensitivity/uncontrolled growth at high frequencies represents a significant practical shortcoming standing in the way of exploring such regimes. Last, it is important to keep in mind that often extensions of GR are obtained from an effective field theory approach. A strong direct energy cascade to the UV takes the solution away from the regime of applicability of the theory and conclusions drawn in such a case need not be connected to the true phenomenology from the putative parent theory from which the EFT is, in principle, derived.

The aforementioned shortcomings are sufficiently delicate that, in a sense, they have prevented significant advances in the study of extensions to GR in nonlinear regimes. However, the detection of gravitational waves by LIGO/VIRGO, together with our current rather robust understanding of compact binary mergers in General Relativity, give hope to shedding light on this enterprise. Namely, current observations are consistent with no significant energy cascade to the UV in these systems. Remarkably, GW150914 has, in a rather spectacular fashion, allowed observing such waves even without invoking a specific theory! Whatever the true theory of gravity is, it seems not to significantly alter the behavior from that of GR—at least in the context of binary mergers—and putative higher-order corrections to GR seem to remain relatively small.

Furthermore, this lack of a strong direct energy cascade to the UV in the context of compact binary mergers is also what is observed in fully nonlinear simulations within General Relativity (see for instance and references cited therein).

These observations present both opportunities and challenges. An exciting opportunity is to study extensions of GR which, as mentioned, even with pathologies might allow one to devise suitable techniques to effectively control them while still capturing the essence of GR departures. After doing so, the challenge is to obtain sufficiently accurate predictions that can be tested or constrained through ever more sensitive detections.

In the current work, we concentrate on assessing proposed techniques to deal with extensions to GR. As the techniques we study are sufficiently general, we will refrain from making contact with any specific extension. Instead, we find it more informative to study a toy model which captures essentially all problems commonly found in extensions to GR. Crucially, being “UV-complete” allows for checking to what extent the solution to the EFT problem stays close to that from the original theory. This is a luxury not yet afforded by extensions to GR.

To fix ideas and briefly anticipate the strategies we will consider, let us assume the equations from a given theory can be schematically expressed as \( G(g) = \lambda S(g) \). Here, and for our current purposes, \( G \) stands for the Einstein tensor and \( S \) some nonlinear, perhaps higher derivative operator, \( \lambda \) is a coupling parameter considered small and \( g \) the metric tensor.

One of these strategies, often referred to as “reduction of order”, extends a successful technique from ordinary differential equations to the less certain partial differential equation arena. In this technique, one first solves for \( G = 0 \) (i.e. GR); the solution to this problem \( g^{(0)} \) is used to evaluate the “source” \( S(g^{(0)}) \) and a new solution \( g^{(1)} \) is obtained from \( G(g^{(1)}) = \lambda S(g^{(0)}) \). This process is iterated to a certain desired tolerance of the difference between two subsequent solutions, i.e. \( |g^{(i)} - g^{(i-1)}| < \epsilon \).

The second strategy, which we refer to as “fixing equations” introduced in Extends a suggestion by Israel and Stewart to address ill-posedness of relativistic hydrodynamics. Here, a new variable \( \Pi \), and its evolution equation (of the form \( \frac{\partial}{\partial \tau} \Pi = -\Pi + \lambda S \)), are introduced such that \( \Pi \) asymptotes to \( \lambda S \) (in a timescale determined by \( \tau \)) and controls higher gradients of \( \Pi \). Thus it remains consistent with the original system in the IR while restraining a runaway direct energy cascade to the UV.

Both the above prescriptions rely on the assumption that the physical system, for the regime of interest, does not naturally transfer energy to shorter wavelengths in a significant way. The former as it assumes \( S \) stays as a small correction and the latter as it controls the shorter wavelengths. As mentioned, observations of gravitational waves by LIGO/VIRGO seem to support such an assumption in at least some regimes and thus, invigorates examining these strategies in detail. This is the purpose of this work.

II. MODEL

We will apply our ideas to a toy model for a complex scalar field \( \phi \) with \( U(1) \) symmetry, as introduced in . Consider the following action:

\[
S = -\int_{\mathbb{R}^4} d^4x \left[ (\partial_{\mu} \phi^* \partial^{\mu} \phi) + V(\phi^* \phi) \right], \tag{1}
\]

with a potential \( V(\phi^* \phi) \) given by

\[
V(\phi^* \phi) = \frac{\lambda}{2} \left( \phi^* \phi - \frac{\mu^2}{2} \right)^2. \tag{2}
\]
At the classical level we will regard the action given by eq. (1) as defining a “UV-complete” theory. Following [30], an EFT from this action parameterized by the mass scale \( M \) can be derived. To this end, one introduces new real fields \( \rho \) and \( \theta \) and exploit that the minimum of the potential lies at \( \phi^* \phi = v^2 / 2 \) to write

\[
\phi(x) = \frac{v}{\sqrt{2}} \left[ 1 + \rho(x) \right] \exp[i\theta(x)].
\]  

(3)

Then the action given by eq. (1) can be expressed as

\[
\frac{S}{v^2} = -\int d^4x \mathcal{L}(\rho, \theta),
\]

(4)

with a Lagrangian density given by

\[
\mathcal{L}(\rho, \theta) \equiv \frac{1}{2} (\partial_{\mu} \rho) (\partial^{\mu} \rho) + \frac{1}{2} (1 + \rho) (\partial_{\mu} \theta) (\partial^{\mu} \theta) + V(\rho),
\]

(5)

and the potential \( V(\rho) \) is defined as

\[
V(\rho) = \frac{M^2}{2} \left( \rho^2 + \rho^3 + \frac{\rho^4}{4} \right).
\]  

(6)

The spectrum of fluctuations about the vacuum described by eq. (5) includes a massive field \( \rho \) and a massless Goldstone boson \( \theta \). The resulting equations of motion are,

\[
\Box \rho = (1 + \rho) (\partial_{\mu} \theta) (\partial^{\mu} \theta) + V'(\rho),
\]

\[\Box \theta = -2 (1 + \rho) \rho \partial_{\mu} \partial^{\mu} \theta.
\]

(7)

(8)

Now, when \( M \) is large compared with the energies of interest, the field \( \rho \) can be integrated out and capture, through an effective action its leading-order effects on \( \theta \). As described in [30], this leads to

\[
\frac{S}{v^2} \simeq -\int d^4x \tilde{\mathcal{L}}(\theta, \partial_\mu \theta, \partial_\mu \partial_\nu \theta),
\]

(9)

where, to \( \mathcal{O}(M^{-4}) \),

\[
\tilde{\mathcal{L}} \equiv \frac{1}{2} (\partial_\mu \theta) (\partial^\mu \theta) - \frac{1}{2M^2} [(\partial_\mu \theta) (\partial^\mu \theta)]^2 + \frac{2}{M^4} (\partial_\mu \partial_\nu \theta) (\partial^\mu \partial^\nu \theta) (\partial_\alpha \partial^\alpha \theta).
\]

(10)

The resulting equation of motion for \( \theta \) can be expressed as

\[
\Box \theta = \frac{2}{M^2} \left[ (\partial_\mu \theta) (\partial^\mu \theta) \Box \theta + 2 (\partial_\mu \partial_\nu \theta) (\partial^\mu \theta) (\partial^\nu \theta) \right] + \frac{4}{M^4} \left[ (\partial^\mu \theta \Box \theta) (\partial_\mu \partial_\nu \theta) + (\partial^\mu \theta \partial^\nu \theta) (\Box \theta) (\partial_\mu \partial_\nu \theta) + (\partial^\mu \theta \partial^\nu \theta) (\partial_\mu \partial_\alpha \theta) + (\partial^\mu \theta \partial^\nu \theta) (\partial_\alpha \partial_\beta \theta) \right. + \left. 2 (\partial^\mu \theta \partial^\nu \theta) (\partial_\mu \partial_\nu \theta) (\partial_\alpha \partial^\alpha \theta) \right].
\]

(11)

Note the appearance of third- and fourth-order derivatives. In the UV-complete theory, these terms do not arise: we see from eq. (7) and (8) that the highest derivatives are second-order.

In general, the presence of time derivatives that are higher than second-order in the equations of motion is a strong indication of ill-posedness. This is because such derivatives usually give rise to the so-called Ostrogradski instability: the Hamiltonian is unbounded from below, leading to runaway solutions [17]. Because of the Ostrogradski instability, it has been argued that keeping all time derivatives in the action at second-order and below should be an important restriction for any physical theory. However, as illustrated in [29], higher-order spatial derivatives are also problematic.

Importantly, as discussed in [17], it is possible to address, to some extent, certain potential sources of ill-posedness arising from EFTs. Indeed, through field redefinitions higher-order derivatives can be removed, or pushed to higher orders. In our case, one can use a field redefinition to push the higher-order time derivatives in eq. (11) to \( \mathcal{O}(M^{-6}) \). As described in [17], let

\[
\theta \to \theta + \frac{2}{M^2} (\partial_\mu \partial_\nu \theta) (\partial^\mu \theta) (\partial^\nu \theta).
\]

(12)

One can then rewrite the EFT action given by eq. (9) as

\[
\frac{S}{v^2} \simeq -\int d^4x \left\{ \frac{1}{2} \partial_\mu \theta \partial^\mu \theta - \frac{1}{2M^2} [\partial_\mu \theta \partial^\mu \theta]^2 \right. \]

\[
\left. - \frac{1}{M^4} \partial_\mu \theta \partial^\mu \theta [\partial_\mu \partial_\nu \theta \partial^\mu \partial^\nu \theta - (\Box \theta)^2] \right\}.
\]

(13)

The equation of motion that follows from the action given by eq. (13) is

\[
\Box \theta = \frac{2}{M^2} \left[ (\partial_\mu \theta) (\partial^\mu \theta) \Box \theta + 2 (\partial_\mu \partial_\nu \theta) (\partial^\mu \theta) (\partial^\nu \theta) \right] + \frac{2}{M^4} \left[ 3 (\partial_\mu \theta) (\partial^\mu \theta) (\partial_\nu \partial^\nu \theta) - (\Box \theta)^3 \right.

\[
\left. - 2 (\partial_\mu \partial_\nu \theta) (\partial_\nu \partial_\alpha \theta) (\partial^\mu \partial^\alpha \theta) \right].
\]

(14)

Note the higher-order derivatives cancelled out, thus sidestepping the Ostrogradski instability. A related approach to remove higher derivatives is to consider adding counterterms in the action [31]. At this point, we find it important to stress once again that the absence of this instability does not guarantee well-posedness. Other problems might be in the way still like: lack of uniqueness, uncertain character of the resulting equations, runaway cascade to the UV, etc.

In the next section, we will illustrate techniques proposed to deal with such issues in the nonlinear regime. Necessarily, this will require a numerical treatment. We stress, however, that the issues we are dealing with are analytical in nature, but they have strong implications on the way we can explore through numerical means the targeted regimes.
III. ASSESSMENT FRAMEWORK

As discussed, our regime of interest is nonlinear, strongly gravitating and highly dynamical. In this enterprise, numerically implementing the equations will be required to at least gain some intuition of the dynamics of the problem under study. As a result, analytical considerations will intertwine with numerical ones.

We first provide an outline of several techniques that can be used for obtaining the numerical solutions to the EFT equations of motion. For simplicity, we limit our considerations will intertwine with numerical ones. As a result, analytical requirements to at least gain some intuition of the dynamics of the problem under study. As a result, analytical conclusions will be required to at least gain some intuition of the dynamics of the problem under study. As a result, analytical considerations will intertwine with numerical ones.

We provide an outline of several techniques that can be used for obtaining the numerical solutions to the EFT equations of motion. For simplicity, we limit our considerations will intertwine with numerical ones. As a result, analytical requirements to at least gain some intuition of the dynamics of the problem under study. As a result, analytical conclusions will be required to at least gain some intuition of the dynamics of the problem under study. As a result, analytical considerations will intertwine with numerical ones.

In one spatial dimension, eq. (11) can be written as the following six coupled equations:

\[
\begin{align*}
(\theta_{xx} - \theta_t) \left[ -1 + \frac{2}{M^2} (\theta_x^2 - \theta_t^2) \right] + \frac{4}{M^2} (\theta_x \theta_t^2 - 2 \theta_t \theta_x \theta_t + \theta_t \theta_t^2) &= 0 . \tag{15}
\end{align*}
\]

Note that eq. (15) contains no more than two derivatives per field. Furthermore, it is also free of further complications that do appear at higher orders. In particular, second derivatives of \( \theta \) (the principal part of the system) appear multiplied at most by first derivatives of the field \( \theta \). Consequently, the equation still has a known PDE hyperbolic character; however, standard arguments indicate characteristics will generically cross \[32\]. This will lead to non-uniqueness and the loss of differentiability of the solution which, in turn, implies that one abandons the regime of applicability of the EFT expansion. Nevertheless, until this happens, eq. (15) can be directly solved numerically and we will do so to compare with those obtained through approximate strategies. Similarly, we can also obtain the solution of \( \{ \rho, \theta \} \) in the full theory (eqs. \[7\] \[8\]) and such solutions will be used as the base ones to compare against. In what follows we outline the basic implementation of each option.

A. “UV-Complete” Problem

Since we restrict to the classical regime, the original problem described by eq. (7) and (8) can be formally studied for any frequency. We thus refer to this as the “UV-complete” problem, and solutions obey

\[
\rho_{tt} - \rho_{xz} = (1 + \rho) (\theta_t^2 - \theta_x^2) - \frac{M^2}{2} (2\rho + 3\rho^2 + \rho^3) ,
\]

\[
\theta_{tt} - \theta_{xx} = \frac{2}{1 + \rho} (\rho_x \theta_x - \rho_t \theta_t) . \tag{16}
\]

For practical reasons, we will write this system in first-order form, by introducing the variables, \( \{ f \equiv \theta_x, g \equiv \theta_t, u \equiv \rho_x, v \equiv \rho_t \} \). Then eq. (16) and (17) can be written as the following six coupled equations:

\[
\begin{align*}
&f_t = g_x , \tag{18} \\
g_t = f_x + \left( \frac{2}{1 + \rho} \right) (fu - gv) , \tag{19} \\
u_t = v_x , \tag{20} \\
v_t = u_x + (1 + \rho) (g^2 - f^2) - \frac{M^2}{2} (2\rho + 3\rho^2 + \rho^3) , \tag{21} \\
\rho_t = v , \tag{22} \\
\theta_t = g . \tag{23}
\end{align*}
\]

Notice that there are no terms with fields multiplying the principal part (first derivatives) of the equations; the system is linearly degenerate and will not lead to discontinuities from smooth initial data.

B. EFT: \( \mathcal{O}(M^{-2}) \) Truncated Theory

In terms of the variables \( \{ f \equiv \theta_x, g \equiv \theta_t \} \), eq. (15) can be reduced to the following three coupled first-order equations:

\[
\begin{align*}
f_t &= g_x , \tag{24} \\
g_t &= f_x \left[ 1 - \frac{4 (f^2 + g^2)}{M^2 - 2 f^2 + 6 g^2} \right] + g_x \left( \frac{8 fg}{M^2 - 2 f^2 + 6 g^2} \right) , \tag{25} \\
\theta_t &= g . \tag{26}
\end{align*}
\]

As already anticipated, eq. (25) contains terms with fields multiplying the principal parts of the equation. The equation is now truly nonlinear and discontinuities will generally develop from smooth initial data \[32\]. Further, note we can consider this method at this order only, as higher-order ones will contain higher derivatives which will require special treatment. Thus, the possibility of “direct integration” will be of limited application. We will primarily contemplate it here for comparison purposes and to bring attention to some specific issues. In general, other options, like the ones discussed next, will be required.

The next two strategies are envisioned to address these shortcomings, and we will assess their behavior with some examples in the next section.

C. Approach 1: “Fixing the Equations”

The first strategy to study the EFT equation of motion is based on the methodology introduced in \[29\], which, in
turn, is inspired by the Israel-Stewart formulation of relativistic viscous hydrodynamics [33] (see also, e.g. [34]). The key idea in this approach is to control terms in the equation of motion that effectively contain higher-order gradients, thereby preventing a runaway towards the UV. In the current problem, this can be achieved in the following way. Modify the $O(M^{-2})$ (or higher) EFT equation by introducing a new variable, $\Pi$ with its own evolution equation as
\[
\square \theta = \Pi ,
\]
\[
\Pi = \frac{2}{M^2} \left[ (\partial_{\mu} \theta) (\partial^\mu \theta) \Pi + 2 (\partial_{\mu} \partial_{\nu} \theta) (\partial^\mu \theta) (\partial^\nu \theta) \right] + \tau \Pi ,
\]
where $\tau$ is an external timescale. Note that if $\tau$ is negative, then the effect of the extra term in eq. (28) is to drive $\Pi$ to
\[
\frac{2}{M^2} \left[ (\partial_{\mu} \theta) (\partial^\mu \theta) \square \theta + 2 (\partial_{\mu} \partial_{\nu} \theta) (\partial^\mu \theta) (\partial^\nu \theta) \right] ,
\]
thus recovering the original equation. The above eqn. (28) is one equation that achieves the desired goal, however, as noted in [35], as long as higher gradient terms remain small, different equations can be envisaged, and the choice is largely irrelevant if no significant direct cascade of energy takes place. In one dimension, eq. (28) can be expressed as
\[
\theta_{xx} - \theta_{tt} = \Pi ,
\]
\[
\Pi = \tau \Pi_t + \frac{2}{M^2} \left[ (\theta_x^2 - \theta_t^2) \Pi + 2 (\theta_x \theta_x^2 + \theta_t \theta_t^2 - 2 \theta_{xt} \theta_x \theta_t) \right] .
\]

Again, introducing $\{ f \equiv \theta_x \text{ and } g \equiv \theta_t \}$, one ends up with the following system of coupled equations:
\[
f_t = g_x ,
\]
\[
g_t = f_x - \Pi ,
\]
\[
\theta_t = g ,
\]
\[
\Pi_t = \frac{1}{\tau} \left\{ \Pi - \frac{2}{M^2} \left[ (f_x^2 - g^2) \Pi + 2 (f_x f_x^2 + (f_x - \Pi) g^2 - 2 g_x f_g) \right] \right\} .
\]

D. Approach 2: “Reduction of Order”

Once again, we would like to solve the one-dimensional $O(M^{-2})$ EFT equation of motion, as given by eq. (15). Inspired by an approach introduced to deal with the Abraham-Lorentz-Dirac equation and avoid runaway solutions [28], an iterative or reduction of order method is operationally introduced in the following way. One rewrites the equation of motion as
\[
\theta_{xx} - \theta_{tt} = S(\theta) ,
\]
where we define,
\[
S(\theta) = \frac{2}{M^2} \left[ (\theta_x^2 - \theta_t^2) (\theta_{xx} - \theta_{tt}) + 2 (\theta_x \theta_x^2 - 2 \theta_{xt} \theta_x \theta_t + \theta_{tt} \theta_t^2) \right] .
\]

Then, one regards this equation as one where the leading part is affected by a subleading source $S$. Since the source depends on the field itself, an iterative procedure is implemented whereby the equation is treated as $\square (\theta^i) = S(\theta^{i-1})$. The assumption here is that the source $S$ will remain subleading, and the strategy of decoupling it from the principal part implies, at first sight, that the “offending” terms in $S$ play only a passive role in determining the solution, thus hopefully controlling a runaway behavior towards the UV (e.g. [36]). Notice that this approach does not introduce further variables, though it does require iterating to some desired tolerance (the difference between successive solutions). Notice that the source $S$ has, in particular, second time derivatives of the field $\theta$ and they need to be evaluated somehow in the iterative scheme. This is no problem at intermediate times, but it is an issue at (and near) end points of the time interval of interest. We will comment on this in the next section.

IV. NUMERICAL IMPLEMENTATION

Having discussed the four options we can pursue, i.e.: (i) Full “UV-complete” problem, (ii) Direct integration of the equations truncated to $O(M^{-2})$ order (TR), as well as adopting the (iii) “Fixed equations” (F) or (iv) the iterative or “Reduction of Order” (RO) approach, we are now in a position to study and contrast the nonlinear behavior described by each option. To this end, we resort to numerical simulations. As a simple setup, we consider a periodic problem with $x \in [0, L]$ discretized uniformly with spacing $\delta x$. Spatial derivatives are approximated through Finite Differences through a standard centered fourth-order accurate expression. Time integration is obtained through the Method of Lines via a Runge-Kutta third-order accurate approximation with $\delta t = \delta x/4$ to satisfy the CFL condition (for further details see, e.g. [37, 38]). The initial data is given by,
\[
\theta(x, t_0) = \begin{cases} 
10^{-5} (x - 3)^4 (x - 7)^4 & \text{if } x \in [3, 7] \\
0 & \text{otherwise}
\end{cases}
\]
and
\[
\theta_t(x, t_0) = 0 ,
\]
where $t_0 = 0$; for concreteness we adopt $L = 10$.

Notice that for options (iii) and (iv), further information must be specified. From the point of view the initial value problem, $\Pi$ (as it is a new variable) or $\theta_{tt}$ (as it is
part of the source) is required at \( t = 0 \). Providing such information is in and of itself a delicate issue, as it can aid in removing spurious modes, as discussed in [30]. In our tests, we adopt \( \Pi(t = 0) = 0, \theta_M(t) = 0 \). In addition, a few “external parameters” are required:

- In case (iii), the parameter \( \tau \). For the most part we will keep it to \( \tau = -40/\tau = -4 \) (though we will also compare with other values).

- For (iv), two extra parameters are required. One is the tolerance \( \epsilon \) (here chosen to be \( 10^{-12} \)). The other is the period of physical time \( \Delta T \) over which the integration will be performed. For instance, if the total physical time of interest is \( t_0 = 0 \) to \( T \), one could choose to integrate the homogenous problem \( \Box \theta^{(0)} = 0 \) over this full range (and iterate over it), or over a smaller interval \( \Delta T \) and advance over this interval to cover the full interval of interest. The introduction of such an interval is natural, and is analogous to what is done, for instance in the integration of the Abraham-Lorentz-Dirac equation [28] (see also [39, 40]) through a reduction of order approach. There it is found, as naturally expected, that more accurate solutions are obtained with smaller intervals.

Last, and with respect to this approach, we note a technical but delicate point. Namely, the evaluation of second-order (or higher) time derivatives that would appear in the application of this scheme when choosing a given \( \Delta T \) has an impact on the practical interval of time where the solution can be found. Indeed, a given interval (say \( \Delta T \)) will not have the final solution (e.g. \( \theta^{(0)} \)) up to \( t = \Delta T \), as the source could only be accurately evaluated to time \( t = \Delta T - \kappa \delta t \) (with \( \kappa \) an integer value related to the accuracy desired in such evaluation, e.g. for fourth order accuracy \( \kappa = 2 \)). As a result, the usable span of time within a given interval is necessarily smaller than \( \Delta T \). Besides this technical detail, the method can be implemented rather straightforwardly.

V. RESULTS

In what follows, we will compare errors between solutions obtained with the three approximate systems of equations \{i(iii)-(iv)\} –which we will collectively refer to as EFT solutions– with respect to those obtained with the “UV-complete” problem (i). Further, we will take different values of \( M \) and study the errors’ dependence on this “coupling” scale. In the context of extensions to General Relativity, beyond the Planck length there is, in principle, no specific value of coupling naturally expected. Consequently, increasingly sensitive detections would help to place bounds on such coupling, provided a clear understanding is obtained for multiple values of it. It will thus be important to understand the extent to which different methods perform under a range of couplings within the regime of validity of the EFT approximation. With respect to the initial data adopted, a reasonable estimate of the size of corrections introduced by the terms at \( O(M^{-2}) \) is \( 2 \times 10^5/M^2 \) (estimated by comparing the magnitude of corrections relative to terms in the principal part of the uncorrected equations). In our tests we thus concentrate on \( M = \{0.01, 0.1, 1\} \) to explore a range. All these values would imply that corrections are, at least initially, small and thus within the regime of validity of the EFT approximation. Naturally, \( M = 0.01 \) will represent the most demanding case.

To begin, we confirmed the expected convergence and stability of the homogeneous problem \( (\Box \theta = 0) \) and convergence of solutions obtained with the truncated problem for \( M = 1 \) until \( T = 5t_c \). With this implementation, we next thoroughly explore the behavior of solutions obtained with the three EFT options and contrast their behavior with the ones obtained with the UV-complete problem. In particular, we monitor (the \( L_2 \) norm of the difference) between the solutions and compute the relative difference defined as,

\[ E \equiv \frac{||\theta_{EFT} - \theta_{UV}||_{L_2}}{||\theta_{UV}||_{L_2}}, \tag{40} \]

where \( \theta_{EFT} \) denotes each of the EFT solutions and \( \theta_{UV} \) the UV-complete one.

For concreteness, we adopt in our production runs \( N_x = 200 \), which shows excellent agreement with the much finer refined grid of \( N_x = 1600 \) in representative cases.

Figure 2 illustrates \( E(t) \) for the truncated (TR), iterative (RO) and fixed (F) schemes in the case \( M = 1 \) (left) and \( M = 0.1 \) (right) versus crossing time \( (t_c \equiv t/L) \) respectively. For this test, we employ a grid which covers the computational domain with \( N = 200 \) points (and for representative cases we compare with solutions obtained with \( N = 100 \) and \( N = 1600 \) to confirm the observed behavior). For the iterative case, we adopted \( \Delta T = 0.015t_c \), and for the “fixed” case \( \tau = -4 \). For early times, all methods perform rather well, with the iterative case showing the smallest differences until \( T \approx 100t_c \), at which point its behavior gradually worsens when compared with the fixed and truncated schemes. Over long integration periods, the fixed method shows the best behavior, with relative errors of order \( \approx 0.6\%, 50\% \) for \( M = 1, 0.1 \) respectively at \( T = \{800, 400\} \) respectively. The truncated method performs somewhere in between the other two, giving errors of order \( \approx 2\%, 100\% \) for \( M = 1, 0.1 \) respectively at \( T = \{800, 400\} \).

As an illustration of the different solutions at relatively early times, figure 2 displays them at \( t = 19t_c \). At this stage, and for this coupling, all methods produce solutions quite similar to the UV-complete one, with the
exception of the iterative one if \( \Delta T \) is not chosen sufficiently small—recall that the iterative method depends on the value chosen for \( \Delta T \).

To examine more closely the dependency of the solutions obtained with the iterative method when varying \( \Delta T \), figure 3 presents the relative errors for different values of \( \Delta T \) with respect to the UV-complete case. For the case \( M = 1 \), significantly smaller errors are obtained for \( \Delta T = 0.015 t_c \) (i.e., the smallest value adopted) than with \( \Delta T = 0.15 t_c \). Additionally, as naturally expected we find that with \( \Delta T = 0.015 t_c \), satisfying the tolerance requirement involves a couple of iterations while for \( \Delta T = 0.15 t_c \) up to 16 iterations are involved. The smallest integration interval also yields the smallest error in the case \( M = 0.1 \), but no clear trend is found for larger values of \( \Delta T \). We further note that for these larger values of \( \Delta T \) there are a number of instances were the tolerance criteria could not be met, though the difference achieved for the largest possible number of iterations was still of order \( 10^{-11} \).

The fixed scheme also depends on the external parameter \( \tau \) and we examine its sensitivity to varying such parameter. The results are presented in figure 4, which shows the relative errors for the case \( M = 0.1 \). Upon varying this parameter by two orders of magnitude,
the solutions obtained are rather similar, though for the larger values of \( \tau \) the obtained solutions are closest.

To significantly stress the methods, we also explore solutions for the value \( M = 0.01 \), focusing on the iterative and fixed methods in particular. Figure 5 shows the norm of the solutions obtained with different methods for this case. The iterative method copes with such a strong coupling for some time, but it eventually leads to an exponentially diverging behavior and by \( t \approx 35.9 t_c \) can no longer yield a solution. The fixed method however, stays bounded, displaying instead a slow exponential decay that is reduced for larger values of \( \tau \).

We illustrate different solutions for this case in figure 6 at \( t = 35.95 t_c \). The solutions obtained with the iterative method clearly diverge at this stage. The fixed method yields well-behaved solutions, though the larger the value of \( \tau \), the closer its solution is to that of the UV-complete problem.

VI. OBSERVATIONS AND FINAL COMMENTS

The unprecedented opportunity to explore gravity in strongly gravitating/highly dynamical regimes has given further input to explore ideas on extensions of General Relativity and hopes to put them to the test. However, most such extensions typically involve pathological features that have hindered a thorough understanding of their implications, especially in regimes where deviations from GR would be largest. Recently, several different ideas have been suggested as possible paths forward in this endeavour [29, 36, 41] and first steps are being taken [41,44]. In this work, weexploited a toy-model which –from an EFT point of view– evidences many of such pathologies, but within a known UV-complete theory. With this model we assessed possible strategies to study such extensions and illustrated their strengths and shortcomings. In particular, we find that the approach that “fixes” the equations [29] to control higher-order gradients performs quite well, both in terms of approximating the UV-complete problem, the robustness it displays with strong couplings and the computational cost associated to implementing it. Faithful solutions are obtained for sufficiently large values of \( \tau \); though we note that for very large values, the associated equation for \( \Pi \) will become stiff. However, the usage of IMEX methods [45] and the fact that the stiff term is quite simple, ensures that one can easily implement the equations without affecting the computational cost. The reduction of order method appears to perform well for sufficiently...
weak couplings and small intervals $\Delta T$. It is relatively straightforward to implement, though it shows a lack of robustness with stronger couplings. However, this behavior could be ameliorated by implementing corrections in a perturbative expansion as introduced recently in [41]—if stronger couplings are of interest. We note that due to the iteration required, its associated computational cost is higher.

Importantly, the faithfulness of all the methods discussed—and in fact the use of EFT to derive a theory and apply it to a given case—requires the physical system to not present a strong cascade of energy to the UV (or to restrict the solution to the regime where such behavior is not present). Otherwise, terms considered and treated as small (in the iterative scheme) or controlled (in the fixed scheme) would yield solutions not necessarily tracking the underlying gravitational theory one wishes to study. Taking GR as an example, many non-trivial scenarios do not display a strong UV cascade. Examples include scenarios ranging from the nonlinear stability of Minkowski [40], and the existence of stable stars (e.g. [47]) to subtle islands of stability in GR in the presence of a negative cosmological constant [48, 49]. Of particular interest, the merger of astrophysical compact objects also seems to be one where such a behavior is satisfied. Both observational evidence, e.g. [24, 50, 51] and analytical studies (through simulations), (see e.g. [25, 27] and references cited), illustrate a rather simple behavior where energy is mainly contained in relatively long-wave modes (commensurate with the size of the objects involved).

However, we know a strong cascade to the UV does indeed take place in the right regimes. For instance, singularity theorems [52] do imply a runaway behavior to the UV—a runaway that is halted if a black hole forms which introduces a natural cutoff. It is not known whether this behavior is also present in extensions to GR, we hope this work helps guide efforts to answer this and related questions through the analysis of solutions to particularly relevant scenarios. In the current work, the adopted toy-model afforded the luxury of having the “UV-complete” solution to compare against. In the gravitational case this is obviously not the case and assessing the faithfulness of obtained solutions is more delicate. This can be done by evaluating the original (EFT) system of equations and studying its residual. Small values would indicate the solution obtained is within the regime of applicability of the EFT expansion. Additionally, one can explore the impact of the “extra parameters” (e.g. $\Delta T, \tau$) and gauge whether the solutions depend on them in a sensitive manner. Regimes not displaying a significant energy cascade to the UV, or manifesting a natural cutoff within the regime of applicability of the EFT should pass these tests.

Acknowledgements- We thank C. deRham, M. Johnson, N. Ortiz, O. Reula, M. Okounkova, F. Pretorius, L. Senatore, T. Sotiriou, L. Stein, A. Tolley, M. Trodden and R. Wald for helpful discussions. This research was supported in part by NSERC, CIFAR and the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada, and by the Province of Ontario through the Ministry of Research and Innovation. LL thanks the Mainz Institute for Theoretical Physics (MITP) and the Centro de Ciencias de Benasque Pedro Pascual for its hospitality and support during the completion of this work.

[1] C. M. Will, “Was Einstein Right? A Centenary Assessment,” arXiv:1409.7871 [gr-qc], arXiv:1409.7871.
[2] P. C. C. Freire, N. Wex, G. Esposito-Farese, J. P. W. Verbiest, M. Bailes, B. A. Jacoby, M. Kramer, I. H. Stairs, J. Antoniadis, and G. H. Janssen, “The relativistic pulsar-white dwarf binary PSR J1738+0333 II. The most stringent test of scalar-tensor gravity,” Mon. Not. Roy. Astron. Soc. 423 (2012) 3328, arXiv:1205.1450 [astro-ph.GA].
[3] T. Baker, P. G. Ferreira, and C. Skordis, “The Parameterized Post-Friedmann framework for theories of modified gravity: concepts, formalism and examples,” Phys. Rev. D 87 no. 2, (2013) 024015, arXiv:1209.2117 [astro-ph.CO].
[4] N. Yunes and X. Siemens, “Gravitational-wave tests of general relativity with ground-based detectors and pulsar-timing arrays,” Living Reviews in Relativity 16 no. 1, (2013) 9.
[5] E. Berti et al., “Testing general relativity with present and future astrophysical observations,” Classical and Quantum Gravity 32 no. 24, (2015) 243001, http://stacks.iop.org/0264-9381/32/i=24/a=243001.
[6] N. Yunes, K. Yagi, and F. Pretorius, “Theoretical physics implications of the binary black-hole mergers gw150914 and gw151226,” Phys. Rev. D 94 (Oct, 2016) 084002, http://link.aps.org/doi/10.1103/PhysRevD.94.084002.
[7] Planck Collaboration, N. Aghanim et al., “Planck 2018 results. VI. Cosmological parameters,” arXiv:1807.06209 [astro-ph.CO].
[8] LIGO Scientific Collaboration and Virgo Collaboration, B. P. A. et al. (LIGO Scientific Collaboration and V. Collaboration), “Binary black hole mergers in the first advanced ligo observing run,” Phys. Rev. X 6 (Oct, 2016) 041015, http://link.aps.org/doi/10.1103/PhysRevX.6.041015.
[9] J. Hadamard, “Sur les problèmes aux dérivées partielles et leur signification physique,” Princeton University Bulletin 13 (1902) 49–52.
[10] J. Healy, T. Bode, R. Haas, E. Pazos, P. Laguna, D. M. Shoemaker, and N. Yunes, “Late Inspiral and Merger of Binary Black Holes in Scalar-Tensor Theories of
[43] A. Ijjas, F. Pretorius, and P. Steinhart. Personal communication.

[44] J. Ripley and F. Pretorius. Personal communication.

[45] L. Pareschi and G. Russo, “Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation,” ArXiv e-prints (Sept., 2010), arXiv:1009.2757 [math.NA]

[46] D. Christodoulou and S. Klainerman, The Global Nonlinear Stability of the Minkowski Space. Martin Classical Lectures. New Series. Princeton University Press, 1993.

https://books.google.ca/books?id=8QlpQgAACAAJ

[47] S. L. Shapiro and S. A. Teukolsky, Black holes, white dwarfs, and neutron stars: The physics of compact objects. 1983.

[48] A. Buchel, L. Lehner, and S. L. Liebling, “Scalar Collapse in AdS,” Phys. Rev. D86 (2012) 123011 arXiv:1210.0890 [gr-qc]

[49] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner, and S. L. Liebling, “Holographic Thermalization, Stability of Antide Sitter Space, and the Fermi-Pasta-Ulam Paradox,” Phys. Rev. Lett. 113 no. 7, (2014) 071601 arXiv:1403.6471 [hep-th]

[50] Virgo, LIGO Scientific Collaboration, B. Abbott et al., “GW170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2,” Phys. Rev. Lett. 118 (Jun., 2017) 221101 https://link.aps.org/doi/10.1103/PhysRevLett.118.221101

[51] LIGO Scientific and Virgo Collaboration

[52] S. W. Hawking and G. F. R. Ellis, The Large Scale Structure of Space-Time Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011.

[53] A. E. Fischer and J. E. Marsden, “Linearization stability of the Einstein equations,” Bull. Amer. Math. Soc. 79 no. 5, (09, 1973) 997–1003. https://projecteuclid.org:443/euclid.bams/1183534988

[54] Though a part of the theoretical physics community argues this is already the case due to the need for dark matter and dark energy.

[55] Note however this assumes linearization stability of the theory. That is, solutions to the linearized problem are consistent with those of the full problem linearized. In the case of General Relativity, the conditions for this being the case were discussed in [53].

[56] Also, further coupled fields can be considered, with their own equations of motion

[57] A second-order form with the Box operator is also possible, see [29]

[58] As well as fluid-gravity duality and related arguments, see [29]