

**Black hole binaries in cubic Horndeski theories**Pau Figueras<sup>\*</sup> and Tiago França<sup>†</sup>*School of Mathematical Sciences, Queen Mary University of London,  
Mile End Road, London E1 4NS, United Kingdom* (Received 21 February 2022; accepted 11 May 2022; published 2 June 2022)

We study black hole binary mergers in certain cubic Horndeski theories of gravity, treating them fully nonlinearly. In the regime of validity of effective field theory, the mismatch of the gravitational wave strain between Horndeski and general relativity (coupled to a scalar field) can be as large as 10%–13% in the Advanced LIGO mass range. The initial data and coupling constants are chosen such the theory remains in the weakly coupled regime throughout the evolution. In all cases that we have explored, we observe that the waveform in the Horndeski theory is shifted by an amount much larger than the smallness parameter that controls the initial data. This effect is generic and may be present in other theories of gravity involving higher derivatives.

DOI: [10.1103/PhysRevD.105.124004](https://doi.org/10.1103/PhysRevD.105.124004)**I. INTRODUCTION**

Detection of gravitational waves produced in stellar-mass black hole binary mergers have revolutionized both the experimental and theoretical studies of gravity. On the one hand, the direct detection of gravitational waves allows for the possibility to perform new tests of general relativity (GR) in the strong field regime<sup>1</sup> and, perhaps, observe deviations from the established theory. To do so, one needs to be able to compute theoretical waveforms from black hole binary mergers in alternative theories of gravity, see e.g., [1]. This necessity has prompted a lot of activity in recent years to study dynamics of black holes in such theories. On the other hand, the current experimental data suggests that the potential deviations from GR are small. Whilst it is unlikely that the present generation of gravitational wave observatories can detect such deviations, the future third generation of detectors and/or the space based detectors will provide more data and with higher precision. Therefore, it is not impossible that even small deviations from GR can be detected in the future. However, there is a lack of full gravitational waveforms in alternative theories of gravity; this implies some deviations from GR may be undetected.

To theoretically compute waveforms produced in the strong field regime one needs to work with a theory that has a well-posed initial value problem. This is not an issue if the corrections to GR are treated perturbatively. In fact, in recent years, significant progress has been made in studies of the strong field dynamics of certain alternative theories of interest using this approach [2–8]. However, it is well known

that perturbation theory may break down over sufficiently long times due to secular effects.<sup>2</sup> In addition, perturbation theory may miss certain nonperturbative effects which, even if very small, may be detectable over sufficiently many orbits of a binary, as future generations of gravitational wave detectors expect to be able to observe. Therefore, it is of interest to treat alternative theories of gravity fully nonlinearly and uncover some of their (perhaps) unique physical effects that may break certain degeneracies.

In this article, when we discuss alternative theories of gravity, we refer to the ‘strongly-coupled regime’ of the theory as the regime in which the new terms in the equations of motion that modify GR are comparable (or even larger) to the original (two-derivative) terms. Conversely, by the ‘weakly-coupled regime’ we will mean the regime of the theory in which the modifications to the GR equations of motion are small. This is compatible with still being in the strong field regime of gravity. It is in the weakly coupled regime that alternative theories of gravity make sense as low-energy effective field theories (EFTs).

Up until recently, only the so-called scalar-tensor and the scalar-vector-tensor theories of gravity had been studied fully nonlinearly [10–13]. The reason is that for this class of theories, it is straightforward to find a well-posed formulation. For other, more general, classes of theories involving higher derivatives and yet second-order equations of motion, such as Horndeski or Lovelock theories, finding a suitable well-posed formulation turns out to be far more difficult. In fact, it has been shown that weak hyperbolicity can fail in Lovelock [14] or Horndeski [15–21] theories if the spacetime curvature and/or the derivatives of the scalar field become too large, i.e., in

<sup>\*</sup>p.figueras@qmul.ac.uk<sup>†</sup>t.e.franca@qmul.ac.uk<sup>1</sup>In this article, by the ‘strong field regime’ we mean the regime in which the nonlinearities of the theory are important.<sup>2</sup>There are recent interesting attempts to resum the perturbative series and hence alleviate these secular effects [9].

the strongly-coupled regime. In a recent breakthrough, [22,23] showed that these theories can be strongly hyperbolic in certain modified generalized harmonic coordinates, also in the weakly-coupled regime, i.e., when the deviations from GR are small. These theoretical developments have led to the first studies of the fully nonlinear dynamics of black holes in a particular subset of these theories; namely, scalar Einstein-Gauss-Bonnet theory [24,25].

A more general approach to find well-posed formulations of general alternative theories of gravity has been proposed by [26,27]. This proposal is inspired by the Müller-Israel-Stewart (MIS) formulation of relativistic viscous hydrodynamics [28–31], and in principle can work even for theories with higher-than-second-order equations of motion, as recently shown in [32] for a certain eight-derivative theory of gravity and in [33,34] for scalar-tensor theories of gravity, albeit with two-derivative equations of motion.

We should remark that even if a well-posed formulation can be found for a certain alternative theory of gravity, it is possible that during the dynamical evolution of certain classes of initial data, the hyperbolicity of the equations of motion is lost due to strong coupling effects. In the context of Einstein-dilaton-Gauss-Bonnet and Horndeski theories, it has been shown that this loss of hyperbolicity is due to Tricomi or Keldysh transitions, in which characteristic speeds go to zero or diverge in finite time respectively, effectively changing the character of the scalar equation of motion from hyperbolic to parabolic or elliptic, after which time evolution cannot proceed [15–19,21].

Therefore, given these recent theoretical developments, it is the right time to start probing the nonlinear regime of alternative theories of gravity and infer predictions for black hole binary mergers. Building on our previous work [21], hereinafter referred to as Paper I, in this paper we study black hole binary mergers in cubic Horndeski theories. Horndeski theories are the most general scalar-tensor theories of gravity with second order equations of motion arising from a diffeomorphism invariant action in four spacetime dimensions [35,36]. Letting  $\phi$  be a real scalar field, then the general action for the Horndeski theories is given by

$$\mathcal{S} = \frac{1}{16\pi G} \int d^4x \sqrt{-g} (\mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5), \quad (1)$$

with

$$\begin{aligned} \mathcal{L}_1 &= R + X - V(\phi), \\ \mathcal{L}_2 &= G_2(\phi, X), \\ \mathcal{L}_3 &= G_3(\phi, X) \square \phi, \\ \mathcal{L}_4 &= G_4(\phi, X) R + \partial_X G_4(\phi, X) [(\square \phi)^2 - (\nabla_\mu \nabla_\nu \phi)(\nabla^\mu \nabla^\nu \phi)], \\ \mathcal{L}_5 &= G_5(\phi, X) G_{\mu\nu} \nabla^\mu \nabla^\nu \phi \\ &\quad - \frac{1}{6} \partial_X G_5(\phi, X) [(\square \phi)^2 - 3 \square \phi (\nabla_\mu \nabla_\nu \phi)(\nabla^\mu \nabla^\nu \phi) \\ &\quad + 2 (\nabla_\mu \nabla_\nu \phi)(\nabla^\nu \nabla^\rho \phi)(\nabla_\rho \nabla^\mu \phi)], \end{aligned} \quad (2)$$

where  $R$  and  $G_{\mu\nu}$  are the Ricci scalar and the Einstein tensor respectively constructed from the spacetime metric  $g_{\mu\nu}$ ,  $X := -\frac{1}{2}(\partial\phi)^2$  and  $V(\phi)$  is an arbitrary potential for the scalar field, which may include a mass term. Here  $G_i(\phi, X)$ ,  $i = 2, \dots, 5$  are arbitrary functions of their arguments. In this notation,  $\mathcal{L}_1$  is the standard Lagrangian corresponding to GR coupled to a scalar field with potential  $V(\phi)$ , and  $\mathcal{L}_i$ ,  $i = 2, \dots, 5$ , can be interpreted as higher-derivative corrections to GR in the matter sector, in this case comprised solely by a scalar field  $\phi$ . The cubic Horndeski theories are given by setting  $G_4 = G_5 = 0$  in (1). The reason why we consider these theories in the present paper and in Paper I is because these theories are known to be well posed in the standard gauges used in numerical GR [37]. As we previously mentioned, the general case has also been shown to be well posed in [22,23], but in a modified version of the generalized harmonic coordinates. While here we only consider the cubic case for simplicity and convenience, one may expect that some of our conclusions hold for more general Horndeski theories.

In Paper I we studied gravitational collapse of a massless scalar in spherical symmetry in certain cubic Horndeski theories<sup>3</sup> given by the choices

$$G_2(\phi, X) = g_2 X^2, \quad (3)$$

$$G_3(\phi, X) = g_3 X, \quad (4)$$

where  $g_2$  and  $g_3$  are arbitrary dimensionful coupling constants that we can tune. This particular choice of  $G_2$  is well motivated by EFT, since it is the leading order correction to GR minimally coupled to a scalar field [38]. On the other hand, this choice of  $G_3$  is a matter of simplicity and convenience. Both choices have been extensively considered in the literature (see e.g., [39–42] and references therein). One of the main results of Paper I was to identify the region in the space of initial conditions and couplings such that the solution in the domain of dependence of the initial data surface remained in the weakly-coupled regime of the theory on and outside black hole horizons if any are present, see Sec. II for more details. This is relevant in the context of EFT to justify that one can consistently keep only the leading order terms beyond GR, i.e., Horndeski, and neglect the otherwise (presumably) infinite number of higher-derivative corrections. At the same time, it is consistent to treat the theory fully nonlinearly, as we do here.

In the present paper we consider the same theories as in Paper I, namely (3)–(4). For the initial data, we choose two boosted lumps of scalar field, with amplitudes chosen so that they quickly collapse into black holes, thus forming a

<sup>3</sup>Whilst in this paper the initial data was chosen to be spherically symmetric, our code did not assume spherical symmetry.

black hole binary. Whilst most of the scalar field is absorbed by the black holes during the initial collapse stage, a scalar cloud remains in their vicinity throughout the lifetime of the binary. This scalar cloud can interact with itself and with the black holes and, over sufficiently long times, give rise to interesting effects. Furthermore, since spacetime curvature can source the scalar field via the Einstein equations, it is conceivable that when the spacetime curvatures are large, i.e., in the merger phase of a binary, one can observe sizeable deviations from GR. We mostly consider massive scalar fields and we restrict ourselves to a choice of potential  $V(\phi) = \frac{1}{2}m^2\phi$ , where  $m$  is the mass of the scalar field. The reason is that the corresponding scalar cloud can remain in the vicinity of the black holes for longer and hence, there is a greater chance of producing larger deviations from GR. The initial separations and velocities of the scalar lumps are tuned so that the black holes that form describe an eccentric binary that merges in five orbits. As we shall see in Sec. III, eccentric binaries seem to be particularly well suited to detect small deviations from GR since the system enters the strong field regime in every close encounter of the binary and not only in the merger phase. It is not clear if circular binaries would also exhibit a build up of the deviations from GR during the inspiral phase and not only in the merger phase. If they do, then one should expect an even larger deviations from GR. We leave this interesting problem for future work.

Finally, we choose the coupling constants  $g_2$  and  $g_3$  such that on the initial data surface, the solution lies well inside the weakly-coupled regime and we monitor that the solution remains in this regime throughout the evolution; this is necessary to ensure the consistency of the truncated EFT.

Note that for the cubic Horndeski theories, the natural frame to consider is the Einstein frame. This would not be the case had we considered more general Horndeski theories such as  $\mathcal{L}_4$  [43]. On top of this, since massive scalar fields cannot propagate to the wave zone, the waveforms presented in Sec. III would look the same in the Einstein and Jordan frames, respectively.

The rest of the paper is organized as follows: In Sec. II we describe our methods, numerical techniques and construction of suitable initial data. Section III contains the main results of the paper. In Sec. III A, we present the waveforms computed in various cubic Horndeski theories and we compare them to the waveforms obtained in GR coupled to a scalar field. In Sec. III B we discuss the properties of the scalar cloud surrounding the black holes and in Sec. III C we show that in our simulations, the weak coupling conditions are satisfied throughout the evolution of the binaries. In Sec. III D we analyze the mismatch between GR and Horndeski. In Sec. IV we summarize the main results of the paper and we discuss future directions for research. The convergence tests are presented in Appendix.

We adopt the following notation: Greek letters ( $\mu, \nu, \rho, \dots$ ) to denote full spacetime indices and Latin letters ( $i, j, k, \dots$ ) for purely spatial indices. We adopt the mostly plus metric signature and  $G = c = 1$ . Mass scales are with respect to the Arnowitt-Deser-Misner (ADM) mass of the spacetime.

## II. METHODS

### A. Equations of motion

The equations of motion for the theories that we consider in this article are given by Eqs. (2.2)–(2.3) in Paper I. In our numerical implementation, we write these equations in the usual  $(3 + 1)$ -conformal decomposition form and use the CCZ4 formulation of the Einstein equations [44,45] (see also [46]). We use the replacement  $\kappa_1 \rightarrow \kappa_1/\alpha$  and fix the constraint damping parameters to  $\kappa_1 = 0.1$ ,  $\kappa_2 = 0$ ,  $\kappa_3 = 1$ .<sup>4</sup> The matter contribution to the Einstein equations as well as the scalar field equations of motion written in the  $3 + 1$  form can be found in Appendix of Paper I.

### B. Initial data

For initial data, we consider the superposition of two equal boosted scalar field bubbles, similarly to what was done in [48]. Each individual scalar bubble is spherically symmetric and has some ingoing momentum, which prompts a quick collapse into a black hole without any outgoing scalar wave while leaving some leftover dynamical scalar hair surrounding the black hole. Whilst the individual scalar bubbles satisfy the Hamiltonian and momentum constraints, the superposition of the two does not; however, we place them sufficiently far apart so that the initial constraint violations due superposing the two scalar bubbles are sufficiently small. Further details and explicit form of the constraint satisfying scalar bubble profiles used can be found in Paper I.

For binary systems, we boost the individual profiles with a Galilean boost with velocity  $\vec{v}$ . This can be implemented in the scalar momentum by adding to it (with the appropriate sign) the Galilean boost<sup>5</sup>

$$\Pi(t, \vec{x})|_{t=0} = \Pi_{\text{original}}(t, \vec{x})|_{t=0} - \frac{1}{\alpha} \vec{v} \cdot \vec{\nabla} \phi(t, \vec{x})|_{t=0}, \quad (5)$$

<sup>4</sup>This is the default in the code used, namely GRChombo, and it greatly improves the stability of black hole spacetimes, see [47] for detailed studies.

<sup>5</sup>This follows simply from a transformation from unboosted to boosted coordinates  $(t', \vec{x}') \rightarrow (t = t', \vec{x} = \vec{x}' + \vec{v}t')$ . With the usual decomposition of the unit normal into the lapse function and shift vector,  $n^\mu = (\frac{1}{\alpha}, -\frac{\beta^i}{\alpha})$ , then  $\Pi = n^\mu \nabla_\mu \phi = \frac{1}{\alpha} (\frac{\partial}{\partial t} - \beta^i \frac{\partial}{\partial x^i}) \phi$ . The shift transforms as a vector under Galilean boosts, whereas  $\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - v^i \frac{\partial}{\partial x'^i}$ , resulting in Eq. (5).

where  $\alpha$  is the lapse function and  $\Pi = \mathcal{L}_n \phi = n^\mu \nabla_\mu \phi$  is the scalar momentum, where  $n^\mu$  is the unit normal 4-vector to spatial hypersurfaces. Such a boost is valid for small velocities and avoids the obstacle of having to evaluate the initial data at different times for distinct points, as one would need to do to implement a proper Lorentz boost. Doing the latter would be unfeasible for nonstatic initial data.

To superpose the initial data of two scalar bubbles,  $A$  and  $B$ , boosted in opposite directions, we use

$$\begin{aligned} \psi &= \psi_A + \psi_B - 1, \\ K_{ij} &= \gamma_{m(i} (K_{j)n}^A \gamma_A^{nm} + K_{j)n}^B \gamma_B^{nm}), \\ \phi &= \phi_A + \phi_B, \\ \Pi &= \Pi_A + \Pi_B, \end{aligned} \quad (6)$$

where  $\psi$  is the conformal factor associated with the induced metric on the initial data surface,  $\gamma_{ij}$ , so that  $\gamma_{ij} = \psi^4 \tilde{\gamma}_{ij}$  and  $\tilde{\gamma}_{ij}$  is flat. The initial data corresponding to each of the individual scalar bubbles is conformally flat, so  $\tilde{\gamma}_{ij} = \tilde{\gamma}_{ij}^A = \tilde{\gamma}_{ij}^B = \delta_{ij}$ , and satisfies  $K^A = K^B = 0$ . For  $K = 0$  and a conformally flat spatial metric, the condition for  $K_{ij}$  reduces to  $\tilde{A}_{ij} = \tilde{A}_{ij}^A + \tilde{A}_{ij}^B$ , where  $\tilde{A}_{ij}$  is the conformal traceless part of the extrinsic curvature given by  $\tilde{A}_{ij} = \psi^{-4} K_{ij}$ . In our simulations, we initialize the lapse and shift vector as  $\alpha = 1$  and  $\beta^i = 0$ .

Motivated by the similarity with previous studies of binary black holes mergers in scalar field environments, we also tried an alternative way of constructing initial data by superposing scalar bubbles as suggested in [49]. The construction of [49] produces initial data which is physically different from (6) but, for the large separations between the initial scalar bubbles as in our paper, the amounts of initial constraint violations are comparable. Furthermore, the evolution of this initial data leads to qualitatively similar results after one orbit. Therefore, henceforth we will only discuss the evolution of the binaries constructed using (6).

Using the notation of Paper I, we use initial scalar bubbles with a profile

$$\phi(r) = A \left( \frac{r^2}{r_0^2 + 2\omega^2} \right) e^{-\frac{1}{2} \left( \frac{r-r_0}{\omega} \right)^2}, \quad (7)$$

where  $r$  is the Euclidean distance from the center of the bubble,  $(A, r_0, \omega) = (0.21, 5, \sqrt{\frac{1}{2}})$  and a scalar mass parameter  $m = 0.5$ . For an isolated scalar bubble, this configuration has an ADM mass of approximately  $M \approx 0.53$ . The mass term in the scalar potential accounts for about 10% of the total ADM mass, while the contribution of the Horndeski terms for a coupling of  $g_2 = 0.02$  or  $g_3 = 0.05$  is of order  $\mathcal{O}(10^{-5})$ .

The binaries with the eccentric orbits presented in this paper are obtained by choosing a “large” initial separation<sup>6</sup> between the centers of the scalar bubbles of  $D = 40$  and an individual scalar boost velocity of  $|\vec{v}| = 0.17$ ; after the initial gravitational collapse, the resulting black holes have an initial velocity of 0.042. We calculate numerically that the superposed data has a total ADM mass of  $M = 1.0346 \pm 0.0001$ . The values quoted above are in code units, and the mass parameter  $m$  and couplings  $g_2, g_3$  will be referred without units henceforth (e.g.,  $g_2 = 0.02$  as opposed to  $g_2 \sim 0.0214 M^{-2}$  after taking into account that the total mass is  $M = 1.0346$ ).

### C. Weak coupling conditions

In order for the Horndeski theories to be valid EFTs, the basic requirement is that corrections to the two-derivative GR terms in the equations of motion on and outside black hole horizons (if there are any in the spacetime) should be small at all times. Inside black holes both GR and the Horndeski theories break down but classically this region is inaccessible to external observers. Therefore, in practice, we only monitor the weak coupling conditions (WCCs) on and outside black hole horizons.<sup>7</sup>

For the cases considered in this paper, the WCCs translate into the requirement (see Paper I for further details),

$$|g_2 L^{-2}| \ll 1, \quad |g_3 L^{-2}| \ll 1, \quad (8)$$

where  $L$  is the typical length scale estimate for the system. This can be computed as

$$L^{-1} = \max\{|R_{\alpha\beta\mu\nu}|^{\frac{1}{2}}, |\nabla_\mu \phi|, |\nabla_\mu \nabla_\nu \phi|^{\frac{1}{2}}, m\}, \quad (9)$$

where  $R_{\alpha\beta\mu\nu}$  is the spacetime Riemann tensor and  $m$  is the mass parameter in the scalar field potential.<sup>8</sup> We consider values for the couplings  $g_2$  and  $g_3$  based on the valid values of  $\eta_2$  and  $\eta_3$  displayed in Figs. 1 and 10 of Paper I. In practice, for massive scalar fields, the scalar clouds that form near black holes tend to be more extended, have lower densities and smaller gradients than in the massless case, thus allowing for larger couplings without violating the WCCs that may lead to a loss of the hyperbolicity of the scalar equations.

<sup>6</sup>This “large” initial separation makes a circular binary unfeasible with our computational resources, but it helps to minimize the errors from the initial data superposition.

<sup>7</sup>While the weak cosmic censorship conjecture [50,51] remains unproven in the astrophysical settings considered in this paper, we will assume that it holds. We do not find any evidence against it in our setting.

<sup>8</sup>It turns out that for our choice of couplings and scalar mass parameter  $m$ , near the black holes the contribution of  $m$  to  $L$  is always smaller than the metric and scalar curvature terms.

### D. Excision

To control the gradients in the interior of black holes and avoid issues such that a potential loss of hyperbolicity of the scalar equations there, we implement the same form of excision as in Paper I. More precisely, inside black holes we modify the equations of motion by smoothly switching off the Horndeski terms (see Appendix C of Paper I for the details of our implementation). This procedure seems justified since the EFT is no longer valid in these regions of the spacetime. We only implement this form of excision during the initial stages of the evolution, while  $t \lesssim 40M$ , namely during gravitational collapse and gauge readjustment phases.<sup>9</sup> Once the black hole has stabilized, the matter density at its center is very small and no loss of hyperbolicity of the scalar equations occurs. Note that as long as the weak coupling conditions hold, the sound horizons should be close to the usual metric horizon and therefore excising a small enough region well inside the metric horizon should not affect the physics in the domain of outer communications.

This procedure of excision is not necessary with the resolutions that we used in this paper if the coupling is sufficiently small. However, for arbitrarily high resolutions, the scalar field gradients near the center of the black holes would grow unbounded and one would expect the code to break down.

### E. Gravitational wave extraction

We extract gravitational waves at finite radii by projecting the Weyl scalar  $\Psi_4$  onto the spin-weighted spherical harmonics on multiple spheres of fixed coordinate radius in the standard way, obtaining the multipoles  $\psi_{\ell m}$ . Data on each integration sphere is obtained from the finest available level in the numerical grid at a given extraction radius, using fourth-order Lagrange interpolation. We calculate the Weyl scalar  $\Psi_4$  using the Newman-Penrose formalism [52] and the electric and magnetic parts of the Weyl tensor,  $E_{ij}$  and  $B_{ij}$  [53]. The latter can be computed from our evolution variables using the following expressions adapted to the (3 + 1)-CCZ4 formulation of the Einstein equations,

$$E_{ij} = [R_{ij} + D_{(i}\Theta_{j)} + (K - \Theta)K_{ij} - K_{im}K^m_j - \frac{\kappa}{4}S_{ij}]^{\text{TF}}, \quad (10)$$

$$B_{ij} = \epsilon_{mn(i}D^m K_{j)}^n, \quad (11)$$

where  $R_{ij}$  is the three-dimensional Ricci tensor,  $K_{ij}$  the extrinsic curvature,  $\Theta := -n_\mu Z^\mu$  is the projection of the Z4 vector,  $Z^\mu$ ,<sup>10</sup> onto the timelike unit normal vector  $n^\mu$ ,  $S_{ij} := \gamma_i^\mu \gamma_j^\nu T_{\mu\nu}$  is the spatial projection of the stress-energy tensor,  $\gamma_{\mu\nu} := g_{\mu\nu} + n_\mu n_\nu$  is the induced metric on the

<sup>9</sup>The initial data is not in puncture gauge, so in the initial stages of the simulation there is a certain amount of artificial dynamics in grid variables due to the adjustment of the coordinates.

spatial hypersurfaces,  $\epsilon_{\mu\nu\rho} = n^\sigma \epsilon_{\sigma\mu\nu\rho}$  is the volume form on such hypersurfaces and  $[\cdot]^{\text{TF}}$  denotes the trace-free part of the expression in square brackets. Note that Eqs. (10)–(11) guarantee that  $E_{ij}$  and  $B_{ij}$  are automatically trace free and symmetric, unlike usual (3 + 1)-ADM expressions [53], which require that the constraints are satisfied.

### F. Gravitational strain

The natural observable measured in detectors and used when constructing templates is the gravitational strain. In the conventions of [53], the strain of a gravitational wave,  $h$ , can be obtained from the  $\Psi_4$  Weyl scalar using the transformation [54]

$$\Psi_4 = -\ddot{h} = -\ddot{h}^+ + i\ddot{h}^\times, \quad (12)$$

where the dot  $\dot{\cdot}$  denotes a time derivative, and  $h^+$  and  $h^\times$  are the usual plus and cross polarizations of the wave. This gives the strain multipoles  $\ddot{h}_{\ell m}^+ = -\text{Re}(\psi_{\ell m})$  and  $\ddot{h}_{\ell m}^\times = \text{Im}(\psi_{\ell m})$ , where  $\psi_{\ell m}$  are the amplitudes of each mode in the multipolar decomposition of the Weyl scalar  $\Psi_4$ . To avoid artifacts from finite length of the wave, discrete sampling and noisy data, we perform the double time integration in the frequency domain using a fixed frequency filter [55], with a cutoff of  $0.01 \text{ M}^{-1}$  for low frequencies.<sup>11</sup> We taper the signal in the time domain with a Tukey window [56] of width  $\sim 40M$  on each side<sup>12</sup> and zero-pad to the nearest power of two. We further zero-pad the waveform to increase the length by a factor of eight before applying the fast Fourier transform [57], in order to increase the frequency resolution of the discrete Fourier transform and reduce the noise for low frequencies introduced by the discretization. Removing the initial junk radiation of the inspiral from the time domain did not result in any significant improvement. Finally, the signal at null infinity can be obtained by extrapolating the results at finite radii assuming a Taylor series expansion in  $\frac{1}{r^*}$  [58], where  $r^* = r + 2M \log|\frac{r}{2M} - 1|$  is the tortoise radius, after first aligning separate extraction radii in retarded time,  $u = t - r^*$ .

### G. Waveform mismatch

In order to estimate the difference between two waveforms, we compute the mismatch between the strain resulting from each wave. First, given the strain of two

<sup>10</sup>The components of the Z4 vector essentially correspond to the Hamiltonian and momentum constraints, assumed to be zero in the BSSN formulation.

<sup>11</sup>This choice affects the noise in the strain, but low frequencies have negligible effects in the final computation of the mismatch. Furthermore, adding a cutoff for the high frequencies resulted in no improvement.

<sup>12</sup>This choice reduces noise, but it does not affect the results in any meaningful way, as the signal is essentially zero in this region.

waves,  $h_1(t)$  and  $h_2(t)$ , one can compute the overlap,  $O$ , using the frequency domain inner product [59–61],

$$O(h_1, h_2) = \frac{\text{Re}(\langle h_1, h_2 \rangle)}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}}, \quad (13)$$

$$\langle h_1, h_2 \rangle = 4 \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_1^*(f) \tilde{h}_2(f)}{S_n(f)} df, \quad (14)$$

where  $\tilde{h}(f)$  denotes the Fourier transform of the function  $h(t)$ ,  $*$  denotes complex conjugation,  $S_n(f)$  is the power spectral density (PSD) of a detector's strain noise as a function of frequency  $f$  (e.g., updated Advanced LIGO sensitivity design curve [62]),  $f_{\min}$  and  $f_{\max}$  is the lowest and highest frequency cutoffs of the PSD of the detector or the frequency minimum/maximum imposed by the time-step and duration of the simulation. Notice that for  $h_1 = h_2$ ,  $\langle h_1, h_2 \rangle$  is real.

Then, we compute the mismatch by maximizing the overlap over time and phase shifts of the second wave,  $h_2^{\delta t, \phi}(t) = h_2(t + \delta t)e^{i\phi}$ ,

$$\text{mismatch} = 1 - \max_{\delta t, \phi} O(h_1, h_2^{\delta t, \phi}). \quad (15)$$

Noticing that  $\tilde{h}_2^{\delta t, \phi}(f) = \tilde{h}_2(f)e^{i\phi}e^{2\pi i\delta t}$  and

$$O(h_1, h_2^{\delta t, \phi}) = \frac{\text{Re}(e^{i\phi}e^{2\pi i\delta t} \langle h_1, h_2 \rangle)}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}}, \quad (16)$$

then maximizing over phase shifts corresponds to simply taking the absolute value of  $\langle h_1, h_2 \rangle$  as opposed to the real part. Maximizing over time shifts is more subtle because the discrete domain implies the Fourier transform changes by more than a mere phase  $e^{2\pi i\delta t}$ . Hence, we perform the time shift maximization numerically. To allow for continuous time shifts, we interpolate the data  $h_{1,2}(t)$  and re-sample appropriately after the time shift. The number of points when resampling the time series does not affect the final result. When comparing two gravitational waves, the length of the time interval used for the Fourier transform and the size of the frequency domains used for integration are the same for all waves (taking into account time shifts). All in all

$$\text{mismatch} = 1 - \frac{\max_{\delta t} |\langle h_1, h_2^{\delta t} \rangle|}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}}. \quad (17)$$

## H. Numerical scheme

The equations of motion are evolved with GRChombo [47,54,63,64], using MPI (Message Passing Interface), OpenMP and templated vector intrinsics/SIMD (Single Instruction, Multiple Data) to obtain a good performance

in the most common architectures. GRChombo uses the Chombo adaptive mesh refinement libraries [65]. We use a tagging criterion that triggers the regridding based on second derivatives of both the scalar field and the conformal factor. GRChombo implements the usual puncture gauge (i.e.,  $1 + \log$  slicing plus Gamma driver [66–68]) for the evolution of the gauge variables and  $N = 3$  Kreiss-Oliger (KO) dissipation, with fixed  $\sigma = 1$  in all our simulations. As of boundary conditions, we use Sommerfeld boundary conditions and take advantage of the reflective/bitant symmetry of the binary problem to evolve only half of the grid. Sixth order spatial stencils are used in order to improve phase accuracy of the binaries [69]. Time updates are still made with a fourth-order Runge-Kutta scheme, which implies that the global convergence order cannot be higher than four.<sup>13</sup> For the results presented in this paper, we have a Courant-Friedrichs-Lewy factor of  $1/4$ , a coarsest level resolution of  $\Delta x = \frac{16}{7}$ , with eight additional refinement levels, and a computational domain of size  $1024^3$ .

## III. RESULTS

In this section we present the results of our numerical simulations for the various Horndeski theories that we have considered and we compare them to GR. To carry out the comparison, we consider standard GR coupled to a massive scalar field (with same mass parameter  $m = 0.5$  as in Horndeski). We comment on the massless scalar field case in Sec. III B. We have also considered the evolution of a black hole binary in vacuum GR with the same total ADM mass and initial velocities for the black holes. In this case, the binary describes many more orbits before merger, as expected, since no energy is transferred to the scalar field. We will not comment any further on this case since it is not relevant for the kind of comparisons that we carry out.

We have constructed superposed initial data for GR coupled to a massive scalar field and for Horndeski theories. One could question whether different results arise from small differences in the initial data. As discussed in Paper I, the effect of the Horndeski terms in the initial data is proportional to  $\frac{g_2 A^2}{r_0^2}$  and  $\frac{g_3 A w^2}{r_0^4}$  depending on the theory. For the values of the couplings  $g_2$  and  $g_3$  that we consider, this results in a difference of order  $\mathcal{O}(10^{-5})$  between the Horndeski and the GR counterpart. To confirm that the small Horndeski corrections in the initial data do not affect the subsequent evolution, we evolved the equations of motion of the Horndeski theories using initial data constructed for GR. Clearly this procedure introduces extra initial constraint violations proportional to the Horndeski couplings. However, our results from the Horndeski

<sup>13</sup>Notice that this allows us to use the usual KO dissipation stencils that are commonly implemented with fourth-order finite differences.

theories initialized with GR initial data and those results obtained using proper Horndeski initial data do not exhibit any significant or quantitative difference. Therefore, we conclude that the differences observed between GR and Horndeski theories are caused by the evolution with distinct evolution equations and not by the extremely small differences in the initial data. Henceforth, for the Horndeski theories we will only present results obtained with Horndeski initial data.

In Fig. 1 we display the trajectories of the punctures on the orbital plane for GR and for a Horndeski theory with  $g_2 = 0.02$ . This figure shows that after the first close encounter of the binary, the trajectories that the black holes follow in GR and in Horndeski are visibly different. Interestingly, the black holes seem to recombine to the same trajectory in the final stages of binary. In the following subsections we will quantify the differences in other observables such as the gravitational strain.

### A. Waveform strain

In this subsection we compare the waveform strain for eccentric binaries in GR and in different Horndeski theories. For the latter, we consider both the  $G_2$  and the  $G_3$  theories for different values and signs of the coupling

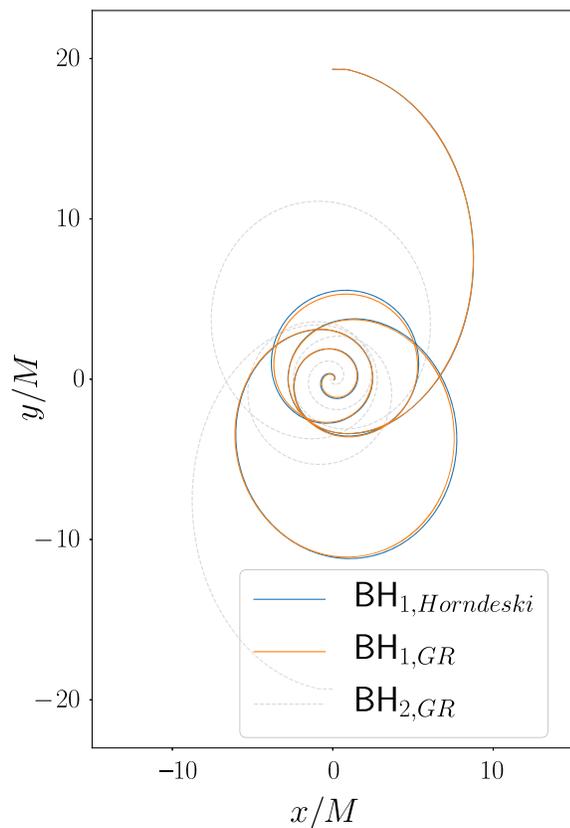


FIG. 1. Orbits of the two black holes in Horndeski for  $g_2 = 0.02$  and GR. For clarity, in the Horndeski case we only show one of the black holes.

constants. In Figs. 2 and 3 we present the  $(\ell, m) = (2, 2)$  mode of the plus polarization of the strain,  $h^+$ , extrapolated to null infinity using six radii between  $50M - 150M$ , for the  $G_2$  and  $G_3$  theories respectively. The strain for higher  $(\ell, m)$  modes exhibits qualitatively similar features.

Referring to Figs. 2 and 3, the two peaks that can be seen at  $t \sim 400M$  and  $t \sim 850M$  correspond to the bursts of radiation emitted during the first two close encounters of the eccentric binary<sup>14</sup> before the final merger phase. The latter starts at around  $t \sim 1100M$  and ends by  $t \sim 1200M$ , depending on the theory and the value and sign of the coupling constants. As for the final state, to the best of our knowledge, it is not known if the class of theories that we consider admit stationary hairy black holes.<sup>15</sup> However, we find evidence that the end state of the evolution is a Kerr black hole surrounded by a scalar cloud that decays in time on a time scale much longer than that set by the initial ADM mass. This evidence is reinforced by the similarity between the final scalar cloud profiles between GR and Horndeski, as described in Sec. III B. For the runs shown in Figs. 2 and 3, the estimated parameters of the final black holes are summarized in Table I. Note that any junk radiation caused by the initial constraint violations or choice of initial data is very small on the scale of these figures, but still visible in the first  $\sim 50M$ .

As Figs. 2 and 3 show, the waveforms obtained in GR and in the various Horndeski theories that we considered, coincide during the initial stages of the binary, but a clear misalignment builds up over time, starting from the second close encounter of the binary and becoming more pronounced in the merger phase. This misalignment is much larger than the smallness parameter controlling the weak coupling conditions of the initial data.<sup>16</sup> In Sec. III C we will provide evidence showing that a suitable local weak-coupling condition remains small during the whole evolution of the binary and hence, in our setting, the Horndeski theories should be valid (and predictive) classical EFTs. The large misalignment that we observe in Figs. 2 and 3 is a cumulative effect arising from the locally small differences between GR and Horndeski, and it gets enhanced whenever the system enters the strong field regime, which happens in each close encounter of the eccentric binary and in the merger phase. This is expected since the corrections to GR are sourced by spacetime and scalar curvature and those

<sup>14</sup>It may be useful for the reader to match the gravitational wave signal in these figures with the visual animation of one of our simulations; <https://www.youtube.com/watch?v=uOed4AG1uIg>.

<sup>15</sup>For the theories considered in this paper and in the shift-symmetric case, Refs. [70–73] have proven that no slowly rotating hairy black holes exist, but in our case the mass term explicitly breaks the scalar shift symmetry in the equations of motion.

<sup>16</sup>Here, we refer to the dimensionless parameters  $\eta_2$  and  $\eta_3$  defined in Paper I Eqs. (2.8)–(2.9), which are  $\mathcal{O}(10^{-5})$ , as previously mentioned in Sec. II B.

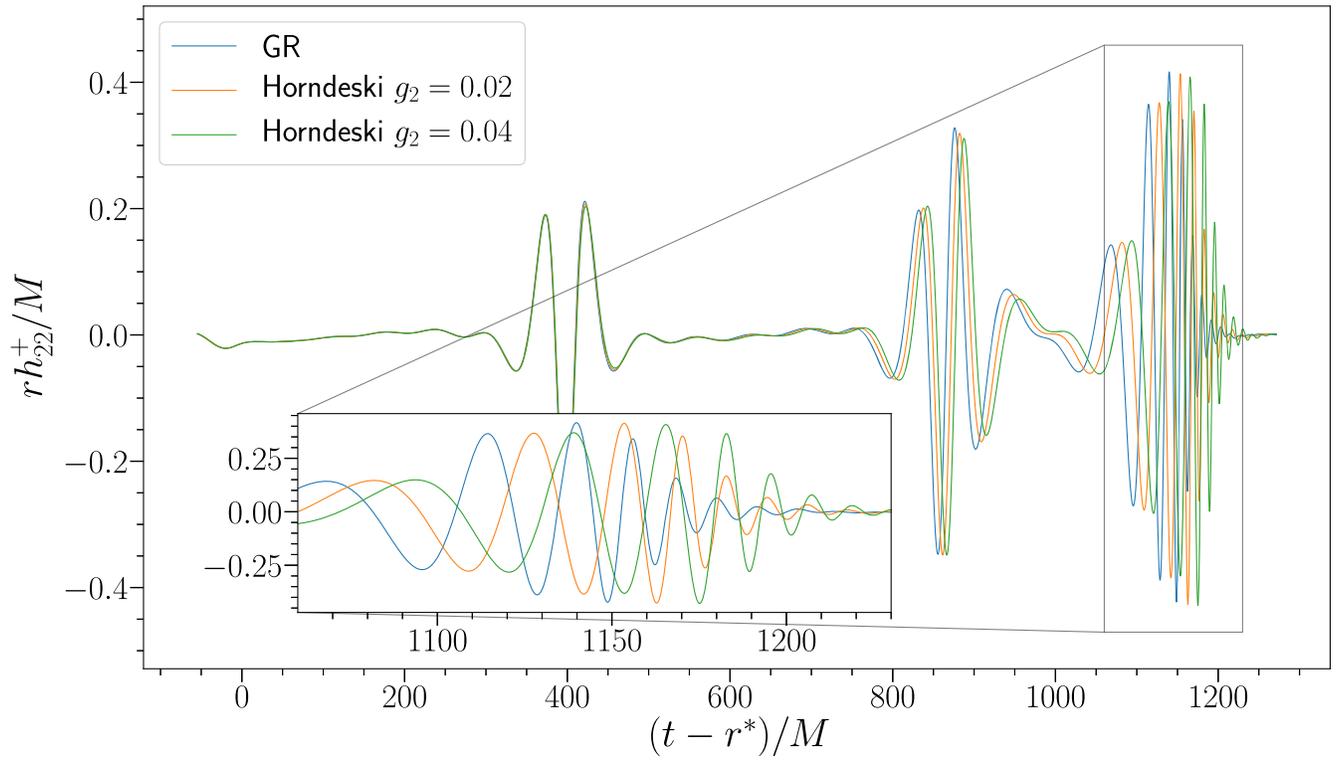


FIG. 2. Comparison of gravitational wave between Horndeski theory with  $g_2 = 0.02$ ,  $g_2 = 0.04$  and GR in retarded time,  $u = t - r^*$ , where  $r^*$  is the tortoise radius. Displaying the  $(\ell, m) = (2, 2)$  mode of the plus polarization of the strain,  $h_{22}^+$ , extrapolated to null infinity. There is a visible misalignment between GR and Horndeski that builds up over time, becoming larger during the merger phase.

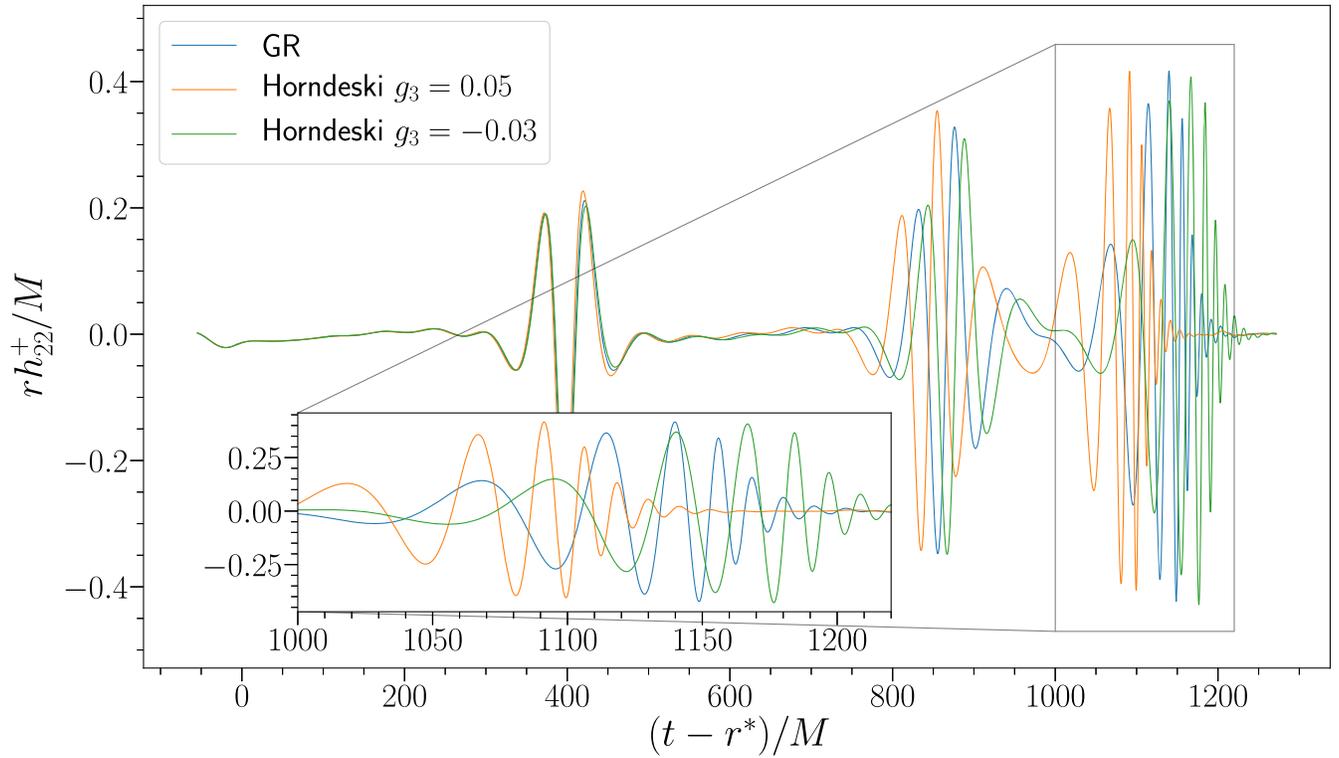


FIG. 3. Comparison of gravitational wave between GR and Horndeski theories with  $g_3 = 0.05$  and  $g_3 = -0.03$ , in retarded time,  $u = t - r^*$ , with  $r^*$  the tortoise radius. Displaying the  $(\ell, m) = (2, 2)$  mode of the plus polarization of the strain,  $h_{22}^+$ , extrapolated to null infinity. As in the  $G_2$  theory, Fig. 2, we observe a misalignment that builds up over time.

become more important precisely in the strong field regime. Therefore, eccentric binaries seem to be useful to potentially detect deviations from GR sourced by curvature through the built up of small cumulative effects and their enhancement in the close encounters. It is conceivable that linearizing the Horndeski theories around GR may allow one to compute some of the misalignment (at least for some small enough couplings) during the merger phase since its duration is relatively short and secular effects may not be an issue. However, it seems unlikely that such an approach would be able to capture the cumulative large deviations that arise from successive close encounters of an eccentric binary, such as in the examples considered here. The relatively long times that we have evolved the binaries require a full nonperturbative treatment of the theory to avoid potential secular effects.

For the  $G_2$  theory (3), a positive  $g_2$  coupling induces a delay of the waveform when compared to GR, whilst a negative  $g_2$  gives rise to an advancement of the signal. On the other hand, for the  $G_3$  theory (4) the effect is the opposite; a positive  $g_3$  coupling leads to an advancement of the signal while a negative  $g_3$  leads to a delay when compare to the GR waveform. In general, the observed misalignment between GR and Horndeski seems to be a rather generic effect that does not depend on the specifics of the theory. Of course, the details such as the amount or the sign of the deviations will depend on the details theory under consideration. Therefore, we are tempted to conjecture that gravitational strain computed in general Horndeski theories of gravity that do not admit equilibrium hairy black holes but with dynamical long-lived scalar clouds surrounding black holes will be misaligned with respect of the GR signals. Finally we note that the peak amplitude of the waveforms seems to be very similar across all theories and couplings. We will point out in the Sec. IV how this misalignment may be potentially detected in gravitational wave observations.

Note that the final state of GR and Horndeski simulations for  $g_2 \leq 0.04$  seems to have the same exact mass and only tiny differences in spin (see Table I). This is counterintuitive given the differences that the gravitational waveforms exhibit and it could be related to the fact that the trajectories of the black holes in the two theories visibly differ in the intermediate stages of the binary, but coincide again near the merger phase (see Fig. 1). The physical mechanisms behind this observation may be related to the frequency shifts analyzed in Sec. III D. The fact that the initial and final state coincide and yet the waveforms are different indicates that, at least for equal mass nonspinning binaries, the degeneracy between the class of Horndeski theories that we have considered and GR is broken. This suggests that the degeneracy between GR and Horndeski may also be broken for unequal mass nonspinning configurations. It would be interesting to study the effects of the intrinsic spins in alternative theories of gravity.

TABLE I. Parameters of the final state Kerr black hole for each coupling  $g_2$  and  $g_3$ . The mass and spin are estimated from the final apparent horizon since the ADM quantities are typically significantly noisier. The errors are estimated from the differences between the medium and high resolution runs.

Coupling	Final mass $M_F/M$	Spin parameter $a/M$
GR	$0.973 \pm 0.001$	$0.676 \pm 0.001$
$g_2 = 0.005$	$0.973 \pm 0.001$	$0.676 \pm 0.001$
$g_2 = 0.02$	$0.973 \pm 0.001$	$0.675 \pm 0.001$
$g_2 = 0.04$	$0.973 \pm 0.001$	$0.673 \pm 0.001$
$g_3 = 0.05$	$0.975 \pm 0.001$	$0.680 \pm 0.001$
$g_3 = -0.03$	$0.972 \pm 0.001$	$0.672 \pm 0.001$

When comparing the waveforms between different theories, one might alternatively want to align the main peaks. However, clearly the misalignment would not disappear; it would simply be translated along the time axis. This can be seen in Fig. 4, where the misalignment is now seen at the early encounters of the inspiral. This shows that the gradual phase shift is a physical effect that cannot be ignored by a constant phase shift and does not depend on how one does the comparison. For long-lived inspirals beyond the strong field regime simulated with numerical relativity, the effect would be enhanced and the misalignment would be present regardless of the time or phase shift considered.

## B. Scalar cloud

In Fig. 5 we display a snapshot of a binary for the  $G_2$  theory with  $g_2 = 0.005$  at a representative instant of time before the merger. This figure shows that the energy density of the scalar field (in blue) is localized in the region near the black holes, being largest near the horizons. It is in this regions where the spacetime and scalar field curvatures are largest, even though the WCCs remain small on and outside the black holes.

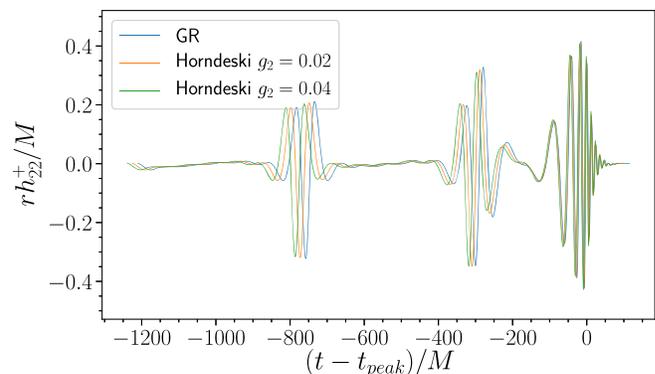


FIG. 4. Comparison of realigned  $h_{22}^+$  between Horndeski theory with  $g_2 = 0.02$ ,  $g_2 = 0.04$  and GR. The waves were aligned so the peak of the amplitude of the complex strain coincides.

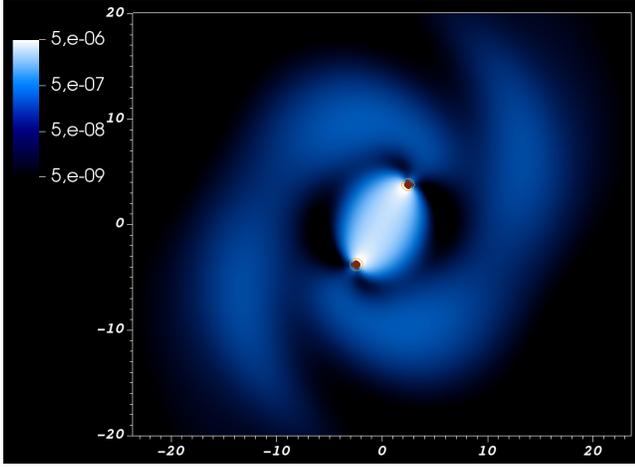


FIG. 5. Energy density (in blue) of the scalar field surrounding the binary black holes for the Horndeski theory with  $g_2 = 0.005$  at a representative instant of time during the inspiral phase,  $t \sim 450$ . The apparent horizon of the black holes is shown in orange. The region where the weak-coupling conditions are larger than one is depicted in brown. Clearly this region is contained well inside the apparent horizon, as required.

For the Horndeski theories that we considered, the accumulation of nonlinear effects is possible due to the presence of long lived scalar cloud surrounding the black holes. This scalar cloud survives all the way up to and well beyond the merger. This is due to the presence of a mass term in the scalar potential, since it is well-known (see e.g., [74–79]) that the effective potential that the scalar field “sees” has a wall that makes it difficult for it to escape to infinity. A scalar mass parameter of  $m = 0.5$  is comparable to what has been seen to give rise to long-lived scalar clouds [79], though the effects observed in this paper did not require any fine tuning. We have also carried out simulations of binaries with massless scalars and the absence of a significant scalar cloud trivially removes any long term effects of the scalar field on the evolution.

In Fig. 6 we display the evolution of the total energy of the scalar field  $E_\phi$  and the evolution of the maximum of the energy density  $\rho$  for GR (green), for the  $G_2$  theory with  $g_2 = 0.005$  (red) and for the same Horndeski theory but with a massless scalar field (blue). These quantities are defined as  $\rho = n^\mu n^\nu T_{\mu\nu}$ , where  $T_{\mu\nu}$  the energy-momentum tensor of the scalar field, and  $E_\phi = \int_V \rho \sqrt{\gamma} dV$ , where  $V$  is the spatial volume on a given spacelike hypersurface. After the initial gravitational collapse, most of the scalar field is absorbed by the black holes, but in the massive scalar cases, a long-lived scalar cloud forms around the black holes. After the first close encounter of the eccentric binary ( $t \sim 400M$ ), the maximum energy density of the scalar cloud is of order  $10^{-5}M^{-2}$  for the massive scalar cases (GR and Horndeski), and it decreases very slowly with time. This long-lived cloud makes it possible for the scalar field to interact with itself and with the geometry and give rise to

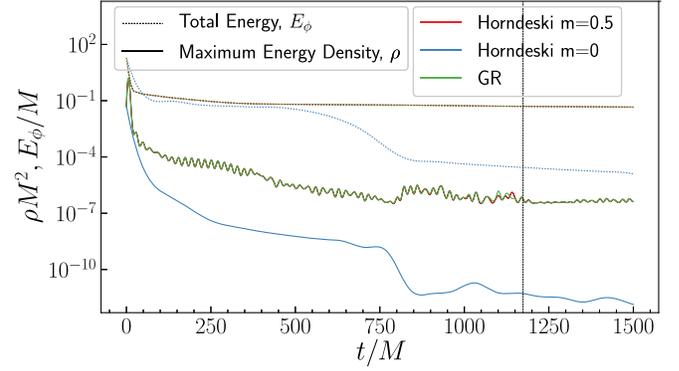


FIG. 6. Total energy of the scalar field  $E_\phi$  and maximum value of energy density  $\rho$  on the spacetime (excluding black holes, by removing from the volume of integration the interior of apparent horizons) for Horndeski with  $g_2 = 0.005$  and a massive scalar field (in red), the same theory with a massless field (in blue) and GR with a massive scalar field (in green). A dashed black line is used to indicate the estimated merger time for the Horndeski run with a massive scalar field.

the build up of significant differences in the physical observables such as the gravitational strain.

On the other hand, in the massless case, Fig. 6 shows that a much larger amount of scalar field is absorbed by black holes during the collapse phase. Furthermore, both the total energy of the scalar field and its energy density show a pronounced dip at  $t \sim 850M$ , namely in the second close encounter of the binary, indicating that any leftover amount of scalar field in the vicinity of the black holes gets absorbed. Beyond this point, the energy density of the scalar field is less than  $10^{-10}M^{-2}$  while the total energy is of the order of  $10^{-5}M$  (corresponding to scalar waves radiated to infinity), and both continue to steadily decrease with time. By the time the merger takes place the maximum energy density of the scalar field is compatible with numerical error. Therefore, we conclude that in the massless case, after the second close encounter of the binary, there is basically no significant amount of scalar field left in the neighborhood of the black holes (as expected from the no-hair theorems of [70–73]) to give rise to any sizeable effect, at least for the duration of our simulations. As a consequence, no noticeable differences between GR and Horndeski are observed in the massless scalar field case.

Comparing Horndeski with GR in Fig. 6 shows that local differences (in time) in the energy density between GR and Horndeski for the massive scalar field are not significant for most of the binary, including the two close encounters; only during the merger phase one can see some differences of order  $10^{-5}M^{-2}$ . These results are expected if the weak coupling conditions are satisfied. Furthermore, the fact that the energy density of the massive scalar field around the black holes is small during the highly dynamical stages of the binary is necessary but not sufficient to ensure that the WCCs are satisfied.

### C. Weak coupling conditions during the evolution

The results reported in Sec. III A can only be trusted as long as the Horndeski theories that we consider are valid (truncated) EFTs. In this subsection we provide evidence that for the initial data and couplings that we considered in the paper, the local WCCs (8) are satisfied at all times, thus ensuring the predictivity of the EFTs.

In Fig. 7 we display the  $L^2$  norm of the WCCs (8) integrated on the black holes' apparent horizons, as a function of time for an eccentric binary evolved with the  $G_2$  theory with different values of the coupling constant  $g_2$  and one value  $g_3$  coupling for the  $G_3$  theory. Excluding the interior of black holes, the apparent horizons are where the WCCs have the largest values in the whole domain. This plot shows that the weak-coupling condition (8) remains approximately constant during the evolution, except in the close encounters of the binary and the final merger phase. The latter events correspond to the peaks in Fig. 7 that can be seen at  $t \sim 400M$ ,  $t \sim 850M$ , and  $t \sim 1100M$ , when the system enters the strong field regime. The constancy of (8) during the inspiral phase is related to the fact the energy density of the scalar field in the vicinity of the black hole remains approximately constant during this phase. The fact that the WCCs exhibits local maxima at the close encounters indicates that in an eccentric binary, we probe the strong field regime during various phases of the binary and not only near and during the merger phase as in a circular binary. It is interesting to see that when one normalizes the WCCs (8) by the coupling constant, the curves for the  $g_2$  couplings collapse onto a single curve, except in regions where the system is in the strong field regime. This indicates that the WCC depends on the coupling constant in a trivial way when the system is not in the strong field regime.

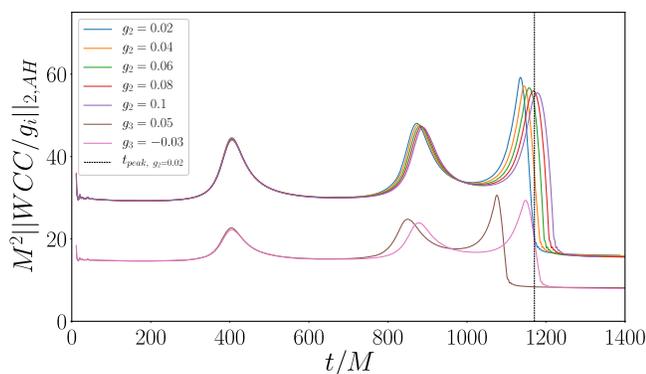


FIG. 7.  $L^2$  norm of the weak coupling conditions (8) integrated over the apparent horizon and normalized by the coupling constant  $g_2$  or  $g_3$ . For the binary that we have evolved, this shows that  $M^2 \text{WCC}/|g_2| \lesssim 50$  and  $M^2 \text{WCC}/|g_3| \lesssim 20$ , which in turn implies that  $|g_2| \lesssim 0.02$  and  $|g_3| \lesssim 0.05$  to guarantee that the WCC (8) is roughly less than one. The dashed black line corresponds to the peak of the amplitude of the strain for  $g_2 = 0.02$ .

Fig. 7 shows that for our choice of initial data,  $M^2 \text{WCC}/|g_2| \lesssim 50$  and  $M^2 \text{WCC}/|g_3| \lesssim 20$  at all times. This implies that if we want the WCCs (8) to be roughly less than one at all times, and hence guarantee that the Horndeski theory is a valid EFT throughout the evolution, then one must choose  $|g_2| \lesssim 0.02$  or  $|g_3| \lesssim 0.05$ . For values larger than these, the WCCs become comfortably larger than one at different (or all) stages of the binary. However, after the initial collapse stage, even for large values of the coupling  $g_2$  well beyond the regime of validity of EFT (e.g.,  $g_2 = 0.1$ ), the equations of motion of the scalar field remain hyperbolic throughout the inspiral and merger phases as long as the scalar density is small enough near the black holes.

When evaluating the WCCs (8) on the apparent horizon to produce Fig. 7, one has to be careful as we are actually dealing with different trapped surfaces. Due to the slicing condition used, each black hole has a trapped surface that during merger shrinks to the puncture, while a larger common apparent horizon forms, surrounding the previous ones [80,81]. This implies that if one is computing the WCCs (8) on the trapped surfaces collapsing to the punctures, it will result in unreasonably large values. To get around this gauge issue, we interpolate the data for the WCC of the original black hole apparent horizons just before merger with the data for the common apparent horizon just after it forms, excluding the unphysically large values right at the merger. The details of how one does the interpolation and which data points are excluded do not affect significantly the bounds  $M^2 \text{WCC}/g_2 \lesssim 50$  and  $M^2 \text{WCC}/g_3 \lesssim 20$ .

We close this subsection emphasizing that our assessment of the regime of validity of EFT is qualitative at best and, up to  $\mathcal{O}(1)$  factors, the unity value of the WCCs is a mere order of magnitude; a more detailed study is needed in order to precisely identify this regime for the cases that we have considered. The conditions (8) are only local; over time, the small effects accumulate giving rise to large deviations in some nonlocal observables such as the waveforms. In the context of complex scalar field with a Mexican hat type of potential, in [82] it is proved that, for sufficiently long times, the truncated EFT will inevitably deviate from the UV theory. Therefore, one has to be cautious when using a truncated EFT for very long times compared to the UV mass scale, even if the local weak-coupling conditions hold (see also [83]).

### D. Mismatch

In this subsection we discuss our results for the mismatch between the GR and Horndeski waveforms. We start focusing on comparing the  $G_2$  theory with  $g_2 = 0.02$  to GR, since this is an example of the limiting coupling that still satisfies the WCCs. Hence, the results for the mismatch presented should be understood as upper bounds. The mismatch depends on the coupling constants in the

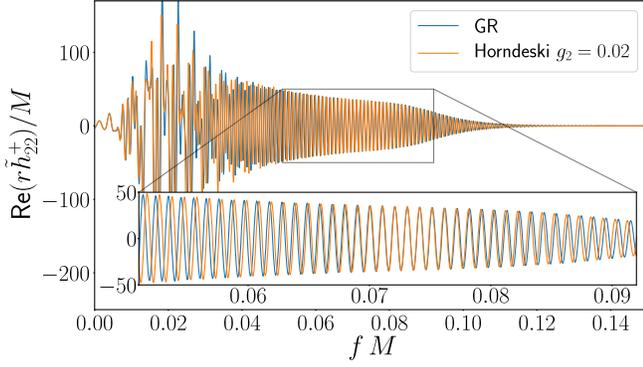


FIG. 8. Real part of  $\tilde{h}_{22}^+$ , the discrete Fourier transform of  $h_{22}^+$  for positive frequencies. Notice that for frequencies around  $f \sim 0.07 M^{-1}$ , Horndeski and GR align, but for lower and higher frequencies they separate in phase in opposite directions. This effect cannot be mitigated by a constant time or phase shifts of the waveform.

expected way, and the results are qualitatively the same for the  $G_3$  theory.

In Fig. 8 we compare the frequencies of the real part of  $\tilde{h}_{22}^+$ , the discrete Fourier transform of the  $(\ell, m) = (2, 2)$  mode of the strain, extrapolated to null infinity. Interestingly, this figure shows that in spite of both theories having approximately the same amplitudes for each frequency in the spectrum, the spectrum of the phase of the complex-valued Fourier transform differs. In range of medium frequencies, i.e.,  $f \sim 0.07 M^{-1} - 0.08 M^{-1}$ , GR and Horndeski theory agree very well. However, for both lower and higher frequencies, a significant discrepancy can be clearly seen. This effect cannot be mitigated by a constant time or phase shift of the time-domain waveform and hence we conclude that it is a physical effect. This discrepancy of both the high and low frequencies suggests that Horndeski theory exhibits both an inverse and a direct energy cascades. It would be interesting to confirm if this is indeed the case and quantify these cascades. Note that because the weak cosmic censorship holds in our scenarios, there is a natural UV cutoff for the frequencies that are accessible to external observers. As long as this cutoff is at lower energies, i.e., larger distances, than the UV cutoff of the theory, then the EFT should be valid; the fact that the WCCs hold in our case, indicates that this is indeed the case.

In Fig. 9 we quantify the mismatch for a detector setup receiving the plus polarization of the strain, extrapolated to null infinity, between GR and the Horndeski theory. We limit ourselves to the  $(\ell, m) = (2, 2)$  mode,  $h_{22}^+$ , as this is the dominant mode by an order of magnitude when compared to higher modes. We use the updated Advanced LIGO sensitivity design curve (*aLIGODesign.txt* in [62], which imposes  $f_{\min} = 5$  Hz) and flat noise ( $S_n = 1$ ) following the procedure described in Sec. II G. We compute the mismatch for black hole masses in the typical range of stellar mass black

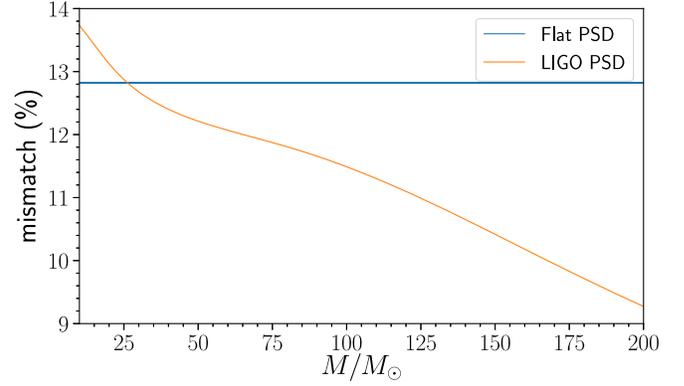


FIG. 9. Mismatch for  $h_{22}^+$  between GR and Horndeski for  $g_2 = 0.02$ , as a function of the final black hole mass (in units of solar masses,  $M_\odot$ ). As power spectral densities, we used the updated Advanced LIGO sensitivity design curve (*aLIGODesign.txt* in [62], which imposes  $f_{\min} = 5$  Hz) and a flat noise mismatches ( $S_n = 1$ ). This allows us to estimate a range for expected mismatches of 10%–13%.

holes binaries observed so far,  $M \in [10, 200] M_\odot$  [84]. As mentioned in Sec. II B, the parameter  $g_2$  is dimensionful, with units of  $M^{-2}$ , and hence this value also varies as we probe different masses in Fig. 9. This figure shows that the mismatch varies between  $\sim 13\%$  at the low mass end and  $\sim 10\%$  at the high mass end. To confirm accuracy of these results, the mismatch between two different resolutions of the same GR evolution ranges between 0.3%–0.5% for the same mass ranges. As a reference, for a signal to noise ratio of 25, similar to GW150914 [85], the minimum expected mismatch for detection is about 0.6% [86,87]. Additionally, [61] estimated that a mismatch of 3.5% would result in a 10% lower detection rate; therefore, the large mismatches obtained for big enough values of the couplings suggest that if the underlying theory of gravity was Horndeski with a massive scalar field, some events could have gone undetected if the black holes had sufficient scalar field surrounding them.

We additionally explore the degeneracy between  $G_2$  and  $G_3$  Horndeski theories considered in this paper by evaluating the mismatch between the waveforms obtained for different values of the coupling  $g_2$  and a fixed value of  $g_3 = -0.03$  respectively.<sup>17</sup> The results are shown in Fig. 10. We can see that the waveforms in the two theories appear to be nearly degenerate for  $g_2 \approx 0.04$  and  $g_3 = -0.03$ . For these values of the couplings, the mismatch is about 0.1%, significantly below any detectability threshold or numerical errors. As an additional remark, notice how the mismatch between the  $G_2$  theory and GR starts decreasing for  $g_2 \gtrsim 0.1$ . This effect happens because for these values of  $g_2$ , the delay of the Horndeski waveform corresponds to approximately one period of the GR wave in the region

<sup>17</sup>We thank the anonymous referee for suggesting this calculation.

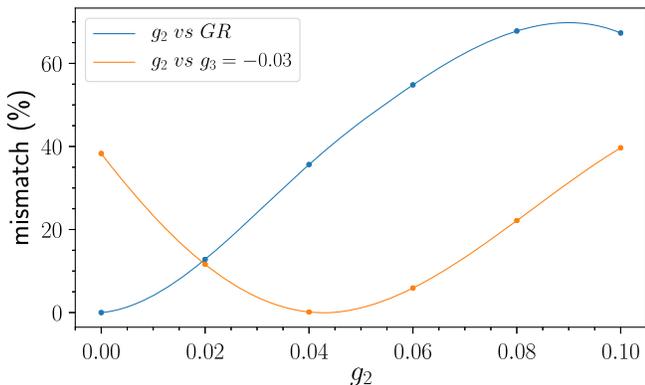


FIG. 10. Mismatch between the  $G_2$  Horndeski theory for several values of  $g_2$  and GR (blue curve) and between the same  $G_2$  theory and the  $G_3$  theory with  $g_3 = -0.03$  (yellow curve), for a flat PSD. The two Horndeski theories appear to be degenerate for certain values of the couplings.

close to the merger phase in the frequency domain, resulting in more aligned peaks and hence smaller mismatch. Hence, we expect the mismatch to oscillate as we vary  $g_2$  due to this effect, with the local minima bounded above zero since the waveforms are significantly different in spite of aligned peaks.

#### IV. SUMMARY AND CONCLUSIONS

In this article we have studied eccentric black hole binary mergers in certain cubic Horndeski theories (3)–(4) with a massive scalar field with mass parameter  $m = 0.5$ . We have chosen initial data and small enough coupling constants such that a certain local weak-coupling condition (8) is satisfied at all times during the evolution. This condition monitors the size of the Horndeski terms in the equations of motion compared to GR terms, and the fact it holds ensures that the EFTs are in their regime of validity and hence we can trust their predictions.

One of the goals of this article was to identify potential deviations from GR in some physical quantities that Horndeski theories of gravity may exhibit. We have observed that locally small deviations from GR build up over time and get enhanced whenever the system enters the strong field regime. In the case of the eccentric binaries, this happens during the successive close encounters of the black holes and in the final merger phase. Since the modifications of GR are locally small, large deviations may still arise in nonlocal observables, such as gravitational waveforms, through a cumulative build up. This cumulative effect gets reflected in the gravitational waveforms as large shifts with respect to the analogous waveforms computed in GR coupled to a massive scalar field with the same mass and angular momentum. Whilst the details, such as its sign and size, of the observed shift in the waveforms depend on the details of the theory and value of the coupling constants, the effect seems to be generic, at least within the class of

Horndeski theories that we have explored here. We conjecture that the same effect should be present in more general Horndeski theories. We have quantified the misalignment of the  $(\ell, m) = (2, 2)$  mode of the plus polarization of the strain,  $h_{22}^+$ , for one of the Horndeski theories that we have considered. We find that the spectrum differs both for low and high frequencies. Furthermore, for large enough values of the couplings, still in the regime of validity of the EFT, we find that the mismatch is around 10%–13% in the whole mass range of current detections. This is quite significant and it suggests that if the underlying theory of gravity differs from GR, some events where the black holes have sufficient scalar field surrounding them may have been and continue to go undetected. For smaller values of the couplings, the mismatch would be smaller.

The misalignment that we have observed is a cumulative effect and hence it only occurs if the black holes are surrounded by a long-lived scalar cloud. In our case this is possible because of the mass term in the scalar potential, which ensures that a nontrivial scalar energy density remains in the vicinity of the black holes for very long times, thus allowing the scalar field to interact with itself and with the geometry. We have also considered massless scalars, but in this case we do not observe any significant difference between Horndeski and GR. This is expected because the theories considered do not admit stationary hairy black holes for massless scalar fields (at least in the slowly rotating limit [70–73]) and, hence, a massless scalar field gets absorbed by the black holes or disperses to null infinity on a time scale much quicker compared to the binary lifetime. In our particular example, the scalar field is essentially completely absorbed in the second close encounter of the binary and by then there has not been enough time to build up any sizeable deviation from GR.

In this article we considered both  $G_2$  and  $G_3$  Horndeski theories and, as we have already mentioned, even though the initial and final states are the same, both lead to misaligned waveforms with respect to GR. Therefore, at least for equal mass nonspinning binaries, the degeneracy between the class of Horndeski theories that we have considered and GR is broken. However, we do not see any visible difference between the waveforms obtained in the  $G_2$  or in the  $G_3$  theories. It would be interesting to investigate if (or how) the degeneracy of the waveforms is broken in Horndeski theories of gravity. It would be interesting to extend our studies to unequal mass and spinning binaries to see if the degeneracy with GR and with the various Horndeski theories is broken when considering different mass ratios and nonzero spins.

We have considered Horndeski theories simply as toy models for EFTs with higher derivatives; in the Horndeski case, the higher derivatives are in the matter (scalar) sector and the equations of motion are of second order. However, more fundamental theories of gravity, such as string theory, predict

higher curvature corrections of the Einstein-Hilbert action. In general, such new terms in the action will result in equations of motion of order higher than two. Refs. [26,27,32] have outlined how the strong field regime of such theories may be probed, but it would be very interesting to do so in the context of a black hole binary. Our work suggests that in the weakly-coupled regime, where these theories are valid EFTs, some of the problems that may arise in general, such as loss of hyperbolicity or shock formation, can be controlled in a physical situation that probes the strong field regimes such as a black hole binary merger.

The main goal of the present paper is to identify what features in the physical observables extracted from black hole binaries in Horndeski theories can allow one to differentiate these theories from GR. Given that the corrections to GR have to be locally small in order for these theories be valid EFTs, nonlocal observables such as gravitational waveforms are particularly useful because small effects can accumulate and, for long enough times, give rise to large deviations from GR. These or other deviations from alternative theories of gravity are potentially being undetected by current gravitational wave observatories. Therefore, our results stress the importance of modelling waveforms in alternative theories of gravity treating them fully nonlinearly. It would be interesting to identify other observables where large deviations show up. In the case of waveforms, until complete waveform templates are built for alternative theories, a potential way to detect the misalignment that we have identified is the following: future space-based gravitational wave observatories such as LISA [88,89] are expected to be able to detect gravitational waves produced in stellar mass black hole binaries during earlier stages of the inspiral phase. From these waveforms one should be able to extract the parameters of the binary and, by using GR, predict the time of merger of the binary. Some binaries should enter the LIGO band in the final stages of the inspiral and merger phase, thus allowing to contrast the GR prediction for the merger time with the observation; a certain advancement or delay of the merger could be attributed to the fact that higher derivative corrections modify GR.

### ACKNOWLEDGMENTS

Part of the work was presented at the workshops “Current Challenges in Gravitational Physics, SISSA (April 2021) and “Mathematical and Numerical Aspects of Gravitation, IPAM, UCLA (October 2021). We thank the GRChombo collaboration [90] for their support and code development work. We also thank Katy Clough, Áron Kovács, Luis Lehner, Miren Radia, Josu Aurrekoetxea, and Harvey Reall for useful discussions. P.F. is supported by the European Research Council Grant No. ERC-2014-StG 639022-NewNGR, and a Royal Society University Research Fellowship Grants No. UF140319, No. RGF/AE\180260, No. URF/R\201026, and No. RFERE\210291.

T.F. is supported by a Ph.D. studentship from the Royal Society RS\PhD\181177. The simulations presented used PRACE resources under Grant No. 2020235545, PRACE DECI-17 resources under Grant No. 17DECI0017, the CSD3 cluster in Cambridge under Projects No. DP128 and No. TC011, the MareNostrum4 cluster at the Barcelona SuperComputing Centre under Grants No. FI-2020-3-0007, No. FI-2020-3-0010, and No. FI-2021-3-0010, and Athena at HPC Midlands + cluster. The Cambridge Service for Data Driven Discovery (CSD3), partially operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility [91]. The DiRAC component of CSD3 is funded by BEIS capital via STFC capital Grants No. ST/P002307/1 and No. ST/R002452/1 and STFC operations Grant No. ST/R00689X/1. DiRAC is part of the National e-Infrastructure. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. [92] for providing computing time on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre [93]. The HPC Midlands cluster was funded by EPSRC on the Grant No. EP/P020232/1, as part of the HPC Midlands + consortium.

### APPENDIX: CONVERGENCE

In this Appendix we provide some details of the convergence tests that we have carried out. As a representative example, we considered the binary in the  $G_2$  Horndeski theory with coupling constant  $g_2 = 0.02$ , and we performed three simulations with different resolutions to study convergence. Our simulations are evolved with a coarsest level resolution of  $\Delta x = \frac{16}{7}$  (medium resolution), with eight additional refinement levels and a computational domain of size  $1024^3$ . To carry out the tests, we used one lower resolution changing  $\Delta x = \frac{8}{3}$  (low resolution) and one higher resolution with  $\Delta x = 2$  (high resolution).

In Fig. 11 we show the error estimates in the quadrupole mode  $h_{22}$  extrapolated to null infinity between low, medium and high resolutions and the estimates for the expected error assuming third- and fourth-order convergence. We decompose the complex strain into its amplitude and phase,  $h_{\ell m} = h_{\ell m}^+ - ih_{\ell m}^- = h_{\ell m}^A e^{ih_{\ell m}^\phi}$ . We compute these expected errors using the continuum limit of the convergence factor of order  $n$ ,

$$Q_n = \frac{(\Delta x_{\text{Low}})^n - (\Delta x_{\text{Med}})^n}{(\Delta x_{\text{Med}})^n - (\Delta x_{\text{High}})^n}. \quad (\text{A1})$$

This indicates the convergence order of  $h_{22}$  is consistent with three.

We also tested convergence of other variables; for instance, the trajectories of the two black holes,  $x_1^i(t)$  and  $x_2^i(t)$ , shown in Fig. 1, can be used to test convergence.

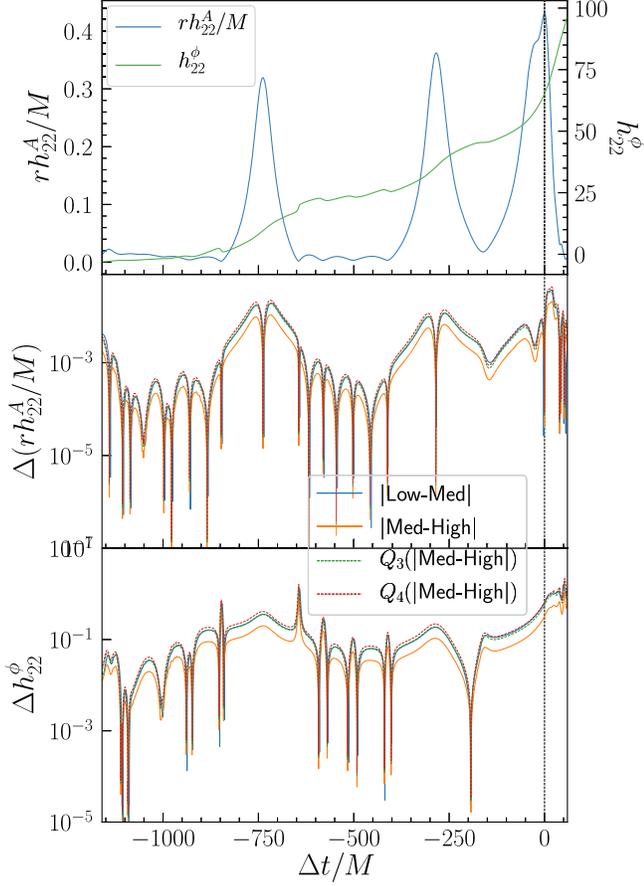


FIG. 11. Convergence test for  $g_2 = 0.02$  with different coarse resolutions: low ( $384^3$ ), medium ( $448^3$ ) and high ( $512^2$ ). Convergence performed on the amplitude and phase of the complex strain,  $h_{22}$ , extrapolated to null infinity.  $\Delta t = 0$  is the peak of the amplitude for the highest resolution. This figure indicates consistency with third-order convergence.

We rewrite these trajectories in terms of the radial distance between the black holes,

$$D(t) = |x_1^i(t) - x_2^i(t)|, \quad (\text{A2})$$

and the phase relative to the initial positions,

$$\theta(t) = \arccos \left[ \frac{(x_1^i(t) - x_2^i(t)) \cdot (x_1^i(0) - x_2^i(0))}{D(t) D(0)} \right]. \quad (\text{A3})$$

These quantities for the Horndeski theory are shown in the top panel of Fig. 12 for the same binary as in Fig. 1. The convergence analysis of these quantities across the three resolutions is shown in the middle and bottom panels of Fig. 12. These figures indicate that both quantities exhibit between third- and fourth-order convergence.

For completeness, in Fig. 13 we show the  $L^2$  norms<sup>18</sup> of the Hamiltonian and the Euclidean norm of the

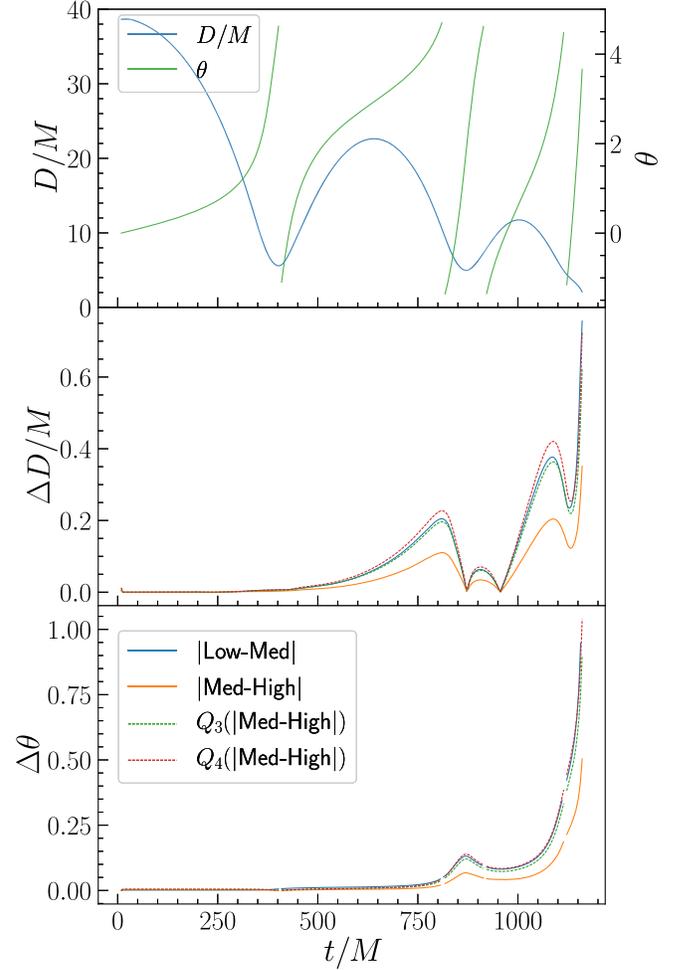


FIG. 12. *Top panel:* Radial distance  $D(t)$  and relative phase  $\theta(t)$  of the black holes' trajectories as functions of time for the  $g_2 = 0.02$  Horndeski theory. Convergence tests for the radial distance  $\Delta D(t)$  (*middle panel*) and relative phase  $\theta(t)$  (*bottom panel*). Both of these quantities exhibit between third- and fourth-order convergence.

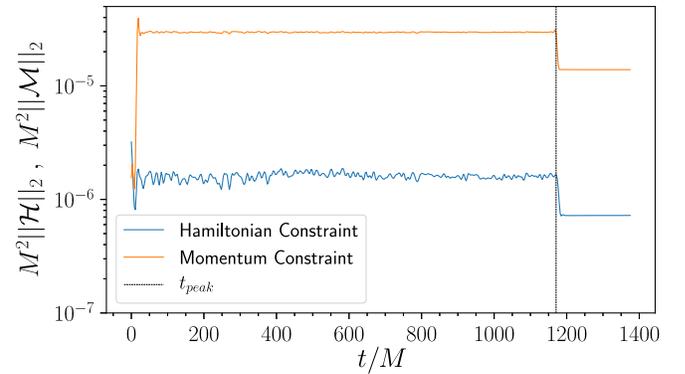


FIG. 13.  $L^2$  norm of the Hamiltonian and the Euclidean norm of the momentum constraints for the medium resolution of the  $g_2 = 0.02$  Horndeski binary.

<sup>18</sup>For a given quantity  $Q$ , we compute  $L^2 Q = \sqrt{\frac{1}{V} \int_V |Q|^2 dV}$ .

momentum constraints over the full computational domain. This figure shows the constraint violations remain stable at the level of  $10^{-6}M^{-2} - 10^{-5}M^{-2}$  respectively throughout the whole evolution, with a

significant and sudden reduction at the merger. Considering the results of our convergence analysis, we conclude our simulations are stable and in the convergent regime.

- 
- [1] N. K. Johnson-McDaniel, A. Ghosh, S. Ghonge, M. Saleem, N. V. Krishnendu, and J. A. Clark, Investigating the relation between gravitational wave tests of general relativity, *Phys. Rev. D* **105**, 044020 (2022).
- [2] M. Okounkova, L. C. Stein, M. A. Scheel, and D. A. Hemberger, Numerical binary black hole mergers in dynamical Chern-Simons gravity: Scalar field, *Phys. Rev. D* **96**, 044020 (2017).
- [3] H. Witek, L. Gualtieri, P. Pani, and T. P. Sotiriou, Black holes and binary mergers in scalar Gauss-Bonnet gravity: Scalar field dynamics, *Phys. Rev. D* **99**, 064035 (2019).
- [4] M. Okounkova, M. A. Scheel, and S. A. Teukolsky, Evolving metric perturbations in dynamical Chern-Simons gravity, *Phys. Rev. D* **99**, 044019 (2019).
- [5] M. Okounkova, L. C. Stein, M. A. Scheel, and S. A. Teukolsky, Numerical binary black hole collisions in dynamical Chern-Simons gravity, *Phys. Rev. D* **100**, 104026 (2019).
- [6] M. Okounkova, L. C. Stein, J. Moxon, M. A. Scheel, and S. A. Teukolsky, Numerical relativity simulation of GW150914 beyond general relativity, *Phys. Rev. D* **101**, 104016 (2020).
- [7] M. Okounkova, Numerical relativity simulation of GW150914 in Einstein dilaton Gauss-Bonnet gravity, *Phys. Rev. D* **102**, 084046 (2020).
- [8] M. Bezares, R. Aguilera-Miret, L. ter Haar, M. Crisostomi, C. Palenzuela, and E. Barausse, No Evidence of Kinetic Screening in Merging Binary Neutron Stars, *Phys. Rev. Lett.* **128**, 091103 (2022).
- [9] J. T. Gálvez Gherzi and L. C. Stein, Numerical renormalization-group-based approach to secular perturbation theory, *Phys. Rev. E* **104**, 034219 (2021).
- [10] J. Healy, T. Bode, R. Haas, E. Pazos, P. Laguna, D. Shoemaker, and N. Yunes, Late inspiral and merger of binary black holes in scalar-tensor theories of gravity, *Classical Quantum Gravity* **29**, 232002 (2012).
- [11] E. Barausse, C. Palenzuela, M. Ponce, and L. Lehner, Neutron-star mergers in scalar-tensor theories of gravity, *Phys. Rev. D* **87**, 081506(R) (2013).
- [12] E. W. Hirschmann, L. Lehner, S. L. Liebling, and C. Palenzuela, Black hole dynamics in Einstein-Maxwell-Dilaton theory, *Phys. Rev. D* **97**, 064032 (2018).
- [13] L. Sagunski, J. Zhang, M. C. Johnson, L. Lehner, M. Sakellariadou, S. L. Liebling, C. Palenzuela, and D. Neilsen, Neutron star mergers as a probe of modifications of general relativity with finite-range scalar forces, *Phys. Rev. D* **97**, 064016 (2018).
- [14] H. Reall, N. Tanahashi, and B. Way, Causality and Hyperbolicity of Lovelock Theories, *Classical Quantum Gravity* **31**, 205005 (2014).
- [15] J. L. Ripley and F. Pretorius, Scalarized black hole dynamics in Einstein dilaton Gauss-Bonnet gravity, *Phys. Rev. D* **101**, 044015 (2020).
- [16] J. L. Ripley and F. Pretorius, Hyperbolicity in spherical gravitational collapse in a Horndeski theory, *Phys. Rev. D* **99**, 084014 (2019).
- [17] L. Bernard, L. Lehner, and R. Luna, Challenges to global solutions in Horndeski's theory, *Phys. Rev. D* **100**, 024011 (2019).
- [18] J. L. Ripley and F. Pretorius, Gravitational collapse in Einstein dilaton-Gauss-Bonnet gravity, *Classical Quantum Gravity* **36**, 134001 (2019).
- [19] J. L. Ripley and F. Pretorius, Dynamics of a  $\mathbb{Z}_2$  symmetric EdGB gravity in spherical symmetry, *Classical Quantum Gravity* **37**, 155003 (2020).
- [20] M. Bezares, M. Crisostomi, C. Palenzuela, and E. Barausse, K-dynamics: Well-posed 1 + 1 evolutions in K-essence, *J. Cosmol. Astropart. Phys.* **03** (2021) 072.
- [21] P. Figueras and T. França, Gravitational collapse in cubic Horndeski theories, *Classical Quantum Gravity* **37**, 225009 (2020).
- [22] A. D. Kovács and H. S. Reall, Well-Posed Formulation of Scalar-Tensor Effective Field Theory, *Phys. Rev. Lett.* **124**, 221101 (2020).
- [23] A. D. Kovács and H. S. Reall, Well-posed formulation of Lovelock and Horndeski theories, *Phys. Rev. D* **101**, 124003 (2020).
- [24] W. E. East and J. L. Ripley, Evolution of Einstein-scalar-Gauss-Bonnet gravity using a modified harmonic formulation, *Phys. Rev. D* **103**, 044040 (2021).
- [25] W. E. East and J. L. Ripley, Dynamics of Spontaneous Black Hole Scalarization and Mergers in Einstein-Scalar-Gauss-Bonnet Gravity, *Phys. Rev. Lett.* **127**, 101102 (2021).
- [26] J. Cayuso, N. Ortiz, and L. Lehner, Fixing extensions to general relativity in the nonlinear regime, *Phys. Rev. D* **96**, 084043 (2017).
- [27] G. Allwright and L. Lehner, Towards the nonlinear regime in extensions to GR: Assessing possible options, *Classical Quantum Gravity* **36**, 084001 (2019).
- [28] I. Müller, Zum paradoxon der wärmeleitungstheorie, *Z. Phys.* **198**, 329 (1967).
- [29] W. Israel and J. Stewart, Thermodynamics of nonstationary and transient effects in a relativistic gas, *Phys. Lett.* **58A**, 213 (1976).
- [30] W. Israel, Nonstationary irreversible thermodynamics: A causal relativistic theory, *Ann. Phys. (N.Y.)* **100**, 310 (1976).
- [31] W. Israel and J. Stewart, Transient relativistic thermodynamics and kinetic theory, *Ann. Phys. (N.Y.)* **118**, 341 (1979).

- [32] R. Cayuso and L. Lehner, Nonlinear, noniterative treatment of EFT-motivated gravity, *Phys. Rev. D* **102**, 084008 (2020).
- [33] M. Bezares, L. ter Haar, M. Crisostomi, E. Barausse, and C. Palenzuela, Kinetic screening in nonlinear stellar oscillations and gravitational collapse, *Phys. Rev. D* **104**, 044022 (2021).
- [34] G. Lara, M. Bezares, and E. Barausse, UV completions, fixing the equations and nonlinearities in  $k$ -essence, *Phys. Rev. D* **105**, 064058 (2022).
- [35] G. W. Horndeski, Second-order scalar-tensor field equations in a four-dimensional space, *Int. J. Theor. Phys.* **10**, 363 (1974).
- [36] A. Nicolis, R. Rattazzi, and E. Trincherini, The Galileon as a local modification of gravity, *Phys. Rev. D* **79**, 064036 (2009).
- [37] A. D. Kovács, Well-posedness of cubic Horndeski theories, *Phys. Rev. D* **100**, 024005 (2019).
- [38] S. Weinberg, Effective field theory for inflation, *Phys. Rev. D* **77**, 123541 (2008).
- [39] S. Brahma and M. W. Hossain, Consistency of cubic Galileon cosmology: Model-independent bounds from background expansion and perturbative analyses, *Universe* **7**, 167 (2021).
- [40] A. Ijjas, F. Pretorius, and P. J. Steinhardt, Stability and the gauge problem in non-perturbative cosmology, *J. Cosmol. Astropart. Phys.* **01** (2019) 015.
- [41] M. Shahalam, S. K. J. Pacif, and R. Myrzakulov, Galileons, phantom fields, and the fate of the Universe, *Eur. Phys. J. C* **76**, 410 (2016).
- [42] C. Deffayet, O. Pujolas, I. Sawicki, and A. Vikman, Imperfect dark energy from kinetic gravity braiding, *J. Cosmol. Astropart. Phys.* **10** (2010) 026.
- [43] V. Faraoni and E. Gunzig, Einstein frame or Jordan frame?, *Int. J. Theor. Phys.* **38**, 217 (1999).
- [44] D. Alic, C. Bona-Casas, C. Bona, L. Rezzolla, and C. Palenzuela, Conformal and covariant formulation of the Z4 system with constraint-violation damping, *Phys. Rev. D* **85**, 064040 (2012).
- [45] D. Alic, W. Kastaun, and L. Rezzolla, Constraint damping of the conformal and covariant formulation of the Z4 system in simulations of binary neutron stars, *Phys. Rev. D* **88**, 064049 (2013).
- [46] S. Bernuzzi and D. Hilditch, Constraint violation in free evolution schemes: Comparing BSSNOK with a conformal decomposition of Z4, *Phys. Rev. D* **81**, 084003 (2010).
- [47] T. Andrade *et al.*, GrChombo: An adaptable numerical relativity code for fundamental physics, *J. Open Source Software* **6**, 3703 (2021).
- [48] F. Pretorius, Evolution of Binary Black Hole Spacetimes, *Phys. Rev. Lett.* **95**, 121101 (2005).
- [49] T. Helfer, U. Sperhake, R. Croft, M. Radia, B.-X. Ge, and E. A. Lim, Malaise and remedy of binary boson-star initial data, *Classical Quantum Gravity* **39**, 074001 (2022).
- [50] R. Penrose, Gravitational collapse: The role of general relativity, *Riv. Nuovo Cimento* **1**, 252 (1969).
- [51] D. Christodoulou, On the global initial value problem and the issue of singularities, *Classical Quantum Gravity* **16**, A23 (1999).
- [52] J. G. Baker, M. Campanelli, and C. O. Lousto, The Lazarus project: A pragmatic approach to binary black hole evolutions, *Phys. Rev. D* **65**, 044001 (2002).
- [53] M. Alcubierre, Introduction to 3 + 1 Numerical Relativity (Oxford University Press, New York, 2008).
- [54] M. Radia, U. Sperhake, A. Drew, K. Clough, E. A. Lim, J. L. Ripley, J. C. Aurrekoetxea, T. França, and T. Helfer, Lessons for adaptive mesh refinement in numerical relativity, [arXiv:2112.10567](https://arxiv.org/abs/2112.10567).
- [55] C. Reisswig and D. Pollney, Notes on the integration of numerical relativity waveforms, *Classical Quantum Gravity* **28**, 195015 (2011).
- [56] D. J. A. McKechnan, C. Robinson, and B. S. Sathyaprakash, A tapering window for time-domain templates and simulated signals in the detection of gravitational waves from coalescing compact binaries, *Classical Quantum Gravity* **27**, 084020 (2010).
- [57] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, Surrogate model of hybridized numerical relativity binary black hole waveforms, *Phys. Rev. D* **99**, 064045 (2019).
- [58] U. Sperhake, B. Bruegmann, D. Mueller, and C. F. Sopuerta, Black-hole binaries go to eleven orbits, [arXiv:1012.3173](https://arxiv.org/abs/1012.3173).
- [59] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith, A surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary black hole mergers, *Phys. Rev. D* **95**, 104023 (2017).
- [60] B. J. Owen, Search templates for gravitational waves from inspiraling binaries: Choice of template spacing, *Phys. Rev. D* **53**, 6749 (1996).
- [61] L. Lindblom, B. J. Owen, and D. A. Brown, Model waveform accuracy standards for gravitational wave data analysis, *Phys. Rev. D* **78**, 124020 (2008).
- [62] L. Barsotti, P. Fritschel, M. Evans, and S. Gras, Updated Advanced LIGO sensitivity design curve, Technical Report No. T1800044-v5, The LIGO Scientific Collaboration, 2018.
- [63] K. Clough, P. Figueras, H. Finkel, M. Kunesch, E. A. Lim, and S. Tunyasuvunakool, GRChombo: Numerical relativity with adaptive mesh refinement, *Classical Quantum Gravity* **32**, 245011 (2015).
- [64] T. Andrade *et al.*, GRChombo: An adaptable numerical relativity code for fundamental physics, *J. Open Source Software* **6**, 3703 (2021).
- [65] M. Adams, P. Colella, D. T. Graves, J. N. Johnson, N. D. Keen, T. J. Ligocki, D. F. Martin, P. W. McCorquodale, D. Modiano, P. O. Schwartz, T. D. Sternberg, and B. Van Straalen, Chombo Software Package for AMR Applications—Design Document, Lawrence Berkeley National Laboratory Technical Report No. LBNL-6616E.
- [66] C. Bona, J. Masso, E. Seidel, and J. Stela, A New Formalism for Numerical Relativity, *Phys. Rev. Lett.* **75**, 600 (1995).
- [67] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Accurate Evolutions of Orbiting Black-Hole Binaries Without Excision, *Phys. Rev. Lett.* **96**, 111101 (2006).
- [68] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, Gravitational Wave Extraction from an Inspiral Configuration of Merging Black Holes, *Phys. Rev. Lett.* **96**, 111102 (2006).

- [69] S. Husa, J. A. Gonzalez, M. Hannam, B. Bruegmann, and U. Sperhake, Reducing phase error in long numerical binary black hole evolutions with sixth order finite differencing, *Classical Quantum Gravity* **25**, 105006 (2008).
- [70] L. Hui and A. Nicolis, No-Hair Theorem for the Galileon, *Phys. Rev. Lett.* **110**, 241104 (2013).
- [71] T. P. Sotiriou and S.-Y. Zhou, Black Hole Hair in Generalized Scalar-Tensor Gravity, *Phys. Rev. Lett.* **112**, 251102 (2014).
- [72] T. P. Sotiriou and S.-Y. Zhou, Black hole hair in generalized scalar-tensor gravity: An explicit example, *Phys. Rev. D* **90**, 124063 (2014).
- [73] A. Maselli, H. O. Silva, M. Minamitsuji, and E. Berti, Slowly rotating black hole solutions in Horndeski gravity, *Phys. Rev. D* **92**, 104049 (2015).
- [74] W. H. Press and S. A. Teukolsky, Floating orbits, super-radiant scattering and the black-hole bomb, *Nature (London)* **238**, 211 (1972).
- [75] J. Barranco, A. Bernal, J. C. Degollado, A. Diez-Tejedor, M. Megevand, M. Alcubierre, D. Nunez, and O. Sarbach, Schwarzschild Black Holes Can Wear Scalar Wigs, *Phys. Rev. Lett.* **109**, 081102 (2012).
- [76] J. Barranco, A. Bernal, J. C. Degollado, A. Diez-Tejedor, M. Megevand, D. Nunez, and O. Sarbach, Self-gravitating black hole scalar wigs, *Phys. Rev. D* **96**, 024049 (2017).
- [77] L. Hui, D. Kabat, X. Li, L. Santoni, and S. S. C. Wong, Black hole hair from scalar dark matter, *J. Cosmol. Astropart. Phys.* **06** (2019) 038.
- [78] K. Clough, T. Dietrich, and J. C. Niemeyer, Axion star collisions with black holes and neutron stars in full 3D numerical relativity, *Phys. Rev. D* **98**, 083020 (2018).
- [79] T. Ikeda, L. Bernard, V. Cardoso, and M. Zilhão, Black hole binaries and light fields: Gravitational molecules, *Phys. Rev. D* **103**, 024020 (2021).
- [80] D. Pook-Kolb, O. Birnholtz, B. Krishnan, and E. Schnetter, Self-intersecting marginally outer trapped surfaces, *Phys. Rev. D* **100**, 084044 (2019).
- [81] J. Thornburg, Event and apparent horizon finders for 3 + 1 numerical relativity, *Living Rev. Relativity* **10**, 3 (2007).
- [82] H. S. Reall and C. M. Warnick, Effective field theory and classical equations of motion, [arXiv:2105.12028](https://arxiv.org/abs/2105.12028).
- [83] A.-C. Davis and S. Melville, Scalar fields near compact objects: Resummation versus UV completion, *J. Cosmol. Astropart. Phys.* **11** (2021) 012.
- [84] R. Abbott *et al.* (LIGO Scientific, Virgo, KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, [arXiv:2111.03606](https://arxiv.org/abs/2111.03606).
- [85] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [86] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, Constructing gravitational waves from generic spin-precessing compact binary inspirals, *Phys. Rev. D* **95**, 104004 (2017).
- [87] M. Pürrer and C.-J. Haster, Gravitational waveform accuracy requirements for future ground-based detectors, *Phys. Rev. Research* **2**, 023151 (2020).
- [88] P. Amaro-Seoane *et al.* (LISA Collaboration), Laser interferometer space antenna, [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).
- [89] K. G. Arun *et al.*, New horizons for fundamental physics with LISA, [arXiv:2205.01597](https://arxiv.org/abs/2205.01597).
- [90] [www.grchombo.org](http://www.grchombo.org).
- [91] [www.dirac.ac.uk](http://www.dirac.ac.uk).
- [92] [www.gauss-centre.eu](http://www.gauss-centre.eu).
- [93] [www.lrz.de](http://www.lrz.de).